Technical Guidance for HIV/AIDS Surveillance Programs

Volume I: Policies and Procedures

HIV Incidence and Case Surveillance Branch
Atlanta, Georgia

All material contained in this document is in the public domain and may be used and reprinted without permission; citation of the source is, however, appreciated.

Suggested Citation

Centers for Disease Control and Prevention and Council of State and Territorial Epidemiologists. *Technical Guidance for HIV/AIDS Surveillance Programs, Volume I: Policies and Procedures*. Atlanta, Georgia: Centers for Disease Control and Prevention; 2005.

The document is available at https://team.cdc.gov/team/cdc/dispatch.cgi/hicsb_TechGuid/folderFrame/100001/0/def/5186.

Contents

Introduction to Policies and Procedures 1-3

Background and Purpose 1-3
Practices and Standards 1-4
Surveillance Using Document-based Data Management 1-5
Organization of Technical Guidance for HIV/AIDS Surveillance Programs 1-5
CSTE Position Statements 1-7

Access to Source Data and Completeness of Reporting 2-3

Structural Requirements 2-3
Definition of Reportable Events and Information 2-3
Process Standards. 2-4
Case Finding and Information Collection 2-4
Outcome Standards 2-7
Bibliography 2-8
Surveillance. 2-8
Capture-recapture Analyses 2-8
Appendixes 2-10
Codes for Report Sources. 2-11
Examples for Calculating Outcome Standards 2-13

Risk Factor Ascertainment 3-3

HIV/AIDS Risk Factor Ascertainment 3-4
Definitions 3-4
Cases of Public Heath Importance (COPHI) 3-5
COPHI Outcome Standard 3-6
COPHI Process Standards 3-6
Investigation and Follow-up. 3-6
NRR Risk Factors 3-7
Pediatric Risk Factors 3-7
HIV-2 and Variant Strain Surveillance. 3-7
Unusual Manifestations. 3-8
Risk Factor Ascertainment Procedures 3-8
Outcome Standard 3-8
Process Standards 3-8
Initial Case Reports. 3-8

Epidemiologic Follow-Up 3-9
Background 3-9
Education and Training of Surveillance Staff 3-13
Background 3-13
Education and Training of Health Care Providers 3-15
Background 3-15
Sampling for Ascertainment of HIV Risk Factors 3-18
Sampling: Advantages and Disadvantages 3-18
Implementation of Sampling 3-19
Feasibility of Sampling within a State or Local Surveillance Program 3-19
Appendixes 3-21
Appendix A 3-23
Appendix B 3-38
Appendix C 3-42
Appendix D 3-69
Appendix E 3-85
Appendix F 3-92
Appendix G 3-95

Death Ascertainment 4-3
Introduction 4-3
Sources of Death Information 4-3
Data from Health Care Providers 4-3
Information from Death Certificates 4-3
State Vital Statistics Records 4-3
National Databases with Death Information 4-3
Data to Be Collected from Death Certificate Records (DCRs) 4-4
Structural Requirements 4-5
Process Standards 4-5
Linkage of HIV Case Records with Death Certificate Records (DCRs) 4-5
Linked Cases 4-5
Nonlinked Cases 4-6
Nonlinked Death Certificate Records 4-6
DCR Not Linked, but Mentions HIV/AIDS 4-7
No Record Linkage to HIV/AIDS Case or Mention of HIV/AIDS on DCR 4-8
Frequency of Linkage of HIV/AIDS Case Records to Death Case Records 4-8
First Time Record Linkage 4-8
Routine Record Linkage 4-8
Computer Log 4-8

Data Exchange 4-9
Exchange of Data between HIV/AIDS Surveillance Programs (Out of State) . . . 4-9
Exchange of Data between Offices of Vital Records (Out of State) 4-9
Handling of Out-of-jurisdiction (OOJ) Case Records 4-9
Outcome Measures 4-10
Proportion of Death-Certificate-Only (DCO) Cases. 4-10
Proportion of Cases for Which the Initial Source of Report Was a DCR. 4-10
Effectiveness of the HIV-Specific Record Linkage in Finding Cases 4-10
Delay between Death and Report of Death. 4-10
Yield from Searching the NDI 4-11
Appendix 4-15
Searching National Death Databases 4-15
Technical Details of Searching the DMF 4-15
National Death Index (NDI) 4-16

Electronic Reporting 5-3
Introduction 5-3
Purpose of Electronic Reporting 5-3
Electronic Case Reporting (ECR) 5-3
Electronic Laboratory Reporting (ELR) 5-4
Considerations for Electronic Reporting 5-6
Policy. 5-6
Security and Confidentiality 5-6
Record Linkage 5-7
Existing ELR Resources for Implementing HIV/AIDS Reporting 5-7
Background 5-7
NEDSS (National Electronic Disease Surveillance System) 5-8
PHIN (Public Health Information Network) 5-8
ELR Requirements 5-9
HL7 (Health Level 7). 5-9
SNOMED (Systematized Nomenclature of Medicine) 5-9
LOINC (Logical Observation Identifiers Names and Codes) 5-9
Electronic Case Reporting 5-10
Review Reporting Laws. 5-10
Develop a System for Reporting 5-11
Identify All Reporting Facilities and Providers 5-11
Documentation for Reporting Facilities and Providers 5-11
Alternative Reporting Methods 5-12
Data Processing of Electronic Case Reports 5-13
Reporting Back to Reporting Facilities and Providers 5-13
Electronic Laboratory Reporting 5-13

Review Reporting Laws 5-14
Develop a Mechanism for Electronic Reporting 5-16
Identify All Reporting Laboratories 5-16
Documentation for Reporting Laboratory 5-17
Alternative Reporting Methods 5-19
Optional Reports 5-19
Resolving Referral Laboratory Problems 5-19
Managing Increased Number of Lab Reports 5-19
Laboratories First Reporting Electronically 5-20
Reporting Back to Laboratories 5-20
HIV and HIV-associated Laboratory Tests 5-21
Diagnosing HIV Infection 5-21
Other Tests Indicative of HIV Infection 5-21
HIV Disease Progression and Markers for Access to Care 5-22
Changes in Laboratory Reporting Initiated by ELR 5-23
Performance Standards 5-23
Case Follow-up from a Lab Report 5-24
The Ideal Complete Lab Report 5-25
Measurable Performance Standards 5-32
References 5-34
Appendix A Electronic Laboratory Reporting Resources 5-36
Appendix B Documentation for Reporting Facilities and Providers 5-39
Appendix C Identifying All Reporting Laboratories 5-51
Appendix D Documentation for Reporting Laboratory 5-56
Appendix E Reporting Back to Laboratories 5-63
Appendix F Changes in Laboratory Reporting Initiated by Electronic Laboratory Reporting (ELR) 5-75
Appendix G Case Follow-up from a Lab Report 5-82

Case Residency 6-3
Case Residency Assignment 6-3
HIV Diagnosis Date 6-3
Structural Requirements to Establish HIV Diagnosis Date 6-3
Process Standards 6-4
HIV Diagnosis Date and AIDS Diagnosis Date Relationship 6-5
Anonymous Test Results 6-5
Patient Self-Reported Date of Prior Positive HIV Test 6-5
Physician-Documented Diagnosis Date 6-5
Laboratory Tests 6-6
Pediatric Cases 6-6

Determination of Case Residency 6-6
Structural Requirements to Determine Case Residency 6-6
Process Standards 6-7
Residency Assignment of HIV Cases in States with Newly Implemented HIV Reporting 6-9
Appendix A. 6-11
Appendix B. 6-12
Appendix C 6-14

Duplicate Review 7-3
Intrastate 7-3
Structural Requirements 7-3
Process Standards 7-4
Outcome Measures 7-4
Interstate 7-4
Routine Interstate Duplicate Review (RIDR) 7-4
Structural Requirements 7-5
Process Standards 7-5
Outcome Measures 7-7
Appendix A. 7-8
Appendix B. 7-9

Data Management 8-3
Data Management 8-3
Structural Requirements 8-3
Process Standards 8-3
Data Entry. 8-3
Record Linkage 8-4
Rules for Managing Multiple Documents per Case 8-4
Edits 8-4
Processing Follow-ups, Corrections, and Deletions 8-4
Documentation 8-5

Record Linkage 9-3
Introduction 9-3
Definition of Record Linkage 9-3
Purposes of Record Linkage 9-4
Structural Requirements and Organizational Barriers to Record Linkage 9-4
Memorandum of Understanding 9-5
Human Subjects Approval 9-5

Information Technology Resources 9-5
Data Files That May Be Useful for Record Linkage to HIV/AIDS Cases 9-5
Identification of HIV-specific Data Files 9-7
Variables of Interest When Updating Existing Cases by Record Linkage 9-9
Variables on Which to Search for Agreement 9-10
Fixing Problems in Data Files before Record Linkage 9-13
Record Linkage Strategies 9-16
Considerations for Selecting a Computer Software Program for Record Linkage . . . 9-23
Visual Review of Indeterminate Records 9-25
Tracking Log 9-26
Measures or Standards for Record Linkage 9-27
Appendixes 9-29
Appendix A 9-30
Appendix B 9-32
Appendix C 9-34
Appendix D 9-37

Data Quality 10-3
Introduction 10-3
Structural Requirements 10-3
Procedure Manuals, Coding Manuals, and Other Documentation 10-3
Edits and Data Processing 10-3
Process Standards 10-4
Adherence to Data Standards 10-4
Training 10-4
Quality Control Activities 10-4
Re-abstraction Studies 10-5
Eligibility Criteria and Study Population 10-5
Sampling Strategy 10-6
Frequency of Re-abstraction Studies 10-6
Data Items Included in Re-abstraction Studies 10-6
Data Collection 10-6
Analysis Plan and Use of Results 10-7
Case Finding 10-7
Outcome Measures 10-8
Appendixes 10-9
Appendix A Sampling Methodology for Re-Abstraction 10-10
Appendix B Taking a Simple Random Sample (SRS) Using the Random Number Table 10-17
Appendix C Desired Number of Strata 10-18

Appendix D Point Estimates and Sample Size Calculation for Simple Random Sampling (SRS) 10-20
Appendix E Point Estimates and Sample Size Calculation for Stratified Random Sampling 10-22
Appendix F Examples of Sampling for Re-abstraction Studies 10-27
Appendix G Variables for Re-Abstraction 10-37
Appendix H Re-abstraction Error Report Form, for Pairs of Reports for One Case 10-38
Appendix I Re-abstraction Report Table, for Diagnosis Year 20XX 10-40

Pediatric HIV/AIDS Surveillance 11-3

Background 11-3
Introduction 11-3
Diagnosing HIV Infection in Children <13 Years of Age 11-4
Case Definition 11-5
The CDC Pediatric Case Report Form 11-5
Structural Requirements 11-6
Data Collection for Children <13 Years of Age 11-6
Structural Requirements 11-6
Data Collection for HIV Exposure Reporting 11-8
Structural Requirements 11-8
Process Standards for Identification and Follow-up of Pediatric HIV/AIDS Cases . . . 11-8
Pediatric HIV/AIDS Case Reporting (Children <13 Years of Age) 11-9
Pediatric HIV Case Finding (Children <13 Years of Age) 11-9
Information Collection and Follow-up 11-9
Process Standards for Data Collection for Children <13 Years of Age 11-10
Pediatric Cases of Public Health Importance (CoPHI) 11-11
Perinatal Exposure Reporting for Infants <18 Months of Age (recommended) . . . 11-11
Exposure Case Finding for Infants <18 Months of Age 11-11
Information Collection 11-11
Process Standards for Data Collection for HIV Exposure Reporting 11-12
Outcome Measures 11-12
Data Collection for Children <13 Years of Age 11-12
Data Collection for HIV Exposure Reporting 11-13
References 11-14

Policies and Procedures for HIV Incidence Surveillance 12-5

Background 12-5
Introduction 12-5
Structural Requirements 12-6
Policies and Procedures 12-6
The BED Assay for STARHS 12-7

Testing and ARV Use History Data 12-8
Staffing Needs 12-8
Process Standards 12-10
Obtaining Testing and ARV Use History Data 12-11
Obtaining Remnant HIV-Positive Specimens for Testing Using STARHS 12-12
Determining the Disposition of Specimens and Communicating with the Public Health and CDC STARHS Laboratories. 12-13
Entering Data into Surveillance Databases 12-14
Creating the HIS Data Set and Transferring Data to CDC 12-15
Ensuring Security and Confidentiality 12-16
Analyzing and Disseminating Data 12-16
Training Staff 12-17
Outcome Standards 12-17
References 12-18
Epidemiologic Flow Chart for HIV Incidence Surveillance 12-19
Operational Public Health Laboratory Flow Chart for HIV Incidence Surveillance . . . 12-20
National Center for HIV, STD, and TB Prevention's Non-research Determination for HIV Incidence Surveillance 12-21
Standard HIV Incidence Surveillance Data Elements. 12-23
Guidance for the Remnant HIV-Positive Specimen Transportation Activities for HIV Incidence Surveillance 12-26
Purpose 12-26
Introduction. 12-26
Laboratory Types 12-26
Laboratory Types. 12-27
Specimen Information. 12-27
Type of Specimens Shipped to CDC STARHS Laboratory 12-27
Specimen Volume 12-27
Sample Storage 12-27
Specimen Numbering 12-28
Specimen Retention 12-28
Packaging and Shipping Procedures 12-28
Shipping Guidance 12-28
Frequency of Shipments 12-28
Shipping Couriers 12-28
Coordinating Shipments 12-29
Additional Information for Commercial Laboratories Only. 12-29
Tracking Shipments 12-29
Additional Information Related to APHL-Contracted Commercial Laboratories Only 12-29
Sample Rejection Criteria 12-30

Confidentiality and HIPAA Regulations 12-30
Funding for Specimen Handling. 12-30
Specimen Transport Options 12-31
Option A: Specimen Originated at Private Laboratory and Is Sent Directly to the CDC STARHS Laboratory 12-31
Transportation Overview 12-31
Procedures for Specimens Sent Directly from a Private Laboratory to the CDC STARHS Laboratory 12-31
Roles of Parties Involved 12-33
Role of Private Laboratories 12-33
Role of CDC STARHS Laboratory. 12-33
Role of the State/Local HIV Incidence Surveillance Coordinator. 12-34
Specimen Numbering 12-34
Theoretical Laboratory Types for This Transportation Model 12-34
Option B: Specimen Originated at or Sent Via State / Local Public Health Laboratory 12-34
Transportation Overview 12-34
Procedures for Specimens Sent from a Private Laboratory through a State Public Health Laboratory to the CDC STARHS Laboratory 12-35
Roles of Parties Involved 12-36
Role of Private Laboratories 12-36
Role of Public Health Laboratories 12-37
Role of State/Local HIV Incidence Surveillance Coordinator. 12-37
Role of CDC STARHS Laboratory. 12-37
Specimen Numbering. 12-38
Theoretical Laboratory Types for this Transportation Model 12-38
Responsibilities 12-38
Private Laboratories 12-38
Selection of a Model Type 12-38
Additional Laboratory Report Information. 12-38
Public Health Laboratories 12-39
Sample Storage and Retention 12-39
Aliquoting and Sample Shipment 12-39
State/Local HIV Incidence Surveillance Coordinator. 12-39
Sample Disposition. 12-39
Data Management 12-39
CDC STARHS Laboratory. 12-39
Sample Rejection Criteria. 12-39
Sample Storage and Retention 12-39
Results Reporting. 12-40
References. 12-40
Guidance for Processing, Storage, and Shipping of Specimens to the CDC STARHS Laboratory 12-44
Purpose 12-44

Introduction 12-44
CDC STARHS Laboratory 12-44
Setting and Personnel for Specimen Processing 12-44
Materials 12-45
Specimen Collection and Processing 12-45
Shipping 12-46
Packing Procedures for Shipping to the CDC STARHS Laboratory 12-46
HICSB Incidence Surveillance STARHS Specimen Submission Form 12-49
Training and Certification for Shipping Infectious Substances 12-50
Appendix A 12-50

Policies and Procedures for Variant, Atypical, and Resistant HIV Surveillance (VARHS) 13-5
Background 13-5
Objectives 13-6
Methods 13-7
Public Health Benefit 13-7
Structural Requirements 13-8
Policies and Procedures 13-8
Laboratory Procedures 13-10
Staffing Needs 13-11
Process Standards 13-13
Determining Eligibility 13-14
Inclusion criteria 13-14
Exclusion criteria 13-14
Identifying Laboratories and Obtaining Remnant Specimens or Sequence Data 13-15
Specimen Handling and Transport Procedures 13-16
Tube type for HIV diagnostic specimens 13-16
Specimen handling 13-16
Serum and plasma processing 13-16
Cryovials and aliquots 13-17
Diagnostic and clinical testing and freezing of serum or plasma aliquots 13-17
Specimen tracking data elements to be recorded in VARHS specimen log or database 13-18
Shipment to the VARHS genotyping laboratory 13-18
Required Data Elements 13-19
Data used to determine VARHS eligibility 13-19
Demographic and clinical data 13-20
Specimen tracking data 13-20
Laboratory data 13-20
Returning Genotyping Results 13-21

Ensuring Security and Confidentiality 13-22
Data Management. 13-23
CDC software 13-23
Identification of VARHS specimens 13-23
Data encryption and transfer 13-24
Data Analysis and Data Dissemination 13-25
Outcome Standards 13-25
References. 13-26
Appendix A. 13-29
Appendix B. 13-30
Appendix C 13-31
Appendix D 13-32
Appendix E. 13-33
Appendix F. 13-45

Data Analysis and Dissemination 14-3
Structural Requirements 14-3
Minimum performance standards 14-3
Sufficient capacity for analyzing HIV/AIDS surveillance data 14-3
Outcome Standards 14-4
Process Standards. 14-4
Statistical report. 14-5
Race/ethnicity 14-5
Age 14-6
Sex 14-6
HIV transmission category 14-6
Geographic area. 14-7
Recommended analyses for annual or more frequent reports 14-7
Suggested additional analyses 14-9
Linking HIV/AIDS surveillance data with other data sources 14-10
Appendixes 14-11
Appendix A Terms, Definitions, and Calculations Relevant to HIV/AIDS Surveillance Data 14-12
Appendix B Techniques for Analyzing and Displaying HIV/AIDS Surveillance Data 14-17
Appendix C Integrated HIV/AIDS Epidemiologic Profile 14-22
Appendix D Tools and Resources for Analyses 14-23
References. 14-24

Notes

Contributors

This document, *Technical Guidance for HIV/AIDS Surveillance Programs*, was developed by the HIV Incidence and Case Surveillance Branch of the Division of HIV/AIDS Prevention, National Center for HIV, STD, and TB Prevention, Centers for Disease Control and Prevention in collaboration with the Council of State and Territorial Epidemiologists. The CDC/CSTE Advisory Committee provided oversight and leadership throughout the entire process. Workgroup contributors consisted of state and local health department representatives. Irene Hall, CDC, and Eve Mokotoff, CSTE, led the development.

Members of the CDC/CSTE Advisory Committee

CDC: Pamela Gruduah, Irene Hall, Martha Miller

CSTE: Gordon Bunch, California; Dena Ellison, Virginia; Jim Kent, Washington; Eve Mokotoff, Michigan; Stanley See, Texas

Chairs of Workgroups, CDC

Michael Campsmith, Data Analysis and Dissemination

Sam Costa, Security and Confidentiality

Irene Hall, Data Quality

Laurie Kamimoto, Electronic Reporting

Lata Kumar, Data Dictionary

Martha Miller, Overview

Kathleen McDavid, HIV Risk Factor Ascertainment

Ruby Phelps, Case Residency Assignment

Richard Selik, Death Ascertainment

Richard Selik, Record Linkage

Suzanne Whitmore, Perinatal and Pediatric Case Surveillance

CDC Contributors

Lori Armstrong, Mi Chen, Betsey Dunaway, John Gerstle, Kate Glynn, Irene Hall, Felicia Hardnett, David Hurst, Jennie Johnston, Danielle Kahn, Tebitha Kajese, Laurie Kamimoto, Kevin Lyday, Martha Miller, Andy Mitsch, Michelle Pan, Richard Selik, Amanda Smith, Damien Suggs, Patricia Sweeney, Kimberly Todd, Will Wheeler, Suzanne Whitmore, Irum Zaidi.

State and Local Health Department Contributors and Reviewers

Alabama: Anthony Merriweather, Danna Strickland; ***California-Los Angeles***: Gordon Bunch, Mi Suk Harlan, Virginia Hu, Ann Nakamura; ***California-San Francisco***: Ling Hsu, Maree Kay Parisi, Sandra Schwarcz; ***District of Columbia***: Gail Hansen, Kompan Ngamsnga; ***Florida***: Becky Grigg, Lorene Maddox; ***Illinois-Chicago***: Margarita Reina; ***Indiana***: Jerry Burkman; ***Iowa***: Randy Mayer; ***Louisiana***: Joseph Foxhood, Greg Gaines, William Robinson, Debbie Wendell, Amy Zapata; ***Massachusetts***: Maria Regina Barros; ***Michigan***: Elizabeth Hamilton, Nilsa Mack, Eve Mokotoff, Yolande Moore;

Minnesota: Luisa Pessoa-Brandao, Tracy Sides; ***New Hampshire***: Chris Adamski; ***New Jersey***: Wogayehu Afework, Linda Dimasi, Abdel Ibrahim, John Ryan; ***New York City***: Melissa Pfeiffer, Judy Sackoff; ***New York State***: Alexa Bontempo, Kathleen Brousseau, Donna Glebatis; ***Ohio***: Sandhya Ramachandran; ***Oklahoma***: Mark Turner; ***Pennsylvania***: Bonnie Krampe, Ming Wei; ***South Carolina***: Dana Giurgiutiu; ***Texas***: Thomas Barnabas, Dianna Highberg, Roy Reyna, Stanley See, Jan Veenstra; ***Virginia***: Dena Ellison; ***Washington***: Maria Courogen; ***Washington-Seattle & King County***: Amy Bauer, Jim Kent; ***Wisconsin***: Loujean Steenberg.

Technical Guidance for HIV/AIDS Surveillance Programs

Introduction to Policies and Procedures

HIV Incidence and Case Surveillance Branch
Atlanta, Georgia

Contents — Introduction

Background and Purpose 1-3
Practices and Standards 1-4
Surveillance Using Document-based Data Management 1-5
Organization of Technical Guidance for HIV/AIDS Surveillance Programs 1-5
CSTE Position Statements 1-7

Technical Guidance for HIV/AIDS Surveillance Programs — Introduction to Policies and Procedures

Technical Guidance for HIV/AIDS Surveillance Programs, prepared by the Centers for Disease Control and Prevention (CDC) in collaboration with the Council of State and Territorial Epidemiologists (CSTE), serves as a guide for managing state and local HIV/AIDS surveillance programs in the United States. The intent of the policies and procedures is to provide the basis for maintaining a national HIV/AIDS surveillance system by using a standardized framework for collecting complete, timely, and high-quality data. At the federal level, the primary functions of an HIV/AIDS case surveillance system are 1) to provide accurate epidemiologic data to monitor the incidence and prevalence of HIV infection and AIDS-related morbidity and mortality and 2) to use these data trends to assist in public health planning and policy.

CDC is authorized to provide federal funding to states and territories through surveillance cooperative agreements to achieve the goals of the national HIV/AIDS surveillance program and also to assist states in developing their own surveillance programs in accordance with state and local laws and practices. The HIV Incidence and Case Surveillance Branch (HICSB) of the Divisions of HIV/AIDS Prevention, National Center for HIV, STD, and TB Prevention, is responsible for national HIV/AIDS surveillance. In addition to financial assistance, HICSB also provides technical assistance to funded areas to ensure that HIV/AIDS reporting systems are comprehensive, timely, accurate, and up-to-date as the epidemic continues to evolve. *Technical Guidance for HIV/AIDS Surveillance Programs*, which supersedes the 1996 *Guidelines for HIV/AIDS Surveillance*, is part of a continuing effort to assist states and territories.

Background and Purpose

AIDS reporting was initiated at the national level in 1981 to help monitor the scope and impact of the HIV/AIDS epidemic. After the Food and Drug Administration (FDA) approved the first HIV antibody test in 1985, several states expanded their AIDS surveillance systems to include surveillance for HIV infection.

Associated with widespread use of highly active antiretroviral therapies (HAART) that began in the mid-1990s, the progression of HIV infection to AIDS has dramatically slowed among treated persons, and an increasing number of persons with HIV infection are living longer and staying healthier. Although AIDS surveillance remains important because it provides information about populations for whom treatment is not accessed or has not succeeded, reliance on AIDS surveillance alone does not adequately describe the burden of HIV disease in the population or current trends in the epidemic. Currently, all 50 states, the District of Columbia, the Commonwealth of Puerto Rico, and the U.S. Territories conduct HIV reporting. Although confidential name-based HIV reporting has been implemented in most areas, as of January 2005, nine areas (7 states and 2 metropolitan areas) use coded identifiers, and 5 states have name-to-code reporting (i.e.,

cases are reported by names that are deleted after a certain amount of time, and only codes are retained). An integrated HIV/AIDS surveillance system, with confidential name-based reporting for both HIV and AIDS, allows for better collection of multiple key events during the case history of a person infected with HIV. In addition, a national, integrated HIV/AIDS surveillance system is better able to monitor the evolving epidemic and to provide useful data about HIV-infected populations to enhance local, state, and federal efforts to prevent HIV transmission, improve allocation of resources for treatment services, and assist in evaluating the impact of public health interventions (see *Integrating HIV and AIDS Surveillance: A Resource Manual for Surveillance Coordinators, 1998* and *MMWR, CDC Guidelines for National HIV Case Surveillance, Including Monitoring for HIV Infection and AIDS, 1999*).

For HIV/AIDS data to be comparable and valuable on a national level, all participating project areas need to collect data with a high level of accuracy and consistency. The purposes of these policies and procedures are to address the importance of maintaining a standardized framework for data collection across all surveillance project areas, to delineate the required components of an effective surveillance system, and to suggest methods and techniques designed to optimize productivity. This document reflects CDC and CSTE current guidance for standards and best practices.

Practices and Standards

CDC and CSTE continue to recommend that all states and territories require reporting of the earliest diagnosis of HIV infection (exclusive of anonymous tests), the earliest diagnosis of AIDS in persons of all ages, deaths among persons with HIV infection, and all cases of perinatal HIV exposure in infants. Recommended practices represent guidance for best public health practices based on scientific data. Because no single set of policies and procedures can address all of the diversity among, and local needs of, individual state and local surveillance systems, state and local programs should develop their own policies and procedures in accordance with *Technical Guidance for HIV/AIDS Surveillance Programs*, but tailored to their specific situations.

To achieve the goal of uniformity in data quality from the various surveillance systems, *Technical Guidance for HIV/AIDS Surveillance Programs* includes structural requirements (what a program must, should, or may HAVE to operate an HIV/AIDS surveillance system), process standards (what a program must, should, or may DO to achieve the objectives), and outcome standards (measurable objectives).

Certain process and outcome standards are listed as requirements; these are activities and measurable objectives that, when met, certify a fully functioning surveillance system with high data quality. Other processes or measures are recommendations of best practices that will help to meet the outcome standards.

Surveillance Using Document-based Data Management

To monitor the HIV/AIDS epidemic, surveillance staff need to collect accurate information on key events during the time a person is first diagnosed with HIV infection until death. This information is collected from many sources and documents, such as health care providers (adult and pediatric case report forms), laboratories (laboratory reports), and vital records offices (birth and death certificates)—see Figure 1. Increased use of diagnostic laboratory testing and the growth of electronic reporting have led to an increased number of documents. Historically, consolidated information from all documents has been used to create a single case record for each HIV-infected person. As technology and the epidemic have evolved, CDC has developed a document-based data management system to better track information received from HIV/AIDS surveillance and to better monitor HIV disease progression.

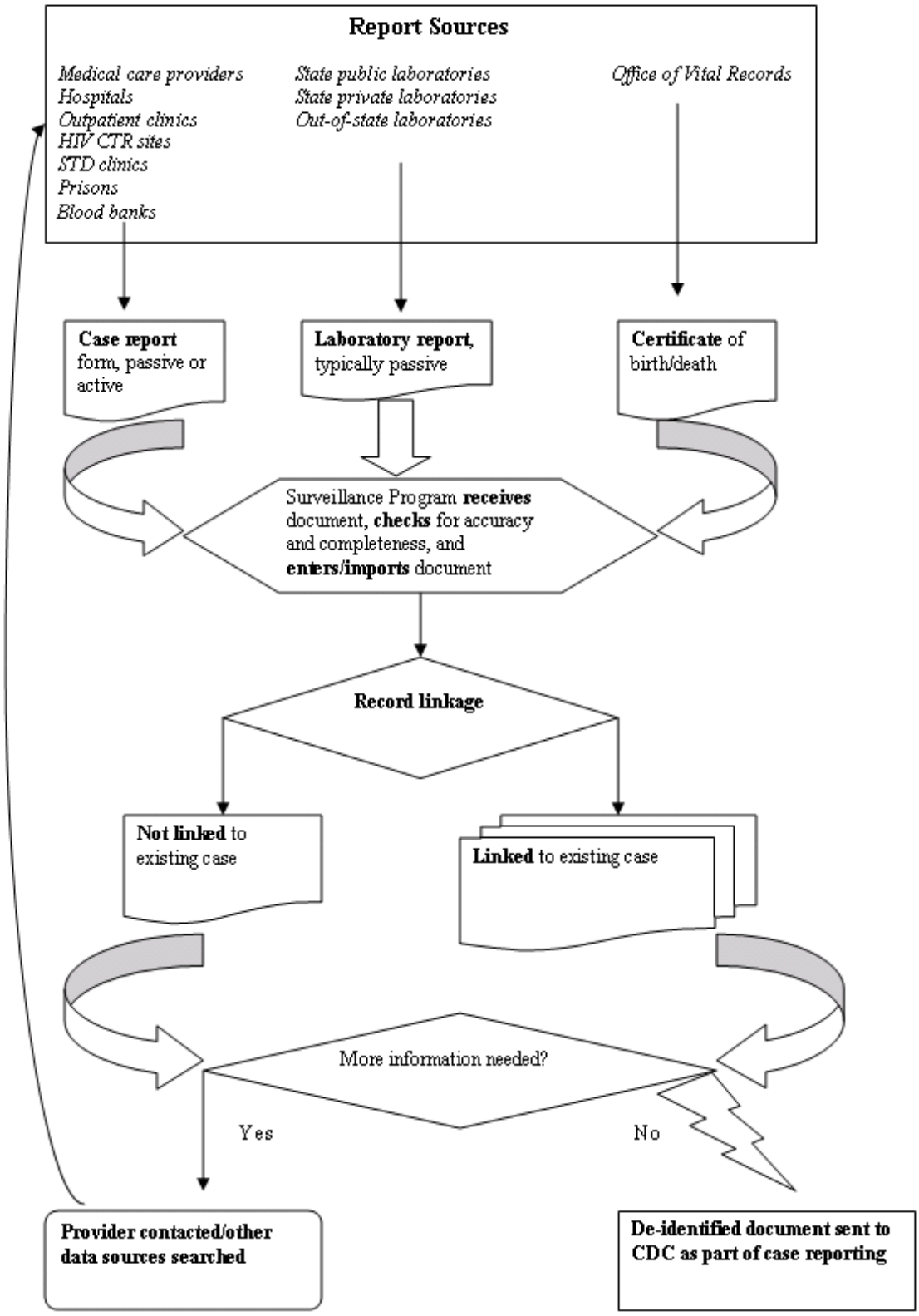
Unlike case-based data management, document-based data management allows all documents to be stored and retained electronically in their original formats. Staff can refer not only to case report records, but also to original source documents. In document-based data management, a predetermined algorithm selects the most appropriate values from incoming documents to create and update case records for analysis while still allowing for review of the overall case information. This approach to data processing requires a shift from storing data in flat rectangular data files to storing data in relational databases. A relational database is one in which data are stored in several tables that are linked by specific key fields or variables, such as a case identifier. By linking multiple documents from multiple sources to each patient within one database, the process of collecting, managing, analyzing, and disseminating data becomes less cumbersome to maintain and requires fewer resources.

This approach requires a software system and programs that support document-based data management by linkage and retention of individual documents. Document-based surveillance requires that all documents received by the surveillance program that contain core surveillance information on HIV/AIDS cases be retained. The tasks and standards for processing these documents are described in other sections.

Organization of *Technical Guidance for HIV/AIDS Surveillance Programs*

Technical Guidance for HIV/AIDS Surveillance Programs is divided into multiple main topic areas. Within each main topic area, standards are subdivided into structural requirements, process standards, and outcome measures. The format of the current surveillance policies and procedures allows for flexibility and ease of responsiveness as changes in methodology and a growing understanding of HIV disease continue to evolve.

Figure 1. Document-based data management



CSTE Position Statements

Year	Topic	Author	Title
2004	Infectious Disease	Gibson, James	Laboratory reporting of clinical test results indicative of HIV infection: new standards for a new era of surveillance and prevention
ID	**Web Address**		
04-ID-07	http://www.cste.org/ps/2004pdf/04-ID-07-final.pdf		
Year	**Topic**	**Author**	**Title**
2004	Infectious Disease	Gibson, James	Support for "Guiding Principles for HIV Prevention"
ID	**Web Address**		
04-ID-10	http://www.cste.org/ps/2004pdf/04-ID-10-final.pdf		
Year	**Topic**	**Author**	**Title**
2004	Infectious Disease	Mokotoff, Eve	Development of population-based HIV/AIDS clinical surveillance
ID	**Web Address**		
04-ID-05	http://www.cste.org/ps/2004pdf/04-ID-05-final.pdf		
Year	**Topic**	**Author**	**Title**
2003	Infectious Disease	Gibson, James J.	The impact of new technologies and therapies on HIV/AIDS surveillance: surveillance of antiretroviral resistance
ID	**Web Address**		
03-ID-09	http://www.cste.org/PS/2003pdfs/2003finalpdf/03-ID-09Revised1.pdf		
Year	**Topic**	**Author**	**Title**
2002	Infectious Disease	Birkhead, Guthrie	Surveillance for perinatal HIV exposure
ID	**Web Address**		
02-ID-04	http://www.cste.org/position statements/02-ID-04.pdf		
Year	**Topic**	**Author**	**Title**
2002	Infectious Disease	Birkhead, Guthrie	Surveillance for HIV Incidence
ID	**Web Address**		
02-ID-07	http://www.cste.org/position statements/02-ID-07.pdf		
Year	**Topic**	**Author**	**Title**
2001	Infectious Disease	Birkhead, Gus MD, MPH	Reciprocal (Inter-state) Notification of HIV cases
ID	**Web Address**		
01-ID-04	http://www.cste.org/ps/2001/2001-id-04.htm		

Year	Topic	Author	Title
2001	Infectious Disease	Birkhead, Guthrie MPH, MD	Improved laboratory surveillance for HIV
ID	**Web Address**		
01-ID-03	http://www.cste.org/ps/2001/2001-id-03.htm		
Year	**Topic**	**Author**	**Title**
1999	Infectious Disease	Birkhead, Guthrie S.	Revised Surveillance Case Definition of HIV Infection
ID	**Web Address**		
99-ID-14	http://www.cste.org/ps/1999/1999-id-14.htm		
Year	**Topic**	**Author**	**Title**
1999	Infectious Disease	Mokotoff, Eve	Funding for HIV/AIDS Surveillance
ID	**Web Address**		
99-ID-12	http://www.cste.org/ps/1999/1999-id-12.htm		
Year	**Topic**	**Author**	**Title**
1998	Infectious Disease	(MacDonald) Moore, Kristine	Definition for Case Surveillance of HIV Infection (including AIDS)
ID	**Web Address**		
98-ID-01	http://www.cste.org/ps/1998/1998-id-01.htm		
Year	**Topic**	**Author**	**Title**
1995	Infectious Disease	Fleming David W.	Pediatric HIV Infection: Addition to the National Notifiable Diseases Surveillance System (NNDSS)
ID	**Web Address**		
95-ID-06	http://www.cste.org/ps/1995/1995-06.htm		
Year	**Topic**	**Author**	**Title**
1995	Infectious Disease	Gensheimer, Kathleen	Reporting HIV and TB Comorbidity
ID	**Web Address**		
95-ID-02	http://www.cste.org/ps/1995/1995-02.htm		
Year	**Topic**	**Author**	**Title**
1994	Infectious Disease	Fleming David W.	HIV Home Collection Kits
ID	**Web Address**		
94-ID-08	http://www.cste.org/ps/1994/1994-08.htm		

Year	Topic	Author	Title
1993	Infectious Disease	MacDonald, Kristine	Pediatric HIV Infection Reporting
ID	**Web Address**		
93-ID-03	http://www.cste.org/ps/1993/1993-03.htm		
Year	**Topic**	**Author**	**Title**
1992	Infectious Disease	Davis, Jeffrey P.	Support for Continued Evaluation of HIV/STD Prevention Program
ID	**Web Address**		
92-ID-04	http://www.cste.org/ps/1992/1992-04.htm		
Year	**Topic**	**Author**	**Title**
1991	Infectious Disease	Foster, Laurence	Surveillance of HIV Infection and Disease
ID	**Web Address**		
91-ID-02	http://www.cste.org/ps/1991/1991-02.htm		
Year	**Topic**	**Author**	**Title**
1990	Infectious Disease	Hadler, Jim	TB-HIV Surveillance and Control
ID	**Web Address**		
90-ID-13	http://www.cste.org/ps/1990/1990-13.htm		
Year	**Topic**	**Author**	**Title**
1989	Infectious Disease	Foster, Larry / Davis, Jeffrey P.	HIV infection reporting
ID	**Web Address**		
89-ID-02	http://www.cste.org/ps/1989/1989-02.htm		
Year	**Topic**	**Author**	**Title**
1988	Infectious Disease	Thompson, Ed	Terminology for HIV infection versus HIV exposure
ID	**Web Address**		
88-ID-01	http://www.cste.org/ps/1988/1988-01.htm		
Year	**Topic**	**Author**	**Title**
1987	Infectious Disease	Davis, Jeffrey P.	Revised case definition for the surveillance of AIDS
ID	**Web Address**		
87-ID-09	http://www.cste.org/PS/1987/1987-09.htm		

Year	Topic	Author	Title
1987	Infectious Disease	Hadler, Jim	Amendment to the revised case definition for the surveillance of AIDS
ID	**Web Address**		
87-ID-10	http://www.cste.org/PS/1987/1987-10.htm		
Year	**Topic**	**Author**	**Title**
1987	Infectious Disease	Hadler, Jim	Tuberculosis and AIDS reporting
ID	**Web Address**		
87-ID-11	http://www.cste.org/PS/1987/1987-11.htm		
Year	**Topic**	**Author**	**Title**
1987	Infectious Disease	Wilder, Michael	AIDS incidence and survival statistics
ID	**Web Address**		
87-ID-06	http://www.cste.org/PS/1987/1987-06.htm		
Year	**Topic**	**Author**	**Title**
1986	Infectious Disease	Chin	National reporting of HIV infections other than AIDS cases meeting the CDC surveillance definition
ID	**Web Address**		
86-ID-01	http://www.cste.org/ps/1986/1986-01.htm		
Year	**Topic**	**Author**	**Title**
1986	Infectious Disease	Chin	Nomenclature of the etiologic retro virus of AIDS
ID	**Web Address**		
86-ID-02	http://www.cste.org/ps/1986/1986-02.htm		
Year	**Topic**	**Author**	**Title**
1986	Infectious Disease	Chin	National indirect fluorescent antibody (IFA) testing capabilities for supplemental HIV antibody testing
ID	**Web Address**		
86-ID-03	http://www.cste.org/ps/1986/1986-03.htm		
Year	**Topic**	**Author**	**Title**
1986	Infectious Disease	Hall	AIDS case reporting: reciprocal notification
ID	**Web Address**		
86-ID-17	http://www.cste.org/ps/1986/1986-17.htm		

Year	Topic	Author	Title
1986	Infectious Disease	Parker	AIDS grants from CDC
ID	**Web Address**		
86-ID-04	http://www.cste.org/ps/1986/1986-04.htm		

Notes

Technical Guidance for HIV/AIDS Surveillance Programs

Access to Source Data and Completeness of Reporting

HIV Incidence and Case Surveillance Branch
Atlanta, Georgia

Contents — Access to Source Data and Completeness of Reporting

Structural Requirements 2-3
Definition of Reportable Events and Information 2-3
Process Standards 2-4
Case Finding and Information Collection 2-4
Outcome Standards 2-7
Bibliography 2-8
Surveillance 2-8
Capture-recapture Analyses 2-8
Appendix A 2-11
Appendix B 2-13

Technical Guidance for HIV/AIDS Surveillance Programs — Access to Source Data and Completeness of Reporting

The primary functions of HIV/AIDS surveillance programs are to monitor the epidemic and provide data for prevention efforts. The tasks to achieve this include soliciting, receiving, reviewing, and filing HIV/AIDS case reports in a timely manner; determining if case reports meet the HIV infection or AIDS case definitions, or both; ensuring the collection of all required information (e.g., through follow-up investigations); and maintaining a complete and accurate HIV/AIDS surveillance database.

The prerequisites (structural requirements), best practices (process standards), and outcome standards for general case finding and data collection are described next, followed by more in-depth guidance on specific topics. CSTE position statements regarding HIV/AIDS data collection can be found in CSTE Position Statements.

Structural Requirements

Permanent, current, and widely distributed documentation of all aspects of HIV/AIDS surveillance activities is essential in establishing standardization, maintaining continuity of meaning, documenting changes over time, and developing training programs. Local project areas should prepare their own policy and procedures manuals tailored to their specific situations. The following information should be included in the background and description of the local surveillance program:

- mission statement
- organizational chart
- description of geographic area
- list of reporting sources
- list of key contacts
- staff training and continuing education
- job descriptions and duties for all staff
- description of hardware/software
- state/local legislation and regulations
- security and confidentiality procedures (see Security and Confidentiality guidelines and Model State Public Health Privacy Act)

Definition of Reportable Events and Information

HIV infection may be detected at various points along the spectrum of disease, and reportable events range from reporting of HIV infection in otherwise asymptomatic persons to death (Figure 2.1). At a minimum, information that defines a case is reportable, as well as demographic and other information that allow for follow-up of case reports and

the determination of the burden of disease in the population. Reportable information (required, recommended, optional) is described in the Instructions for Completing the Data Collection Form; additional data items may be collected by local areas and may vary by state. The 2000 case definition for HIV infection is described in the December 1999 *MMWR*.

Process Standards

Case Finding and Information Collection

Surveillance for HIV/AIDS cases involves the following:

- identifying reporting sources,
- establishing communication with personnel who provide services to persons with HIV/AIDS or who compile related information,
- educating providers about the importance of reporting to the HIV/AIDS surveillance system, and
- reviewing medical records to ensure more complete case finding and better ascertainment of risk factor information.

New cases of HIV infection are found by both passive and active surveillance. The definitions of active and passive surveillance are

- Active surveillance: Health department staff regularly contact reporting facilities (hospitals, clinics, physician offices, laboratories) to identify potential/suspect HIV/AIDS cases (or confirm no cases) (Teutsch SM, Churchill RE (eds.). Principles and practice of public health surveillance, 2nd edition. New York: Oxford University Press; 2000). Health department staff review medical records at provider sites or receive information over the telephone, by fax, e-mail, US mail, etc. to establish an HIV/AIDS case and to elicit information for HIV/AIDS case report forms.
- Passive surveillance: The health department receives HIV/AIDS case reports from physicians, laboratories, or other individuals or institutions without regularly contacting the reporting sources.

In either active or passive surveillance, case reports are usually mandated by state law. In many instances, case information may be contained in documents obtained by both active and passive surveillance. For example, the health department may receive an electronic laboratory report through passive reporting. However, for complete case information, health department staff may contact or travel to the diagnostic or treatment facilities to obtain the additional data items not provided on laboratory reports. Instructions for data collection can be found in the Adult HIV/AIDS Confidential Case Report, Instructions for Completion and the Pediatric Confidential Case Report Form, Instructions for Completion and in other specific sections of this chapter.

Identifying Report Sources—Each surveillance program should establish mechanisms for identifying reporting sources and obtaining case reports. Sources of information on HIV/AIDS patients include hospitals and hospital-based physicians, physicians in nonhospital practices, public and private clinics, record systems (e.g., TB registries, death certificates), laboratories, counseling and testing sites at community-based organizations, and other sources. Surveillance programs need to develop lists of all potential sources of information on HIV/AIDS patients within their jurisdictions. State licensing agencies may be able to supply a list of providers and laboratories, and information on providers of diagnostic and treatment services can be collected from documents reviewed at facilities or received at the health department. Potential sources of information and codes are listed in Appendix A. Once report sources are identified, they can be sorted in order of priority for visits or contacts based on their HIV/AIDS case load or availability of information about the cases.

Educating Providers—Surveillance programs should implement outreach activities to all reporting sources of HIV/AIDS cases within their jurisdictions. Such activities may include visits to health care facilities, telephone contacts, listserves, or mailings. Potential providers of case information should be educated regarding reporting requirements, including laws and regulations, specific data elements, document types (e.g., case report forms, laboratory reports), report medium (e.g., paper forms or electronic reporting), and timeliness of reporting. Report sources must identify a liaison(s) or contact person(s) who will be responsible for case reporting (e.g., infection control officer in a hospital). Instruction materials, such as the Adult HIV/AIDS Confidential Case Report, Instructions for Completion and the Pediatric Confidential Case Report Form, Instructions for Completion or information on how to elicit risk factors (see Risk Factor Ascertainment), can be mailed to providers, posted on the health department Web site, and distributed at professional meetings.

Priorities for outreach activities can be set to maximize the case yield for the expenditure of staff time. Depending on the size of the surveillance program and the reporting sources involved, local programs should decide their own priorities for case finding. Some factors to consider when setting priorities for contact or visit of facilities include HIV/AIDS patient case load, availability of information (e.g., information on risk factors), and prior reporting history (e.g., providers who reliably report complete and timely case information may need only quarterly feedback while nonreporting providers would need more frequent contact).

Collecting Information—Information abstracted from medical records (e.g., inpatient or outpatient records) is reported on the HIV/AIDS Confidential Case Report forms; depending on the patient's age, either the pediatric PCRF or the adult ACRF case report form is used. Instructions on how to complete the case report form are found in Instructions for Completing the Data Collection Form. In document-based surveillance, recording the sources of information is an essential part of data collection. Therefore, a separate case report form is used for information collected from separate sources (see sources in Appendix A). Information may also arrive at the health department in documents used by other agencies, such as death certificates, birth certificates, or laboratory reports, and may be received on paper forms or by electronic reporting (see

Electronic Reporting). See Data Management for instructions on the management of paper and electronic documents, record linkage, and documentation of procedures. Instructions on how to safeguard the information collected can be found in Security and Confidentiality.

The surveillance program should conduct follow-up investigations of cases with insufficient information (e.g., missing risk factor information, see also Risk Factor Ascertainment) or potential cases with insufficient information to establish an HIV diagnosis. Follow-up is conducted by contacting health care providers or by reviewing medical records at the source of the original report or elsewhere.

Identifying Missed Cases—Indicators of problems of missed cases may be anecdotal reports of cases not being reported, timeliness assessments that show cases reported more than 6 months after the diagnosis date, cases found by matching to death certificate or other databases, and assessments of completeness of reporting from particular providers. Education of providers should be geared to those who fail to report and those for whom reporting is not timely (see previous Educating Providers). Examples of report cards for providers can be found in Electronic Reporting.

Monitoring Completeness and Timeliness of Reporting—The completeness and timeliness of reporting should be monitored on a regular basis for both individual providers and the surveillance program as a whole. Assessments of provider reporting should occur at least quarterly (but may be as frequent as monthly and depend on the size of the HIV patient population) to determine whether the expected number of cases have been reported. The expected number of cases may be based on historical reporting or estimates from visits to the facility (see examples in Electronic Reporting and Appendix B). Such assessments of provider reporting will quickly reveal potential problems with reporting (e.g., staff change at a laboratory has resulted in failure to send monthly case reports to the program). Providers should also receive feedback on their reporting, including completeness, timeliness, and error rates for critical variables.

Annually, the program assesses the completeness and timeliness of case ascertainment as specified by the outcome standards (see next section). Completeness of HIV/AIDS case ascertainment can be calculated by comparing the number of cases diagnosed and reported to the surveillance program for a diagnosis year with the number of cases expected to be diagnosed during that year. Capture-recapture methods are used to estimate the expected number of cases, provided that document-based surveillance has been correctly implemented. In capture-recapture methods, the overlap of reporting of cases from different sources is determined, as well as the number of cases reported solely by each source. From this, the total number of cases that were not reported by any source is estimated. The total number of cases expected to be diagnosed in the population is the sum of the number of cases that were not reported from the capture-recapture analyses and the number of cases reported (see Bibliography for in-depth descriptions of capture-recapture analyses). Programs to conduct capture-recapture analyses should be programmed into the software system or may be obtained from CDC (**Note**: small case counts will likely preclude such analyses for individual providers). If information from document-based surveillance is not available, case-finding audits will need to be conducted to determine reporting completeness (see Data Quality).

Outcome Standards

Outcome standards, such as completeness of case ascertainment and timeliness of reporting, are used to assess the quality of HIV/AIDS data collected and reported to the national HIV/AIDS surveillance system. Results of these evaluations should be used to improve the system.

Completeness can be assessed for cases that meet the case definition (including cases that may have some demographic information missing) and for cases that have all information needed to include them in the national database. Checks on case ascertainment include

- Case-finding audits (Data Quality: Re-abstraction Studies)
- Capture-recapture analyses (prerequisite: correct document-based data management)
- Regular assessments of reported vs. expected number of cases received from a reporting source, e.g., electronic lab reporting (determine frequency at least quarterly), with feedback to the reporting source (see Electronic Reporting).

Completeness of case ascertainment for a diagnosis year is measured at 12 months after the diagnosis year, comparing the number of cases diagnosed and reported to the surveillance system for a given year to the number of cases expected to be diagnosed during that year. The standards for completeness are

- Minimum performance standard: ≥85% of expected number of cases for a diagnosis year are reported by 12 months after the diagnosis year.
- Target performance standard: ≥95% of expected number of cases for a diagnosis year are reported by 12 months after the diagnosis year.

Timeliness should be assessed in conjunction with completeness, as the number of expected cases that have not been reported at a given time point add to the proportion of cases that are not reported in a timely manner. Timeliness can be measured in two ways, (1) as the time from diagnosis to report to surveillance program and (2) as the number of expected cases reported at the time completeness is assessed. Timeliness of reporting for a diagnosis year is assessed at 12 months after the diagnosis year.

The timeliness standards are described next. (1) Using the time from diagnosis (diagnosis date) to report to surveillance program (report date-date first document was received at the health department, or if not available, date first document was entered into the surveillance system), the minimum performance standard is ≥66% of cases for a diagnosis year are reported within 6 months of diagnosis, assessed at ≥85% completeness 12 months after the diagnosis year. (2) Timeliness assessed as part of the completeness standard:

- Minimum performance standard: ≥85% of expected number of cases for a diagnosis year are reported by 12 months after the diagnosis year.
- Target performance standard: ≥95% of expected number of cases for a diagnosis year are reported by 12 months after the diagnosis year.

Timeliness can also be assessed according to legislative mandates or for specific providers, such as laboratories. Timeliness calculations for selected sources can suggest which sites may benefit from visits by surveillance staff. Examples are provided in Appendix B.

Bibliography

Surveillance

Brookmeyer R, Stroup DF. *Monitoring the health of populations: statistical principles and methods for public health surveillance*. New York: Oxford University Press; 2004.

Teutsch SM, Churchill RE. *Principles and practice of public health surveillance*, 2nd edition. New York: Oxford University Press; 2000.

Capture-recapture Analyses

Cormack RM. Loglinear models for capture-recapture. *Biometrics* 1989;45: 395–413.

Hook EB, Regal RR. Capture-recapture methods in epidemiology: methods and limitations. *Epidemiol Rev* 1995;17:243–264.

Hook EB, Regal RR. The value of capture-recapture methods for apparent exhaustive surveys. The need for adjustment for source of ascertainment intersection in attempted complete prevalence studies. *Am J Epidemiol* 1992;125:1060–1067.

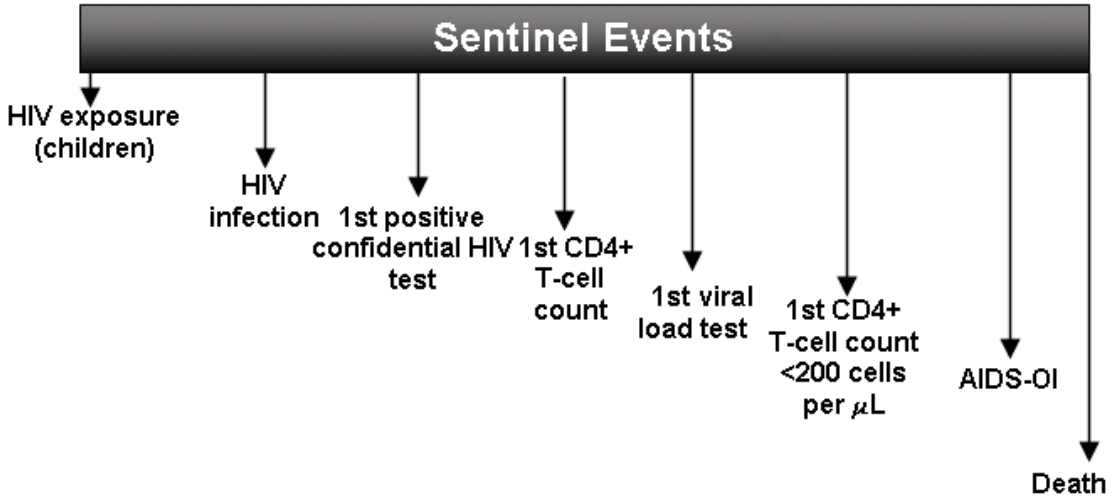
International Working Group for Disease Monitoring and Forecasting. Capture-recapture and multiple-record systems estimation. I. History and theoretical development. *Am J Epidemiol* 1995;142:1047–1058.

Papoz L, Balkau B, Lellouch J. Case counting in epidemiology: limitations of methods based on multiple data sources. *Int J Epidemiol* 1996;25:474–478.

Stephen C. Capture-recapture methods in epidemiologic studies. *Infect Control Hosp Epidemiol* 1996;17:262–266.

Figure 2.1

HIV/AIDS Surveillance
Monitoring the Spectrum of Infection



Appendixes

Appendix A Codes for Report Sources

Appendix B Examples for Calculating Outcome Standards

Appendix A

Codes for Report Sources

First level source <Source 1>	Second (more detailed) level source <Source 2>
A01.= Inpatient	A01.01 = IP/Acute care facility A01.01.02 = IP/ACF/OBGYN records A01.01.03 = IP/ACF/Pediatric records A01.01.04 = IP/ACF/Birth records A01.02 = IP/VA A01.03 = IP/Military hospital A01.03.02 = IP/Military/OBGYN records A01.03.03 = IP/Military/Pediatric records A01.04 = IP/Long-term care facility A01.04.03 = IP/LTCF/Drug TX program A01.05 = IP/Hospice
A02.= Outpatient	A02.01 = OP/HMO A02.02 =OP/VA A02.03 = OP/Private physician A02.04 = OP/Adult HIV clinic A02.05 = OP/Infect. Dis. clinic A02.06 = OP/County HD clinic A02.07 = OP/Maternal HIV clinic A02.08 = OP/Prenatal clinic or records A02.09 = OP/Pediatric HIV clinic A02.10 = OP/OBGYN clinic (not HIV related) A02.11 = OP/Pediatric clinic A02.12 = OP/TB clinic A02.14 = OP/IHS clinic A02.15 = OP/Early intervention nurse A02.16 = OP/Visiting nurse service A02.17 = OP/Hemophilia TX clinic A02.18 = OP/Hospice A02.19 = OP/Drug TX center A02.20 = OP/Rehab center A02.25 = OP/Other clinic
A03. = Emergency room	A03 = Emergency room

First level source **<Source 1>**	**Second (more detailed) level source** **<Source 2>**
A04. = Screening, diagnosis, and referral agencies	A04.01 = Scr, Dx, Ref/Blood bank A04.02 = Scr, Dx, Ref/Drug TX program A04.03 = Scr, Dx, Ref/Family planning clinic A04.04 = Scr, Dx, Ref/HIV case management agency A04.05 = Scr, Dx, Ref/HIV counseling & testing site A04.06 = Scr, Dx, Ref/Immigration report A04.07 = Scr, Dx, Ref/Insurance report A04.08 = Scr, Dx, Ref/Job corps A04.09 = Scr, Dx, Ref/Military A04.10 = Scr, Dx, Ref/Partner referral & counseling service A04.11 = Scr, Dx, Ref/STD clinic
A05. = Laboratory	A05.01 = Lab/Hosp. A05.02 = Lab/State A05.03 = Lab/Private
A06. = Other databases	A06.01 = Other DB/ADAP A06.02 = Other DB/ASD A06.03 = Other DB/Birth certificate A06.04 = Other DB/Birth defects registry A06.05 = Other DB/Cancer registry A06.06 = Other DB/Database from coroner A06.07 = Other DB/Death certificate review A06.08 = Other DB/EHRAP database A06.09 = Other DB/EPS database A06.10 = Other DB/HARS database A06.11 = Other DB/Health department records A06.12 = Other DB/Hepatitis registry A06.13 = Other DB/Hosp billing summary or discharge data A06.14 = Other DB/HRSA HIV Care database A06.15 = Other DB/Immunization registry A06.16 = Other DB/Medicaid records A06.17 = Other DB/NDI A06.18 = Other DB/Out-of-state report A06.19 = Other DB/Prison, jail, or other correctional facility database A06.20 = Other DB/PSD A06.21 = Other DB/State disease registry A06.22 = Other DB/SHAS A06.23 = Other DB/SHDC database A06.24 = Other DB/STD registry A06.25 = Other DB/TB registry A06.50 = Other DB/Other database or report
A07. = Other facility records	A07.01 = Other facility records/Prison, jail, or other correctional facility A07.02 = Other facility records/Coroner, not associated with IP facility
A10 = Other source	A10 = Other source (specify)_____
Unknown	

Appendix B
Examples for Calculating Outcome Standards

Timeliness of Case Reporting

Step 1: Calculate completeness of reporting at 12 months after the diagnosis year. If completeness is ≥85%, then go to Step 2.

Step 2: Calculate time (number of months) from diagnosis to report:

(report date) - (diagnosis date) OR

[(year of report)*12) + month] - [((year of diagnosis)*12) + month]

For example, report date is May 2004 and diagnosis date is November 2003. The time interval (in months) is

[(2004*12) + 5] - [(2003*12) + 11] = 6 months

Step 3: Determine the number of cases with a time to report ≤6 months.

Step 4: Calculate timeliness of case reporting:

$$\frac{\textit{Number of cases diagnosed within a year and reported within 6 months of diagnosis}}{\textit{Number of cases diagnosed and reported for that diagnosis year}}$$

Note: At less than 100% completeness, this measure overestimates timeliness from diagnosis to reporting.

Timeliness of Laboratory Reporting

When timeliness of reporting is assessed from a specific source, for example, mandated laboratory reporting within a specific time frame, completeness is not taken into account. An example of such a calculation would be (use tests reported from specific source):

$$\frac{\textit{Number of (reportable HIV-related) tests received within (3) days of test date during specified time period}}{\textit{All tests reported during same period}}$$

Completeness of Case Ascertainment

Completeness of case ascertainment for the most recent diagnosis year is calculated at 12 months after the diagnosis year, and expressed as a percentage ($p \times 100$). P is estimated by:

$$\frac{\textit{Number of cases diagnosed and reported for a diagnosis year}}{\textit{Number of cases expected to be diagnosed in that diagnosis year}}$$

Where the expected number of cases diagnosed during that diagnosis year is derived from capture-recapture analyses using the surveillance software, or request programs from CDC (see Bibliography for references on capture-recapture methods).

Technical Guidance for HIV/AIDS Surveillance Programs

Risk Factor Ascertainment

HIV Incidence and Case Surveillance Branch
Atlanta, Georgia

Contents — Risk Factor Ascertainment

HIV/AIDS Risk Factor Ascertainment 3-4
Definitions 3-4
Cases of Public Heath Importance (COPHI) 3-5
COPHI Outcome Standard 3-6
COPHI Process Standards 3-6
Investigation and Follow-up 3-6
NRR Risk Factors 3-7
Pediatric Risk Factors 3-7
HIV-2 and Variant Strain Surveillance 3-7
Unusual Manifestations 3-8
Risk Factor Ascertainment Procedures 3-8
Outcome Standard 3-8
Process Standards 3-8
Initial Case Reports 3-8
Epidemiologic Follow-Up 3-9
Education and Training of Surveillance Staff 3-13
Education and Training of Health Care Providers 3-15
Sampling for Ascertainment of HIV Risk Factors 3-18
Sampling: Advantages and Disadvantages 3-18
Implementation of Sampling 3-19
Appendix A Terminology, classification, and history of the collective terms used by CDC for risk factors for HIV infection and for their summary classification categories 3-23
Appendix B Prototype Forms for Collecting Risk Factor Data 3-38
Appendix C Procedures for Investigations of HIV/AIDS Cases of Epidemiologic, Diagnostic, or Special Public Health Importance 3-42
Appendix D Educational Tools and Materials for Risk Factor Ascertainment 3-69
Appendix E Table shells (A–G) 3-85
Appendix F Training and Dissemination Options 3-92
Appendix G Sampling 3-95

Technical Guidance for HIV/AIDS Surveillance Programs — Risk Factor Ascertainment

Risk factor ascertainment surveillance policies and procedures were revised by the Exposure Category Workgroup, whose members include Chris Adamski (New Hampshire), Dena Ellison (Virginia), Maree Kay Parisi (San Francisco), Judy Sackoff (New York City), and Centers for Disease Control and Prevention (CDC) representatives Michael Campsmith, Mi Chen, Lorena Espinoza, John Gerstle, Felicia Hardnett, David Hurst, Jennie Johnston, Kathleen McDavid, Richard Selik, and Irum Zaidi.

Throughout this text, use of the word "must" implies a required action and use of the word "should" suggests a recommended action.

There has been much inconsistency and ambiguity in the terms used for the routes of HIV exposure/transmission and for the classifications of HIV-infected persons that summarize those routes. Multiple terms have been used to refer to the same thing, and the same terms have been used to refer to different things. Terms that end in the word "risk" are especially ambiguous because they may refer to categorical (MSM, IDU) rather than numerical (relative risk=1.2) variables. Therefore, to avoid confusion, a standard terminology should be used. After consideration of a variety of alternatives (see Appendix A for background on risk factors and the history of the terminology), the following standard terms and definitions are recommended:

- **Risk factors**—the collective term for the individual routes of exposure/transmission on which data are routinely collected for surveillance of HIV/AIDS cases. They require a "yes," "no," or "unknown" be checked on the case report form. See Appendix A for a list of risk factors.
- **Transmission category**—the term for summarizing the multiple risk factors (defined in Appendix A) that an individual may have had by selecting the one through which HIV was most likely to have been transmitted. The selection of the most likely route of transmission is based on a presumed hierarchical order of transmission that was developed in the early years of the AIDS epidemic. See Appendix A for a list of transmission categories.
- **Exposure category**—the term for a new classification (or any of the categories in it) that summarizes the multiple risk factors that an individual may have had by including combination categories of the three most common ones (MSM, IDU, HTC). Exposure category is less hierarchical than the transmission category. See Appendix A for a list of exposure categories.

HIV/AIDS Risk Factor Ascertainment

Definitions

Active surveillance: Health department staff regularly contact reporting facilities (hospitals, clinics, physician offices, laboratories) to identify potential/suspect HIV/AIDS cases (or confirm no cases) (Teutch SM, Churchill RE, eds. *Principles and practice of public health surveillance*, 2nd edition. New York: Oxford University Press; 2000). Health department staff review medical records at provider sites or receive information over the telephone, by fax, e-mail, US mail, etc. to establish an HIV/AIDS case and to elicit information for HIV/AIDS case report forms. All communication should follow security and confidentiality guidelines.

Passive surveillance: The health department receives HIV/AIDS case reports from physicians, laboratories, or other individuals or institutions without regularly contacting the reporting sources.

Cases of public health importance (COPHI): These are cases initially reported with a rare or unusual risk factor for HIV infection. Examples of these types of cases, which should be the first priority for follow-up, are occupational exposure, a human bite or tattoo, blood transfusion, or transplant recipient. See Section Cases of Public Heath Importance (COPHI) and Protocol 776 (Appendix C).

Date of initial case report: This is the date on which the public health department receives the first report on a potential case of HIV or AIDS. The document can be a laboratory report, a case report form, a birth certificate, a death certificate, etc. This date is not the date on which the surveillance area enters the reported information into a surveillance system.

Epidemiologic follow-up: This is the investigative process for obtaining additional information on a reported HIV/AIDS case.

Complete epidemiologic follow-up: A case reported with HIV/AIDS is considered to have undergone complete epidemiologic follow-up if 1) one or more risk factors described in Appendix A are identified or confirmed, and 2) all data sources available for a person (those listed in Section 2 of Epidemiologic Follow-Up) have been reviewed and/or contacted for risk factor information; or 3) 12 months have elapsed since the date of the initial case report and no risk factor has been identified.

No reported risk (factor) (NRR): A case is classified as an NRR if it is reported without any risk factor information or with unconfirmed COPHI risk factor information.

No identified risk (factor) (NIR): This is an NRR case for which an HIV risk factor cannot be identified or confirmed 1) although all available data sources have been reviewed or contacted, or 2) epidemiologic follow-up was either not initiated or not completed, but 12 months have elapsed since the date of the initial case report.

The Health Insurance Portability and Accountability Act (HIPAA): This is the law enacted by Congress in 1996 (Public Law 104-191) to protect health insurance coverage for workers and their families when they change or lose their jobs. It was also enacted to improve the efficiency and effectiveness of the nation's health care system by standardizing the electronic data for specific administrative and financial transactions, while protecting the security and confidentiality of patient data from inappropriate disclosure or use.

Cases of Public Heath Importance (COPHI)

Since 1981, the need to collect high-quality risk factor data has been a priority for HIV/AIDS surveillance. Included in this section are cases with unusual transmission circumstances that should be brought to the attention of CDC and should be a top priority for follow-up (followed by cases reported without a risk factor). As part of routine public health practice and routine HIV/AIDS surveillance, investigations of cases of public health importance (COPHI) meet at least one of the following criteria:

1. Clusters of unusual clinical, laboratory, or geographic occurrences that have potential public health significance
2. Possible unusual transmission circumstances where scientific evidence can confirm or refute the possibility of transmission (where possible)
3. Cases without detectable antibody response on standard testing
4. Cases of HIV-2 and non-B subtypes in the United States
5. Infections in children (< 13 years) not attributed to perinatal mother-to-child exposure.

In addition, data on occupational exposures to HIV are useful in evaluating the extent of HIV transmission in various health care and public safety settings and have practical value to CDC, the National Institute for Occupational Safety and Health (NIOSH), the Occupational Safety and Health Administration (OSHA), and health professional organizations in formulating more effective prevention strategies. Investigations of cases of public health importance often provide evidence to confirm or refute the possibility of HIV transmission in certain settings or under certain circumstances and, as such, contribute to the health of the public at large.

Initial case reports with information suggesting a COPHI should be the first priority for follow-up by surveillance areas. For COPHI the surveillance program should initiate a thorough investigation using trained personnel and following CDC Protocol 776 in Appendix C.

COPHI Outcome Standard

Annual progress reports (prepared 12 months after a reporting year) should list all COPHI from the previous year with status and include the following:

- Number of cases reported (for example, the yearly progress report for 2003 would reflect COPHI cases reported between 1/1/2003 and 12/31/2003)
- Number of cases where investigation was initiated
- Number of cases with a final disposition or status
- Number of CDC-confirmed cases (if any)
- Percentage of cases still open (number of cases reported minus the number of cases with a final disposition/number of cases reported)

COPHI Process Standards

- The state or jurisdiction should have documentation describing its legal authority to investigate HIV cases of public health importance regardless of reporting method (name, code, or name-to-code).
- All HIV and AIDS cases of public health importance must be investigated to confirm the reported exposure. There is no minimum performance standard.
- Investigation of cases should be initiated within 3 months of date of initial case report or at the time of notification from the patient or provider if sooner.
- A COPHI risk factor for a case can only be called confirmed by CDC in consultation with the health department, after its investigation, based on criteria as outlined in Protocol 776 (Appendix C).
- All cases should either be in 'active' investigation status or closed with a final disposition.
- Cases should be closed after one year if no further information becomes available, but can be reopened to confirm risk factors at a later date.
- The state or jurisdiction should run reports of all nonconfirmed COPHI on at least a quarterly or more frequent basis depending on morbidity, using case data from the HIV/AIDS surveillance system or equivalent software that is reported to CDC.

Investigation and Follow-up

Investigation and follow-up of cases of public health importance continue to be a priority at CDC. The COPHI coordinator in the HIV Incidence and Case Surveillance Branch/CDC should be notified about any potential special investigation. For detailed instructions on conducting investigations and risk factor definitions, see Protocol 776 in Appendix C.

NRR Risk Factors

Risk factors that occur in persons who are NRR that must be investigated include

- Received transfusion of blood/blood components after March 1985
- Received clotting factor injection for hemophilia or another chronic coagulation disorder, and date of birth is after March 1985
- Received transplant of tissue/organs or artificial insemination
- Worked in a health care or clinical laboratory setting with possible exposure to human blood or other body fluids
- Women whose only sexual contact has been with another woman
- Other exposure to human blood or body fluids including but not limited to the following:
  - Household or other ‘casual’ contact
  - Patient exposure in a health care setting
  - Physical interaction where blood or body fluids were exchanged
  - Occupational exposure other than in the health care industry
  - Tattoo, piercing, other cosmetic exposures
  - Intentional self-inoculation or intentional inoculation by another person
  - Human bite
  - Other unusual circumstance (not previously identified as a risk factor)

Pediatric Risk Factors

Risk factors for pediatric cases that must be investigated regardless of whether the reported person has another risk factor include

- Sexual contact with a male
- Sexual contact with a female
- Injected illicit or nonprescribed drugs
- Where mother was known seronegative after the child's birth

HIV-2 and Variant Strain Surveillance

- Idiopathic CD4 lymphocytopenia
- HIV antibody negative or other laboratory test inconsistent with the clinical picture
- HIV-2
- Unusual strains including Group O

Unusual Manifestations

Unusual manifestations of clinical, laboratory, and geographic clusters of public health importance that relate to the HIV epidemic

Risk Factor Ascertainment Procedures

Outcome Standard

At least 85% of reported cases or a representative sample (explained later) for a diagnosis year has an identified HIV risk factor within 12 months of the date of the initial HIV/AIDS case report, measured at 12 months after the close of the diagnosis year.

Process Standards

- 75% of all initial HIV/AIDS case reports (if a laboratory report was the initial report, then 75% of the first subsequent reports) have at least one HIV risk factor identified.
- All initial HIV/AIDS case reports are reviewed by a designated local surveillance program person to prioritize and determine if epidemiologic follow-up is warranted, and to ensure follow-up is conducted and appropriate cases are correctly reclassified.
- NRR cases are followed up within one month of the date of the initial case report by contacting medical providers, case managers, and/or reviewing medical charts to ascertain patient risk factors.
- The surveillance program conducts audits on a regular basis (at least annually) to assess the degree to which the outcome standard (see Outcome Standard) was achieved.
- The surveillance program has a written manual, which is reviewed annually and updated as necessary, containing training plans for its staff and providers, as well as policies and procedures for ascertainment, collection, and evaluation of HIV risk factors.

Initial Case Reports

Initial case report refers to the first report, whatever the source, which alerts the surveillance program to a case or possible case of HIV or AIDS. Local health departments will often receive initial reports from a variety of sources; but the reporting sources will depend on state-mandated reporting requirements, local practices, and financial resources. Sources of initial reports may include laboratories that conduct HIV-related tests, HIV diagnostic and care providers, matches with other registries such as vital statistics and tuberculosis, and active surveillance by field staff. The initial report may be relatively complete and require little or no follow-up, for example, when surveillance staff conduct active surveillance at a reporting site and conduct a thorough investigation to complete the case report. Alternatively, the initial report may contain

only minimal information and serve as a trigger for a complete investigation. Laboratory reports of HIV-related events often fall into the latter category.

Epidemiologic Follow-Up

Background

Core surveillance activities include epidemiologic follow-up to obtain complete information on reported HIV and AIDS cases (demographics, risk factors, and clinical information). A critical component of epidemiologic follow-up is to identify all known risk factors. The risk factors correspond to the period "before the first positive HIV test or AIDS diagnosis," and no longer correspond to the period "after 1977 and preceding the first positive HIV antibody test or AIDS diagnosis" as stated on the adult case report form. Many young adults today were either very young or not yet born in 1977. Those who may have been exposed before 1977 may not remember if it was before or after the cutoff date. Removing the cutoff date may be more reliable than what is currently on the adult case report form.

These data are essential to accurately describe epidemiologic trends. After first priority COPHI cases or suspected COPHI cases have been addressed, the second priority should be cases reported to the health department without an HIV risk factor. The process of risk factor ascertainment should begin when the surveillance program receives a report of an actual or suspected HIV or AIDS case report without any risk factor information (NRR). The risk factor ascertainment should include usual follow-up activities, such as calling a reporting facility or delegating field staff to inquire on a laboratory or provider report received by the surveillance program. Routine case follow-up should include inquiry about all HIV risk factors for each case, or a sample of cases, for surveillance programs that use the sampling protocol. The investment of time and resources to educate providers/reporters and surveillance staff regarding proper risk factor ascertainment should reduce the number of cases on which follow-up is needed and help to achieve complete and accurate information on all cases reported to the surveillance system.

The surveillance program should prioritize for investigation those cases reported with the unusual risk factors specified in Protocol 776 (Appendix C) for cases of public health importance (COPHI) over NRR cases; third priority should be cases reported with a risk factor, but with incomplete risk factor information. The surveillance program should initiate a thorough investigation using trained personnel and following CDC Protocol 776. The surveillance program should conduct follow-up on all NRR cases, if feasible, prioritizing facilities with the largest number of NRRs. Alternatively, the surveillance program should select a representative sample and conduct follow-up on that sample of cases. All NRR cases should be reclassified as NIR cases if 1) all reasonable data sources have been reviewed or contacted and no risk factor was identified, or 2) epidemiologic follow-up has not been initiated or completed, and 12 months have elapsed since the date of the initial case report. All cases or all cases in the selected representative sample will be used to measure the progress toward achieving the national outcome standard.

1. All risk factor variables listed in Appendix A should be given an answer of "yes," "no," or "unknown"; no risk factor variable should be left blank.

   This is the goal for HIV/AIDS surveillance. If a provider does not indicate a "yes," "no," or "unknown" for a risk factor (in other words leaves something blank), it cannot be assumed that it is a "no." HIV/AIDS surveillance programs should not modify original documents unless the program contacts the source of the document and receives additional information.

2. Within one month of the date of the initial case report, if an HIV risk factor has not been documented (i.e., an NRR), follow-up should begin by conducting a medical chart review at the initial case report facility; or, if a laboratory was the initial facility of report, then begin at the next subsequent facility of case report. Surveillance areas should prioritize new cases over old cases for follow-up. Within three months of the date of the initial case report, epidemiologic follow-up should expand beyond the reporting facility to include other available data sources or other facilities at which the individual has received care and other reporting sources, such as counseling and testing sites (CTS) and sexually transmitted disease (STD) databases. These sources should be prioritized based on where risk factors are most likely to be found.

   Surveillance program staff should develop standardized, systematic procedures for searching data sources, including database matches, beyond the initial reporting facility. Based on data from the Ascertainment of Transmission Risk (ATR) project in 10 states, risk factor information was most likely to be found from the following six sources:

   - Review of medical charts at health care provider who did not test or report the patient for HIV but for whom the reported person is a patient (according to the ATR project sites, this provider is most likely the current HIV treatment provider for the patient).
   - Review of medical charts at health care provider who tested the patient for HIV.
   - Telephone calls or visits to the health care provider who did not test or report the patient for HIV but for whom the reported person is a patient (most likely the current HIV treatment provider for the patient), but where review of the medical charts was not done.
   - Telephone calls or visits to a social service case manager providing physical and emotional assistance to the patient with HIV.
   - Review of medical charts at health care provider who reported, but did not test, the patient for HIV.
   - Telephone calls or visits to the health care provider who reported, but did not test, the patient for HIV, but where review of the medical charts was not done.

3. The procedure for investigating NRR cases should continue until all reasonable sources of information have been exhausted or until one year after the date of initial case report.

(The risk factors correspond to the period "before the first positive HIV test or AIDS diagnosis" and no longer correspond to the period "after 1977 and preceding the first positive HIV antibody test or AIDS diagnosis.")

- Cases that remain without risk factors after one year from the date of the initial case report must be counted as no identified risk factor (NIR) cases for the purposes of measuring progress toward the national outcome standard.
- If risk factors are found on an NIR case after one year from the date of the initial case report, the case's risk factors must be updated, resulting in the case being reclassified.

4. Surveillance programs should use every opportunity, such as telephone contact and site visits, to reemphasize to health care providers the importance of collecting all known HIV risk factors using CDC- and locally defined terms.

- Keep a copy of the tools included in Appendix 4 to remind staff to educate providers during telephone conversations (Appendix D).
- Remind staff to carry CDC definition of HIV risk factors with them on site visits to use for educating providers during routine follow-up (Appendix D.3).

5. On a quarterly basis, surveillance programs should send each reporting facility statistical reports summarizing the percentage of cases reported by that facility with missing risk factors.

Surveillance programs may choose to prioritize facilities with the largest number of cases reported without any risk factor information. The idea is to identify mistakes early and correct them (assuming the surveillance program has made the effort to train providers and surveillance staff as recommended in Appendix F: Training and Dissemination Options, one would want to catch any incomplete entries or mistakes in reporting not long after they happen so the problems can be corrected). Good news should be stressed in the feedback, and it should indicate what aspects of reporting need work. Once a facility has attained a desirable level of complete risk factor reporting, less frequent feedback is acceptable; however, at least annual feedback should be given. Create a facility-specific statistical HIV/AIDS report showing recent trends (Appendix E, Table shell C and Table shell D).

- Create a facility-specific report of missing variables (Appendix E, Table shell E).

- Identify resources developed to assist providers in collecting missing variables. For example, indicate that resources such as the risk factor ascertainment materials (Appendix D) will be available on Internet Web sites.
- Include local HIV/AIDS statistics in a packet of information to a reporting facility.
- All nonlaboratory reporting facilities with 10% or greater NRR (of their total) cases should be contacted (e.g., a letter with the appropriate tables attached) about their obligation to ascertain and document patient risk factors. When explaining how the epidemiologic information is used in the surveillance program, you should first point out the positive aspects of the facility's reporting and then explain the deficiencies. Assistance in the form of tools/materials and instruction should be provided to facilities. The reporting facilities should be monitored until less than 10% of their initially reported cases are NRR.

6. On a quarterly basis, surveillance programs should examine the distribution of NRR cases by reporting facility to find out which facilities require follow-up. Once a facility has attained a desirable level of complete risk factor reporting, less frequent monitoring is acceptable.

- Create a table showing the total number of HIV/AIDS cases and percentage that are NRR by reporting facilities (excluding laboratories) (Appendix E, Table shell F).
- For training, prioritize reporting facilities that reported the largest number of NRR cases (Appendix E, Table shell F).
- Designated surveillance program staff should prepare and review a summary of cases by reporting facility and compare the completeness of risk factor reporting at the time of initial report to the process standard of 75% for all reporting facilities (see Process Standards). Reporting facilities with less than 75% completeness should be identified, and a printout specific to the respective facility should be generated that shows the exposure category and transmission category distribution of cases reported by that facility, highlighting the NRR proportion (Appendix E, Table shell C and Table shell D).
- For active surveillance sites, designated field staff should follow up by completing another chart review and/or a follow-up with the provider to ascertain risk factors of patients. If no risk factor information is available, request that the provider ascertain risk factor information and inform the provider that a follow-up contact will be made with him/her regarding risk factor information for the NRR patients.
- For passive surveillance sites, contact should be made with providers to explain that cases that they have reported lacked sufficient epidemiologic information and to emphasize the importance of risk factor information for

public health purposes. If the provider does not respond within three weeks, another contact should be made to the provider requesting the information (Appendix D, "Dear Colleague" letter, e-mail, phone call, etc.).

7. Surveillance programs should include routine data quality assurance procedures as described in Data Quality as part of routine case reporting and epidemiologic follow-up.

Note: Re-abstraction pertains to the information on a case collected before the report date. However, during the re-abstraction process, staff will also obtain information recorded in the medical record that may have been added since the case was first reported to the state health department. Risk factors are among those fields that must be re-abstracted routinely. Through re-abstraction, discovery of previously unknown or missing risk factor information may help reclassify an NRR or NIR case.

Basic elements of data quality assurance for risk factor ascertainment include

- Conduct visual editing (proofreading) of hard copy case report forms on 100% of all cases before data entry.
- Monitor NRR and NIR cases for active surveillance cases by each field staff (Appendix E, Table shell G).
- Re-abstract the first 10 to 20 records of all new surveillance program staff as a quality assurance measure [see *Training, Data Quality* or *Training, Exposure Category*].

Education and Training of Surveillance Staff

Background

State and local HIV/AIDS surveillance coordinators are viewed as having experience and expertise in collecting, analyzing, and disseminating population-based data on the HIV epidemic. Therefore, the surveillance coordinators and experienced surveillance staff should provide training on risk factor ascertainment to internal staff and to staff at external reporting sources, both laboratory (when appropriate) and nonlaboratory (e.g., health care provider), on a routine basis.

In most areas, surveillance staff do not have direct contact with HIV-infected persons, and the opportunity to interview the person or otherwise ascertain his/her risk factors rests with the provider. Nonetheless, the surveillance staff has a pivotal role in risk factor ascertainment through their relationship with providers. This role includes ensuring that providers have a clear understanding of risk factors as defined by CDC and the local health department, that they are actively ascertaining risk factors from clients, and that they are providing the information to surveillance staff through verbal communication or chart notes.

1. The surveillance program should have a written manual of procedures for risk factor ascertainment.

The manual should be specific and include

- the sources of information that are to be searched and the order in which they are to be searched
- the evidence/documentation that is acceptable for a particular risk factor
- the importance of collecting information on multiple risk factors and not stopping when a single risk factor has been identified
- the performance expectations in terms of turn-around time, communication with providers at assigned sites, etc.

2. The surveillance program should conduct risk factor ascertainment training for new surveillance staff within one month of their start date and make available the written manual of procedures for risk factor ascertainment.

3. New surveillance staff should receive an orientation and on-the-job training.

These sessions provide an early and important opportunity to teach new staff about their critical role in risk factor ascertainment and provide them with the tools to implement this process. New staff should be told that ascertaining risk factors is as important as other elements of case reporting, such as obtaining demographic information, and examples of how risk factor data are used and why they are important should be provided.

4. As part of the orientation/training, staff should be given a package of written reference materials that includes definitions of risk factors (Appendix D.3, CDC-defined HIV risk factors), and a written manual of procedures defining how risk factor ascertainment is conducted at the local level.

The orientation/training should also include a presentation or discussion of risk factor ascertainment by a senior surveillance staff member. The presentation can address the "big picture" in terms of how risk factor information is used by the program and the consequences of inadequate information.

Sample risk factor ascertainment exercises can provide a context for applying policies and procedures, reinforcing key points, and answering questions.

- The on-the-job trainer should address risk factor ascertainment as a separate element of the review of the new staff person's case reports. This is a prime opportunity to ensure that definitions of risk factors are well understood and to identify strategies for actively eliciting this information from providers.
- Surveillance programs may also use re-abstraction to test training efficacy. For example, a specific number of cases a new staff member has recently completed can be re-abstracted to ensure standards are being met. The re-abstraction should be done by a different, preferably experienced, staff person.

5. The surveillance program should conduct retraining and evaluate all appropriate staff at least annually.
    - Presentations at local, regional, and statewide meetings or conferences on HIV/AIDS can ensure continued education and training on risk factor ascertainment procedures. In addition, risk ascertainment information published in health department newsletters and in surveillance reports would advance the message.
    - Although training in risk factor ascertainment should be treated as an ongoing activity, an annual training should also be conducted. This will ensure that all staff is presented a single consistent message. The annual training is also an opportunity to update staff on new initiatives or materials related to risk factor ascertainment, and to review the progress of the program toward outcome and process objectives.
    - Use role-playing with staff on talking with/educating providers regarding risk factor documentation.
    - Protocol drift—a slow but steady unintentional movement away from the original protocol—is a potential problem in all surveillance activities. Avoiding this drift requires regular feedback to staff about the consistency of their work within the requirements of the protocol. Therefore, after the initial training of staff, opportunities must be found to integrate feedback on risk factor ascertainment into the routine activities of the surveillance program. These might include supervisors reviewing submitted case reports, giving staff group and individual feedback on the success of risk factor ascertainment activities, and discussing difficult aspects of the protocol at staff meetings.
    - Create a summary report illustrating progress made by each surveillance staff member who is responsible for risk factor ascertainment. The report should provide information on the total number of NRR cases the staff is responsible for following up, the percentage of NRR cases successfully reclassified, the percentage of such cases still under investigation, the percentage reclassified as NIR because all available sources were contacted/reviewed, and the percentage reclassified as NIR because epidemiologic follow-up was incomplete but 12 months had elapsed since the date of the initial case report (Appendix E, Table shell G).

Education and Training of Health Care Providers

Background

Providers are busy and may not share or prioritize the health department's goal of ascertaining HIV risk factors. Therefore, we are in jeopardy of losing their attention if our communications are not clear, realistic, and concise (without excessive use of acronyms and terminology not well understood by providers). When necessary, reporting aids should be developed for providers (Appendix D); an example of such an

aid is a laminated card with the algorithm for reporting risk factors on a new case. Many providers may not be clear about what comprises CDC-defined "heterosexual contact" and may note "multiple sex partners" in the medical record and believe they are providing risk factor information when in fact they are not.

Communications regarding risk factor ascertainment must be directed to the appropriate person or persons at the reporting facility. Frequently more than one person has this responsibility, or one person is responsible for eliciting risk factor information from clients and a second person for communicating with the health department. No single job title adequately describes the appropriate contact person (e.g., administrators, social workers, counselors, nurses, or physicians may be the contact person). Further, at many facilities staff turnover is high, and it is important to keep abreast of these staff changes. Every surveillance program should have access to and the use of current educational tools/materials to disseminate to staff at new and existing HIV/AIDS reporting facilities. The training and education of staff at reporting facilities should assist surveillance programs in maintaining a heightened awareness and visibility, and will serve to foster collaboration with health care providers. This training of staff at reporting sources, if done well, can reduce the initial time, money, and other resources expended by a surveillance program to conduct risk factor ascertainment.

1. All nonlaboratory reporting facilities should have at least one person trained in the surveillance program's procedures for ascertaining HIV risk factors (including knowledge of CDC-defined risk factors).

- Surveillance staff should identify the designated reporter or primary contact at each reporting facility known to the surveillance program.
- Ideally, an on-site training should be conducted at all nonlaboratory reporting facilities. However, surveillance program areas may not have field staff or resources to conduct active surveillance (regular site visits, field staff to follow-up on each case). For that reason, at a minimum, a trained surveillance staff member should complete a telephone call and follow-up with a mailing (can be electronic distribution) of educational materials to the reporting facility (Appendix D).
- Surveillance staff should maintain a reference file (electronic or hard copy) of reporting facilities, which includes training and materials provided to each facility (Appendix D).

2. Identify and work with reporting facilities that may not be documenting or reporting risk factors because of concerns or misperceptions about HIPAA, confidentiality, or related issues.

- Discussion about confidentiality and the reporting process should be included in all provider contacts (telephone calls, site visits, mailing, e-mail) and as part of core surveillance activities. Surveillance programs should use available general disease reporting resources to discuss with providers their concerns about the collection, management, and dissemination (data release) of confidentially reported information.
- For reticent reporting sites, surveillance programs should use "Dear Colleague" letters signed by public health officials describing HIPAA as it relates to public health disease surveillance (Appendix D.4, HIPAA Letter).

3. The surveillance program should communicate with health care providers at least annually to educate them on their responsibilities for reporting risk factors. As much as is possible, communication related to risk factors should be incorporated into existing calls, e-mails, visits, etc.

- At a minimum, all providers should be contacted annually to ensure that they understand their responsibility of reporting HIV/AIDS cases and the importance of reporting complete information. The form of this communication (written, telephone calls, face-to-face, Internet) and its frequency will depend on a number of factors. These include the number of different sources reporting to the health department, their volume of reports, and the opportunities they present for conducting active surveillance. For example, it may be cost-effective to contact high-volume providers at more regular intervals and schedule face-to-face meetings.
- For low-volume reporters, such as private medical doctors who report only one or two cases annually, a mailing or Web-based communication annually may be a better use of resources.
- Presentations to professional associations at national, regional, and state meetings and publications in professional journals would target a larger audience of health care providers.
- Incorporate risk factor ascertainment material into the education and training centers' programs for health care providers treating persons with HIV/AIDS. Present epidemiologic data that include but are not limited to risk factor data.

4. All newly identified HIV/AIDS care providers/facilities should have at least one person trained in HIV risk factor ascertainment (including knowledge of CDC-defined risk factors) within one month of being identified as a reporting source.

- The surveillance program should identify any previously unknown HIV/AIDS care provider and initiate contact to both inform about and

discuss HIV/AIDS reporting rules, process, and general programmatic information (available statistics, care services).

- Surveillance program staff will most likely benefit by establishing face-to-face contact with a newly identified HIV/AIDS care provider. Surveillance programs should extend the offer for program staff to conduct a site visit at first contact to establish a relationship with the care provider. This expenditure of time and resources at first report (newly identified) will likely maximize the surveillance program/provider relationship and facilitate complete case reporting in the future.
- A reference list (electronic or hard copy) with the names, contact information, and training dates of the contact person should be maintained to facilitate regular communication and track reporting sources who have received training.

Sampling for Ascertainment of HIV Risk Factors

Consultation with your CDC core surveillance epidemiology technical assistant (Epi TA) (list available from the secure HIV Incidence and Case Surveillance Branch Web site, under "home") is strongly recommended throughout the sampling process, particularly during implementation and data analysis. State and local surveillance programs should also enlist the aid of their own researchers and/or statisticians for implementing sampling and analysis of state data.

Sampling: Advantages and Disadvantages

Sampling is a statistical method that collects information from a subset of subjects (sample) selected as representatives of a larger population of interest. The intent is to infer knowledge about the entire population from the sample. If information is collected from every subject in the population, the study is called a census. The Division of HIV/AIDS Prevention at CDC has used sampling to seek additional information on reported HIV/AIDS cases for a variety of projects. In particular, sampling has been used to validate previously reported HIV-related risk factor information and to ascertain risk factors on cases reported without such information.

Limiting data collection to a sample rather than the entire population has major advantages. Selecting a representative sample on which to complete epidemiologic follow-up can reduce cost, provide more timely data for monitoring risk factors among infected persons, and improve accuracy. Furthermore, a properly employed sampling scheme can allow state and local surveillance programs to analyze the data using well-established statistical principles and extrapolate results to the entire population of interest.

Although statistical sampling is a powerful and useful technique, it does require users to have prior knowledge of the population to be sampled; it may also introduce complexity in implementation and in analysis of the data. Erroneous use of sampling designs may result in samples that are not representative of the population. Furthermore, data analyses

that do not account for the sample design will greatly affect the accuracy and validity of results.

Implementation of Sampling

Feasibility of Sampling within a State or Local Surveillance Program

It is strongly recommended that all state and local surveillance programs carefully consider sampling HIV/AIDS cases on which to perform epidemiologic follow-up. There are, however, a number of scientific and practical factors to consider when weighing the feasibility of sampling within a given program. These factors include

- Number of HIV/AIDS cases reported per year.
- Resources available to design and conduct a sampling protocol.
- Number and percentage of cases with unknown risk factors.

Each of these factors has a significant impact on the practicality and scientific validity of sampling within a surveillance program. For example, sampling would be recommended for programs in high-morbidity areas with limited field resources and a large volume of cases with no reported risk factors (NRRs). On the other hand, if a program has low or moderate morbidity or a low percentage of NRRs, epidemiologic follow-up could possibly be performed on all NRRs.

Consultation with your CDC Epi TA is strongly recommended as part of the decision-making process for whether or not to sample. However, before consulting with your Epi TA, you should conduct the following evaluation using local HIV/AIDS surveillance program data:

- Determine the number of reported HIV/AIDS cases in the past year.
- Determine the number and percentage of cases with unknown risk factors.

Also, for planning purposes, it would be useful to determine the level of detail that is needed for the analysis. For example, is precise risk factor ascertainment at the sex, race, or county level desirable? If so, frequency tables using these factors to determine the number of cases that fall into each category should also be generated. Numbers in these frequency tables will not only help to determine if sampling is feasible, but also assist in deciding the appropriate sampling design and associated sample sizes. Appendix G.1 and Appendix G.4 demonstrate how such frequency tables are used.

Once this evaluation is done, the state or local program area may choose not to sample if

- The program has fewer than 1,000 HIV/AIDS cases reported in the past year.
- If the percentage of NRRs within your program is less than 15%. According to the minimum performance standards outlined in the CDC National HIV/AIDS surveillance guidelines[1], each program should have

> risk factor information for a minimum of 85% of all reported cases or of a representative sample of reported cases after epidemiologic follow-up is completed.

Otherwise, you should consult with your CDC Epi TA to properly assess the feasibility of sampling within your surveillance program. Surveillance areas that have 85% of cases or more reported with at least one risk factor may choose to sample to ascertain multiple risk factors.

If your surveillance program chooses to sample, see Appendix G for recommendations.

List of Appendixes

A. Terminology, classification, and history of the collective terms used by CDC for risk factors for HIV infection and for their summary classification categories 3-23

B. Prototype forms for collecting risk factor data 3-38

C. Protocol 776 for cases of public health importance (COPHI) 3-42

D. Educational tools and materials for risk factor ascertainment 3-69

- **1.** Document with key discussion points for provider
- **2.** Risk factor assessment tools: one self-administered patient questionnaire and one provider-administered questionnaire
- **3.** (1) Definition of CDC HIV Risk Factors
  (2) Key to CDC HIV Risk Factor Definitions
- **4.** Dear Colleague letter about HIPAA and public health disease surveillance
- **5.** Dear Colleague letter about reporting HIV infection risk factor ascertainment—to motivate, inform, and educate
- **6.** Q's and A's about reporting individual risk factors—for provider/field staff
- **7.** Provider letter to accompany tool kit (all tools and materials to be packaged and sent to providers)

E. Table shells 3-85

- **A.** Transmission category among HIV/AIDS cases reported during the previous five quarters
- **B.** Exposure category among HIV/AIDS cases reported during the previous five quarters
- **C.** Transmission category among HIV/AIDS cases by nonlaboratory reporting facilities reported during the previous five quarters
- **D.** Exposure category among HIV/AIDS cases by nonlaboratory reporting facilities reported during the previous five quarters
- **E.** Completeness of risk factor reporting among adult and adolescent HIV/AIDS cases by nonlaboratory reporting facilities reported during the previous five quarters
- **F.** Adult and adolescent HIV/AIDS cases by nonlaboratory reporting facilities reported with incomplete risk factor information during the previous five quarters; facilities are sorted in descending order of the total number of cases with incomplete risk factor information during the five quarter time period
- **G.** Status of epidemiologic follow-up among adult and adolescent HIV/AIDS cases by surveillance staff reported with incomplete risk factor information during the previous 10 Quarters

F. Training and Dissemination Options 3-92

G. Sampling 3-95

1. Point Estimates and Sample Size Calculation for Simple Random Sampling (SRS)

2. Taking a Simple Random Sample (SRS) Using the Random Number Table

3. Desired Number of Strata

4. Point Estimates and Sample Size Calculation for Stratified Random Sampling

Appendix A

Terminology, classification, and history of the collective terms used by CDC for risk factors for HIV infection and for their summary classification categories

Background

For state HIV/AIDS surveillance, data are reported to the local health department on possible cases either directly from health care providers or laboratories or by health department staff through abstracting in- and out-patient records. Case information is subsequently stripped of key personal identifiers (e.g., name, address, social security number) before being reported to the Centers for Disease Control and Prevention (CDC) for national surveillance. Variables reported to CDC include demographic characteristics, clinical and laboratory information, and all identified risk factors. Based on individual risk factors for a case, CDC assigns the most probable mode of transmission (the "transmission category," Section Transmission Category), which is based on a presumed hierarchical order of probability of transmission. Developed in the 1980s, the hierarchy was based on information available at that time about how HIV was transmitted. National surveillance data on HIV/AIDS transmission category, as well as demographic characteristics (geographic location of residence, age, sex, race/ethnicity), are used to allocate funds for HIV infection prevention programs and care services, and to evaluate prevention interventions and programs, among other uses. Therefore, having complete and accurate data for these variables is crucial.

To promote the collection of high-quality, timely data, CDC established minimum performance standards for HIV/AIDS surveillance systems in 1999. According to these guidelines, state and local HIV/AIDS surveillance systems should have information regarding risk factors that is at least 85% complete (from reported cases or a representative sample) after completion of epidemiologic follow-up (CDC. Guidelines for national human immunodeficiency virus case surveillance, including monitoring for human immunodeficiency virus infection and acquired immunodeficiency syndrome. *MMWR* 1999; 48(RR-13):13).

For AIDS cases diagnosed in 2002, approximately 27% (24% of males and 36% of females) were reported to CDC without risk factor information (CDC. *HIV/AIDS Surveillance Report* 2002; 14:30). For comparison, in 1994, 15% of all AIDS cases (13% for males, 22% for females) were reported without risk factor information (CDC. *HIV/AIDS Surveillance Report*, 1995;7(no.2):10). The reasons for the increase in the proportion of cases without a risk factor may be due to fatigue in reporting, an increase in the turnover of surveillance or provider staff, reduced resources in both health departments and provider settings, inadequate training and retraining of health department staff and of providers in taking sex and drug risk factor histories, and/or an increase in the proportion of physicians not conducting risk factor assessment for fear of offending patients with sexual behavior questions (Klevens et al. *Am J Epidemiology* 1999;149:75–84). In addition, another likely reason is the increase in heterosexual transmission among women, who often do not know the risk factors of their heterosexual partners. Initiation of HIV case reporting in the early 1990s, which generated a large increase in the number of

prevalent cases followed by increasing numbers of new reported cases thereafter, is also a factor. As a rule, HIV cases usually have a higher percentage reported without risk factor information than do AIDS cases, perhaps because AIDS patients are more likely to be in care and thus have more opportunities for providers to solicit risk factor information.

In the early 1990s, CDC developed a method to address the problem of increasing proportions of cases reported to CDC without a risk factor by using a statistical model to estimate the unreported risk factors. Risk factor redistribution weights were calculated based on two assumptions:

1. The distribution of risk factors among cases initially reported without risk factors (no reported risk factors or NRRs) does not change over time; and
2. Reclassified NRR cases are representative of all NRR cases.

Both of these assumptions were increasingly unlikely to be met. The pattern of risk factors has changed since the beginning of the epidemic (CDC. *MMWR* 1992; 41:28–29 and CDC. *MMWR* 2003; 52(47); 1145–1148), and reclassified cases usually represent a convenience sample of the cases on which risk factors are easiest to find.

Since the calculated redistribution procedure based on these assumptions is not an adequate solution to the growing problem of unreported risk factors, in December 2001 CDC convened a group of HIV/AIDS and behavioral research experts from academia, the private sector, and federal and state health departments, including state surveillance representatives, to generate ideas about how to best tackle this problem. The group recommended that CDC adopt a probable heterosexual (PH) subcategory of the no identified risk factor (NIR) category in addition to the existing risk factor categories in the hierarchical transmission categorization. The group further recommended that CDC stop using the current hierarchy to present the most likely transmission route, but instead report mutually exclusive categories of all HIV risk factors (Lee et al. *Public Health Rep* September–October 2003;118:400–407).

Subsequently, Schmidt and Mokotoff (*Public Health Rep* May–June 2003;118:197–204) conducted a record review and transmission category validation study using Michigan data to examine whether additional risk factors were recorded but were not captured by the CDC transmission category. Using data from two sources in addition to the surveillance data, the authors determined that 3% of males and 15% of females were reported as NRR, and 26% of males and 22% of females had multiple risk factors. This led the authors to conclude that the current CDC hierarchy may be inadequate to describe recent trends in HIV transmission in women. The authors went on to recommend that CDC add a second dual transmission category (in addition to male-to-male sex (MSM) and injection drug use (IDU)) of IDU and heterosexual, and that CDC add a PH subcategory for women.

In an effort to assess the feasibility and utility of the recommendations from the December 2001 meeting and the analysis by Schmidt and Mokotoff, Lee and colleagues (*Public Health Rep* September–October 2003; 118:400–407) used national HIV/AIDS surveillance data from 1999–2001, reported through September 2002, to assess whether adding PH as a category in the hierarchy would help to further describe HIV transmission. PH was assigned as a risk factor for an individual if the case report form

for that individual indicated "yes" for sex with one or more persons of the opposite sex, and said "no" for all other risk factors (see Sections Risk Factors and Definition of CDC Risk Factors for a list and definitions of risk factors). The authors compared risk factors

1. as reported;
2. as estimated by the statistical model for the redistribution of unreported risk factors, as modified by using the above criteria for PH to reclassify some NRR cases as PH cases; and
3. as reclassified into mutually exclusive categories (nonhierarchical).

The results indicated that the proportion of AIDS cases without risk factors decreased by 1.4 percentage points (PP) for males and 2.0 PP for females when some NRR cases were reclassified as PH cases. For HIV (not AIDS) cases, the results were similar, with a 1.8 PP reduction for males and a 3.1 PP reduction for females after some NRR cases were reclassified as PH cases. The authors (Lee et al. *Public Health Rep* September–October 2003; 118:400–407) concluded that "Our inability to reduce substantially the proportion of NRR cases in the national database simply by creating additional categories or assigning cases to PH based on currently available data suggests a measurement or information ascertainment issue." Based on these results, more should be done to improve risk factor ascertainment rather than add additional risk factor categories. For example, with risk factors well documented in the medical record, additional investigations (involving more time and other resources) would be unnecessary.

CDC has funded several projects aimed at improving ascertainment of risk factors. One project is evaluating a standard protocol that gathered information from existing records to determine the most efficient sources of risk factors. Another project funded HIV/AIDS surveillance programs in high-morbidity areas to pilot sampling methods in an effort to make inferences about risk factors from a representative sample to the population of all persons diagnosed with HIV infection.

Terminology and Classification

Recommendations for Terminology Routinely Used in HIV/AIDS Case Surveillance and Risk Factor Classification

There has been much inconsistency and ambiguity in the terms used for the routes of HIV exposure/transmission and for the classifications of HIV-infected persons that summarize those routes. Multiple terms have been used to refer to the same thing, and the same terms have been used to refer to different things. Terms that end in the word "risk" are especially ambiguous because they may refer to categorical (MSM, IDU) rather than numerical (relative risk=1.2) variables. An example appeared in the *MMWR* of January 2, 2004, in which two articles about HIV used the same term— "transmission risk"—in entirely different ways. Therefore, to avoid confusion, a standard terminology should be used. After consideration of a variety of alternatives, the following standard terms and definitions are recommended:

Risk Factors

"Risk factors" should be the collective term for the individual routes of exposure (before the person found out he/she was HIV positive or diagnosed with AIDS) on which data are routinely collected for surveillance of HIV/AIDS cases. They are the following variables:

- Male sexual contact with another male (MSM). This is actually based on two variables: male sex and sex with a male.
- Receipt of nonprescribed drugs by injection, intravenously, intramuscularly, or subcutaneously ("injection drug use": IDU).
- Heterosexual contact with a person known to have HIV infection or at least with a person at increased risk of HIV infection (based on a history of MSM, IDU, or receipt of blood products). This could be based on any of several variables that describe the risk factor of the heterosexual sex partner:
  - sex with a partner who had received nonprescribed injections of drugs
  - sex with a bisexual male [applies only to females]
  - sex with a partner who had received blood products for treatment of a coagulation disorder such as hemophilia
  - sex with a transfusion recipient
  - sex with a recipient of a transplanted organ or tissue
  - sex with a partner not known to have any of the above risk factors for HIV infection but who is known, nonetheless, to have HIV infection
- Perinatal mother-to-child contact: birth to a woman who was known to have HIV infection or was at least at increased risk of HIV infection. This could be based on any of several variables that describe a risk factor for HIV infection that the child's mother had before the child's birth or that confirm her infection status:
  - mother received injection of nonprescribed drugs
  - mother had heterosexual contact with a male who injected nonprescribed drugs
  - mother had sexual contact with a male who had sexual contact with another male [MSM]
  - mother had sexual contact with a male who had received blood products for treatment of a coagulation disorder
  - mother had sexual contact with a male known to have HIV infection who had received a blood transfusion
  - mother had sexual contact with a male known to have HIV infection who had received an organ or tissue transplant
  - mother had sexual contact with a male known to have HIV infection but who was not known to have any of the above risk factors for HIV infection
  - mother had received a transfusion of blood or blood components and was known to have HIV infection
  - mother had received a transplant and was known to have HIV infection

The above variables describing the mother's potential routes of exposure to HIV are used in conjunction with the variable: mother's HIV infection status. In the absence of knowledge that the mother had one of the above risk factors for HIV infection, a value for mother's HIV infection status indicating that the mother had HIV infection is necessary to infer that the child experienced perinatal exposure. Conversely, a value for mother's HIV infection status indicating that the mother did not have HIV infection after the child's birth would mean that the child did not have a perinatal exposure.

- receipt of an infusion of clotting factor blood product for treatment of hemophilia or other chronic coagulation disorder
- receipt of a transfusion of blood or blood components
- receipt of a transplant of organ or tissue, or of artificial insemination
- exposure to HIV-contaminated human body fluids by some other route. Currently this variable exists on the pediatric case report form but not the adult case report form.

"Supplemental risk factors" should be the term for other behaviors or proxies that may be associated with various routes of transmission, such as number of sex partners, condom usage, noninjection drug usage, selling sex in exchange for money or drugs, a history of other sexually transmitted diseases, having spent time in prison, and diagnosis of viral hepatitis. Unless specified otherwise, "risk factors" should be assumed to refer only to those listed in the preceding paragraph.

Transmission Category

"Transmission category" should be the term for summarizing the multiple risk factors (as defined in risk factors) that an individual may have had by selecting the one through which HIV was most likely to have been transmitted. The selection of the most likely route of transmission is based on a presumed hierarchical order of transmission that was developed in the early years of the AIDS epidemic, and was based on what was known at the time about how HIV was transmitted. The hierarchy has not changed even though our understanding of the most efficient ways of HIV transmission has changed. The expanded transmission category variable has 5 categories of heterosexual contact (HTC), which differ by the risk factor of the sex partner, and 8 categories of perinatal (mother-to-child) exposure, which differ by the risk factor of the mother. In the transmission category variable (not expanded), the 5 categories of HTC are combined into a single category and the 8 categories of perinatal exposure are combined into a single category. For cases in which there were multiple risk factors, the hierarchical nature of this classification may conceal some risk factors. For example, with a combination of IDU and HTC, only the IDU would be selected and the HTC would be hidden. An exception to the hierarchy is made for the combination of MSM and IDU, in which one of those two risk factors is not selected over the other and both are presented in a combination category. The transmission category and expanded transmission category variables include some categories that come in pairs—one for adults/adolescents and another for children—for 1) receipt of blood transfusion/transplant, 2) receipt of blood products for treatment of hemophilia, 3) "other" risk factors, and 4) absence of reported/identified

risk factors. The list of categories in the transmission category classification is shown in Table 1:

Table 1: Distribution of US AIDS cases reported through 2003 by transmission category

Transmission Category	Frequency	Percent
MSM	373257	44.22
IDU	203226	24.08
MSM & IDU	49915	5.91
Adult Hemophiliac	4633	0.55
HTC	105616	12.51
Adult Transfusion*	8777	1.04
Adult Other	29	0.00
Adult Undetermined	88964	10.54
Pediatric Hemophiliac	563	0.07
Mother w/HIV	8353	0.99
Pediatric Transfusion*	415	0.05
Pediatric Other	20	0.00
Pediatric Undetermined	299	0.04
Total†	844067	100.00

* Receipt of a blood transfusion, donated organ or tissue, or artificial insemination.
† Excluding cases reported from more than one health department (for comparison with Table 2).

Exposure Category

"Exposure category" should be the term for a new classification (or any of the categories in it) that summarizes the multiple risk factors that an individual may have had by including combination categories of the three most common ones (MSM, IDU, HTC). The exposure category classification was developed in response to the "Risk Consultation" of December 3–4, 2001, as an alternative to the transmission category classification. The consultants stated that the assumption on which the current hierarchical classification ("transmission category") is based—that sufficient information is collected to allow accurate selection of the most likely mode of transmission from among multiple possible routes of exposure—is probably not true, and that the resulting concealment of routes of exposure lower in the hierarchy by those that are higher is therefore unjustified and misleading. They therefore recommended a classification that would be less hierarchical, particularly so that HTC would not be concealed as a risk factor when it occurred in combination with MSM or IDU.

The exposure category still is hierarchical with respect to risk factors other than the primary three groups (e.g., receipt of a blood transfusion), which appear only in single categories ranked lower hierarchically than the combinations of MSM, IDU, and HTC.

Unlike transmission category, exposure category adheres strictly to the rules for "cases of public health importance" (COPHI), which require that certain rare routes of exposure not be accepted unless they are confirmed by a special investigation, and that, if they are confirmed, should then take hierarchical precedence in the classification over other possible routes of exposure in the same individual.

Thus, if the COPHI investigation confirms that HIV transmission probably resulted from sexual abuse of a child or from a blood transfusion that became contaminated with HIV despite the screening of blood donors for HIV antibody, then this rare route of exposure should be selected as the single most likely mode of transmission. In contrast, transmission category, as it exists currently, does not adhere strictly to the COPHI rules. For example, transmission category attributes some cases to a risk factor like receipt of blood transfusions after screening of blood donors for HIV antibody had begun, despite lack of confirmation that the blood donor had been HIV-infected. For that reason, many cases that are classified in the transmission category for transfusion recipients are classified in the exposure category for no identified risk factor. Also, unlike transmission category, exposure category eliminates the age distinction between adult/adolescent and pediatric categories for receipt of blood products, transfusion, or transplant, and for absence of reported/identified risk factors.

Finally, the exposure category classification distinguishes between cases reported without any risk factor within the preceding 12 months and those remaining with no identified risk factor for 12 or more months. The list of categories in the exposure category classification is shown in Table 2:

Table 2: Distribution of US AIDS cases reported through 2003 by proposed exposure category

Exposure Category	Frequency	Percent
MSM Only	359553	42.60
IDU Only	158096	18.73
HTC Only	105821	12.54
MSM + IDU	42662	5.05
IDU + HTC	45128	5.35
MSM + HTC	13708	1.62
MSM + IDU + HTC	7254	0.86
Perinatal exposure	8380	0.99
Other*	12136	1.44
NIR†	82231	9.74
NRR‡	9098	1.08
Total§	844067	100.00

* Includes receipt of a transfusion, infusion, or injection of donated blood, or blood products, receipt of an organ or tissue transplant, artificial insemination, and occupational exposure.

† Cases with no risk factor identified, with investigation completed/abandoned (18,217; 2.16%) or with investigation ongoing but at least 12 months elapsed since report (64,014; 7.58%).
‡ Cases with no risk factor reported and less than 12 months since initial report.
§ Excludes cases reported from more than one health department (because risk factor information was available for this analysis from only the primary state a case was reported from because of limitations in the analytic file used by CDC).

Risk Factor Variables to Be Collected

Volume II of *Technical Guidance for HIV/AIDS Surveillance Programs* includes a listing of the required, recommended, and optional variables each surveillance area should collect for HIV/AIDS surveillance. Although critical to collect, the following risk factor variables are classified as "recommended." The current case report forms (CDC 50.42A Rev 01/2003 for adults and CDC 50.42B Rev 01/2003 for pediatrics) are limited by space such that the information requested about risk factors is not always clear. This section is meant to clarify exactly when a "yes" should be checked for each of the variables listed below.

Risk factors for adults/adolescents:

Currently, the introduction to the patient history section of the adult case report form (CDC 50.42A) states "After 1977 and preceding the first positive HIV antibody test or AIDS diagnosis, this patient had…." *Technical Guidance for HIV/AIDS Surveillance Programs* changes the introductory statement to "Before the earliest known positive HIV test or AIDS diagnosis, this patient had…."

If the person was exposed to HIV infection by any of the following routes before his/her first positive HIV test or AIDS diagnosis, then it is considered a CDC-defined risk factor for HIV infection:

1. **Male who had sex with another male**: This wording is intended to avoid the issue of sexual orientation or identity (whether the man considered himself homosexual, bisexual, "gay," "on the down-low," or basically heterosexual but occasionally having sex with other men). The important consideration here is not how this male perceived himself, but simply whether he had sex with another man. In this context, "having sex" or "sexual contact" means penis-to-mouth, penis-to-anus, or mouth-to-anus contact (but not mouth-to-mouth contact) regardless of which role (insertive or receptive) is played by the male in question. It does not include contact only with skin (not a body orifice). However, if explicit information on whether the man had sex with another man is unavailable, the man may be assumed to have done so if he stated that he was "homosexual" or "bisexual" or described himself with a similar term that implies that he had sex with another man. In addition, male-to-male sex may be inferred if he was diagnosed with any rectal STD (e.g., gonorrhea) before or at the time of HIV/AIDS diagnosis.

2. **Injected illicit or nonprescribed drug(s)**: This means receiving an injection, either self-administered or given by another person, of a drug that was not prescribed by a physician for this person. It generally includes illicit drugs used for

producing euphoria, but it may also include prescription drugs that were not prescribed (e.g., estrogen, testosterone, anabolic steroids, or human growth hormone). It does not include injection of prescribed drugs (e.g., insulin for treating diabetes). The drug itself is not the source of the HIV infection, but the context of it being taken illicitly (i.e., without a prescription) is likely to be associated with sharing of injection equipment (e.g., syringes, needles, cookers), which can result in transmission of bloodborne pathogens, such as HIV. The case report form does not include a separate question asking if injection equipment was shared.

3. **Had sex with someone of the opposite sex who had either of the two risk factors listed previously in items 1 or 2 (MSM and/or IDU)**: In this context, "having sex" or "sexual contact" means contact of one person's penis or mouth with the vagina or anus or penis of another person. It does not include mouth-to-mouth contact or contact of the penis or mouth with skin (not a body opening). Because the prevalence of HIV infection is high among men who have had sex with men and among injection drug users, such sex partners will be presumed to have HIV infection and it will not be necessary to document their infection.

4. **Had sex with someone of the opposite sex who was diagnosed with HIV infection or AIDS after having any risk factor for HIV infection listed below in items 6 (receipt of clotting factor for coagulation disorder), 7 (receipt of blood transfusion), or 8 (receipt of transplant or artificial insemination)**: Having sex is defined as previously in item 3. Because HIV infection has an extremely low prevalence among recipients of clotting factors who were born in March 1985 or after, recipients of blood transfusions, and recipients of transplants or artificial insemination, such sex partners will not automatically be presumed to have HIV infection. Instead, their HIV infection must be documented (e.g., by the history given by the person being reported as a case to his/her health care provider), but it does not require confirmation by a special investigation. For the sex partner's HIV infection to be attributed to any of these risk factors (e.g., receipt of clotting factor, transfusion, or transplant), the sex partner's diagnosis of HIV infection or AIDS should have occurred after the sex partner's exposure to these risk factors (the exposure must not have occurred only after the diagnosis). If the time relationship between the sex partner's HIV infection diagnosis and the sex partner's exposure to the risk factor cannot be ascertained, the risk factor may be assumed to have preceded the sex partner's diagnosis.

5. **Had sex with someone of the opposite sex who was diagnosed with HIV infection or AIDS but was not known to have had any of the risk factors of sex partners described above in items 3 and 4**: Having sex is defined as previously in item 3. The sex partner's HIV infection or AIDS must be documented (e.g., by the history given by the person being reported as a case), but it does not require confirmation by a special investigation.

6. **Received clotting factor injection for hemophilia or another coagulation disorder**: This mainly involves Factor VIII and Factor IX. In the United States, screening of blood donors for antibody to HIV began in March 1985, which reduced the likelihood that clotting factor obtained after that month would be contaminated with HIV. In addition, clotting factor blood products for hemophilia began to be more effectively heat-treated around that time. More recently, clotting factors have been synthesized without using donated blood. Therefore, it would be unexpected for persons with hemophilia born in or after March 1985 to acquire HIV infection by this route, and such an occurrence must be confirmed by an investigation under the protocol for cases of public health importance (COPHI).

7. **Received transfusion of blood or blood components (e.g., platelets)**: Because screening of blood donors in the United States for antibody to HIV began in March 1985, it would be unexpected for persons who received a transfusion in or after March 1985 to acquire HIV infection by this route. Therefore, such an occurrence must be confirmed by an investigation under the protocol for cases of public health importance (COPHI).

8. **Received a transplant of tissue or organ or artificial insemination**: Because of its rarity and public health implications, any acquisition of HIV by this route must be confirmed by an investigation under the protocol for cases of public health importance (COPHI).

9. **Worked in a health care or clinical laboratory setting with possible exposure to human blood or other body fluids**: This has been reworded to clarify that it includes work that involves physical contact with patients, blood, or body fluids. For example, it does not include work as a clerk, secretary, or administrator who does not have physical contact with patients. It may include a custodian, however, who could be exposed to contaminated materials that have been discarded. Because of its rarity and public health implications, any acquisition of HIV by this route must be confirmed by an investigation under the protocol for special cases of public health importance (COPHI).

10. **Other exposure to human blood or body fluids**: This could include other forms of occupational exposure, such as that experienced by a police officer or fire fighter, or non-occupational exposures, such as contact with another person's blood or an open wound as a result of providing informal health care to another person or as a result of a fight. Any acquisition of HIV by this route must be confirmed by an investigation under the protocol for special cases of public health importance (COPHI).

Risk factors for children:

1. **Perinatal (mother-to-child) exposure with specified maternal risk factors**: The child's mother had any of the risk factors described previously in items 2–5 or 7, or 8 in adult risk factors before her first positive HIV test or AIDS diagnosis, and the mother was not known to be uninfected after the child's birth:

   - Maternal injection drug use
   - Maternal sexual contact with a man known to have had HIV infection or who was at high risk because he
     - was an injection drug user, or
     - was a man who had sex with other men, or
     - had HIV infection after having received clotting factor blood products, a transfusion, or a transplant
   - Maternal receipt of a transfusion (and mother known to have had HIV infection)
   - Maternal receipt of a transplant (and mother known to have had HIV infection)

     Maternal receipt of clotting factors (item 6 of adult risk factors) is not included among these possible maternal risk factors because hemophilia tends to be an X-linked hereditary disorder that does not occur among females. If the mother's experience of one of these risk factors is known to have occurred only after her diagnosis of HIV infection or AIDS, then it should not be accepted as the route by which she became infected.
   - Maternal perinatal exposure

2. **Perinatal (mother-to-child) exposure without specified maternal risk factors**: The child's mother has or had (if deceased) HIV infection or AIDS, but she was not known to have any of the risk factors described previously in items 2–5 or 7, or 8 of adult risk factors before her first positive HIV test or AIDS diagnosis, and she was not known to have been uninfected after the child's birth.

3. **Risk factors for children, other than the mother's HIV infection**: These all require confirmation under the protocol for special cases of public health importance (COPHI).

   - Received clotting factor injection for hemophilia or another coagulation disorder: This is defined as previously in adult risk factor item 6.
   - Received transfusion of blood or blood components (e.g., platelets): This is defined as previously in adult risk factor item 7.
   - Received a transplant of tissue or organ or artificial insemination: This is defined as previously in adult risk factor item 8.

- Sexual abuse by an HIV-infected adult: This is defined as sexual contact of the child with a man or woman who had HIV infection or AIDS. The case report form asks if the child had sexual contact with a male or a female, but does not ask if that sex partner had HIV infection, because it is expected that the infection status of the perpetrator of the sexual abuse may not be known when the form is initially completed. A "yes" answer to the question about sexual contact will then result in an investigation under the COPHI protocol, which should reveal the infection status of the sex partner. So far, no case of pediatric HIV infection or AIDS has been reported in which a child had sex with a female; so in practice, this risk factor is limited to sexual abuse by an HIV-infected man.
- Injected illicit or nonprescribed drug(s): This is defined the same as previously in adult risk factor item 2, except that its rarity in children requires it not be accepted as such until it is confirmed by COPHI investigation.
- Other exposure to human blood or body fluids: This is similar to adult risk factor item 10. Such exposures among children may involve physical contact between children while playing or fighting, in which blood or serum from an injury of one child may come into contact with an open wound of the other child.

The History of the Collective Terms Used by CDC for Risk Factors for HIV Infection and for Their Summary Classification Categories

Note: This appendix provides a brief history of the development of the terms that have been used to refer *collectively* either to what we now call the "risk factors" for HIV infection (the individual routes/modes of HIV exposure/transmission) or to the categories of classifications that summarize them. It does not include the history of the terminology and definition for each individual risk factor.

Through 2003, the routes by which HIV-infected persons may have become infected with HIV have been examined in three different ways in HIV/AIDS surveillance case data:

1. As all the different potential routes of exposure (now called "risk factors"), each represented by a separate (yes/no/unknown) variable (See section Risk Factors for a list of them).

2. As a classification, represented by a single variable with many categories, that summarizes the multiple routes of exposure that an individual may have experienced by selecting the one through which HIV was most likely to have been transmitted, based on a presumed hierarchical order of probability. This is referred to as the "transmission category" in *Technical Guidance for HIV/AIDS Surveillance Programs*. For a more complete description of it, see section Transmission Category.

3. As a classification, represented by a single variable with many categories, that summarizes the multiple routes of exposure an individual may have experienced by representing them all in a combination category. Theoretically, this nonhierarchical classification should have a category for every possible combination of risk factors.

In the early 1980s, all the various routes of HIV exposure ["(1)" above and the section on Risk Factors] were collectively called "risk factors," a term that has also been used for describing the epidemiology of hepatitis B, which shares routes of transmission in common with HIV.

The categories in the hierarchical classification ["(2)" above and the section on Transmission Category] summarizing these risk factors for HIV infection were collectively called "high-risk groups" originally. Most of them were, in fact, groups at increased risk of HIV/AIDS, but one of them, transfusion recipients, was at relatively low risk, and the category with no apparent route of exposure ("undetermined") certainly was not at increased risk. The term "mode of transmission" was adopted, because it was already established in public health. Something could be done to interrupt a mode of transmission, but innate characteristics are much more difficult to change.

Thus, the collective term for the various routes of exposure became "modes of transmission," and the hierarchical classification was called the "transmission category" in the HIV/AIDS Surveillance Reports in 1986. Later, the fact that the modes of transmission were generally presumed based on types of experiences without strong evidence linking individuals to one another, and the fact that it was impossible to be sure which was the actual route of transmission for individuals who experienced multiple potential routes of exposure, led to the change of terminology to "modes of exposure" and "exposure category," which was implemented in the *HIV/AIDS Surveillance Report* in 1988 and has continued through the latest *Surveillance Report* (for 2002, published in 2003).

Recently, it has been argued that the word "transmission" should be used instead of "exposure" for the hierarchical classification, because the "transmission category" represents the route thought most likely to have been the one through which transmission actually occurred, whereas the various routes from which it is selected are only potential exposures through which transmission could have occurred. However, in cases involving multiple potential routes of exposure, the category selected by the hierarchy is not necessarily the most probable route of transmission. For example, in a combination of MSM and receipt of a transfusion of blood that was known to have been donated by an HIV-infected donor, the transfusion would be the more likely route of transmission, but the hierarchy would select MSM and ignore the transfusion. The hierarchical classification is mainly a convenient but somewhat arbitrary way to summarize the data into a relatively small number of categories.

Different combinations and permutations of the above terminology, together with a tendency to abbreviate by dropping one or two words (especially "factor," "group," and "category"), led to a great deal of inconsistency and ambiguity, in which the following terms have been used interchangeably with the ones above: "risk," "mode,"

"transmission," "exposure," "behavior," "behavioral risk," "risk behavior," "exposure risk," "transmission risk," "transmission mode," "exposure mode," and "risk mode." These terms have been used inconsistently across separate documents and also within the same document (e.g., the Ascertainment of Transmission Risk project protocol). It is sometimes not clear whether they are used as a collective term for the various potential exposure routes (1) or as the term for the single variable representing the hierarchical classification of exposure categories (2).

The above terms that end in the word "risk" introduce an additional element of confusion unique to HIV/AIDS epidemiology, because they refer to categorical rather than numerical variables. For all other diseases, and even sometimes for HIV/AIDS, terms ending in the word "risk" refer to numerical or at least semiquantitative variables (proportions, percentages, ratios, rates, or indications of higher or lower). An example appeared in the *MMWR* of January 2, 2004, in which two articles about HIV used the same term—"transmission risk"—in entirely different ways. In the article on "Implementation of named HIV reporting…," "transmission risk" was used to mean the HIV exposure category, whereas in the article on "Prenatal HIV testing and antiretroviral prophylaxis…," "transmission risk" was used to mean the percentage of HIV-infected children among the total number of children born to HIV-infected mothers.

The term "supplemental risk factors" is used for other risk factors (including proxies or promoters of HIV infection) that may be associated with various routes of transmission, such as number of sex partners, condom usage, noninjection drug usage, selling sex in exchange for money or drugs, a history of other sexually transmitted diseases, and viral hepatitis.

In 2004, the Exposure Category Workgroup was established to develop improved policies and procedures for routine ascertainment of risk factors for HIV infection in surveillance for HIV/AIDS cases. The workgroup recommended the term "transmission category" for the hierarchical multicategory classification. The workgroup also recommended the term "risk factors" for the collective term for the binary variables that each describe a route of exposure. Because of the nonspecificity of the term "risk factors," it was recommended that this term should be defined in every document in which it was used by specifying the risk factors that it included or indicating that these risk factors were limited to the ones included in the determination of the transmission category.

The Exposure Category Workgroup also considered a proposal for a new classification as an alternative to the "transmission category" to satisfy the recommendations from the "Risk Consultation" that took place on December 3–4, 2001. The consultants in 2001 said that the assumption on which the current hierarchical classification ("transmission category") is based—that sufficient information is collected to allow accurate selection of the most likely mode of transmission from among multiple possible routes of exposure—is probably not true, and that the resulting concealment of routes of exposure lower in the hierarchy by those that are higher is therefore unjustified and misleading. They therefore recommended a classification that would be nonhierarchical and mutually exclusive, at least with respect to the presentation of data for all possible combinations of the more common routes of exposure (MSM, IDU, and heterosexual

contact). Other routes of exposure would be only in single categories ranked lower hierarchically than the combinations of MSM, IDU, and heterosexual contact.

In addition to satisfying the recommendations of the 2001 consultation, the new classification also would differ from the current one by more strictly adhering to the rules for cases of public health importance (COPHI), which require that certain rare routes of exposure not be accepted unless they are confirmed by a special investigation, and that, if they are confirmed, then they should take hierarchical precedence in the classification over other possible routes of exposure in the same individual. Thus, if the COPHI investigation confirms that HIV transmission probably resulted from sexual abuse of a child or from a blood transfusion that became contaminated with HIV despite the screening of blood donors for HIV antibody, then this rare route of exposure should be selected as the single most likely mode of transmission.

The consensus of the Exposure Category Workgroup was to recommend the use of this new classification and to name it the “exposure category.” The rationale for choosing this name was that it conveys the idea that the category describes many of the combinations of risk factors by which a person may have been exposed to HIV, rather than presuming, as with the “transmission category,” to select the single one responsible for the actual transmission. Unfortunately, it may be easy to become confused between the “exposure category” and the “transmission category” because these terms have been used interchangeably in the past.

Appendix B

Prototype Forms for Collecting Risk Factor Data

Form for Risk Factor History for Adults/Adolescents

Sex of Sex Partners

Before the first positive HIV test or AIDS diagnosis, did this person have?

(Respond to all categories)

	Yes	No	Unk
Sex with a male?	___	___	___
Sex with a female?	___	___	___

Risk Factors for HIV Infection

Before the first positive HIV test or AIDS diagnosis, did this person have any of the following risk factors for HIV infection? (Respond to all categories):

	Yes	No	Unk
• Male who had sex with another male	___	___	___
• Injected illicit or nonprescribed drug(s)	___	___	___
• Had sex with someone of the opposite sex who:			
◆ Injected illicit or nonprescribed drug(s)	___	___	___
◆ Was a male who had sex with another male (*applies only to cases among females*)	___	___	___
◆ Was diagnosed with HIV or AIDS after receiving clotting factor infusion for hemophilia or another coagulation disorder	___	___	___
◆ Was diagnosed with HIV or AIDS after receiving a transfusion of blood or blood components (e.g., platelets)	___	___	___
◆ Was diagnosed with HIV or AIDS after receiving an organ or a tissue transplant or artificial insemination	___	___	___
◆ Was diagnosed with HIV or AIDS but was not known to have any of the above risk factors	___	___	___

	Yes	No	Unk
◆ Received clotting factor infusion for hemophilia or another coagulation disorder	___	___	___
If yes, specify disorder: ___ Factor VIII (Hemophilia A) ___ Factor IX (Hemophilia B) ___Other (Specify): _____			
• Received transfusion of blood or blood components (e.g., platelets)	___	___	___
First Mo. ___/___ Year ___/___ Last Mo. ___/___ Year ___/___			
• Received a transplant of tissue or organ or artificial insemination	___	___	___
• Worked in a health care or clinical laboratory setting with possible exposure to human blood or other body fluids (Specify occupation):_____	___	___	___
• Other exposure to human blood or body fluids (Specify): _____	___	___	___

Form for Risk Factor History for Children

Mother's Risk Factors for HIV Infection

Did the child's biologic mother have any of the following risk factors for HIV infection (before her first positive HIV test or AIDS diagnosis, if dates known)? (Respond to all):

	Yes	No	Unk
• Mother injected illicit or nonprescribed drug(s)	___	___	___
• Mother had sex with a man who:			
◆ Injected illicit or nonprescribed drug(s)	___	___	___
◆ Was a male who had sex with another male	___	___	___

	Yes	No	Unk
• Mother was diagnosed with HIV or AIDS after having sex with a man who:			
◆ Received clotting factor infusion for hemophilia or another coagulation disorder	___	___	___
◆ Received a transfusion of blood or blood components (e.g., platelets)	___	___	___
◆ Received an organ or a tissue transplant or artificial insemination	___	___	___
◆ Was diagnosed with HIV or AIDS but was not known to have injected drugs, had sex with another man, or received clotting factor, a transfusion, or a transplant	___	___	___
◆ Received clotting factor infusion for hemophilia or another coagulation disorder	___	___	___
• Mother was diagnosed with HIV or AIDS after receiving a transfusion of blood or blood components (e.g., platelets)	___	___	___
• Mother was diagnosed with HIV or AIDS after receiving an organ or a tissue transplant or artificial insemination	___	___	___
• Mother was diagnosed with HIV or AIDS after perinatal exposure from her own biological mother	___	___	___

Child's Risk Factors for HIV Infection Unrelated to Having an Infected Mother

Before this child's first positive HIV test or AIDS diagnosis, did the child have any of the following risk factors for HIV infection? (Respond to all categories. A “yes” answer to any of these unusual risk factors should be reported promptly to the state/local NIR coordinator as a COPHI case requiring investigation):

	Yes	No	Unk
• Received clotting factor infusion for hemophilia or another chronic coagulation disorder	___	___	___

If yes, specify disorder:

___ Factor VIII (Hemophilia A) ___ Factor IX (Hemophilia B) ___Other (Specify): _____

	Yes	No	Unk
• Received transfusion of blood or blood components (e.g., platelets)	___	___	___
Mo. Year Mo. Year First ___/___ ___/___ Last ___/___ ___/___			
• Received a transplant of tissue or organ	___	___	___
• Sexual contact with male	___	___	___
• Sexual contact with female	___	___	___
• Injected illicit or nonprescribed drug(s)	___	___	___
• Other exposure to human blood or body fluids	___	___	___

Appendix C

Procedures for Investigations of HIV/AIDS Cases of Epidemiologic, Diagnostic, or Special Public Health Importance

August 2005

Table of Contents

I. BACKGROUND . . . 3
1) Clusters of unusual clinical, laboratory or geographic occurrences that have potential public health significance. . . . 3
2) Cases without detectable antibody response on standard testing (including blood donation screening) . . . 3
3) Cases of HIV-2 and non-B subtypes in the United States . . . 4
4) Possible unusual transmission circumstances where scientific evidence can confirm or refute the possibility of transmission, where possible . . . 5
5) Non-perinatal exposure infections in children . . . 5
6) Occupational exposures to HIV in health care and other settings . . . 5
II. REPORTING REQUIREMENTS . . . 5
III. CASE DEFINITIONS AND PROCEDURES . . . 6
Seronegative HIV/AIDS Cases. . . . 6
Divergent strain Cases and HIV-2 Cases. . . . 6
Other Specific Cases of Public Health Importance. . . . 6
Female-to-female transmission: Exclusive female to female sex where the partner is known HIV positive and no other risk factors are identified . . . 7
Sexual contact in pediatric cases . . . 7
Needle use in pediatric cases. . . . 7
Pediatric cases where the mother was known to be HIV-negative at the time of the child's birth. . . . 7
Receipt of blood components that screened negative for HIV antibodies . . . 7
Other suspected unusual circumstances . . . 7
Occupational exposure. . . . 7
IV. SPECIMEN COLLECTION PROCEDURES . . . 8
V. DATA MANAGEMENT. . . . 8
PROCEDURES FOR RECORDING/REPORTING RISK INFORMATION. . . . 8
DATA HANDLING, STORAGE AND DISPOSITION OF CASE RECORDS. . . . 9
VI. INVESTIGATION CLOSURE . . . 9
VII. PROCESS AND PERFORMANCE STANDARDS . . . 9
VIII. PROTECTION OF HUMAN RESEARCH PARTICIPANTS . . . 10
RISKS AND BENEFITS . . . 10
INFORMED CONSENT . . . 10
CONFIDENTIALITY . . . 10
IX. INVESTIGATORS AND CONTACT INFORMATION . . . 11
X. REFERENCES . . . 14
APPENDIX A. . . . 16
GENERAL GUIDELINES FOR INVESTIGATION OF CASES OF SUSPECTED UNUSUAL TRANSMISSION . . . 16
APPENDIX B. . . . 19
POST-SCREENING TRANSFUSION FOLLOW-UP INVESTIGATIONS. . . . 19
APPENDIX C. . . . 20
TRANSPLANT INVESTIGATIONS. . . . 20
APPENDIX D. . . . 21
INVESTIGATION OF SEXUAL TRANSMISSION IN CHILDREN . . . 21
INFORMATION TO COLLECT IN A PEDIATRIC CASE INVESTIGATION . . . 22
APPENDIX E. . . . 25
INVESTIGATION OF SUSPECTED OCCUPATIONAL EXPOSURES. . . . 25
APPENDIX F. . . . 26
HIV-2 CASE DEFINITION[1]. . . . 26

I. BACKGROUND

Since 1981, CDC has systematically collected data regarding probable mode of HIV risk exposures. These data provide the largest population-based collection of HIV risk factor/exposure data available in the U.S. and are used to: 1) monitor trends in transmission, 2) identify new or unusual transmission circumstances, 3) identify behavioral risks to target risk reduction interventions and allocate prevention resources, 4) provide evidence to support the lack of transmission in certain settings or under certain circumstances, and 5) provide data for projection analysis. The need to continue to collect high quality risk factor information remains a priority in order to adequately reflect the current epidemic.

Cases of public health importance are those cases where either the risk factor reported is unusual and should be brought to the attention of CDC or where other circumstances exist, such as unusual strains of HIV, that necessitate investigation, intervention or communication with CDC. These include (but are not limited to):

1) Clusters of unusual clinical, laboratory or geographic occurrences that have potential public health significance

People who become infected with HIV may have multiple high risk exposures to other infectious agents. The HIV/AIDS surveillance system may identify and investigate unusual manifestations of clinical, laboratory, and geographic clusters of public health importance that relate to the HIV epidemic. These unusual occurrences may include, but are not limited to: clusters of co-infections with tuberculosis, hepatitis, or other sexually transmitted diseases (STD) and HIV; adverse outcomes of long-term antiretroviral therapy; changes in the ability of licensed tests to detect HIV antibody; threats to the blood supply due to defective test kits; and resistant strains of an STD in an immunosuppressed person.

2) Cases without detectable antibody response on standard testing (including blood donation screening)

Cases or suspected cases of HIV infection without detectable antibody response are probably rare events, but should be considered cases of highest public health priority because they could be associated with screened-negative HIV-infected blood or plasma donations, tissue or organ transplants, or because they could represent a newly emerging agent factor with implications for HIV diagnostics.

Clinical diagnosis of HIV-1 infection and screening of blood donations relies on detection of antibodies to HIV-1 in the serum or plasma of patients. Initial screening is done using a licensed enzyme immunoassay (EIA) screening kit; for specimens with reactive EIA results, HIV infection is confirmed using a Western blot or IFA test.

There may be several reasons for negative HIV screening results in patients who are HIV-infected. The most well recognized of these is “window period” collection (i.e. collection of a blood specimen during the time between infection with HIV and the advent of a detectable antibody response). In 2005, this window period is generally between 7–22 days following exposure, depending on the specific EIA reagents or other screening tests used.[1,2] Since May 1996, screening of blood donations has also included testing for the presence of HIV-1 p24 antigen[3]. In October 2004 the FDA approved Nucleic Acid testing (NAT)[4] for screening. Since NAT has been found to detect HIV earlier in the window period than p24 antigen, many blood centers are using NAT and eliminating the p24 antigen test.

Other possible causes of negative antibody screening results in HIV-infected patients include infection with non group M subtype B HIV-1,[5] conversion from antibody seropositive to antibody seronegative status (seroreversion), and laboratory or attributional errors. Although HIV-infected patients in late

stages of disease may have decreasing antibody titers to some HIV antigens, with current, highly sensitive antibody EIA kits, true seroreversion is very rare.[5]

Several case reports have described patients with recognized HIV infection and persistently seronegative antibody screening results. In some of these cases, the clinical progression of HIV disease has been very rapid,[6] leading to the hypothesis that some patients may fail to develop antibodies following primary HIV infection, and that this lack of antibodies allows rapid progression of disease. In two cases,[7,8] HIV-infected partners who were infected with the same viral strain as the seronegative partner had detectable antibody responses; this supports the hypothesis that lack of detectable antibody response is a result of a defect in the host immune system (a host factor), rather than of a viral characteristic (agent factor).

3) Cases of HIV-2 and non-B subtypes in the United States

Comprehensive surveillance of the HIV/AIDS epidemic includes surveillance of other variant strains that may not react to standard FDA approved HIV-1 tests. HIV type 1 (HIV-1) and type 2 (HIV-2) both cause AIDS. Efforts to prevent transmission of HIV-1, particularly through the blood supply, led to the development of diagnostic tests for HIV-1 antibodies, in 1985. Because HIV-2 infections are not always detected by HIV-1 tests, antibody tests for HIV-2 infections were developed. The FDA licensed an enzyme immunoassay (EIA) test kit for HIV-2 on April 25, 1990. Since 1990, each revision of the HIV/AIDS surveillance case report form has included data collection fields for HIV-2 test results. Beginning June 1992, the FDA recommended that all donated blood, blood products, and plasma be screened for HIV type 2 (HIV-2). Although considerable serologic cross-reaction occurs between HIV-1 and HIV-2, HIV-2 infection may not be diagnosed when screening is done exclusively with HIV-1 tests.

Numerous HIV strain types have been identified, based primarily on variations in the env genes, which code for the viral envelope proteins. The different strain types have been classified as HIV 1 group M (subtypes A J), HIV 1 group N, and group O, and HIV 2, according to the degree of their relatedness to each other as determined by phylogenetic analysis. Although subtype B isolates were first identified in North America and are believed to represent the majority of isolates in North America and Western Europe, the other subtypes have varying geographic distributions throughout the world. HIV 2 strains, primarily found in Western Africa, have not been as well characterized.

HIV 1 group O strains,[9] are a highly divergent group of HIV not reliably detected by routine enzyme immunoassay (EIA) testing methods.[10] The group O strains are primarily found in Central Africa, particularly in Cameroon (estimated to account for less than 10% of HIV infection in Cameroon). Although uncommon, group O strains have begun to appear outside of Central Africa. In the U.S., the first case of HIV 1 group O infection was identified in a woman originally from West Africa in April 1996,[11] and more recently, a second case was confirmed.[12] As of June, 2005 there are two licensed EIAs that included testing for Group 0. Group O infections usually test positive on HIV-1 western blots although they may have diminished envelope reactivity or might register an indeterminate result.

Less is known concerning detection of HIV 1 group M, non B subtypes by currently licensed EIA kits to detect antibodies to HIV 1. Recent data indicate that, among HIV and AIDS cases identified through national HIV/AIDS surveillance, for some communities a significant proportion of HIV 1 infections may be with group M, non B subtypes of HIV-1.[13] Individuals who are group O or HIV-2 infected and receive negative HIV test results may represent unrecognized infections or have inaccurate results when tested using viral load testing,[15] and, based on in vitro data, may have decreased response to therapy with certain routinely used antiretroviral medications.[16] Therefore, CDC will continue to work with state and local health departments to investigate potential cases of HIV-2 and to institute mechanisms for identifying and monitoring infections with group O and other unusual HIV strains.

4) Possible unusual transmission circumstances where scientific evidence can confirm or refute the possibility of transmission, where possible
Examples of unusual transmission circumstances include (but are not limited to): transplants (including artificial insemination), female-to-female transmission, transmission of HIV from a health care worker to a patient, household transmissions, and other suspected unusual circumstances (e.g. tattoo, human bite, intentional self-inoculation). Several examples of these are cited in the references.[17–19]

5) Non-perinatal exposure infections in children
These include suspected sexual contact in pediatric cases, suspected non-prescription needle use and cases where the mother was known HIV negative at the time of the child's birth.

6) Occupational exposures to HIV in health care and other settings
Health care workers reported without another known risk factor are a high priority for follow-up risk investigations. These cases are handled by the Prevention and Evaluation Branch, Division of Healthcare Quality Promotion at NCID. Transmission of HIV from a health care worker to a patient is also a public health priority and is handled by the Division of HIV/AIDS Prevention.

Data on occupational exposure to HIV are useful in evaluating the extent of HIV transmission in various health care and public safety settings and has practical value to CDC, NIOSH, OSHA, and health professional organizations in formulating more effective prevention strategies in these areas.[20] These data are also useful to manufacturers of medical devices and personal protective equipment in designing safer products and to occupational groups for education and review of recommended infection control practices.

II. REPORTING REQUIREMENTS

All project areas receiving HIV/AIDS surveillance funding are expected to report such cases as part of their core surveillance requirement (50 states, any separately funded cities, the District of Columbia and U.S. Trusts and Territories). The follow-up of such cases by State health department and CDC staff will be coordinated and prioritized within each organization. Communication between the organizations is essential. A staff person at both the state/local level and at CDC should be identified and function as the coordinator for follow-up investigations.

State and local health departments should inform all care providers who report HIV/AIDS cases to be alert to possible unusual HIV transmission circumstances and to call the State health department when such a possibility exists. The State health department in turn should call CDC to inform us that an investigation is proceeding.

Any person who meets a case definition as outlined in Section III is eligible for investigation through medical record review and discussion with health care providers. All persons, except prisoners and persons less than 18 years of age are eligible for interview and blood draw. Children under 18 years of age can be eligible if parent or guardian signs appropriate consent form.

Epidemiologic investigation of cases infected with divergent HIV strains will be conducted on a case-by-case basis, based on the public health importance of the strain in question and the laboratory characteristics of the infection.

Participants may choose to discontinue voluntary participation in an investigation at any time. If a participant decides that they no longer want to participate in the investigation or allow further testing of laboratory specimens, they will contact the HIV/AIDS surveillance staff at the State health department who will then contact all appropriate persons to inform them of the decision to discontinue the investigation and/or laboratory analysis.

III. CASE DEFINITIONS AND PROCEDURES

Cases cannot be called "confirmed" without CDC review of available documentation and laboratory results. It is better to contact CDC prior to beginning any investigation that will require additional effort beyond that of routine surveillance case follow-up. We anticipate that the majority of cases will be closed for a variety of reasons: another more common risk factor is identified, persons refuse interview and/or blood draw, persons are lost to follow-up, or, while the claim may be legitimate, insufficient documentation of the event makes confirmation impossible. Timely reporting and follow-up of cases while documentation and memory of events by affected persons is fairly recent can make the difference between a confirmed case and one closed due to insufficient information.

Case definitions are defined on an as-needed basis when a case or group of cases, identified through routine HIV/AIDS surveillance reporting, indicates the need for additional follow-up investigations to address questions having potential public health implications. Therefore, some categories listed below may not have a standard definition.

1 Seronegative HIV/AIDS Cases
   a Adults and adolescents (≥13 years of age) are seronegative HIV/AIDS cases if they have infection with HIV-1, documented by positive DNA or RNA-PCR or viral culture drawn from two specimens on different dates; and with at least two negative antibody screening test results from independent laboratories on two specimens collected on different dates.
2 Divergent strain Cases and HIV-2 Cases
   a Adults and adolescents (≥13 years of age) are divergent strain cases if they show laboratory evidence of humoral response to HIV-2 or non-subtype B antigens, or if sequence analysis of clinical isolates give results indicative of infection with a divergent HIV strain.
   b This can also include cases that are antibody positive but with persistently undetectable viral load tests without therapy
   c See Appendix F for further information on HIV-2 testing and the case definition.

For cases listed in 1) and 2), contact the CDC COPHI coordinator to arrange possible shipment of specimens to CDC for laboratory testing and HIV-2 or variant strain sub-typing. Some demographic information will be requested including: Sex, Race/ethnicity, date of birth, country of origin, CD4 T lymphocyte counts and any other previous test results relating to HIV.

3 Other Specific Cases of Public Health Importance
   a Transplants
      i) Documented: receipt of a transplanted organ that screened negative for HIV-antibody in a person for whom no other risk factor was identified and for whom an HIV seropositive donor was identified.
   b Artificial insemination:
      i) Documented: women with a seropositive sperm donor and no other risk for HIV infection
      ii) Suspected: 1) women with a history of artificial insemination for whom no other risk has been identified, and for whom a seropositive donor has not been identified, and 2) women with a history of artificial insemination who have other risk factors for HIV, as well as an identified seropositive donor.

c Female-to-female transmission: Exclusive female to female sex where the partner is known HIV positive and no other risk factors are identified

d Sexual contact in pediatric cases

i) Documented: The child must have

(1) A diagnosis of HIV/AIDS,

(2) A diagnosis of sexual abuse based on physical examination or physician diagnosis; child, care-giver, or perpetrator disclosure; previous or concurrent non-congenital sexually transmitted disease; prosecution of perpetrator; or pregnancy, and

(3) Documentation of an HIV-infected or high-risk perpetrator. A perpetrator is considered to be at high risk if he engages in male-to-male sexual contact or if he/she is an injection drug user or has hemophilia or other coagulation disorder.

ii) Suspected: The child must have

(1) Confirmed HIV infection or AIDS,

(2) A diagnosis of sexual abuse, and

(3) No other risk factor for HIV.

(i) These cases are suspected and not confirmed because the following information is unavailable: the identity, HIV serostatus, or HIV risk of the perpetrator.

e Needle use in pediatric cases.

f Pediatric cases where the mother was known to be HIV-negative at the time of the child's birth.

g Receipt of blood components that screened negative for HIV antibodies (after 3/85) and p24 antigen (3/96), by a person for whom no other behavioral risk factors were reported and for whom a seropositive blood donor has been identified

i) Documented: Receipt of blood or blood products that screened negative for HIV antibodies (after 3/1985) and p24 antigen (after 3/15/1996), in a person for whom no other behavioral risk factor information was reported or identified and for whom a seropositive blood donor was identified.

ii) Persons reported whose only risk factor for HIV is the receipt of a blood transfusion after March, 1985 are investigated to determine the likelihood of the transfusion as the mode of transmission. Cases that cannot be re-classified to another risk or for whom documentation of a post-screening transfusion risk is unavailable, are closed as incomplete, either because the serostatus of all donors is not known or because receipt of the transfusion cannot be verified (e.g., transfusions received in a foreign country) and therefore no look-back study can be done to determine the serostatus of the donors.

h Other suspected unusual circumstances (e.g. household transmission, tattoo, bite, intentional self-inoculation).

For cases listed in 3 a – h, contact the CDC COPHI coordinator once you have confirmed with the provider that one of the risk factors listed above is being considered the primary mode of exposure. For adults/adolescents complete the demographics and NIR modules of the questionnaires. When the purported source of infection is known, specimens may be requested to confirm or refute the hypothesis of infection from one individual to the other.

i Occupational exposure

i) Documented: Evidence of HIV seroconversion in temporal association with an occupational exposure and no other known exposure to HIV during the same time period. Seroconversion is defined as a serum specimen that has been negative for HIV antibody up to one year before or one month after the occupational exposure. Also included in this category are persons infected with HIV strains that are shown to be related to the occupational source through DNA sequencing techniques. A health care worker (HCW) with a history of high-risk behaviors for HIV infection is classified as having documented

occupationally acquired infection if the health department's epidemiologic investigation concluded that an occupational contact was the only exposure to HIV during the period when seroconversion occurred, or if the viruses of the HCW and the occupational source of exposure were shown to be related through DNA sequencing.

ii) Possible: Person who reports no behavioral risk for HIV infection or transfusion of blood or blood products before March, 1985, but reports past percutaneous or mucocutaneous occupational exposures to blood or body fluids or to HIV-containing laboratory specimens. Seroconversion as a result of an occupational exposure cannot be established, and the time and source of HIV infection for that HCW cannot be documented.

Occupational exposures involving infection of a health care worker from an HIV-infected source patient are the primary responsibility of the Division of Health Quality promotion (DHQP) (see contact information in Section IX of this document). However, an investigation of a health care worker exposure usually begins with an assessment of other possible risk factors for HIV necessitating involvement of the COPHI coordinator. Therefore, investigations are usually handled jointly. For HIV exposed health care workers who received post-exposure prophylaxis and subsequently tested positive for HIV, please notify DHQP for possible entry in the PEP Failure protocol.

Occupational exposures involving occupational exposures not in a health care setting or health care settings where the worker has potentially infected a patient should be brought to the attention of the COPHI coordinator.

IV. SPECIMEN COLLECTION PROCEDURES

For some cases, blood specimens may be collected for laboratory studies and future testing. Blood specimens will be collected if the epidemiologic investigation determines there is a public health need to do testing beyond that which has already been conducted and documented. Personnel from the State health department and from the HIV Incidence and Case Surveillance Branch (HICSB) and the Laboratory Branch, DHAP, will be consulted to determine if there is a public health need for the collection of specimens. **HICSB staff at CDC should be consulted before obtaining blood specimens.**

Laboratory test results will be reported to the State health department, which in turn will contact the participant or their provider to inform them of the results.

V. DATA MANAGEMENT

PROCEDURES FOR RECORDING/REPORTING RISK INFORMATION

Risk factor information is routinely collected by the State/local health departments using the Adult and Pediatric HIV/AIDS Confidential Case Report Forms. Risk factor information is entered into the computerized HIV/AIDS surveillance database at the State and sent to the CDC monthly or bi-monthly for upload into the national database.

If detailed risk factor information is not obtained through discussion with health care providers or review of medical records and the case is a priority for follow-up, the person with HIV/AIDS (PWHA) will be asked to participate in an interview. Approaching an individual to discuss their risk of HIV infection should be based on the recommendations of the health care providers involved and on the experience of the health department staff.

A proxy interview will only be considered in the very rare event that 1) the case is a high priority for investigation, 2) the PWHA is dead or unable to participate in an interview because of their mental or physical health condition, and 3) an appropriate proxy can be identified. An appropriate proxy would be

someone who knows the PWHA, knows their infection status, and would probably know the kinds for information asked on the questionnaire.

Standard worksheets/questionnaires for COPHI are completed by State/local health department personnel and may be sent to the CDC coordinator for review once patient identifiers have been removed for confidentiality. Further discussion of confidentiality of forms is contained in the Section VIII of this document. Contact the CDC COPHI coordinator if there are questions about which form(s) to use.

There are six worksheets currently available as electronic documents:

1) Module A - Demographics
2) Module B - No Identified Risk
3) Module C - Occupational exposure
4) Module D - HIV-2 and variant strains
5) Module E - Female to female transmission
6) Pediatric sexual abuse assessment

DATA HANDLING, STORAGE AND DISPOSITION OF CASE RECORDS

Once a case has been closed, the questionnaires and all accompanying documentation are filed in a secure storage area at CDC by State and State number and without patient identifiers. Records are eventually archived to off-site storage.

VI. INVESTIGATION CLOSURE

The following guidelines will be used for closing investigations:

1) If other risk information becomes available or the case is confirmed as a COPHI, the HIV/AIDS Reporting System software will be updated with the risk information.
2) If the case is interviewed, the appropriate modules of the CDC standardized questionnaire should be completed and forwarded to CDC.
3) If no risk could be confirmed and no interview was conducted:
   - Cases open for at least one year will be reviewed for possible closure. If no further follow-up is anticipated and information is deemed unavailable, the case will be closed. The case will be left open if there is a specific reason to do so (e.g., health department believes patient will discuss risk at later time or high priority status). Cases left open will be reevaluated for closure every 6 months until closed.
   - Follow-up on a case will be completed and the case closed within a reasonable amount of time following receipt by the State/city health department of death notification unless there is a specific reason to keep the case open
   - Cases may be re-opened at any time that new information becomes available.

VII. PROCESS AND PERFORMANCE STANDARDS

There is no minimum performance standard. All cases of public health importance should be investigated to the extent possible and closed.

Any case with a risk factor or variant strain (as described in Sections I and III) that is reported to the health department should be identified at data entry and referred to the appropriate staff member. The staff member should contact the provider or reporter to ascertain whether the case requires investigation. If it is determined that the case has another more common risk factor and that this is not a COPHI, the HIV/AIDS Reporting System software should reflect this change. Reports should be run every six months at a minimum to identify cases requiring investigation.

A yearly report of COPHI and the final disposition of each case should be run using the HIV/AIDS Reporting System software

VIII. PROTECTION OF HUMAN RESEARCH PARTICIPANTS

RISKS AND BENEFITS

The risk to participants of these investigations is minimal. Participants who consent to an interview may experience some degree of emotional anxiety or distress when asked personal questions of a highly sensitive nature. Also, participants who consent to a blood draw may feel a “pinch” associated with having a venipuncture. They may also experience some bruising and or soreness at the site of the needle stick. In rare instances, they may experience fainting associated with the procedure.

The participant could benefit from blood testing performed as part of the study. Verification of HIV infection status and the results of subtyping of the HIV virus may have direct application to the clinical management of the patient. Both the interview and blood draw component of the investigation will help CDC and the health care community better understand the risks of HIV and how better to prevent its transmission as well as allowing for a more complete epidemiologic and laboratory profile of HIV-infected patients without detectible antibody responses. Both components will also allow for the important collection of standardized data on the occurrence of HIV-2 and non-B subtypes of HIV-1 in the United States.

INFORMED CONSENT

Written informed consent will be obtained from all persons with HIV/AIDS (and/or their proxies) prior to an interview or blood draw. **The CDC approved consent form is now required to be used to obtain participant's consent in any cases where CDC will be performing laboratory tests.** The person conducting the interview or blood draw (usually a State or local health department staff person) will discuss the consent with the participant, explaining how the information gathered will be used and answering any questions the participant may have. The investigator will review the consent form with the participant and answer any questions that may arise. The participant will be asked to sign the consent form. **These consent forms will be kept by the State, local, or territorial health department and neither the original nor any copy will be sent to CDC. A copy of the consent form will be given to the participant.**

State health departments may also elect to use their own departmental consent forms. While this is allowed, the CDC consent form must also be used for any cases where information and/or laboratory specimens are sent to CDC for review.

CONFIDENTIALITY

No CDC HIV/AIDS surveillance information that could be used to identify individuals, either directly or indirectly, will be made available to anyone for non-public health purposes. In particular, such information will not be disclosed to the public, parties involved in civil, criminal, or administrative litigation (within the limits of the law), or non-health agencies of the federal, state, or local government. Results of these investigations, without any personally identifying information, may be described in a written document explaining the outcome of the investigation.

Information on these individuals that is provided to CDC should be identified by the health department state number. The consent form and any personal identifying information contained on the first page of the interview form should be retained by the State, local, or territorial health department. All personal identifiers including name, address and social security number should be removed from the body of the interview form and any other documentation (such as copies of medical records) before being mailed to CDC. All information collected is covered under an Assurance of Confidentiality at CDC. The Assurance of Confidentiality is guaranteed under Sections 306 and 308(d) of the Public Health Service Act (42 U.S.C. 242K and 242m(d)). Persons involved in these investigations should maximize their efforts to protect this confidentiality.

Surveillance information reported to the CDC will be used without identifiers for statistical and analytic summaries in which no individual can be identified. When necessary for confirming surveillance information or in the interest of public health and disease prevention, the CDC may confirm information contained in case reports or may notify other medical personnel or health officials of such information; in each instance, only the minimum information necessary will be disclosed.

IX. INVESTIGATORS AND CONTACT INFORMATION

Centers for Disease Control and Prevention (CDC), National Center for HIV, STD, and TB Prevention (NCHSTP), Division of HIV/AIDS Prevention, Surveillance Branch:

Adelisa Panlilio, MD, MPH, Co-investigator (Healthcare worker exposures)
Kate Glynn, DVM, MPVM, Co-investigator
Lata Kumar, Coordinator for investigations of HIV/AIDS Cases of epidemiologic, diagnostic, or special public health importance

Investigations and interviews will be conducted by State, local, and/or CDC personnel responsible for HIV/AIDS surveillance activities as deemed appropriate by State, local, and CDC staff. Laboratory analyses will be conducted by CDC staff in the National Center for HIV, STD and TB Prevention, Division of HIV/AIDS Prevention Laboratory Branch. Investigations of potential health care worker occupational exposures will be handled by the Division of Health Quality Promotion, National Center for Infectious Diseases.

Although several people may be involved in any one investigation, one individual in each State health department and one individual at CDC will serve as primary coordinator and contact for any investigation being conducted.

CDC Contact Persons

All questions about investigation of cases of suspected unusual transmission (including possible post-screening transfusion and organ/tissue transplant risks) will be directed to:

HICSB COPHI Coordinator
As of 1/2006: Lata Kumar, MPH
(404) 639-2050
Centers for Disease Control and Prevention
NCHSTP/DHAP
HIV Case and Incidence Surveillance Branch
MS E-47
1600 Clifton Road, NE
Atlanta, GA 30333
LKumar@cdc.gov

For questions about other investigations of public health importance, initial contact can be made to the person listed above. Depending on the circumstances of the investigation, the most appropriate Surveillance Branch staff person will be designated as the primary contact at CDC.

Questions about investigations of persons with potential health care worker occupational exposures will be directed to:

Adelisa Panlilio, MD, MPH
Centers for Disease Control and Prevention
NCID/Division of Health Quality Promotion
MS E-68
1600 Clifton Road, NE
Atlanta, GA 30333
(404) 498-1265
APanlilio@cdc.gov

There is also a Needlestick Resource Center: (404) 498-1250 or (toll free) (800) 893-0485
For post-exposure prophylaxis failure contact the PEP failure hotline: (888) 448-4911

Contact for questions on divergent HIV strains, HIV-2 and testing methodologies is:

Bernie Branson, MD
Centers for Disease Control and Prevention
Associate Director for Laboratory Diagnostics
MS D-21
1600 Clifton Road, NE
Atlanta, GA 30333
(404) 639-6166
BBranson@cdc.gov

Blood transfusion safety involving HIV:

Matthew Kuehnert, MD, MPH
Associate Director of Blood Safety
Division of Viral and Rickettsial Diseases, NCID
CDC

1600 Clifton Road, NE
MS A-30
Atlanta, GA 30333
(404) 639-4350
MKuehnert@cdc.gov

Laboratory contact for cases requiring genotypic analysis:

Marcia Kalish
Associate Chief For Science
DHAP Laboratory Branch
CDC
1600 Clifton Road, NE
MS G-19
Atlanta, GA 30333
(404) 639-3957
MKalish@cdc.gov

Laboratory contact for cases requiring testing for HIV-2 or divergent strains:

Michele Owen
DHAP Laboratory Branch
CDC
1600 Clifton Road, NE
MS A-25
Atlanta, GA 30333
(404) 639-1046
MOwen@cdc.gov

CDC Surveillance staff should be notified prior to the shipment of any laboratory specimens to CDC. The address for shipment of specimens is:

CDC Serum Bank Branch
602 Webb Gin House Road
Lawrenceville, GA 30245-5427
Phone: (770) 339-5917
FAX: (770) 339-5915
SBartley@cdc.gov
DHemmerlein@cdc.gov

Primary contacts: Suzette Bartley and Dollene Hemmerlein

X. REFERENCES

1. Busch MP, Kleinman SH. Nucleic acid amplification testing of blood donors for transfusion-transmitted infectious diseases. Transfusion 2000; 40: 143–59.

2. Tabor E, Epstein JS. NAT screening of blood and plasma donations: evolution of technology and regulatory policy. Transfusion 2002; 42: 1230–7.

3. Centers for Disease Control and Prevention. U.S. Public Health Service guidelines for testing and counseling blood and plasma donors for human immunodeficiency virus type 1 antigen. MMWR. 1996; 45 (RR-2): 1–9.

4. Food and Drug Administration. Center for Biologics Evaluation and Research. Guidance for Industry. Use of nucleic acid tests on pooled and individual samples from donors of whole blood and blood components (including source plasma and source leukocytes) to adequately and appropriately reduce the risk of transmission of HIV-1 and HCV. October 2004.

5. Roy MJ, Damato JJ, Burke DS. Absence of true seroreversion of HIV-1 antibody in seroreactive individuals. JAMA 1993; 269(22): 2876–2879.

6. Martin-Rico P, Pedersen C, Skinhej C, et al. Rapid development of AIDS in an HIV-1 antibody-negative homosexual man. [Letter] AIDS 1995; 9(1): 95–96.

7. Reimer L, Mattice S, Schable C, et al. Absence of detectable antibody in a patient infected with human immunodeficiency virus. Clin Infec Dis 1997; 25: 98–100.

8. Michael NL, Brown AE, Voigt RF, et al. Rapid disease progression without seroconversion following primary human immunodeficiency virus type 1 infection—evidence for highly susceptible human hosts. J Infect Dis 1997; 175 (6): 1352–9.

9. De Leys R, Vanderborght B, Haesevelde MV, et al. Isolation and partial characterization of an unusual human immunodeficiency retrovirus from two persons of west-central Africa origin. J Virol. 1990; 64 (3): 1207–1216.

10. Schable C, Zekeng L, Pau CP, et al. Sensitivity of United States HIV antibody tests for detection of HIV-1 group O infections. Lancet. 1994; 344: 1333–1334.

11. Rayfield MA, Sullivan P, Bandea CI, et al. HIV-1 group O virus identified for the first time in the United States. Emerging Infect Dis. 1996; 2(3): 209–212.

12. Sullivan PS, Do AN, Robbins K, et al. Surveillance for variant strains of HIV: subtype G and group O HIV-1. JAMA. 1997; 278(4): 292.

13. Sullivan PS, Do AN, Ellenberger D, et al. Human immunodeficiency virus (HIV) sub-type surveillance of Africa-born persons at risk for group O and group N HIV infections in the United States. JID 2000; 181: 463–469.

14. Dunne AL, Crowe SM. Comparison of branched DNA and reverse transcriptase polymerase chain reaction for quantifying six different HIV-1 subtypes in plasma. AIDS. 1997; 11 (1): 126–127.

15. Schutten M, von der Ende ME, Antiretroviral therapy in patients with dual infection with human imunodeficiency virus types 1 and 2. NEJM 2000; 342 (23): 1758–60

16. Simonds RJ, Holmberg SD, Hurwitz RL, et al. Transmission of human immunodeficiency virus type 1 from a seronegative organ and tissue donor. NEJM 1992; 326: 726–32.

17. Centers for Disease Control and Prevention. HIV transmission between two adolescent brothers with hemophilia. MMWR. 1992; 42 (49): 948–951.

18. Centers for Disease Control and Prevention. HIV infection in two brothers receiving intravenous therapy for hemophilia. MMWR. 1992; 41 (14): 228–231.

19. Fitzgibbon JE, Gaur S, Frenkel LD, et al. Transmission from one child to another of human immunodeficiency virus type 1 with a zidovudine-resistance mutation. NEJM 1993; 329 (25): 1835–1841.

20. Centers for Disease Control and Prevention. Update: investigations of persons treated by HIV-infected health-care workers—United States. MMWR. 1993; 42 (17): 329–331, 337.

APPENDIX A

GENERAL GUIDELINES FOR INVESTIGATION OF CASES OF SUSPECTED UNUSUAL TRANSMISSION

Objective

The primary focus of an investigation is to determine whether or not the person with HIV/AIDS has contracted HIV through a circumstance not previously documented in the scientific literature and/or through circumstances where the safety of the public might be in jeopardy if an investigation (and appropriate follow-up) were not done.

Resources

1. Health care providers (including):
   - Infectious disease physicians
   - Infection control practitioners
   - Nurses
   - Primary physicians
   - Social workers

2. Medical records:
   - Inpatient records (acute care hospital)
   - Outpatient clinic records
   - Physician's office records

3. Other records (including but not limited to):
   - Sexually transmitted disease (STD) records
   - Hepatitis B registry

4. Patient interview

Approach

Keep an open mind when investigating a case. Often when a provider calls the health department to report a case, the patient may already have some theory as to how he/she may have contracted HIV and this information will be passed on to health department officials. To obtain a complete picture collect the following information (go back as far as the last negative HIV test or the last 10 years up to the time of HIV diagnosis):

1) Demographic
- Age
- Country of birth—if not from the U.S. how long has the person been in the U.S.? Any potential exposures in the home country?
- Date of birth
- Gender
- Level of education
- Marital status—how long married, HIV status of spouse,
- Any close relatives with HIV
- Race

2) Drug Use History
- Skin popping of drugs
- Use of crack, cocaine, heroin, crank, speed, barbiturates, PCP or any other drug
- Use of any drug intramuscularly or intravenously
- Visits to a drug ‘shooting gallery’
- Visits to a ‘crack house’ or sex in a 'crack house’

3) Medical History—Any of the following: [collect dates of procedures, physician's name, facility where procedure performed]
- Activities involving contact with needles, knives, blood
- Acupuncture or other alternative medical therapy
- Artificial insemination
- Dental work
- Hepatitis, hepatitis vaccine, hepatitis immune globulin
- History of any sexually transmitted diseases
- Hospitalizations
- Injections by non-medical care personnel
- Medical conditions requiring frequent injections (e.g. diabetes, steroids, etc.)
- Organ or tissue transplants
- Receipt of blood or blood products
- Surgeries

4) Military History (in the last 10 years or before the diagnosis of HIV infection)
- Drug use while in the military
- Sexual contacts while in the military (if contacts occurred outside the U.S., collect details regarding country, date, gender and risk of partner, etc.)

5) Occupational history
- Dates of jobs
- Employment in a health care facility, clinical laboratory or as a public safety worker (policeman, fireman, emergency medical technician)
- Location of jobs
- Type of jobs

6) Prison/Jail History
- Drug use while incarcerated
- Sexual contacts while incarcerated
- Other possible exposures (receipt of tattoos, bloody brawls, etc)

7) Sexual History
- Age at first sexual intercourse
- Any sex partners with HIV or at high risk for HIV (heterosexual and same-sex partners)
    - Person with HIV/AIDS

- Man who had sex with another man
- Person who has used injection drugs
- Person who has received blood or blood products
- Person who has had a transplant or received artificial insemination
- Person born outside the U.S. in a country where heterosexual transmission of HIV predominates
- Person who has ever accepted money or drugs for sex (prostitute, commercial sex worker)

- Exchange of sex for money or drugs
- History of rape
- Number of opposite and same sex partners
- Kinds of sex participated in
  - Anal
  - Oral
  - Vaginal
- Whether a person's sex partners have had sex with a person in a high-risk category as listed above

8) Shared Residence

- Shared residence with a person with HIV infection or known to have AIDS
- Home health care administration to a person with HIV or AIDS
- Shared residence with a person at high risk for HIV infection as listed in 7).

9) Travel History

Travel outside of the U.S. and receipt of any medical care, injections, or sexual contacts while outside the country

APPENDIX B

POST-SCREENING TRANSFUSION FOLLOW-UP INVESTIGATIONS

Persons reported with HIV/AIDS whose only risk of HIV infection is receipt of a blood or blood product after March, 1985 should be investigated to: identify other risk factors for HIV, initiate blood center lookback studies if appropriate, and to confirm and document the transfusion risk.

Medical record reviews will often determine that persons reported with a post-screening transfusion risk received the transfusion at a date earlier than that initially reported, or that the date of the transfusion was incorrect. When the initial date reported is correct and it is after March, 1985, check with the blood center to determine exactly when antibody testing (screening) of the blood supply at that center was begun. Although the antibody test was licensed in early March 1985, some blood centers did not begin using the test until several months later.

If it is determined that the case only received screened blood, investigate for other HIV infection risks using the standard NIR procedures. If extensive follow-up fails to yield other risk factors, then the blood donors should be evaluated for HIV infection. The unit numbers for each of the blood components should be contained in the medical records. In some cases, the facility where the transfusion took place will contact the appropriate blood center, but in most cases, the health department will contact the blood center directly. The blood center conducts a look-back study to determine if any of the donors were found to be HIV positive during a subsequent donation. An attempt is made to contact those donors who have not returned for subsequent donation. These persons are asked to consent to HIV testing.

Regulations pertaining to the Social Security Administration's (SSA's) Blood Donor Locator Service were published in the Federal Register on December 24, 1991. Based on these regulations, if a State agency or authorized blood donation facility have been unable to locate blood donors who may be infected with HIV, the SSA will furnish to participating State agencies, at their request, the donor's last known personal mailing address.

CDC assistance with these investigations may be requested and should be requested in the case of interstate circumstances. If requested, an epidemiologist from the HIV Incidence and Case Surveillance Branch will be assigned to assist with the investigation. Because cases involving documented post-screening transfusions are tracked and reported, a written summary of the investigation, including the status of all blood donors if a look-back study was conducted, should be sent to the CDC point-of-contact.

APPENDIX C

TRANSPLANT INVESTIGATIONS

1) Conduct a routine NIR investigation using Modules A and B.

2) Obtain a transfusion history.

3) Verify receipt of the transplant through the facility where the transplant took place.

4) Contact the transplant coordinator at the facility.

5) Identify source of transplant. Contact the organ harvesting center.

6) Check other recipients of tissues/organs from the same donor for HIV serostatus.

APPENDIX D

INVESTIGATION OF SEXUAL TRANSMISSION IN CHILDREN

1) HIV/AIDS Surveillance Definition for **Confirmed** Sexual Exposure to HIV Infection in Children.
To meet this definition, a child must have:

- Confirmed HIV infection or AIDS, AND
- A diagnosis of sexual abuse (i.e., genital-genital, genital-anal contact) based on at least one of the following:
  - Physical examination or physician diagnosis
  - Child, caregiver, or perpetrator disclosure
  - Previous or concurrent non-congenital sexually transmitted disease (e.g., syphilis, gonorrhea, condyloma, etc)
  - Prosecution of perpetrator
  - Pregnancy, AND
- Documentation in perpetrator of:
  - HIV infection or AIDS, OR
  - Risk for HIV infection:
    - male-to-male sexual contact
    - injecting drug use
    - hemophilia/coagulation disorder
    - transfusion/transplant recipient with documented HIV infection

2) HIV/AIDS Surveillance Definition for **Suspected** Sexual Exposure to HIV Infection Among Children.
To meet this definition a child must have:

- Confirmed HIV infection or AIDS, AND
- A diagnosis of sexual abuse as noted above, AND
- No other risk factor for HIV

These cases may not meet the confirmed definition of sexual abuse because the following information was not available: the identity, HIV serostatus, or HIV risk of the perpetrator.

INFORMATION TO COLLECT IN A PEDIATRIC CASE INVESTIGATION

1) Basic Demographics
   - Age
   - Race/Ethnicity
   - Parental (biologic) race
   - Country of birth
   - Parental (biologic) country of birth
2) School/Living/Occupational History (since birth)
   - Day care attendance
   - School attendance
   - Foster care placement
   - Types, dates, and location of any job held
3) Medical history (since birth)
   - Date HIV infection confirmed
   - Dates and results of all HIV testing
   - Reason for HIV testing
     - Documented or suspected HIV infection in the parent(s)
     - Presentation with symptoms consistent with HIV
     - Recipient of blood/blood products or transplant
     - Dialysis recipient
     - Parental/caregiver request
     - Social service agency initiated
     - Positive neonatal drug screen
     - Routine screening
   - Hospitalizations (emphasis on description of neonatal course)
   - Surgeries (including surgeon's name)
   - Dental surgery or treatment (including dentist's name)
   - Receipt of blood or blood products/components
   - Receipt of any immune globulin
   - Organ transplant
   - Dialysis
   - Acupuncture
   - Medical conditions requiring injections (i.e., growth hormone, steroids)
   - Injections by non health care personnel
   - Any activities involving needles, knives, blood
   - Sexually transmitted diseases (GC, syphilis, chlamydia, herpes, venereal warts, urethritis, genital ulcers)
   - Hepatitis, hepatitis vaccine, hepatitis immune globulin
   - Hepatitis B serology on parents and child
   - Chronic skin disease (eczema, seborrhea)
   - Bites or blood contact (any setting: foster care placement, school, day care)
4) Travel history (since birth) outside the US (medical care, injections, sexual contacts)
5) Illicit drug use history (this checklist should apply to parent or child with illicit drug use history and include time period in the past 10 years and before diagnosis of child's HIV infection)
   - Ever visited a drug 'shooting gallery'
   - Ever visited a 'crack house'
   - Use of crack, cocaine, heroin, crank, speed, barbiturates, PCP, any other drug
   - Ever used any drug intravenously

- Ever used drugs by skin popping

6) Sexual history (since time period for which child acknowledges sexual contact)
- Number and name of opposite and same sex contacts
- Ever accepted drugs, money for sex
- History of sexual abuse, rape
  - Report filed with police or children's protection agency
  - Physician record with documentation of genital-genital or genital-anal sex (e.g., hymenal distortion as reported by a physician familiar with pediatric introitus anatomy, anal lacerations, vaginal lacerations, perineal bruising, etc)
  - Child's verbal affirmation of sexual assault/penetration
  - Adult caregiver's or perpetrator's verbal affirmation of sexual assault/penetration
  - Documentation of concomitant or previous non-congenital sexually transmitted disease (GC, syphilis, venereal warts, genital ulcers, herpes, chlamydia, urethritis, trichomonas)
  - Resultant pregnancy
  - Prosecution of perpetrator for genital-genital or genital-anal sexual contact
- Any sex partners/perpetrators:
  - HIV/AIDS
  - Bisexual
  - Intravenous drug user
  - Received blood or blood product/components
  - Received transplant or artificial insemination
  - Born outside US
  - Ever accepted drugs, money for sex
- Any sex partners/perpetrators had sex with any person in the above high risk groups (section VII/D)

7) Parent's history (time period after last negative HIV test or last 10 years and before child's diagnosis of HIV infection)
- Illicit drug use
- Receipt of blood or blood products
- Organ transplant
- Dialysis
- Medical conditions requiring injections (i.e., pituitary hormone, steroids)
- Artificial insemination
- Injections by non health care personnel
- Any activities involving needles, knives, blood
- Sexually transmitted diseases (GC, syphilis, chlamydia, herpes, venereal warts, urethritis, genital ulcers)
- Employed in any health care facility or clinical laboratory
- Accepted money or drugs for sex
- Sexual partners at high risk
  - HIV/AIDS
  - Bisexual
  - Intravenous drug user
  - Received blood or blood product
  - Received transplant or artificial insemination
  - Born outside the US
  - Ever accepted money or drugs for sex
- History of rape
- Travel outside of the US (medical care, injections, sexual contacts)

8) Prison/jail history of the parent (in the last 10 years and before the child's diagnosis of HIV infection)
- Specify reason for incarceration
- Drug use
- Sexual contacts including rape

9) Social history of the parent: Was there ever a referral to a children's protective agency for neglect or abuse?
- Hospital or physician referral
- School referral
- Neighbor referral
- Relative referral

10) Household contacts
- Any HIV-infected household members/caretakers; any known HIV-infected visitors to household (describe contact with child)
- Any home health care taking place (needles or other medical sharps in household)
- Exposure to open wounds/cracked skin of HIV-infected household members
- Any sharing of previously chewed food
- Bowel and bladder control
- Household cleansing practices (linen, medical care, physical care, shared toothbrush, shared razors)
- Shared eating utensils
- Shared sleeping facilities
- Shared cleansing utensils (oral, skin, hair, razor, etc)
- Shared clothing

APPENDIX E

INVESTIGATION OF SUSPECTED OCCUPATIONAL EXPOSURES

Reports to CDC of occupational exposures are generated through several different mechanisms:

1) Post-exposure prophylaxis (PEP) failures
2) State health department notification
3) Health care provider calls
4) Health care worker calls
5) Health care worker identified in HIV/AIDS Reporting System (HARS)

Investigations of suspected occupational exposures are handled as 'routine No Identified Risk' cases but should be investigated using Module C, "Assessment Form for HIV/AIDS Cases Reported with Possible Occupationally Related Exposure to HIV."

APPENDIX F

HIV-2 CASE DEFINITION[1]

The CDC algorithm for testing recommends HIV-2 testing only in cases where:

HIV-1 Western blot results are negative or indeterminate following a positive HIV-1 & 2 EIA or with an unusual indeterminate pattern after a positive HIV-1 EIA with *gag* (p55, p24, or p17) plus *pol* (p66, p51 or p32) bands in the absence of *env* (gp160, gp120, or gp41) bands

OR

A patient tests positive on a licensed EIA test (including licensed rapid tests) that discriminates HIV-1 from HIV-2 (HIV-2 EIA)

AND

The patient has epidemiologic risk factors associated with HIV-2 infection listed below:

If a patient tests positive on HIV-1 western blot, no further testing for HIV-2 is recommended based on current CDC guidelines.

Persons at risk for HIV-2 include:

1. Sex partners of a person from a country where HIV-2 is endemic (this includes persons originally from such countries)
2. Sex partners of a person known to be infected with HIV-2
3. Persons who received a transfusion of blood or a non-sterile injection in country where HIV-2 is endemic[2]
4. Persons who shared needles with a person from a country where HIV-2 is endemic or with a person known to be infected with HIV-2
5. Children of women who have risk factors for HIV-2 infection or who are known to be infected with HIV-2.

Additionally, testing for HIV-2 is indicated when there is clinical evidence for or suspicion of HIV disease (such as an AIDS-associated opportunistic infection) in the absence of a positive test for antibodies to HIV-1 and in cases of an indeterminate HIV-1 Western blot as described above. Or, persons with documented HIV infection and clinical deterioration with persistently undetectable viral load.

[1]Testing for Antibodies to Human Immunodeficiency Virus Type 2 in the United States. MMWR July 17, 1992/ 41 (RR12); 1–9

At this time, no Western Blot tests are licensed in the U.S. for the diagnosis of HIV-2

CDC confirms a case using the following:

- Positive result on the Canadian Select HIV 1 & 2 test—the Canadian Select uses separate plates for HIV-1 and HIV-2
- If the HIV-2 plate is positive, an HIV-2 Western Blot* is done if the kit is available

- Or, confirmation with the Biorad Multispot Rapid Test if WB kit not available

*Genlab—purchased through Singapore (not available in the U.S.)

If requested a qualitative HIV-2 RNA test (unlicensed) can be performed in the CDC laboratory.

Probable case:

Both must apply:

- Statement in the medical record by a physician that the patient tested HIV-2 positive (and details of how that diagnosis was made)
- An epidemiologic picture consistent w/ HIV-2 risk factors listed above

Potential cases not covered under the criteria listed above will be handled on a case-by-case basis.

[2]Countries where HIV-2 is prevalent (+ indicates more than 1% HIV-2 prevalence in the general population):

West African nations:

Benin	Mauritania+
Burkina Faso	Niger
Cape Verde+	Nigeria+
Code d'Ivoire+	Sao Tome
Gambia+	Senegal
Ghana	Sierra Leone+
Guinea	Togo
Guinea-Bissau+	
Liberia	Other African Nations: Angola+, Mozambique+
Mali+	Other countries where HIV-2 has been reported: Portugal, France, India, Brazil, Canada and the United States

Appendix D

Educational Tools and Materials for Risk Factor Ascertainment

Appendixes D.1 – D.2

Appendix D.1

Talking with Your Patients about Behavioral Risk Factors for HIV and AIDS

TALKING WITH YOUR PATIENTS ABOUT BEHAVIORAL RISK FACTORS FOR HIV AND AIDS

Patients may be uncomfortable disclosing personal risk factors and hesitant to respond to questions about sensitive issues, such as sexual behaviors and illicit drug use. However, evidence suggests that when asked patients will often discuss behaviors that increase their risk of acquiring HIV. Evidence also suggests that some patients have greater confidence in their clinician's ability to provide high-quality care when asked about sexual and STD history during the initial visits. Of course, the more comfortable you are with discussing these issues the more comfortable your patients will be.

Below are some ideas for talking with your patients about these subjects.

Put your patients at ease.

- Reassure your patients that their responses will remain confidential.
- Let them know that you ask all of your patients these types of questions.
- Tell them that the information they provide about their sexual and drug-use behaviors will help you provide the best possible care.
- Use open-ended questions to avoid simple "yes" or "no" responses. This encourages patients to discuss personal risks and the circumstances in which risks occur. Open-ended questions also help you gather enough detail to understand potential transmission risks and make more meaningful recommendations for prevention of secondary transmission.
- Respect a patient's choice to not answer a question. This increases the chance that she/he will provide the information at a later date.

At the end of the session,

- Summarize the patient's responses to make certain that both you and your patient understand what was said.
- Encourage the patient to ask questions about any issues he or she might not have understood, and, if needed, schedule a follow-up appointment.

Some things to remember when speaking with your patients about risk factors:

Honest responses may be more forthcoming if the question is worded in such a way to "normalize" the behavior:

- "Some people inject drugs. Have you ever done that?"
- "Some people have anal intercourse. Have you ever done that?"
- "Some people exchange sex for drugs or money. Have you ever done that?"

Labels can be misleading.

- Some men do not consider themselves "gay" if they practice same sex anal insertive intercourse, but their receptive partners may be considered to be "gay".
- The question, "Are you a homosexual?" may be answered with "no" by a person who has had only a few same sex encounters or who considers him/herself to be bisexual.
- Describe behaviors instead of assigning labels to the behavior. Use terms "drug user", "men who have sex with men", "women who have sex with women", or "sex worker".

Source:

Gerbert B, Bronstone A, Pantilat S, et al. When asked, patients tell: disclosure of sensitive health-risk behaviors. Med Care 1999;37:104--11.

MountainPlains AIDS Education and Training Center

Printed in the USA. April 2005.

Appendix D.2

Risk Factor Assessment Tools: One Self-Administered Patient Questionnaire and One Provider-Administered Questionnaire

PATIENT QUESTIONNAIRE

In order to understand your risk factors for HIV, we have to ask you some very personal questions. You may be embarrassed but your answers are very important. Knowing your risk factors for HIV may help keep you and others you care about healthier. We encourage you to talk to the medical staff about your concerns and ask any questions you may have. All information is kept strictly confidential. Be sure to answer each question using a check mark.

SECTION I-The questions in this section are about you <u>before</u> you found out you were HIV positive:	***Yes***	***No***	***Do not know***
1. Did you have sex with a male?	❑	❑	❑
2. Did you have sex with a female?	❑	❑	❑
3. Did you use needles to inject heroin, cocaine, steroids or any other drug that was not prescribed by a doctor?	❑	❑	❑
4. The following are currently unlikely ways to get HIV. We would like to know if you have had any of the following happen to you since February, 1985. *Please check all that apply:* ❑ transfusion of blood or blood products ❑ hemophilia or other bleeding disorder ❑ organ/tissue transplant ❑ artificial insemination	❑	❑	❑
5. Did you work in a health care or laboratory setting where you might have been exposed to human blood or other body fluids? If you checked yes, please state your occupation_____	❑	❑	❑
6. How do you think you got infected with HIV? _____			

ONLY PEOPLE WHO HAVE HAD HETEROSEXUAL SEX SHOULD COMPLETE SECTIONS II AND III.

SECTION II-Please answer these questions only if you had a sex partner of the <u>opposite sex</u> before you found out you were HIV positive:	***Yes***	***No***	***Do not know***
7. Before you found out you were HIV positive, did any of your opposite sex partners use needles to inject heroin, cocaine, steroids or any other drug that was not prescribed by a doctor?	❑	❑	❑
8. Women only: Before you found out you were HIV positive, did any of your male sex partners have sex with other men?	❑	❑	❑
9. Before you found out you were HIV positive, did any of your sex partners receive a transfusion of blood/blood products or organ/tissue transplant before they found out they had HIV or AIDS?	❑	❑	❑
10. Before you found out you were HIV positive, did any of your opposite sex partners have hemophilia or any other bleeding disorder?	❑	❑	❑
11. Before you found out you were HIV positive, did any of your opposite sex partners have HIV or AIDS?	❑	❑	❑
12. Before you found out you were HIV positive, were any of your opposite sex partners born outside of the United States? If yes, where_____	❑	❑	❑
13. Before you found out you were HIV positive, did any of your opposite sex partners live or work outside the U.S.? If yes, where_____	❑	❑	❑

SECTION III-Please answer these questions only if you had a sex partner of the opposite sex before you found out you were HIV positive:	***Yes***	***No***	***Do not know***
14. Before you found out you were HIV positive, did you have a sexually transmitted disease, such as chlamydia, gonorrhea, herpes or syphilis?	❑	❑	❑
15. Before you found out you were HIV positive, did you trade money, drugs or gifts for sex?	❑	❑	❑
16. Before you found out you were HIV positive, did you use crack, cocaine or crystal meth?	❑	❑	❑
17. Did you have more than one sex partner in the year before you found out you were HIV positive?	❑	❑	❑

Please fill out the patient information below:

Last name: _____ First name: _____ Middle initial: ___

Date of birth: month __ __ day __ __ year __ __ __ __

Sex: Male _____ Female _____

Are you transgendered? If yes, please check one: Male to female _____ Female to Male _____

Provider Information (to be completed by medical staff)

Provider's Name: _____

Facility Name: _____
(Your current location)

Printed in the USA. April 2005.

PROVIDER-ADMINISTERED QUESTIONNAIRE

The questions below are provided as a guide to obtaining complete risk factor information for HIV/AIDS cases. The guide roughly follows the Centers for Disease Control and Prevention's (CDC) criteria for reporting HIV risk factors. Most of the CDC risk factor criteria are straightforward. In cases of heterosexual contact, the CDC requires that the risk factors of the sex partner be reported. If your patient had heterosexual sex, please ask the questions in all sections.

SECTION I-Questions for all patients before they found out they were HIV positive:	***Yes***	***No***	***Do not know***
1. Did you have sex with a male?	❑	❑	❑
2. Did you have sex with a female?	❑	❑	❑
3. Did you use needles to inject heroin, cocaine, steroids or any other drug that was not prescribed by a doctor?	❑	❑	❑
4. The following are currently unlikely ways to get HIV. I would like to know if you have had any of the following happen to you since February, 1985. *Please check all that apply:* ❑ transfusion of blood or blood products ❑ hemophilia or other bleeding disorder ❑ organ/tissue transplant ❑ artificial insemination	❑	❑	❑
5. Did you work in a health care or laboratory setting where you might have been exposed to human blood or other body fluids? If you checked yes, please state your occupation_____	❑	❑	❑
6. How do you think you got infected with HIV? _____			

SECTION II-Questions only for patients who had a sex partner of the opposite sex before they found out they were HIV positive:	***Yes***	***No***	***Do not know***
7. Before you found out you were HIV positive, did any of your opposite sex partners use needles to inject heroin, cocaine, steroids or any other drug that was not prescribed by a doctor?	❑	❑	❑
8. Women only: Before you found out you were HIV positive, did any of your male sex partners have sex with other men?	❑	❑	❑
9. Before you found out you were HIV positive, did any of your sex partners receive a transfusion of blood/blood products or organ/tissue transplant before they found out they had HIV or AIDS?	❑	❑	❑
10. Before you found out you were HIV positive, did any of your opposite sex partners have hemophilia or any other bleeding disorder?	❑	❑	❑
11. Before you found out you were HIV positive, did any of your opposite sex partners have HIV or AIDS?	❑	❑	❑
12. Before you found out you were HIV positive, were any of your opposite sex partners born outside of the United States? If yes, where_____	❑	❑	❑
13. Before you found out you were HIV positive, did any of your opposite sex partners live or work outside the U.S.? If yes, where_____	❑	❑	❑

SECTION III-Questions only for patients who had a sex partner of the opposite sex before they found out they were HIV positive:	***Yes***	***No***	***Do not know***
14. Before you found out you were HIV positive, did you have a sexually transmitted disease, such as chlamydia, gonorrhea, herpes or syphilis?	❑	❑	❑
15. Before you found out you were HIV positive, did you trade money, drugs or gifts for sex?	❑	❑	❑
16. Before you found out you were HIV positive, did you use crack, cocaine or crystal meth?	❑	❑	❑
17. Did you have more than one sex partner in the year before you found out you were HIV positive?	❑	❑	❑

Please fill out the patient information below:

Last name: _____ First name: _____ Middle initial: ___

Date of birth: month __ __ day __ __ year __ __ __ __

Sex: Male _____ Female _____

Are you transgendered? If yes, please check one: Male to female _____ Female to Male _____

Provider Information

Your Name: _____
(Health Professional who is completing this form)

Facility Name: _____
(Your current location)

Printed in the USA. April 2005.

Appendix D.3

(1) Definition of CDC HIV Risk Factors *(page 1 of 2)*

	Please respond to all Categories		
Before the first positive HIV test or AIDS diagnosis, did this person have:	**Yes**	**No**	**Unk.**
• Sex with a male?			
• Sex with a female?			
Risk Factors for HIV Infection (See reverse side for definitions.)			
Before the first positive HIV test or AIDS diagnosis, did this person have any of the following risk factors for HIV infection? (Respond to all categories):	**Yes**	**No**	**Unk.**
1. Male who had sex with another male			
2. Injected illicit or nonprescribed drug(s).			
Had sex with someone of the opposite sex who:	**Yes**	**No**	**Unk.**
3. Injected illicit or nonprescribed drug(s).			
4. Was a male who had sex with another male (applies to females only).			
5. Was diagnosed with HIV or AIDS after receiving a transfusion of blood or blood components (e.g. platelets), an organ or a tissue transplant, artificial insemination, a clotting factor infusion for hemophilia or another coagulation disorder after February 1985.			
If yes, check which one: blood transfusion blood components organ/tissue transplant artificial insemination clotting factor for hemophilia clotting factor for another coagulation disorder			
6. Was diagnosed with HIV or AIDS but was not known to have any of the above risk factors.			
Before the first positive HIV test or AIDS diagnosis, did this person have any other of the following risk factors for HIV infection? (Respond to all categories):	**Yes**	**No**	**Unk.**
7. Received clotting factor for hemophilia or another coagulation disorder, transfusion of blood/blood components (other than clotting factor), transplant of tissue or organ, or artificial insemination after February 1985.			
8. Worked in a health care or clinical laboratory setting with possible exposure to human blood or other body fluids (specify occupation and exposure): _____			
9. Other exposure to human blood or body fluids (specify): _____			

HIV Risk Factor Definitions — Page 2 (back of page 1)

Risk Factors for HIV Infection	Definition
Before the first positive HIV test or AIDS diagnosis, did this person (have)	
1. Male who had sex with another male.	Sexual contact with another male. This wording is intended to avoid the issue of sexual orientation (whether the man considered himself homosexual, bisexual, "gay," "on the down-low," or basically heterosexual but occasionally having sex with other men). The important consideration here is not how this male perceived himself, but simply whether he had sex with another man. In this context, "having sex" or "sexual contact" means penis-to-mouth, penis-to-anus, or mouth-to-anus contact (but not mouth-to-mouth contact) regardless of which role (insertive or receptive) is played by the male in question. It does not include contact only with skin (not a body orifice). However, if explicit information on whether the man had sex with another man is unavailable, the man may be assumed to have done so if he stated that he was "homosexual" or "bisexual" or described himself with a similar term that implies that he had sex with another man. In addition, male-to-male sex may be inferred if he was diagnosed with any rectal STD (e.g., gonorrhea) before HIV/AIDS diagnosis.
2. Injected illicit or nonprescribed drug(s).	This means receiving an injection, either self-administered or given by another person, of a drug that was not prescribed by a physician for this person. It generally includes illicit drugs used for producing euphoria, but it may also include prescription drugs that were not prescribed (e.g., estrogen, testosterone, anabolic steroids, or human growth hormone). It does not include injection of prescribed drugs (e.g., insulin for treating diabetes). The drug itself is not the source of the HIV infection, but the context of it being taken illicitly (i.e., without a prescription) is likely to be associated with sharing of injection equipment (e.g., syringes, needles, cookers), which can result in transmission of bloodborne pathogens, such as HIV. The case report form does not include a separate question asking if injection equipment was shared, but sharing may be presumed.
Before the first positive HIV test or AIDS diagnosis, did this person have sex with someone of the opposite sex who	
3. Injected illicit or nonprescribed drug(s).	Had heterosexual contact with someone who injected illicit or nonprescribed drugs (as defined above for Risk Factor #2). Sexual contact means contact of a male's penis with the mouth, vagina, or anus of a female. It does not include mouth-to-mouth contact or contact of the penis or mouth with skin (not a body orifice).
4. Was a male who had sex with another male (applies only to females).	Had heterosexual contact with a male who had sexual contact with another male (as defined above for Risk Factor #1).
5. Was diagnosed with HIV or AIDS after receiving a transfusion of blood or blood components (e.g. platelets), an organ or a tissue transplant, artificial insemination, a clotting factor infusion for hemophilia or another coagulation disorder after February 1985.	Had heterosexual contact with a transfusion recipient who was known to have HIV infection or a transplant recipient who was known to have HIV infection, or heterosexual contact with a woman who had artificial insemination and was known to have HIV infection, or a sex partner who received clotting factor for treatment of hemophilia/coagulation disorder and was known to have HIV infection, or receipt of clotting factor by the sex partner may be presumed if he had hemophilia.
6. Was diagnosed with HIV or AIDS but was not known to have any of the above risk factors.	Had heterosexual contact with a person with AIDS or HIV infection but who was not known to have any of the above risk factors listed in items #3 through #5 (injection drug use, male-to-male sex, transfusion, transplant, or clotting factor/hemophilia).
Before the first positive HIV test or AIDS diagnosis this person	
7. Received clotting factor for hemophilia or another coagulation disorder, transfusion of blood/blood components (other than clotting factor), transplant of tissue or organ, or artificial insemination after February 1985.	Receipt of clotting factor for treatment of hemophilia may be assumed if the person had hemophilia. Blood components could include platelets or packed red cells. Person received a transplant of tissue/organs or, if the patient was female, artificial insemination.
8. Worked in a health care or clinical laboratory setting with possible exposure to human blood or other body fluids.	Worked in a health care or clinical laboratory setting as a health care worker. This does not include work as a clerk, secretary, or administrator who does not contact patients or body fluids. It may include a custodian, however, who could be exposed to contaminated materials that have been discarded.
9. Had other exposure to human blood or other body fluids.	This could include other forms of occupational exposure, such as that experienced by a policeman or fireman, or nonoccupational exposures, such as contact with another person's blood or an open wound as a result of providing informal care to a family member/friend or as a result of a fight.

(2) Key to CDC HIV Risk Factor Definitions

KEY TO CDC HIV RISK FACTOR DEFINITIONS

An increasing number of case reports for HIV infected individuals lack risk factor information as defined by the Centers for Disease Control and Prevention (CDC). HIV risk factor information is crucial for:

- Allocating resources
- Monitoring trends in transmission of HIV
- Planning prevention programs and targeting risk factor reduction interventions

This document outlines current requirements for providing complete HIV/AIDS risk factor information.

RISK FACTORS FOR HIV INFECTION MORE THAN ONE MAY APPLY	GUIDELINES Before the first positive HIV test or AIDS diagnosis, did this person have any other of the following risk factors for HIV infection?
Male sex with male	• Sexual contact with another male, whether or not the man identifies himself as "gay" or "bisexual" • Male-to-male sex may be inferred if male was diagnosed with rectal infection or any rectal STD
Injection of non-prescription drug(s)	• May include use of intravenous needles or "skin popping" illegal drugs • May include injection of non-prescribed drugs, (e.g., anabolic steroids or hormones)
Receipt of clotting factor for hemophilia or another coagulation disorder	• Only if received after February, 1985
Receipt of blood/blood components other than clotting factors	• Only if received after February, 1985 • Blood components can include platelets or packed red cells
Receipt of transplanted tissue/organs or artificial insemination	• Only if received after February, 1985
Worked in a health care or clinical laboratory setting with possible exposure to human blood or other body fluids	• Specify occupation
Other exposure to human blood or other body fluids	• Occupational exposure, other than healthcare worker: such as a police officer, firefighter, or emergency medical technician • Non-occupational exposures, such as contact with another person's blood or an open wound as a result of providing information care to a family member/friend or during a fight

To document heterosexual contact you must record at least one of the risk factors of the person's sex partner. The risk factors are listed below.

MORE THAN ONE MAY APPLY

- Intravenous/injection drug user
- Bisexual male (applies to females only)
- Person with hemophilia/coagulation disorder
- Transfusion recipient with documented HIV infection
- Transplant recipient with documented HIV infection
- Person with AIDS or documented HIV infection, where risk factor is not specified

"Unprotected sex", "multiple sex partners" or "sex with a prostitute" alone does not meet the definition for heterosexual contact used here

Printed in the USA. April 2005.

Appendix D.4

Dear Colleague Letter about HIPAA and Public Health Disease Surveillance

Dear Colleague,

Implementation of the Health Information Portability and Accountability Act (HIPAA) has resulted in questions from health care providers regarding what information may be provided to health departments. Such questions have arisen in relation to the reporting of notifiable diseases such as HIV and AIDS. Although HIV/AIDS reporting is name-based in most states, no names are sent to CDC.

The HIPAA Privacy Rule explicitly allows reporting of public health information for the purposes of public health. Therefore, if your state requires reporting of certain diseases (including HIV and AIDS), HIPAA does not supersede that requirement. This is because the state's interest in monitoring rates of notifiable diseases to protect the public health outweighs the privacy rights granted by HIPAA.

HIV and AIDS are reportable diseases in all 65 U.S. states, territories, and protectorates, also known as jurisdictions. U.S. jurisdictions collect this information and forward it to the Centers for Disease Control and Prevention (CDC). Case information, including the risk factor(s) associated with transmission, is vital to CDC's efforts to track the epidemic and construct an accurate picture of the numbers of cases attributable to each known HIV risk factor. Only with such information can CDC precisely target prevention program funds to the groups hardest hit by the epidemic and achieve greater accuracy in evaluation of existing HIV/AIDS prevention programs.

Thank you for your efforts to collect HIV and AIDS case information and forward it to your state or local health department. Please contact [insert name and phone number] at [insert state/city] HIV surveillance coordinator for more information.

Sincerely,

XXXX

Appendix D.5

Dear Colleague Letter about Reporting HIV Infection Risk Factor Ascertainment—To Motivate, Inform, and Educate

Dear Colleague,

In this third decade of the HIV/AIDS epidemic, it is still very important to collect and document the risk factor(s) associated with each case of HIV transmission. Complete risk factor information accompanying each reported case of HIV helps states and CDC accurately and efficiently allocate prevention resources. Yet CDC has noted a decline in risk factor information accompanying states' reported cases of HIV. In fact, the number of cases forwarded to CDC without complete risk factor information has doubled in the past decade. Currently, of the estimated xxx people diagnosed with HIV or AIDS in 2003 in the United States, approximately 40% of HIV cases and 30% of AIDS cases in the National HIV/AIDS Reporting System have no identified risk factor.

In this era of level or declining funding for HIV/AIDS prevention [and treatment?], it is critical that we work more effectively with fewer resources. To do this, we must construct an accurate picture of the numbers of cases attributable to each known HIV risk factor. Only with such information can we precisely target prevention program funds to the groups hardest hit by the epidemic and achieve greater accuracy in evaluating existing HIV/AIDS prevention programs. Complete risk factor information is the critical, but missing, component of the greater efficiency these times demand.

To construct an accurate picture of the epidemic, CDC expects all jurisdictions to have complete HIV risk factor data on at least 85% of all HIV and AIDS cases by completion of follow-up investigation. There are two options for collecting risk factor information. One is to document it yourself on the reportable infectious diseases form approved by your state. The second option is to open your records to someone from the state or local health department who will come in and obtain the information from patient records.

Research with colleagues at the state level as well as with individual providers has shown that the reasons for the increased number of case reports arriving at CDC with incomplete risk factor information are varied. They can range from provider discomfort with asking patients about risk factors to confusion over how CDC defines risk factors or how CDC anticipates providers will fill in the state reportable infectious disease form. For example, research respondents told us it was not clear that the provider is asked to note all risk factors elicited from the patient rather than stopping after eliciting one risk factor. Also, it was not common knowledge that checking “unknown” on the form is preferable, in terms of CDC’s data gathering efforts, to not completing the form in the first place.

To clear up any confusion about documenting HIV risk factors, CDC has created the enclosed reference card. This card defines CDC HIV risk factors and steps you through all the information CDC needs to properly categorize a reported case. It can be stored in a coat pocket, posted on a bulletin board, or made available in any setting where you might find yourself interviewing a patient on HIV risk factors.

As a health care provider, you are in a unique position to obtain the risk factor information vital to a better, clearer understanding of those most affected by the HIV/AIDS epidemic. At the same time, if your practice includes individuals with HIV and AIDS, you may be on the receiving end of funds directed to HIV and AIDS services. Therefore, you not only have the opportunity during patient visits to document risk factors but may also receive the benefits of accurate reporting through monies and services targeted to the populations who need them the most.

[insert state/city jurisdiction information] looks to you for assistance in this challenge. Please join with us to expand and refine our understanding of where the epidemic is going and which groups are most disproportionately affected.

If you have questions or would like to provide feedback on this issue, please contact XXXXX.

Sincerely,

XXXXX

Enclosure

Appendix D.6

Q's and A's about Reporting Individual Risk Factors—For Provider/Field Staff

Q: Why is it important to collect risk factor information on each diagnosed case of HIV or AIDS and forward this information to my state health department?

A: Complete risk factor information accompanying each reported case of HIV helps CDC and the states accurately and efficiently allocate prevention resources. Yet CDC has noted a decline in risk factor information accompanying states' reported cases of HIV. In fact, the number of cases forwarded to CDC without complete risk factor information has doubled in the past decade. In this era of level or declining funding for HIV/AIDS prevention [and treatment?], it is critical that we work more effectively with fewer resources. To do this, we must construct an accurate picture of the numbers of cases attributable to each known HIV risk factor.

Q: If I am interviewing a patient newly diagnosed with HIV or AIDS, and that person answers "yes" to one of the risk factor questions, what should I do?

A: Check the Yes box next to that question. But do not stop there. Continue the interview, working your way through all the risk factor questions to see if you can elicit any other risk factors. In some cases, a patient will present with more than one risk factor. Capturing all the applicable risk factor information associated with each diagnosed case of HIV or AIDS is critical to building an accurate picture of the impact of the epidemic in various communities.

Q: My female patient newly diagnosed with HIV says that she has had multiple heterosexual partners, but she does not know about the drug use or sexual histories of these partners. Would her risk factor be "heterosexual contact"? If not, why?

A: A female with multiple heterosexual partners is not classified as "heterosexual contact" because that does not explain if those sex partners included people who are HIV-positive or who engage in behaviors that put them at higher risk for HIV (including injection drug use or male-male sexual contact). Instead, you would note that the patient has had sex with a male and check the "Unknown" box for all other questions regarding sex partners. Do not leave these boxes blank.

To be classified as a heterosexual contact case, a patient must have had heterosexual contact with one of the following:

- Intravenous/injection drug user
- Bisexual male (for a female only)
- Person with hemophilia/coagulation disorder
- Transfusion recipient with documented HIV infection
- Transplant recipient with documented HIV infection
- Person with AIDS or documented HIV infection, risk factor not specified.

Q: What box do I check if I am interviewing a patient newly diagnosed with HIV or AIDS and that person cannot or will not answer one of the questions on the form (sex with any of the sexes listed on the form; sex with users of illicit or nonprescription drugs; sex with a male who has sex with men; diagnosis with HIV or AIDS after receiving a transfusion, clotting factor, organ transplant, or artificial insemination after March 1985; worked in a health care or laboratory setting; had other exposure to human blood or body fluids)?

A: If the patient cannot or will not answer a question, check the Unknown box next to each category after asking the question. Do not leave any category with all blank boxes. But do not stop there. Continue the interview, asking the patient about any other possible risk factors.

Q: Because of constraints (for example, patient volume, off-site laboratory forwards HIV test results, or discomfort with talking to patients about sex and drug use), I am often not able to get risk factor information on my patients diagnosed with HIV or AIDS. Are there tools that can help me?

A: Yes. Contact your state health department for newly developed tools/materials that were created to help providers document risk factors of newly diagnosed HIV patients.

Q: I have a case of HIV to report to the state, but I don't have risk factor information. What should I do?

A: Your state health department surveillance program may follow up with your office on a case if it is reported without a risk factor. If no risk factor is found after follow-up, the case is called a case without an identified risk factor and reported to the CDC as such.

Q: How do the state health department and CDC use these data? What if I come up with more than one risk factor from a patient newly diagnosed with HIV or AIDS; how are they counted?

A: The health department will tabulate cases to examine the distribution of risk factors among cases. Providers should document all known risk factors. The states and CDC count each person as one case. When CDC receives a case report from a state, it puts all reported risk factors into a hierarchy, a summary of which follows. Even though an individual can have multiple risk factors, he or she will be classified into only one of the following categories. It is this category that is the most likely mode of transmission for a particular person.

The current transmission categories CDC uses are

- Male-to-male sex (MSM)
- Injection drug use (IDU), which includes only nonprescription drugs
- MSM + IDU, which is a category for men who have sex with men and also inject drugs.
- Hemophilia/coagulation disorder
- Heterosexual contact with high-risk individuals
- Blood, blood component, or transplant recipient
- Others/Not reported/not identified

Appendix D.7

Provider Letter to Accompany Tool Kit (All Tools and Materials to Be Packaged and Sent to Providers)

Dear Colleague Letter—cover letter for risk factor information/tools

Dear Colleague,

Over the past decade, CDC has noted a decline in risk factor information accompanying states' reported cases of HIV and AIDS. In fact, the number of cases forwarded to CDC without complete risk factor information has doubled during the past 10 years. In an era of level or declining funding for HIV/AIDS prevention, it is more important than ever to collect and document the risk factor(s) associated with each case of HIV transmission. Only with this information can we accurately apportion prevention funds to the groups hardest hit by the epidemic. We are asking for your renewed efforts to address this important issue and enclosing some tools that may assist.

As you know, HIV and AIDS risk factors are ideally collected by the provider, at the time of diagnosis. CDC recognizes that for a variety of reasons these risk factors are not always ascertained at the time of diagnosis. Research with colleagues at state or local health departments and in the provider community has shown that, in some cases, the missing or incomplete risk factor information is due to confusion about how CDC defines risk factors, confusion about how to document all risk factors on the reportable disease form, or provider discomfort at addressing risk factor-related behavior with patients. To address these potential barriers, we have developed some aids that may help in ascertaining all known risk factors for patients diagnosed with HIV or AIDS.

Enclosed you will find these four aids:

- A reference card, suitable for a coat pocket, posting on a bulletin board, or displaying in any setting where you might find yourself interviewing a patient on HIV risk factors. This card defines CDC HIV risk factors and steps you through all the information CDC needs to properly categorize a reported case.
- A patient-administered questionnaire. If you would rather have patients self-report on risk factor behaviors or self-report as an aid to starting a conversation about risk factor behaviors, this may be useful.
- A provider-administered questionnaire. This may assist in starting, and continuing, the conversation necessary to elicit all the risk factors associated with a case of HIV or AIDS.
- A fact sheet, “Talking with Your Patients about Behavioral Risk Factors for HIV and AIDS.” This fact sheet provides some sound tips for starting and continuing a conversation about risk factor behaviors and completing a case report form for a case of HIV or AIDS.

A common misconception uncovered in our conversations with providers is that it is only necessary to document one risk factor associated with a case of HIV or AIDS. Actually, we want to capture all risk factors associated with a transmission—and there can be more than one. For that reason, when interviewing a patient diagnosed with HIV or AIDS, it is necessary to work through the entire form rather than stopping after eliciting a single risk factor. Even if all the patient's responses to subsequent questions are “No” or “Unknown,” by filling in these boxes you provide helpful and complete information. Further, you may discover additional risk factors as you continue through all the questions on the form. That additional information is of critical importance to us.

As a health care provider, you are in a unique position to obtain the risk factor information vital to a better, clearer understanding of those communities most affected by the HIV/AIDS epidemic. CDC looks to you for assistance in this challenge. Please join with us to expand and refine our understanding of where the epidemic is going and which communities are the most affected.

If you have questions or would like to provide feedback on this issue, please contact XXXXX.

Sincerely,

XXXXX

Enclosures

Appendix E

Table shells (A–G)

Table shell A

Table Shell A. Transmission Category Among HIV/AIDS Cases Reported During the Previous Five Quarters.

	2003			2004	
	Quarter 2	Quarter 3	Quarter4	Quarter 1	Quarter2
Transmission Category	N (%)	N (%)	N (%)	N (%)	N (%)
(01) Men having sex with men (MSM)					
(02) IV drug use (IDU)					
(03) MSM/IDU					
(04) Adult hemophilia/Blood products					
(05) Heterosexual contact					
(06) Adult transfusion/transplant					
(07) Older perinatal					
(08) Adult other					
(09) Adult undetermined					
NIR					
NRR					
(11) Pediatric hemophilia					
(12) Mother with HIV					
(13) Pediatric transfusion/transplant					
(18) Pediatric other					
(19) Pediatric undetermined					
NIR-No identified risk factor					
NRR-No risk factor reported					
Total					

Table shell B

Table Shell B. Exposure Category Among HIV/AIDS Cases Reported During the Previous Five Quarters.

	2003			2004	
	Quarter 2	Quarter 3	Quarter4	Quarter 1	Quarter2
Exposure Category	N (%)	N (%)	N (%)	N (%)	N (%)
(01) Men having sex with men only (MSM)					
(02) IV drug use only (IDU)					
(03) Heterosexual contact only (HTC)					
(04) MSM and IDU					
(05) IDU and HTC					
(06) MSM and HTC					
(07) MSM, IDU, and HTC					
(08) Perinatal exposure					
(09) Other					
(10) NIR					
(11) NRR					
Total					

Table shell C

Table Shell C. Transmission Category Among HIV/AIDS Cases By Non-Laboratory Reporting Facilities Reported During the Past Five Quarters.

		2003			2004	
		Quarter 2	Quarter 3	Quarter4	Quarter 1	Quarter2
Reporting Facility	Transmission Category	N (%)	N (%)	N (%)	N (%)	N (%)
Facility A	(01) Men having sex with men (MSM)					
	(02) IV drug use (IDU)					
	(03) MSM/IDU					
	(04) Adult hemophilia/Blood Products					
	(05) Heterosexual contact					
	(06) Adult transfusion/transplant					
	(07) Older perinatal					
	(08) Adult other					
	(09) Adult undetermined					
	NIR					
	NRR					
	(11) Pediatric Hemophilia					
	(12) Mother with HIV					
	(13) Pediatric transfusion/transplant					
	(18) Pediatric other					
	(19) Pediatric undetermined					
	NIR					
	NRR					
	Total					
		N (%)	N (%)	N (%)	N (%)	N (%)
Facility B	(01) Men having sex with men (MSM)					
	(02) IV drug use (IDU)					
	(03) MSM/IDU					
	(04) Adult hemophilia/Blood Products					
	(05) Heterosexual contact					
	(06) Adult transfusion/transplant					
	(07) Older perinatal					
	(08) Adult other					
	(09) Adult undetermined					
	NIR					
	NRR					
	(11) Pediatric Hemophilia					
	(12) Mother with HIV					
	(13) Pediatric transfusion/transplant					
	(18) Pediatric other					
	(19) Pediatric undetermined					
	NIR					
	NRR					
	Total					

Table shell D

Table Shell D. Exposure Category Among HIV/AIDS Cases By Non-Laboratory Reporting Facilities Reported During the Past Five Quarters.

		2003			2004	
		Quarter 2	Quarter 3	Quarter4	Quarter 1	Quarter2
Reporting Facility	Exposure Category	N (%)	N (%)	N (%)	N (%)	N (%)
Facility A	(01) Men having sex with men only (MSM)					
	(02) IV drug use only (IDU)					
	(03) Heterosexual contact only (HTC)					
	(04) MSM and IDU					
	(05) IDU and HTC					
	(06) MSM and HTC					
	(07) MSM, IDU, and HTC					
	(08) Perinatal exposure					
	(09) Other					
	(10) NIR					
	(11) NRR					
	Total					
		2003			2004	
		Quarter 2	Quarter 3	Quarter4	Quarter 1	Quarter2
		N (%)	N (%)	N (%)	N (%)	N (%)
Facility B	(01) Men having sex with men only (MSM)					
	(02) IV drug use only (IDU)					
	(03) Heterosexual contact only (HTC)					
	(04) MSM and IDU					
	(05) IDU and HTC					
	(06) MSM and HTC					
	(07) MSM, IDU, and HTC					
	(08) Perinatal exposure					
	(09) Other					
	(10) NIR					
	(11) NRR					
	Total					

Table shell E

Table Shell E. Completeness of Risk Factor Reporting Among Adult and Adolescent HIV/AIDS Cases By Non-Laboratory Reporting Facilities Reported During the Previous Five Quarters.

			2003												2004							
			Quarter 2				Quarter 3				Quarter 4				Quarter 1				Quarter 2			
			"Yes"	"No"	"Unk"	Missing	"Yes"	"No"	"Unk"	Missing	"Yes"	"No"	"Unk"	Missing	"Yes"	"No"	"Unk"	Missing	"Yes"	"No"	"Unk"	Missing
Reporting Facility	Risk Factor Variable	Variable Label	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)
Facility A	SEX	Sex of patient																				
	SEX_MALE	Sex with male																				
	SEX_FMLE	Sex with female																				
	IV	Injected nonprescription drugs (IDU)																				
	BLDPRD	Received clotting factor																				
	S_IV	HC with IDU																				
	S_BI	HC with bisexual male																				
	S_HEMO	HC with person with hemophilia																				
	S_TX	HC with HIV+ transfusion recipient																				
	S_TRNPLT	HC with HIV+ transplant recipient																				
	S_HIV	HC with HIV+, risk factor unspecified																				
	TRANSFUS	Received blood or blood components																				
	TRANPLNT	Transplant recipient																				
	HCW	Worked in health-care/laboratory setting																				
	OTH_RISK	Other exposure to human blood/body fluids																				
	Total																					
			Quarter 2				Quarter 3				Quarter 4				Quarter 1				Quarter 2			
			"Yes"	"No"	"Unk"	Missing	"Yes"	"No"	"Unk"	Missing	"Yes"	"No"	"Unk"	Missing	"Yes"	"No"	"Unk"	Missing	"Yes"	"No"	"Unk"	Missing
			N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)
Facility B	SEX	Sex of patient																				
	SEX_MALE	Sex with male																				
	SEX_FMLE	Sex with female																				
	IV	Injected nonprescription drugs (IDU)																				
	BLDPRD	Received clotting factor																				
	S_IV	HC with IDU																				
	S_BI	HC with bisexual male																				
	S_HEMO	HC with person with hemophilia																				
	S_TX	HC with HIV+ transfusion recipient																				
	S_TRNPLT	HC with HIV+ transplant recipient																				
	S_HIV	HC with HIV+, risk factor unspecified																				
	TRANSFUS	Received blood or blood components																				
	TRANPLNT	Transplant recipient																				
	HCW	Worked in health-care/laboratory setting																				
	Total	Other exposure to human blood/body fluids																				

Table shell F

Table Shell F. Adult and Adolescent HIV/AIDS Cases By Non-Laboratory Reporting Facility Reported With Incomplete* Risk Factor Information During the Previous Five Quarters, Facilities are Sorted in Descending Order of the Total Number of Cases with Incomplete Risk Factor Information During the Five-Quarter Time Period.

Reporting Facility	Year of Report	Quarter of Report	Total Number of Cases Reported (C1)	Cases With Incomplete Risk Factor Information at Time of Report (C2)	Status of Epidemiologic Follow-Up Among Cases in C2	
					NIRs	NRRs
			N	N (% of C1)	N (% of C2)	N (% of C2)
Facility A	2003	Quarter 2				
		Quarter 3				
		Quarter 4				
	2004	Quarter 1				
		Quarter 2				
		Total				
			N	N (% of C1)	N (% of C2)	N (% of C2)
Facility B	2003	Quarter 2				
		Quarter 3				
		Quarter 4				
	2004	Quarter 1				
		Quarter 2				
		Total				

*Incomplete risk factor information: any case that does not have a "Yes", "No", or "Unk." documented for each risk factor variable.

Table shell G

Table Shell G: Status of Epidemiologic Follow-Up Among Adult and Adolescent HIV/AIDS Cases By Surveillance Staff Reported With Incomplete* Risk Factor Information During the Previous Ten Quarters.

Surveillance Staff	Year of Report	Quarter of Report	Total Number of Cases Reported (C1)	Cases With Incomplete Risk Factor Information (C2)	Status of Epidemiologic Follow-Up Among Cases in (C2)			
					Cases With Complete Risk Factor Information Within 12 Months of Date of Initial Case Report	NIR, Because All Available Sources Have Been Reviewed/Contacted	NIR, Because 12 Months Have Elapsed Since Date of Initial Case Report	Cases Still Under Investigation (NRR)
			N	N (% of C1)	N (% of C2)	N (% of C2)	N (% of C2)	N (% of C2)
Staff A	2002	Quarter 1						
		Quarter 2						
		Quarter 3						
		Quarter 4						
	2003	Quarter 1						
		Quarter 2						
		Quarter 3						
		Quarter 4						
	2004	Quarter 1						
		Quarter 2						
		Total						
			N	N (% of C1)	N (% of C2)	N (% of C2)	N (% of C2)	N (% of C2)
Staff B	2002	Quarter 1						
		Quarter 2						
		Quarter 3						
		Quarter 4						
	2003	Quarter 1						
		Quarter 2						
		Quarter 3						
		Quarter 4						
	2004	Quarter 1						
		Quarter 2						
		Total						

*Incomplete risk factor information: any case that does not have a "Yes", "No", or "Unk." documented for each risk factor variable.

Appendix F

Training and Dissemination Options

Training is an essential element in developing the skills required for conducting surveillance activities. It is important in establishing and maintaining the expertise necessary for staff and health care providers to perform the required functions for risk factor ascertainment.

The goal is to improve risk factor ascertainment through structured training and dissemination of tools and materials. Training should be routinely conducted to improve and maintain skills proficiency and overall expertise in all surveillance field staff, coordinators, and health care providers performing HIV/AIDS related surveillance activities.

Target Audience

The target audience includes two groups:

1. surveillance field staff and coordinators, and
2. health care providers.

   Health care providers include physicians, nurses, physician assistants, or any other person documenting HIV risk factors. This plan addresses initial training on the revised policies and procedures for HIV/AIDS risk factor ascertainment.

To target health care providers here are some suggestions:

Health Care Provider Training

- Presentations at professional associations/organizations
- National meetings
  - AMA, APHA, IDSA
- State, regional, local meetings
- Cover memoranda that accompany quarterly statistical reports
- Publications in professional journals
  - JAMA, JAIDS, state journals, newsletters
- “Dear colleague” letters
  - Signed by public health officials describing changes/additions to surveillance policies and procedures, including risk factor ascertainment procedures

- Three Training Centers
  - National Network of STD/HIV Prevention Training Centers (PTCs)
    - Funded by CDC.
    - A group of 13 regional centers created in partnership with health departments and universities.
    - Provides increased knowledge and skills for health care providers in the areas of sexual and reproductive health with an emphasis on prevention.
    - http://depts.washington.edu/nnptc/
  - Title X Family Planning Regional Training Centers
    - Funded by the Health Resources and Services Administration (HRSA).
    - Located in 10 DHHS Regional Offices.
    - Dedicated to family planning and reproductive health with a mandate to provide a broad range of acceptable and effective family planning methods and services.
    - http://opa.osophs.dhhs.gov/titlex/ofp-training-grantees-listing.html
  - AIDS Education and Training Centers (AETCs)
    - Funded through HRSA and consisting of 11 regional centers and 130 local performance sites.
    - Provides multidisciplinary education and training programs for health care providers treating persons with HIV/AIDS.
    - Health care providers include physicians, advanced practice nurses, physician assistants, nurses, oral health professionals, and pharmacists.
    - Training activities are based on assessed local needs.
    - Surveillance staff could meet with local AETC and explain risk factors as defined by CDC.
    - http://www.aids-ed.org/
    - Quarterly AETC director's meeting
      - Present information on the importance of risk factor ascertainment or possibly incorporating our message into one of their courses.

The three training centers have an annual meeting.

- Presentation at annual meeting
- Presentation via a quarterly/monthly conference call

The three training centers present an opportunity to integrate risk factor ascertainment materials or key discussion points into their training curriculum. The health care providers decide what courses they will attend, so it may be best to integrate our information into an existing course.

- Integration with Advancing HIV Prevention (AHP) activities
  - The Capacity Building Branch/CDC is developing training materials for AHP activities. We may integrate our message into their materials.
- Presentation at large facility staff meetings
  - If a medical facility is identified as having a large proportion of HIV/AIDS patients, or identified as having a large proportion of cases reported without risk, a presentation of the importance of risk factor ascertainment and current procedures would be suggested.
- Training of newly identified HIV/AIDS health care providers in the importance of HIV risk factor ascertainment and procedures.
- Counseling and Testing sites (CTS)
  - Services provided by state and local health departments with CDC funding.

To target state and local health care providers, we make the following recommendations:

State/local surveillance staff members

- Presentations at Surveillance Workshop
- Integration with Surveillance Guidance training activities
- State/local biannual/annual training
- New employee training
- Ryan White I, II, III programs

Appendix G

Sampling

If a surveillance program chooses to sample, this information should prove helpful.

Choosing a Sampling Frame

The sampling frame is the list of cases from which a sample is drawn. For example, if the target population is all newly reported HIV/AIDS cases in a given calendar year in State X, the frame will be constructed from all HIV/AIDS cases reported to the HIV/AIDS surveillance system during that specific year within State X.

When sampling for risk factor ascertainment, a program may choose to sample from all cases or from only NRR cases. If a program chooses to sample from all reported cases regardless of risk factor ascertainment, its sampling scheme will serve to validate previously recorded risk factor data as well as obtain previously unreported risk factor data. If the sampling frame is restricted to only NRR cases, then all of the sampling resources within the surveillance program will focus on risk factor ascertainment and not data validation. A third sampling option (which is a hybrid of the previous two) is to sample from all cases and only perform epidemiologic follow up on NRR cases. This third option simplifies the construction of the sampling frame. However, for those surveillance programs where the percentage of NRR is relatively low, this approach could yield a sample with no or very few NRRs. Consultation with your CDC Epi TA will help you choose which sampling frame is best for your program.

Simple Random Sampling

Simple random sampling (SRS) is a sampling technique where each member of the sampling frame has an equal chance of being selected. It should be noted that throughout this document, we assume that sampling is done without replacement. In other words, a particular case can appear only once in a given sample.

The advantage of SRS is that it is simple and easy to apply. However, the fact that every case in the population has an equal chance of being selected into the sample may not be acceptable for those surveillance programs that would like to focus their NRR investigations on certain subgroups that compose a small proportion of the NRR population. For example, if a state has an NRR population of 351 cases made up of 74 (21%) whites, 207 (59%) blacks, 63 (18%) Hispanics, and 7 (2%) others, a simple random sample of 100 NRR cases may by chance contain 45 (45%) whites, 30 (30%) blacks, and 25 (25%) Hispanics—a racial/ethnic distribution that would not be representative of the NRR population in that state.

Sample Size Calculation for SRS

Sample size refers to the number of cases to be sampled from the sampling frame. Sample size consideration should be based both on statistical criteria and on available resources. In general, the smaller the sample, the less reliable the resulting estimates of

the population parameters. CDC recommends that a minimum of 30 cases total be sampled. If a surveillance program has less than 30 NRR cases, then risk factors should be ascertained for all NRR cases.

The statistical criteria for determining sample size include the specified level of confidence and the desired level of precision (e.g., being 95% sure that our sample estimate falls within 3% of the true value). Appendix G.1 illustrates an example of sample size calculation for SRS.

Consequently, low response rates can compromise estimates of the population parameters with respect to the prespecified level of confidence and precision. To guard against this, state and local surveillance programs should estimate response rates and inflate sample sizes by the reciprocals of the response rates. For instance, if one estimates the response rate to be 80%, then one would inflate the sample size by a factor of 1/0.8 = 1.25.

Selecting the Sample for SRS

There are several ways a surveillance program can take an SRS from the sampling frame. One approach is to first assign a number from 1 to N to each case in the sampling frame, then pick a sample of n of these numbers by using random numbers generated by a computer or a random number table. Once the numbers are chosen, cases in the sampling frame that correspond to these numbers are selected to form the sample. Appendix G.2 demonstrates how to take a SRS using the random number table.

Surveillance programs may also consult with your CDC Epi TA in developing SAS programs to generate an SRS. The SURVEYSELECT procedure available in SAS version 8 or 8.2 offers a variety of methods for selecting probability-based random samples, including selecting an SRS.

Stratified Random Sampling

If the entire population of interest can be partitioned into mutually exclusive homogeneous subpopulations (called strata), then it is advantageous to generate the sample by selecting an SRS within each stratum independently. This powerful and commonly used technique is called stratified random sampling (StRS). Unlike SRS, StRS allows inferences to be made about HIV/AIDS risk factors at a specified level of reliability and precision for subpopulations as well as for the whole population. StRS has the added benefit of controlling for confounding when the stratification factor is associated both with the ascertainment of risk factors as well as the risk factor itself.

Stratification Variables

If StRS for ascertainment of risk factors is chosen, surveillance programs should consider stratifying on the following variables and/or variables for which they would like to have risk factor estimates:

- Disease Status

  Completeness of risk factor ascertainment differs between persons with HIV infection that has not progressed to AIDS (HIV infection, not AIDS) and those with AIDS. The *HIV/AIDS Surveillance Report* 2002[2] showed that 31% of adult/adolescent HIV (not AIDS) cases reported through 2002 from areas with confidential name-based HIV reporting had "other" (not MSM, not IDU, and not heterosexual contact) or no reported/identified risk factors. The comparable percentage was 13% for adult/adolescent AIDS cases reported during the same time period from these areas.

- Recency of Diagnosis

  Current trends in the epidemic with respect to risk factors are reflected most accurately among those newly infected with HIV. Therefore, it is important to identify cases newly diagnosed (an approximation for those newly infected) with HIV/AIDS for monitoring recent trends in transmission and for developing timely and effective prevention strategies. Cases newly reported to the HIV/AIDS surveillance system can be classified into one of two strata: 1) recently diagnosed cases and 2) prevalent cases. A recently diagnosed case could be defined as a diagnosis made within a short time (e.g., 6 to 12 months) before the report date.

- Sex

  The distribution of risk factors differs by sex. As an example, the exposure category "men who have sex with men (MSM)" does not apply to women. Furthermore, a greater proportion of women are infected through heterosexual contact than by injection drug use, whereas the reverse is true among men.

- Race/Ethnicity

  Ascertainment of risk factors for reported HIV/AIDS cases has differed by racial/ethnic group. The *HIV/AIDS Surveillance Report* 2002[2] reported that 4% of white (non-Hispanic) male adult/adolescent AIDS cases reported through December 2002 in the United States had "other" (not MSM, not IDU, and not heterosexual contact) or no reported/identified risk factors. The comparable percentages were 14% and 10% among black (non-Hispanic) and Hispanic male adult/adolescent AIDS cases reported during the same time period, respectively.

- County of Residence at Diagnosis

  To guide the allocation of state and local funding for prevention services, estimates at the county level are often desired. To ensure adequate

representation of NRR cases from all areas, states may consider stratifying counties into three morbidity strata: high, medium, and low.

Surveillance programs may consider excluding from the sampling frame cases for which information on the selected stratification variables is missing. However, programs should examine the distribution of exclusions by strata to be aware of their effect on the ability to generalize the results. If the proportion of cases excluded varies substantially by level of the stratification variables and the number excluded is a sizeable proportion of the total in a given stratum, then the stratification scheme and/or exclusion criteria should be modified.

Number of Strata

Programs may wish to consult with CDC in defining the desired number of strata. Although using more strata ensures more complete coverage of the population of interest, a large number of strata may not be beneficial because they may result in strata that are too small to yield reliable rate estimates (small denominators).

For some states, when choosing the probability that a county will be included in a sample, it may be desirable to stratify the counties by morbidity, so that those with the highest morbidity would be selected with certainty (assigned a probability of 1) because they would be considered counties that must be represented to get an accurate overall picture of what is happening in the state. In that case, each of the few high-morbidity counties may be considered as separate stratum.

In general, it is recommended that if only an overall estimate is desired, little is gained in the precision of the estimate by having more than five or six strata. If estimates are wanted for subsets of the population, a larger number of strata may be necessary. Appendix G.3 contains an example that illustrates the process of deciding the desired number of strata.

Sample Size Calculation for StRS

Statistical criteria for sample size determination in StRS are similar to those discussed for SRS.

Once the strata are specified and the total number of cases, n, to be sampled is known, an important step in StRS is to determine how many cases should be taken from each stratum under the constraint that a total of n is to be selected from all strata. The four methods in which the sample can be allocated to strata are

1. equal allocation,
2. proportional allocation,
3. Neyman allocation, and
4. optimal allocation.

In equal allocation, the same number of cases are sampled from each stratum. Under proportional allocation, the sample is allocated to the strata based on the proportion of the sampling frame that each stratum represents. With the Neyman allocation approach, the sample is distributed proportional to the standard deviation of the random variable in the strata. If the cost of sampling is not the same for each stratum, optimal allocation allows allocation proportional to the relative cost by variability in the strata. You should consult with your CDC Epi TA to decide the most appropriate method of allocation for your surveillance program. Appendix G.4 illustrates an example for sample size calculation in StRS with proportional allocation.

Similar to sample size determination for SRS, state and local surveillance programs should estimate response rates within each stratum and inflate sample sizes by the reciprocals of the response rates.

Selecting the Sample for StRS

CDC has developed SAS programs to generate StRSs from cases reported to the HIV/AIDS surveillance system. The SAS program uses the SURVEYSELECT procedure available in SAS versions 8 or 8.2. The program, designed to assist state and local surveillance programs with selecting samples for the Ascertainment of Transmission Risk (ATR) Project, is available upon request from your CDC Epi TA. CDC can review the program and make necessary modifications to fit the specific needs of the states. The state or local surveillance program, in consultation with CDC, will generate the sample of cases for each stratum using the SAS program developed by CDC.

References

1 CDC. Guidelines for national human immunodeficiency virus case surveillance, including monitoring for human immunodeficiency virus infection and acquired immunodeficiency syndrome. *MMWR* 1999;48(RR13):1–28.

2 CDC. HIV/AIDS surveillance report 2002;14:[A2].

3 Cochran WG. *Sampling techniques*. 3rd edition. New York: John Wiley & Sons; 1977.

4 Levy P.S., Lemeshow S. *Sampling of populations: methods and applications*. 3rd edition. New York: John Wiley & Sons; 1999.

Appendix G.1

Point Estimates and Sample Size Calculation for Simple Random Sampling (SRS)

Let us suppose that State X wishes to perform epidemiologic follow-up on an SRS of adult/adolescent HIV/AIDS cases reported without risk factor (NRR) information in 2003. Their goal is to estimate the proportion of injection drug users (IDUs) among the NRR cases.

Sampling Frame:	Adult/adolescent HIV/AIDS NRR cases reported by State X between January and December 2003.
Sampling Method:	SRS

We shall use the following notation throughout this example and assume from this point on that we are referring to adult/adolescent HIV/AIDS cases reported during the specified time period.

- N = the total number of NRR cases
- n = the number of NRR cases to be sampled
- d = the number of sampled NRR cases that are determined to be IDUs after epidemiologic follow-up
- P = the true proportion of NRR cases that are IDUs

Then the point estimate for P is

$$p = \frac{d}{n}$$

and

$$var(p) = \frac{N-n}{N-1} \times \frac{P(1-P)}{n}$$

where $\frac{N-n}{N-1}$ is the finite population correction (fpc) for the variance. A 95% confidence interval for *P*, based on asymptotic normality, is

$$p \pm 1.96\sqrt{var(p)}$$

Let $\delta = 1.96\sqrt{var(p)}$, δ is called the precision or margin of error. Using the formula for *var(p)*,

$$\delta = 1.96\sqrt{\frac{N-n}{N-1} \times \frac{P(1-P)}{n}}$$

To calculate sample size based on a specified precision, we simply solve the δ-formula for *n*.

$$n = \frac{1.96^2 NP(1-P)}{\delta^2(N-1) + 1.96^2 P(1-P)}$$

Note that in the absence of knowledge about *P*, we assume *P*=0.5, which results in the most conservative (largest) sample size necessary.

The table below shows frequencies of adult and adolescent HIV/AIDS cases reported in 2003 in State X by transmission category. Data are based on June 2004 data reported to the HIV/AIDS surveillance system.

Transmission Category	Frequency	Percent	Cumulative Frequency	Cumulative Percent
MSM	467	26.2	467	26.2
IDU	212	11.9	679	38.0
MSM/IDU	63	3.5	742	41.6
Other	7	0.4	749	41.9
HTC	239	13.4	988	55.3
NRR	798	44.7	1,786	100.0

With the total number of NRR cases $N = 798$, the next table shows sample sizes required for various designated level of precision, δ, assuming 95% confidence limit. The population proportion, P, is assumed to be 0.5.

Precision, δ	Sample Size, n	Sample Size, n, Inflated for 20% Non-Response
0.02*	599	719
0.025	525	631
0.03	457	548
0.035	396	475
0.04	343	411
0.045	298	357
0.05	261	311

* Sample estimate is within 2 percentage points of the true value.

Appendix G.2

Taking a Simple Random Sample (SRS) Using the Random Number Table

The following example illustrates how to take an SRS using a random number table. Suppose State X has 313 adult and adolescent HIV/AIDS cases reported with no risk factor (NRR) information between July 1, 2003, and June 30, 2004. State X decided to randomly select 157 (50%) of the 313 NRR cases to perform epidemiologic follow-up to ascertain risk factors.

The first step in selecting an SRS is to assign a number from 1 to 313 to each of the 313 NRR cases. The next step is to pick a sample of 157 numbers using a random number table. To do this, we need to ensure that the numbers selected from the table are all different and none are equal to 0 or greater than 313.

Using the random number table in Table A.1 in reference 4, starting at some arbitrary point (say row 3) and taking three-digit numbers (first three digits in column 1), we proceed down the column and take different numbers between 001 and 313 until 157 such numbers are selected. In this case, we get 241, 289, 099, 103, 071, 023, 010, 070, 293, 024, 296, 007, 053, 005, 259, 097, 179, 145, 089, 156, 050, 069, etc. Note, when we come to the bottom of a particular column, we proceed to the top line of the very next column on the same page.

To form the sample of 157 NRR cases for risk factor investigation, we take from the 313 NRR cases the 241st, the 289th, the 99th, the 103rd case, and so on.

Appendix G.3

Desired Number of Strata

The following example illustrates construction of the number of strata. Suppose State X has decided to use the stratified random sampling approach to sample from all adult and adolescent HIV and AIDS cases with no risk factors or with a heterosexual contact risk factor reported in 2003. State X will consider county of residence at diagnosis, disease status, recency of diagnosis, and sex as possible stratification variables. There are 5 counties in State X, A, B, C, D, and E, and the proportion of cases in County E is much smaller relative to those in the remaining counties. State X combined disease status and recency of diagnosis into one variable, namely Epi_Dx. Epi_Dx categorized HIV (not AIDS) cases diagnosed between June 1, 2002, and December 31, 2003, as recently diagnosed and cases diagnosed before June 1, 2002, as prevalent cases; these same date parameters were used for reported AIDS cases.

Stratification Variables	Number of Levels	Values
County	5	• A • B • C • D • E
Epi_Dx	4	• Recently diagnosed HIV (non-AIDS) • Prevalent HIV (non-AIDS) • Recently diagnosed AIDS • Prevalent AIDS
Sex	2	• Male • Female

The following are 5 possibilities for stratification schemes:

Scheme	Stratification	Number of Strata
1	(County) by (Epi_Dx) by (Sex)	40
2	(County) by (Recently diagnosed HIV, Prevalent HIV) by (Sex) + (All AIDS) by (Sex)	22
3	(County E, Other Counties) by (Recently diagnosed HIV, Prevalent HIV, All AIDS) by (Sex)	12
4	(County E, Other Counties) by (Recently diagnosed HIV, Prevalent HIV) by (Sex) + (All AIDS) by (Sex)	10
5	(Recently diagnosed HIV, Prevalent HIV, All AIDS) by (Sex)	6

Scheme 1 has a large number of strata. This increases the probability that a number of strata will have small cell sizes and consequently increases the potential for requiring the entire stratum to be analyzed instead of only a sample of it. Furthermore, estimates based on small cell sizes are unreliable, which is a major drawback to having a highly stratified design.

Scheme 2 reduces the number of strata somewhat, but the number is still considerable and probably unmanageable. Recently diagnosed and prevalent AIDS cases are collapsed into one category because of the small number of newly reported prevalent AIDS cases. This scheme reflects the desire to allow the distribution of risk factors among HIV (non-AIDS) cases to guide the allocation of funding for prevention at the county level.

The number of strata in Scheme 3 begins to approach a manageable number. This scheme collapses all counties except County E into one category, but allows for oversampling in the County E stratum to ensure adequate number of cases from County E.

Scheme 4 refines Scheme 3 in a manner similar to the way in which Scheme 2 refines Scheme 1.

Scheme 5 offers a very manageable number of strata. This proposal would require that county information be collected on all the sampled cases. Post-stratification (stratification after sampling) by county would then allow estimation at the county level. However, this does not address the problem of a county with a low number of cases (County E). Inference for that county may be difficult because of the small size of the sample in that stratum.

Appendix G.4

Point Estimates and Sample Size Calculation for Stratified Random Sampling (StRS)

Let us suppose that State X wishes to estimate the proportion of injection drug users (IDUs) among adult/adolescent HIV/AIDS cases reported in their state between 1999 and 2003. They would like the estimates to have a certain level of precision at the sex level. To do this, State X decided to perform epidemiologic follow-up on an StRS of adult/adolescent cases reported without risk factor (NRR) information and with heterosexual contact (HTC) between 1999 and 2003.

Sampling Frame:	Adult/adolescent HIV/AIDS NRR and HTC cases reported by State X between January 1, 1999, and December 31, 2003.
Sampling Method:	StRS, with precision at the sex level
Stratification Variable:	Sex (male, female)

We shall use the following notations throughout this example and assume from this point on that we are referring to adult/adolescent HIV/AIDS cases reported during the specified time period. Note that point estimates and sample size derivations are only provided for males, as the same methods apply for females.

- N = the total number of males
- N_1 = the number of NRR and HTC males
- D = the number of males reported with IDU risk factor
- n_1 = the number of males to be sampled from N1
- d_1 = the number of sampled males that are determined to be IDUs after epidemiologic follow-up
- P = the true proportion of male IDUs
- P_1 = the true proportion of IDUs among NRR and HTC males
- p_1 = the estimated proportion of IDUs among NRR and HTC males

Then the estimate for P is

$$p = \frac{D + N_1 \times p_1}{N} = \frac{D + N_1(d_1/n_1)}{N}$$

and (to simplify matters, the fpc is ignored)

$$var(p) = \left(\frac{N_1}{N}\right)^2 \frac{P_1(1-P_1)}{n_1}$$

A 95% confidence interval for *P*, based on asymptotic normality, is

$$p \pm 1.96\sqrt{var(p)}$$

Let $\delta = 1.96\sqrt{var(p)}$, δ is called the *precision* or *margin of error*. Using the formula for *var(p)*,

$$\delta = 1.96\left(\frac{N_1}{N}\right)\sqrt{\frac{P_1(1-P_1)}{n_1}}$$

To calculate sample size based on a specified precision, we solve the δ-formula above for n_1.

$$n_1 = \left(\frac{N_1}{N}\right)^2 \frac{P_1(1-P_1)1.96^2}{\delta^2}$$

Note that in the absence of knowledge about P_1, we assume P_1=0.5, which results in the most conservative (largest) sample size necessary.

The table that follows shows frequencies of adult/adolescent HIV/AIDS cases reported between 1999 and 2003 in State X by sex, race/ethnicity, and transmission category. Data are based on June 2004 data reported to the HIV/AIDS surveillance system.

		Transmission Category						
Sex	Race/Ethnicity	MSM	IDU	MSM/ IDU	Other	HTC	NRR	Total
Male(M)	White (W)	1,001	424	91	34	180	446	2,176
	Black (B)	1,073	1,556	130	34	867	1,405	5,065
	Hispanic (H)	564	590	44	9	319	434	1,960
	Other (O)	39	15	2	2	16	36	110
	Unknown	7	3	1	0	3	54	68
	Total	2,684	2,588	268	79	1,385	2,375	9,379
Female(F)	White (W)		267		6	296	189	758
	Black (B)		914		58	1,447	1,028	3,447
	Hispanic (H)		189		20	390	213	812
	Other (O)		2		2	18	18	40
	Unknown		0		0	5	24	29
	Total		1,372		86	2,156	1,472	5,086

With N_1 = 3,760 and 3,628 for males and females, respectively, and N = 9,379 and 5,086 for males and females, respectively, the next table shows sample sizes required for various designated level of precision, δ, at the sex level, assuming 95% confidence limit. The population proportion, P_1, is assumed to be 0.5.

Precision, δ	Sample Size, n_1	
	M	F
0.02*	386 (463**)	1,222 (1,466)
0.025	247 (296)	782 (938)
0.03	172 (206)	543 (652)
0.035	126 (151)	399 (479)
0.04	96 (116)	305 (367)
0.045	76 (91)	241 (290)
0.05	62 (74)	195 (235)

* Sample estimate is within 2 percentage points of the true value, at the sex level.

** Numbers in parentheses are sample size inflated for 20% nonresponse.

Let us further suppose that State X wants estimates at the race/ethnicity level as well, with precision remaining at the sex level. Then the sample size in the previous table can be proportionally allocated to race/ethnicity categories. Under this type of sampling scheme, it is recommended that 1) cases with unknown race/ethnicity be excluded from the sampling frame, 2) 100% sampling be implemented in the race/ethnicity=other category, and 3) a minimum of 30 cases be sampled.

The table below shows sample sizes for sex-race/ethnicity categories at varying levels of precision, δ, assuming a 95% confidence limit. The population proportion, P_1, is assumed to be 0.5 to maximize sample size. Precision is calculated at the sex level based on frequencies with race=unknown excluded. Total sample size for each sex category is proportionally allocated to race/ethnicity categories, with the exception of race/ethnicity=other in which 100% sampling is recommended.

Precision, δ	Sample Size							
	MW	MB	MH	MO	FW	FB	FH	FO
0.02*	77	280	93	52	197	1,004	244	36
0.025	49	179	59	52	126	642	156	36
0.03	34	124	41	52	87	446	109	36
0.035	30	91	30	52	64	328	80	36
0.04	30	70	30	52	49	251	61	36
0.045	30	55	30	52	39	198	48	36
0.05	30	45	30	52	31	161	39	36

* Sample estimate is within 2 percentage points of the true value, at the sex level.

Suppose State X wants estimates at the race/ethnicity level, but with precision at the sex-race/ethnicity level. The table below shows sample sizes for sex-race/ethnicity categories at varying levels of precision, δ, assuming a 95% confidence limit. The population proportion, P_1, is assumed to be 0.5 to maximize sample size. Precision is calculated at the sex-race/ethnicity level. Sample sizes are inflated for 20% nonresponse and * indicates 100% sampling.

Precision, δ	Sample Size								
	MW	MB	MH	MO*	FW	FB	FH	FO*	Total
0.02	238	580	425	52	485*	1,485	603*	36	3,904
0.025	153	371	272	52	485*	951	603*	36	2,923
0.03	106	258	189	52	485*	660	603*	36	2,389
0.035	78	189	139	52	385	485	519	36	1,883
0.04	60	145	106	52	295	371	397	36	1,462
0.045	47	115	84	52	233	293	314	36	1,174
0.05	38	93	68	52	189	238	254	36	968

* Sample estimate is within 2 percentage points of the true value, at the sex-race/ethnicity level.

Comparing the two previous tables, one notices that larger sample sizes are required if precision is preferred at a finer level (at the sex-race/ethnicity level as compared with at the sex level only).

Notes

Technical Guidance for HIV/AIDS Surveillance Programs

Death Ascertainment

HIV Incidence and Case Surveillance Branch
Atlanta, Georgia

Contents — Death Ascertainment

Introduction 4-3
Sources of Death Information 4-3
Data from Health Care Providers 4-3
Information from Death Certificates 4-3
Data to Be Collected from Death Certificate Records (DCRs) 4-4
Structural Requirements 4-5
Process Standards 4-5
Linkage of HIV Case Records with Death Certificate Records (DCRs) 4-5
Frequency of Linkage of HIV/AIDS Case Records to Death Case Records . 4-8
Computer Log 4-8
Data Exchange 4-9
Outcome Measures 4-10
Proportion of Death-Certificate-Only (DCO) Cases 4-10
Proportion of Cases for Which the Initial Source of Report Was a DCR 4-10
Effectiveness of the HIV-Specific Record Linkage in Finding Cases 4-10
Delay between Death and Report of Death 4-10
Yield from Searching the NDI 4-11
Appendix 4-15
Searching National Death Databases 4-15
Technical Details of Searching the DMF 4-15
National Death Index (NDI) 4-16

Technical Guidance for HIV/AIDS Surveillance Programs — Death Ascertainment

Introduction

The following are policies and procedures for HIV/AIDS surveillance programs on how to collect data on the deaths of persons with HIV/AIDS. The recommended method of ascertaining death data on persons diagnosed with HIV/AIDS is the linking of death certificate records (DCRs) and HIV/AIDS case records, which allows surveillance programs to

- update vital status and
- identify unreported HIV/AIDS cases.

Sources of Death Information

Data from Health Care Providers

Surveillance programs may obtain death data by requesting that health care providers report the deaths of patients previously reported with HIV infection or AIDS or by having surveillance staff abstract death information from medical records. Instructions for completing the relevant data elements on the case report form are in Instructions for Completing the Data Collection Form. However, these methods do not provide reliable ways for obtaining death information and do not eliminate the need to obtain death certificate information from the state Office of Vital Statistics. Deaths may not be reported by health care providers, or death information may not be included or complete (e.g., the International Classification of Diseases (ICD) coded cause of death may be missing) in medical records.

Information from Death Certificates

State Vital Statistics Records

The state vital statistics records (SVSRs) are the primary source of death certificate information. This information should be obtained through electronic record linkage.

National Databases with Death Information

In many instances, state death records will not have information for persons with HIV infection who moved to and/or died in another state. The following data sources are available to ascertain death information for such cases:

- Social Security Administration's Death Master File (DMF)
- National Death Index (NDI)
- National Death Index Plus (NDI-Plus)

The information obtained from these sources, the advantages of each source, and the limitations of these data sources are listed in Table 1; additional information on how to access these sources is provided in the Appendix.

Data to Be Collected from Death Certificate Records (DCRs)

The data items to be collected from DCRs are listed below. The list represents data items that are required information for each case if present on the DCR as well as information needed to conduct follow-up investigations of deaths related to HIV/AIDS among persons not previously known to the surveillance program. For information on what data items are sent to CDC, see *Instructions for Completing the Data Collection Form.*

- Full name (last, first, and middle or initial)
- Social Security number
- Date of death
- Causes of death, both underlying and multiple
- Sex
- Race
- Ethnicity
- Date of birth (month, day, and year)
- Place of death (state, city, county, facility, type [e.g., hospital, residence]) (Note: Identification of a health care facility may be important if follow-back is needed.)
- Residence at death (street address, city, county, state, ZIP Code)
- Death certificate identification number (important for interstate communication or further searching for record linkages to obtain the International Classification of Diseases [ICD] codes for causes of death in records identified only by death certificate number)
- State or country of birth
- Mother's name if the decedent is a child (to allow searching for the mother's records, but do not send to CDC)
- Certifier (name, address, and license number, to allow follow-back if needed; do not send to CDC)
- Age (in years) at death (not necessary to record because it should be calculated by the software from the date of birth).
- Socioeconomic information (optional)
  - Education
  - Occupation

Note: A specially trained nosologist should select the underlying cause of death from among multiple causes that may be reported on a death certificate record. In making this selection, the nosologist should follow the standardized rules established by the World Health Organization with input from the National Center for Health Statistics (NCHS). The rules are part of the International Classification of Diseases (ICD) (ICD-9 or ICD-10). The rules may be assumed to have been followed reliably if the ICD codes for underlying causes were in an electronic data set covering a full year of data received from the state Office of Vital Records or from the National Death Index. Underlying cause data should not be assumed to be reliable, however, if the information was entered by an HIV/AIDS surveillance staff member based on inspection of text on hard copies of death certificates. A clear distinction must be made between these reliable and unreliable sources of data. Only those sources considered reliable should be used in the analyses of causes of death described in subsequent sections.

Structural Requirements

The structural requirements for conducting death ascertainment are accessibility to death certificate records and the ability to electronically link HIV cases to the death certificate records. Electronic linking of records is described in Data Management, Record Linkage.

Process Standards

Linkage of HIV Case Records with Death Certificate Records (DCRs)

From the process of linking HIV/AIDS cases to the state death certificate records, three categories of linked and nonlinked records will be created (see Figure 1.). The first category, linked cases, includes records that contain information from both databases (HIV case and DCR information). The second category, nonlinked cases from the HIV/AIDS surveillance system, consists of records that contain HIV case information only (that is, they did not link with any DCRs). The third category, nonlinked DCRs, consists of records that contain death certificate information only. Determination of whether records should be linked is described in Data Management, Record Linkage. The follow-up procedures for each of these categories are described next.

Linked Cases

When death certificate records are accepted as true links to previously reported HIV/AIDS cases, the information (see data items listed previously in Data to Be Collected from Death Certificate Records (DCRs)) is imported and stored in death documents linked to the respective cases. At a minimum, the date of death, place of death, and causes of death are obtained in addition to the identifying information used to conduct the record linkage.

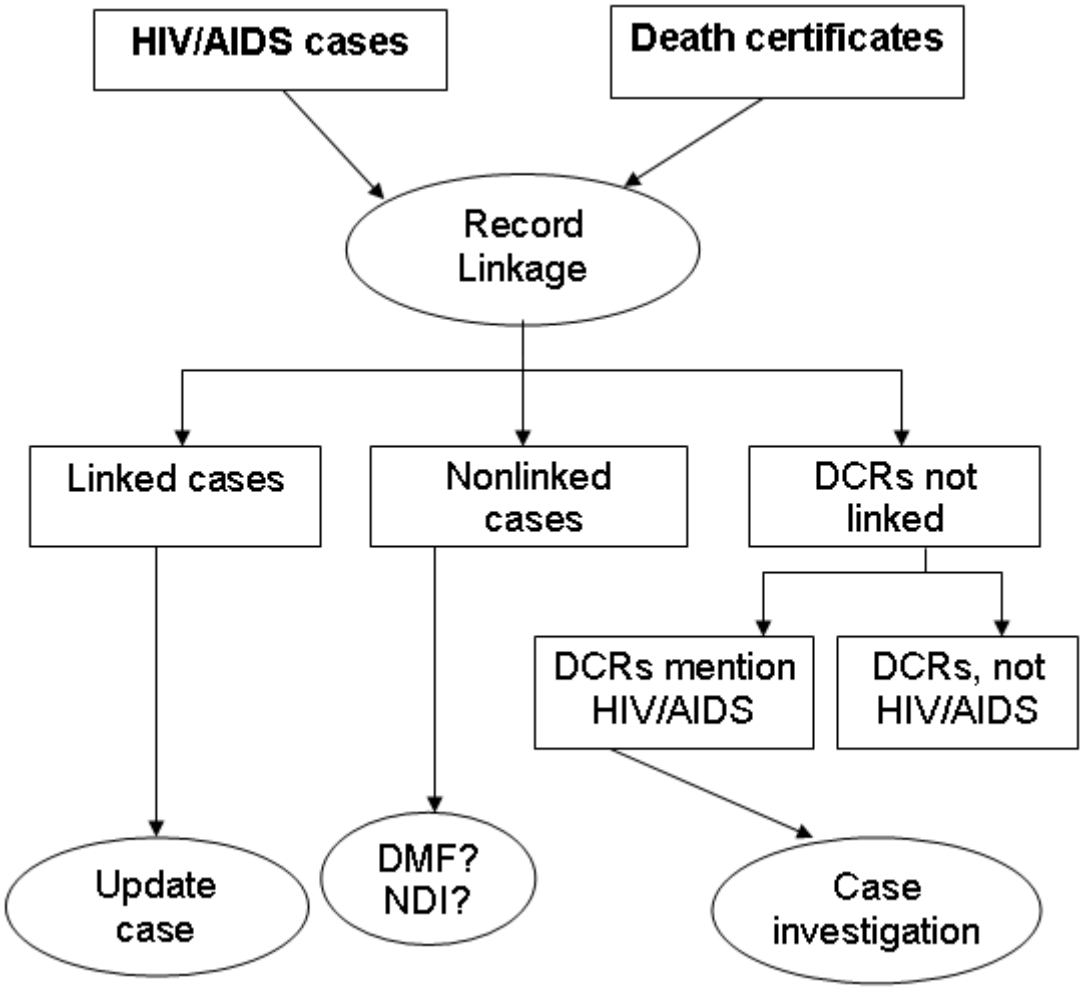
Nonlinked Cases

HIV/AIDS case records that did not link with any DCRs may belong to persons who are still alive. To further assess vital status for these cases, surveillance programs have the following options: conduct a search in another database (e.g., NDI-Plus, see National Databases with Death Information) and/or link again the following year to the state's DCRs.

Nonlinked Death Certificate Records

The DCRs that did not link with HIV/AIDS cases must be reviewed to check for mention of HIV/AIDS as described next to identify any potential HIV cases. This process will divide these records into DCRs that mention HIV/AIDS and those that do not.

Figure 1. Linkage between HIV/AIDS cases and death certificate records (DCRs)



DCR Not Linked, but Mentions HIV/AIDS

Death certificate records that do not link with the HIV/AIDS case records, but mention HIV infection as a cause of death, may be identified by referencing the ICD codes in Table 2. A Statistical Analysis System (SAS) program to identify DCRs with the pertinent ICD codes can be obtained from CDC, including a program to convert the underlying causes determined from ICD-9 rules (for deaths before 1999) to the underlying causes as they would have been determined from ICD-10 rules. Deaths that mention HIV/AIDS are entered in the surveillance software, and follow-up should be done on these cases as described subsequently.

Note: Records may be identified as those in which any of the data fields for causes of death (underlying cause or up to 20 entity-axis multiple causes of death) have an ICD-9 code in the range of 042.0–044.9 (most commonly 042.9) or 795.8, or an ICD-10 code in the range of B20.0–B24 (most commonly B23.8) or R75. The yield of unreported cases is not likely to be high enough to justify searching for a link to death certificates that do not have a code for HIV/AIDS itself but contain only a code for an opportunistic disease consistent with AIDS (Table 2). If the surveillance program chooses to search for links to death certificates that mention such diseases, the most useful disease to use would be pneumocystosis (ICD-9 code 136.3 or ICD-10 code B59). In most Offices of Vital Records in the United States, ICD-10 codes have been used for deaths that occurred in 1999 or later, and ICD-9 codes have been used for deaths that occurred during 1979–1998. Codes specific for HIV infection (042.0–044.9) were not introduced until data year 1987. In the absence of a code specific for HIV infection for deaths before 1987, the underlying cause of deaths due to AIDS was commonly indicated by the ICD-9 code for immunodeficiency (279.1) or any of the AIDS-defining opportunistic diseases (e.g., 136.3 for pneumocystosis, 173 for Kaposi's sarcoma). Because there was no code specific for HIV infection, it is not essential for the program to collect data on underlying causes for deaths that occurred before 1987. In most ASCII text files of death certificate data, the decimal point in the ICD code may not appear, but may be implied (i.e., a decimal point may be assumed to exist between the 3rd and 4th digits of the code [unless the 4th digit does not exist, as in code B24], but the decimal point may not actually appear in the data file). For example, ‘B238’ in the data file would mean ‘B23.8’.

Follow-up—The surveillance program should conduct follow-up investigations of such nonlinked DCRs that mention HIV/AIDS by contacting health care providers (identified as the facility where death occurred or the physician/coroner who signed the death certificate [the “certifier”]) or by reviewing their medical records to determine if the death certificate records are truly for unreported HIV/AIDS cases. If they are, the surveillance program should collect all data required for an HIV/AIDS case as well as for the death.

Determination of whether death certificate records represent HIV or AIDS cases should be resolved within 3 months after the search for linkages by appropriate follow-back to medical records and health care providers to confirm diagnoses and ascertain diagnosis dates.

Cases Based on Death Certificates Only—The question arises whether it is appropriate to count HIV/AIDS cases based solely on death certificate information. If laboratory documentation of HIV infection were obtained from another source, the ICD codes on the death certificate record could provide sufficient information for those AIDS-defining diseases for which a presumptive diagnosis would meet the criteria of the AIDS case definition Table 3. Laboratory documentation or some other evidence that the person was HIV-infected should be obtained from follow-up investigation so a death with an HIV ICD code may be counted as a case of HIV infection. In the absence of laboratory documentation of HIV infection it is unknown whether some of the AIDS indicator diseases that may also be listed on the death certificate constitute definitive or presumptive diagnoses, and an AIDS diagnosis could not be inferred solely from a death certificate in these instances without additional information from other sources. In the absence of evidence of HIV infection other than the death certificate, knowledge about local circumstances (e.g., coroner) should guide the decision on whether to count this as an HIV case. Death certificates not designated as pertaining to HIV cases should be retained in the surveillance system because follow-up information that confirms the HIV infection may become available later.

No Record Linkage to HIV/AIDS Case or Mention of HIV/AIDS on DCR

No action required.

Frequency of Linkage of HIV/AIDS Case Records to Death Case Records

First Time Record Linkage

- All cases and all death years available

Routine Record Linkage

- Once a year (minimum)
- HIV/AIDS case reports (all cases without minimum death information) and DCRs (most recent year of deaths available).

Computer Log

To facilitate and evaluate the processing of recent death certificate records, the surveillance program may generate a computer log during the record linkage and follow-up process that would include the date each record was received by the surveillance program, the date the HIV/AIDS surveillance database was searched for a record link to it, the date it was accepted as a true link or, if not, the date that follow-back investigation was completed to determine if it was a newly found HIV/AIDS case. This log for recent deaths may include some that, after investigation, are found not to be HIV/AIDS cases. Cases whose residence at death was in another jurisdiction (e.g., another state) should be reported to the surveillance program for that jurisdiction, and the date of this report should also be recorded in the log.

Data Exchange

Exchange of Data between HIV/AIDS Surveillance Programs (Out of State)

HIV/AIDS surveillance programs should develop policies and procedures for exchanging death data with different states (see CSTE Statement). If a surveillance program discovers during the review of DCRs that the decedent's state of residence at death was another state (state B), the surveillance program of the state where death occurred (state A) should forward the information to the program of the state of residence at death (state B). That surveillance program for the state of residence at death (state B) may already have the case in its database but be unaware of the death. If that program (state B) does not yet have the case in its database, the surveillance program for the state where death occurred (state A) should collect relevant information (e.g., diagnosis date, risk factors) available from the facility where death occurred or from the certifier of death to create the initial case record and forward the information to the surveillance program of the person's state of residence at death.

Exchange of Data between Offices of Vital Records (Out of State)

Record linkage to the state death certificate database may include records for deaths of residents of the program's jurisdiction who died in other states, and whose death certificate data were forwarded from the Offices of Vital Records of those other states to the Office of Vital Records of the jurisdiction of the decedents' residence. Such deaths may be indicated by a special out-of-state identifier, but the record may simply say that the decedent died in another state without identifying the other state. Surveillance programs should find out which of the state death certificate databases (underlying cause, multiple causes, or other) contain information on out-of-state deaths. However, such exchange of death data between the Offices of Vital Records of different states may not be complete. Therefore, surveillance programs need to have other ways to discover deaths that occurred outside their jurisdictions.

Handling of Out-of-jurisdiction (OOJ) Case Records

Record linkages to the state death certificate database may reveal decedents for whom the residence at the time of death was in another state. To avoid duplication at the national level, the surveillance program of the other state should be notified about these cases, and an interstate agreement should be reached about the residency at first diagnosis of HIV and/or AIDS. For decedents with a residence at first diagnosis outside the surveillance program's jurisdiction (e.g., in another state) but who received care in its jurisdiction, surveillance programs should collect routine surveillance data and include them in record linkages to the state death certificate database. This information should be forwarded to the state where the decedent resided at first diagnosis. For further guidance, see Case Residency Assignment.

Outcome Measures

It is impractical to evaluate death ascertainment itself because no gold standards are available against which to assess its quality or quantity. However, the data collected through death ascertainment should be used to evaluate other components of surveillance (e.g., gaps in reporting by particular sources). The primary outcome standard is the proportion of death certificate only cases; all other outcome measures are recommendations to assess program performance.

Proportion of Death-Certificate-Only (DCO) Cases

The proportion of cases among all cases that have only a death certificate document is calculated (percent DCO = 100* [number of cases with only a death certificate document/all cases]). The proportion should be calculated for each diagnosis year and should be less than 5%. A higher proportion would indicate the need to strengthen reporting from other sources of information. The standard is assessed at 24 months after the death year (**Note**: death year is also the diagnosis year for these cases), assuming files with death information are available within 18 months of the diagnosis year (otherwise, the standard is assessed 6 months after death files become available from the Office of Vital Records).

Proportion of Cases for Which the Initial Source of Report Was a DCR

The proportion of cases initially reported through death certificate records (first document received is a death certificate) among all cases, by year of report, is calculated. A high proportion (10% or more) would indicate the need to strengthen reporting from other sources of information.

Effectiveness of the HIV-Specific Record Linkage in Finding Cases

It may also be of interest to measure how many DCRs with mention of HIV (or AIDS) are confirmed as newly identified cases (% confirmed = number of cases confirmed/number of DCRs with mention of HIV/AIDS). The number of confirmed cases excludes DCOs; confirmed cases are those for whom follow-up was successful, and additional information from other sources has been obtained to confirm the case meets the AIDS and/or HIV case definitions. It may be impossible to follow back some of the potential unreported cases found among the nonlinked records by review of medical records or contact with health care providers to assess whether they meet the criteria of the case definitions for AIDS or HIV (non-AIDS). No target goal has been set for this outcome measure because it is unknown how many DCRs have complete information to allow follow-up of cases. However, surveillance programs may want to determine the reasons for failure to follow-up on DCRs.

Delay between Death and Report of Death

Conducting annual searches of death certificate records will probably eventually lead to an annual death-to-report interval ranging from 1 to 30 months, with a mean delay of 1 to 2 years, depending on when the multiple-cause database becomes available and when the search is conducted. This death-to-report interval should be evaluated by calculating the

annual mean and median interval between death and report (or data entry) and the proportion of deaths for which this delay exceeded 36 months. Trends in this interval should be examined by year of death and may be assessed by source of information (death reports from providers/medical records vs. death certificate linkage [all cases will have the same delay time]). The target goal is to obtain death information for cases within 24 months after the year of death. However, availability of death certificate files varies by state. When calculating the reporting delays for recent years, the delay will be underestimated if any death files are not yet linked to the HIV/AIDS Reporting System (HARS).

Yield from Searching the NDI

Because the value of searching the NDI is not well established for HIV/AIDS surveillance, if surveillance programs search the NDI to find deaths, it is recommended that the yield from it for a given year of death be evaluated. If X represents the number of deaths newly discovered through the NDI, W represents the number of deaths in the same year already known from other sources (including searches of the state/local death certificate database), regardless of whether they were also found in the NDI, then the yield from the NDI may be expressed as $Y = X/(X + W)$. If the yield is low (< 0.05), then it may not be worthwhile for the surveillance program to search the NDI in the future.

Table 1. National databases with death information

	NDI	NDI-PLUS	SSDI/DMF
INFORMATION AVAILABLE	• Date of death • Date of birth • State of residence at death • Death certificate ID number	• Date of death • Date of birth • State of residence at death • Cause of death • Death certificate ID number	• Name • Social Security # • Date of death • Date of birth • Death in foreign country
ADVANTAGES	• Provides date of death • State where death occurred • Death certificate ID number	• Provides cause of death	• Less expensive (no cost) • More recent deaths • May include deaths in foreign countries • Over 65 million records of death since 1962 • ASCII text file
LIMITATIONS	• Most recent data are 2 years old • Cost	• Most recent data are 2 years old • More costly than routine NDI search	• No cause of death • No reliable residence at death (last address is that for receipt of benefits only) • Not available if decedent had no SS# or no benefits were claimed[1]

[1] The Social Security Administration's DMF includes decedents who claimed benefits before death and therefore had a Social Security number (mostly persons who worked at least 10 years and were older than 65 years or disabled before death) and decedents who did not claim benefits before death but whose spouse or children claimed survivor or dependent benefits. This database does not include decedents for whom no benefits were claimed. Thus, persons who did not qualify for benefits—because they died before age 65, had not been disabled, were unmarried, and had no minor children—are likely to be missing from this database.

Table 2. ICD-9 & ICD-10 codes for AIDS-defining illnesses

Illness	ICD-9	ICD-10
Candidiasis	112	B20.4, B37
Candidiasis of lung (pulmonary)	112.4	B37.1
Candida esophagitis or esophageal candidiasis	112.84†	—
Carcinoma of the cervix of the uterus	180	C53
Coccidioidomycosis	114	B38
Meningitis or disseminated coccidoidomycosis	114.2, 114.3	B38.4, B38.7
Cryptococcosis	117.5	B45
Cryptococcal meningitis	—	B45.1
Cryptosporidiosis	007.2, 007.8	A07.2
Cytomegalovirus disease	078.5	B20.2, B25
Dementia/encephalopathy	043.1, 049.8, 049.9, 290.1, 294.9, 310.9, 323.9, 331.9, 341.9, 348.3, 348.9, 349.9	A85.8, A86, A88.8, A89, F02.4, B22.0, F03, F06.7, G04.9, G31.8, G31.9, G37.9, G93.4, G93.9
Histoplasmosis	115	B39
Herpes simplex	054	B00
Isosporiasis	007.2	A07.3
Kaposi's sarcoma or other nonmelanoma skin cancer	173	B21.0*, C46* (KS), C44 (other ca.)
Kaposi's sarcoma*	176*†	B21.0*, C46*

Table 2. ICD-9 & ICD-10 codes for AIDS-defining illnesses (continued)

Lymphoid interstitial pneumonitis (in children only)	516.8	B22.1*, J84.8
Lymphoma, non-Hodgkin's	200.0–200.8, 202.8	B21.1, B21.2, C83, C85
Mycobacteriosis, nontuberculous	031	A31
Mycobacteriosis at sites other than lung or skin*	031.8*	A31.8*
Pneumocystosis (codes 136.3 & B59 lack information on anatomic site, but it is probably pneumonia)	136.3*	B20.6*, B59*
Pneumonia, bacterial	481–482	J13–J15
Pneumonia, organism unspecified	485–486	J18

Table 2. ICD-9 & ICD-10 codes for AIDS-defining illnesses

Illness	ICD-9	ICD-10
Progressive multifocal leukoencephalopathy	046.3	A81.2
Salmonella septicemia (problem: code lacks information on recurrence)	003.1	A02.1
Toxoplasmosis	130	B58
Meningoencephalitis due to toxoplasmosis*	130.0*†	B58.2*
Tuberculosis*	010–018*	A15–A19*
Wasting/cachexia	261, 263.9, 783.2, 799.4	B22.2, E41, E46, R63.4, R64

* Diagnosis of this disease is accepted by the AIDS case definition as indicative of AIDS, even if the disease was diagnosed presumptively rather than definitively. For surveillance purposes, a death certificate record with the ICD code for this disease is sufficient evidence of AIDS if it is accompanied by independent evidence of HIV infection from another source document, but not if the death certificate record is the sole source of information.

† This code exists only in the ICD-9-CM, which is used for morbidity data, such as hospital discharge diagnoses, but not for causes of death on death certificates.

Table 3. ICD-9 & ICD-10 codes for AIDS-defining illnesses sufficiently specific for the criteria of the AIDS case definition* as to make it unnecessary to search medical records for additional information about these diseases, given that HIV infection is known

Illness	ICD-9	ICD-10
Kaposi's sarcoma	—†	B21.0, C46
Lymphoid interstitial pneumonitis (in children only)	—†	B22.1
Mycobacteriosis at sites other than lung or skin	031.8	A31.8
Pneumocystosis	136.3	B20.6
Toxoplasmosis of brain (meningoencephalitis)	—†	B58.2
Tuberculosis	010–018	A15–A19

* These diseases are accepted as indicative of AIDS even if they are diagnosed presumptively rather than definitively. An ICD code for one of these diseases means that it was diagnosed at least presumptively, if not definitively, and therefore is sufficient to meet the surveillance criteria for AIDS, but only if the diagnosis of HIV infection was supported by finding it also in another document besides the death certificate record.

† Unlike the ICD-9-CM codes, the corresponding ICD-9 codes for these conditions are not sufficiently specific for an AIDS diagnosis to be assumed based solely on these codes.

Appendix

Searching National Death Databases

After a program has conducted the annual record linkage between cases and the state multiple-cause death certificate database, some cases will remain without acceptable links to death records. Some of the persons represented by these nonlinked cases may still be alive, but others may not have linked death records because they died outside the jurisdiction of the program (e.g., in another state). Nationally, by year of report, the proportion of deaths of persons with AIDS that occurred in a state other than the one in which they were first diagnosed and reported as such to CDC has been small, ranging from 3% to 9% for all cases reported through 2003. However, the proportion of deaths occurring out of state has varied greatly by state or territory, ranging from <1% for Puerto Rico to 35% for Wyoming, and tending to be higher for states with smaller populations. Longer survival could result in out-of-jurisdiction deaths with no mention of HIV on the death certificates. This possibility makes searching for links in national death databases an important way to discover out-of-jurisdiction deaths. Two national databases can be searched for deaths: the National Center for Health Statistics' (NCHS) National Death Index (NDI) and the Social Security Administration's (SSA) Death Master File (DMF) (some private companies have derived from it the "Social Security Death Index" [SSDI]). These databases are compared in Table 1 of the main text of this section.

For programs with a low proportion of deaths out-of-state (based on data exchange with other states or prior NDI searches), there may be less need to use the NDI-Plus search to ascertain causes of death because the proportional distribution of cases by underlying cause of death can be estimated based on the large sample of deaths for which the causes will probably be ascertained from the in-state record linkages. Health departments might consider letting enough time (e.g., 5 to 10 years) pass between HIV/AIDS diagnosis and NDI searches so that the yield will be high enough to justify the cost. For programs with a large proportion of deaths out-of-state, it may be more important to conduct an NDI or NDI-Plus search to ascertain vital status/causes of deaths that occurred out-of-state. Fees and other information about NDI are available at http://www.cdc.gov/nchs/ndi.htm.

Technical Details of Searching the DMF

The DMF is an ASCII text file that can be batch-searched as easily as the death certificate database of a state Office of Vital Records. The complete DMF contains over 65 million records of deaths that have been reported to SSA since 1962. This file includes the following information on each decedent if the data are available to SSA: Social Security number (SSN), name, date of birth, date of death, state or country of residence (only for deaths in February 1988 or earlier), ZIP Code of last residence, and ZIP Code of lump sum payment. Some documentation of the DMF is available at http://www.ntis.gov/products/ssa-dmf.aspx and https://dmf.ntis.gov/recordlayout.pdf.

The DMF will be sent to programs in the form of 5 or 6 CD-ROMs provided by CDC, which may be transferred to the hard drive of a LAN server or a free-standing PC. The names, date of birth, and SSN can be used as identifiers when searching for links, and the

date of death is the information the program should seek from the links found. The ZIP Codes mentioned previously may be missing, particularly if no benefits were claimed before death, or they may not accurately reflect the residence at death. The NDI would be more complete and reliable for ascertainment of state of residence at death.

Versions of the DMF and SSDI are available that allow searches to be conducted online. Online searches can even be done for free at some Web sites. However, CDC experts in data security have advised against this practice because of the possibility of electronic interception of uploaded confidential information about HIV/AIDS cases. Although the probability of such electronic eavesdropping may be small, CDC recommends that surveillance programs protect the security of personally identifying case data by not conducting online searches.

National Death Index (NDI)

The NDI-Plus search may be used for linking to two types of cases:

1. cases representing persons not yet known to be dead, diagnosed more than 10 years ago, and
2. cases representing persons known to be dead but for which the causes of death are unknown (which may happen if the deaths were discovered from a search of the DMF).

Detailed information about how to search the NDI and the application form can be obtained from this Web site: http://www.cdc.gov/nchs/ndi.htm. Before completing the application form, you should obtain guidance by contacting the NDI at (301) 458-4444 or by e-mail through the NDI Web site.

- Attachment A of the application should be completed because this is required of a disease registry, which is what the HIV/AIDS surveillance program is. Completion of a different attachment is required for a research study.
- In the application for using the NDI-Plus, CDC (and any other agency with which the data will be shared) must sign a supplemental confidentiality agreement to permit cause-of-death data to be shared with it.
- A need for follow-back could arise if the decedent's status last known to the program was as an HIV (non-AIDS) case, and a determination is sought about whether the person had an AIDS diagnosis. The death certificate identification number can be used to request a hard copy of the death certificate from the Office of Vital Records of the state where the death occurred, which would identify the facility where death occurred and the identity of the certifier. These health care providers could then be contacted or their records could be reviewed to find the missing information. However, determining whether the person had AIDS may be difficult to accomplish because the information sought may be located in another state (assuming that the NDI is mainly used to find out-of-state deaths). Help in obtaining the follow-back information should be sought from the program in the other state. If the program anticipates the need for follow-back, it must complete Attachment B of the application. Follow-back to health care providers or

next-of-kin will not be permitted by the NDI agreement unless it is approved by an institutional review board (IRB), as described in the NDI application. Surveillance programs need to determine whether such resource-intensive follow-up is of high enough priority to conduct.

How to encrypt data for sending to the NDI

A database should be prepared consisting of personal identifiers for the case records for which the program seeks potentially linking death certificate records, as described in the instructions for the NDI from NCHS. An acceptable practice is to use an express courier service (e-mail is not sufficiently secure) to mail encrypted files to the NDI and to have the NDI mail the encrypted results back to the programs, provided that the encryption used is strong (at least 128-bit). Simple password protection, such as that afforded by Microsoft Word, PKZip, or WinZip, is not adequate. The general practice has been for the data sender to create a "self-decrypting" encrypted file, which means that the recipient needs to be provided with only a password to decrypt it, but does not need to have a particular decryption program installed on his/her computer. The sender needs encryption software, however, and must communicate the password to the recipient by telephone or e-mail. This decipherability must work in both directions, because the NDI staff must encrypt the results of the search (which includes the personal identifiers of the submitted cases) and send them back to the program, which then must decrypt them to collect the death data. Encryption/decryption software packages that programs may consider include, but are not limited to, the following:

- Advanced Encryption Package (AEP) 2004 (version 2.0.5) or Cryptocrat 2004 (http://secureaction.com)
- DigiSecret (www.tamos.com)
- FineCrypt (www.finecrypt.net)
- Kryptel (www.kryptel.com)
- PGP Mail (part of PGP 8.0) (www.pgp.com)

Features to look for include

- strong encryption (≥128 bit);
- the ability to create self-decrypting (or self-extracting) files (requiring only a password, not requiring the software package); and
- compression capability (to allow large files to fit on a single disk).

An alternative to the use of self-decrypting files (which require that the sender tell the receiver the password) would be the use of encryption "keys" (in which the sender encrypts the file to the receiver's public key). However, that would require both the sender and the receiver to use the same encryption/decryption software. The NDI staff have been using PGP. After a file has been encrypted, it is important to remember that the unencrypted or decrypted version of it still exists, and that version may need to be erased

to leave only the encrypted file if its location is insecure. It is important to erase a confidential file in a special secure way that the encryption software should be able to do (e.g., which will over-write the file at least 3 times).

How to select the true link from multiple potential links in the NDI

Demographic variables, such as sex, race, and the state of residence at death, as well as identifiers, such as name, SSN, and date of birth, are not available from NDI data. Instead, the NDI data will tell only whether (yes or no) each of these variables in the NDI matched the corresponding ones in the HIV/AIDS case data that had been submitted for linking. For this reason, it may be more difficult to distinguish a true link from among multiple potential links in the NDI data than it would be in data from the state Office of Vital Records. As a result, a special algorithm may need to be developed to distinguish a true link from several potential NDI links. Even then, a true link may not be recognized because names classified as "different" by the NDI may actually be only slightly misspelled or one may be a nickname for the other. These names would have been recognized as equivalent if their exact spellings had been visible, as they would have been on a search for matches in the state death certificate database. See Record Linkage for additional information.

Jurisdictions that withhold information on causes of death from the NDI-Plus output

NCHS has an agreement with each independent Office of Vital Records regarding the terms under which NCHS is permitted to disclose data on deaths in their jurisdictions in the results from NDI searches. Illinois prohibits NCHS from disclosing data on causes of death in the results of NDI-Plus searches for persons who died in Illinois if the decedent died of HIV infection or any other sexually transmitted disease (STD). Consequently, programs should expect that, if the death occurred in Illinois and HIV or another STD was mentioned on the death certificate, NDI-Plus records linking their HIV/AIDS case records will have missing information on causes of death (the underlying cause will be replaced by "N/A" and multiple causes will be blank). A small proportion of NDI records from every state will also have similarly missing data on causes of death if those data arrived too late to be included in the most recent year's data. In addition, New York City prohibits NCHS from disclosing data on causes of death, regardless of the cause, if this information will be linked to personal identifiers for an indefinite period of time, as would be true for the HIV/AIDS surveillance registry (but not necessarily for a time-limited special research project). Even if the causes of death are missing, other information on the NDI-Plus record, such as the date of death, may still be useful. Finally, it should be mentioned that Florida permits NCHS to disclose data on causes of death only if the recipient agrees to keep personal identifiers in a data file separate from the file containing causes of death. Thus, if programs have not already done so, they should keep identifiers such as names, birth dates, and SSNs in a file separate from the file that contains the data that will be used for epidemiologic analysis. The two files may be linked, however, by a case identification number (e.g., STATENO, CITYNO). Programs must agree to this stipulation when they apply for an NDI search if they want to obtain data on causes of death in Florida.

Technical Guidance for HIV/AIDS Surveillance Programs

Electronic Reporting

HIV Incidence and Case Surveillance Branch
Atlanta, Georgia

Contents — Electronic Reporting

Introduction 5-3
Purpose of Electronic Reporting 5-3
Electronic Case Reporting (ECR) 5-3
Electronic Laboratory Reporting (ELR) 5-4
Considerations for Electronic Reporting 5-6
Policy 5-6
Existing ELR Resources for Implementing HIV/AIDS Reporting 5-7
Background 5-7
NEDSS (National Electronic Disease Surveillance System) 5-8
PHIN (Public Health Information Network) 5-8
ELR Requirements 5-9
Electronic Case Reporting 5-10
Review Reporting Laws 5-10
Develop a System for Reporting 5-11
Identify All Reporting Facilities and Providers 5-11
Documentation for Reporting Facilities and Providers 5-11
Alternative Reporting Methods 5-12
Data Processing of Electronic Case Reports 5-13
Reporting Back to Reporting Facilities and Providers 5-13
Electronic Laboratory Reporting 5-13
Review Reporting Laws 5-14
Develop a Mechanism for Electronic Reporting 5-16
Identify All Reporting Laboratories 5-16
Documentation for Reporting Laboratory 5-17
Alternative Reporting Methods 5-19
Optional Reports 5-19
Resolving Referral Laboratory Problems 5-19
Managing Increased Number of Lab Reports 5-19
Laboratories First Reporting Electronically 5-20
Reporting Back to Laboratories 5-20
HIV and HIV-associated Laboratory Tests 5-21
Diagnosing HIV Infection 5-21
Other Tests Indicative of HIV Infection 5-21
HIV Disease Progression and Markers for Access to Care 5-22
Changes in Laboratory Reporting Initiated by ELR 5-23
Performance Standards 5-23
Case Follow-up from a Lab Report 5-24
The Ideal Complete Lab Report 5-25
Measurable Performance Standards 5-32
References 5-34
Appendix A 5-36
Appendix B 5-39
Appendix C 5-51
Appendix D 5-56
Appendix E 5-63
Appendix F 5-75
Appendix G 5-82

Technical Guidance for HIV/AIDS Surveillance Programs — Electronic Reporting

Introduction

Purpose of Electronic Reporting

With the growing sophistication of information technology, surveillance data can now be transmitted, stored, manipulated, and analyzed electronically in a secure environment. These advancements are changing the way HIV/AIDS surveillance activities are conducted and what can be achieved. Electronic reporting of data for both cases and lab results should enhance current surveillance practices, as long as technical capability is in place to integrate electronic reporting into routine surveillance practices. As manual entry or re-entry of data diminishes, the efficiency of surveillance activities should also improve. Information can be transmitted to and received by HIV/AIDS surveillance programs in a timelier manner. When case confirmatory lab results and case data from confluent data streams are linked and complementary data combined, a number of beneficial effects follow: case ascertainment should be enhanced; surveillance data received passively should be more complete; progression of disease in cases should be identified; and unmet health care need should be easier to ascertain. Accuracy of data should also improve as errors that used to occur when entering case information or lab results from paper copies will now be limited to the initial point of data entry, and conflicting information from various data sources can be identified and reconciled. As less paper is used, storage and security of confidential reports will also be improved.

Electronic Case Reporting (ECR)

Definition: An HIV/AIDS case report that arrives in a pre-established electronic format (e.g., HL7 data stream, ASCII, spreadsheet) and does not require extensive human intervention (e.g., data entry, cutting and pasting, or translation) to add it to a database.

Electronic case reporting, ECR, is used by <10% of HIV/AIDS reporting areas (1). The emphasis on electronic laboratory reporting, coupled with lack of resources, has forced a number of areas to lower their prioritization for implementation of ECR. Nevertheless, in the Electronic Case Reporting section of this chapter, steps that have been used by those conducting ECR are included as a guide for those areas considering implementing ECR. As an integrated nationally notifiable electronic disease surveillance system is established, states may be able to build upon a communicable disease reporting infrastructure for ECR.

The use of electronic case reports from reliable sources can help identify new cases and facilitate reporting of existing cases, complete case report information of HIV disease, and improve the accuracy and completeness of patient identifiers and demographic and risk

information. With careful planning and proper implementation, ECR can also redistribute some portion of staff time currently dedicated toward medical record abstraction and data entry for HIV/AIDS surveillance.

Electronic Laboratory Reporting (ELR)

Definition: Laboratory reporting that arrives in a pre-established electronic format (e.g., HL7 data stream, ASCII, spreadsheet) and does not require extensive human intervention (e.g., data entry, cutting and pasting, or translation) to add it to a database.

Electronic laboratory reporting, or ELR (also known as electronic laboratory-based reporting), will markedly enhance passive identification of cases and enrich surveillance data from diagnosis to death. While most areas currently receiving ELR are receiving inadequate demographic surveillance information to complete a case report, thus requiring case follow-up, efforts are underway to encourage laboratories to submit this needed demographic information to public health departments, especially for newly diagnosed persons.

> *"ELR is an important component of CDC's public health information and surveillance systems integration efforts [that must be coordinated] with related efforts at CDC and in state and local public health agencies."*
>
> *Electronic Reporting of Laboratory Information for Public Health*
> *Centers for Disease Control and Prevention, Atlanta, GA*
> *January 7–8, 1999*

Lab data for HIV/AIDS surveillance can be used to

- find cases
- monitor care and treatment
- identify disease stage at diagnosis and progression to AIDS
- audit laboratories
- measure unmet health care need

Passive case reporting depends on a motivated reporting source and rarely results in complete case ascertainment. Active case finding at hospitals or specialty clinics is very resource-intensive but has been the most reliable method to identify cases, supplement case ascertainment, and collect complete surveillance case information. A great benefit of ELR will be to help speed the identification of new cases and provide a method for more complete case ascertainment when laboratories report diagnostic or other HIV testing results directly to state or local HIV/AIDS surveillance programs. These test results will be sent to surveillance programs at approximately the same time as to providers. As such, based on the type of lab test, the priority and method for follow-up will differ. Following a confirmatory HIV test result, surveillance programs will probably continue to conduct

follow-up to obtain complete case information from a provider; but other HIV-associated testing, such as CD4+ T-lymphocyte (CD4) counts, may need to be reserved until case identification has occurred. One potentially efficient way to ensure that CD4 count results are collected and reported at the time of HIV diagnosis is to use the opportunity for medical record abstraction from provider's offices. At that time, case information can be completed for *all* HIV-infected persons who are cared for by that provider, including initial CD4 count. Follow-up of additional CD4 results may, however, be delayed depending on the number and level of CD4 results that are reported and the available resources for follow-up.

The presence of lab testing for a newly diagnosed HIV-infected person can be a marker of access to health care, reflecting the successful linkage between diagnosis and entry into care. Furthermore, the timing and frequency of lab testing is a marker for meeting HIV treatment objectives and providing optimum quality of care. Decreasing CD4 counts and increasing viral load reflect the progression of HIV disease. Population-based CD4 counts at the time of initial HIV diagnosis provide a cross-sectional view of immune suppression and stage of disease among newly diagnosed HIV-infected persons. CD4 count at the time of diagnosis is an indicator of how well education, risk recognition, testing, and access to care are working. If early testing and detection for HIV is occurring, CD4 counts should be normal or near normal. Severely deficient initial CD4 counts, obtained at initial HIV diagnosis, that meet the current AIDS immunologic criteria, reflect a significant delay and missed opportunity for early treatment and prevention.

ELR can also be used to enhance completeness of surveillance activities. The volume of lab results received from a laboratory during a specified period of time establishes a baseline rate; this rate can be compared with that of lab reports received in the future to monitor for any discrepancies in laboratory reporting.

Twenty-eight (46%) of 61 HIV/AIDS surveillance areas reported in a 2004 survey that they received lab data electronically (data that arrived in a pre-established electronic format—e.g., HL7 data stream, ASCII, spreadsheet—and did not require extensive human intervention such as data entry, cutting and pasting, or translation to add it to a database). Of those, only two areas reported they have a state law/regulation/code that requires laboratories to report electronically, but neither specified a volume of testing for which reporting was mandated. Among the 33 surveillance areas not currently conducting ELR, 52% expected to begin ELR within the next 12 months. Lack of financial resources and lack of dedicated informatics personnel were the most frequently cited reasons among areas not planning on beginning ELR anytime soon.(1)

The Electronic Laboratory Reporting section of this chapter outlines steps that should be considered when implementing electronic laboratory reporting.

Considerations for Electronic Reporting

Electronic reporting will necessitate some planning and implementation of safeguards by the HIV/AIDS surveillance program. Structural requirements for supporting electronic reporting are included in the section on Existing ELR Resources for Implementing HIV/AIDS Reporting. Practices and standards are also covered in detail in the Introduction to Policies and Procedures and Record Linkage chapters, and in Vol. III: Security and Confidentiality Guidelines.

Policy

Before implementing electronic reporting, HIV/AIDS surveillance programs should review existing state laws regarding requirements for communicable disease reporting, electronic reporting, and specifically HIV/AIDS reporting. Health Insurance Portability and Accountability Act (HIPAA) policies are covered in Vol. III: Security and Confidentiality Guidelines.

Security and Confidentiality

Security and confidentiality are covered in detail in Vol. III: Security and Confidentiality Guidelines. There are five guiding principles upon which all HIV/AIDS surveillance program requirements and security considerations are derived. They are

Guiding Principle 1 HIV/AIDS surveillance information and data will be maintained in a physically secure environment. Refer to sections Physical Security and Removable and External Storage Devices.

Guiding Principle 2 Electronic HIV/AIDS surveillance data will be held in a technically secure environment, with the number of data repositories and individuals permitted access kept to a minimum. Operational security procedures will be implemented and documented to minimize the number of staff that have access to personal identifiers and to minimize the number of locations where personal identifiers are stored. Refer to sections Policies, Training, Data Security, Access Control, Laptops and Portable Devices, and Removable and External Storage Devices.

Guiding Principle 3 Individual surveillance staff members and persons authorized to access case-specific information will be responsible for protecting confidential HIV/AIDS surveillance information and data. Refer to sections Responsibilities, Training, and Removable and External Storage Devices.

Guiding Principle 4 Security breaches of HIV/AIDS surveillance information or data will be investigated thoroughly, and sanctions imposed as appropriate. Refer to section Security Breaches.

Guiding Principle 5 Security practices and written policies will be continuously reviewed, assessed, and as necessary, changed to improve the protection of confidential HIV/AIDS surveillance information and data. Refer to sections Policies and Security and Confidentiality Program Requirement Checklist.

Note: See Vol. III: Security and Confidentiality Guidelines for each of the sections specified above.

Record Linkage

Linkage of new electronic lab records and case reports to an HIV/AIDS case can fulfill many purposes for surveillance: find cases, update patient status (e.g., verify the living status of a patient, monitor disease progression from HIV to AIDS to death, etc.), assess access to care, verify accuracy of information (e.g., race, state of residence, etc.), complete surveillance information (e.g., risk factors, presence of co-morbid conditions, etc.), and facilitate care and treatment services. Surveillance programs use this information to more accurately describe the epidemic, characterize epidemiologic changes, and assess effectiveness of intervention programs.

Processing large volumes of electronic case reports or electronic lab reports requires linkage with existing case surveillance records to determine whether the report should be linked to an existing case or whether a new case record should be created. Strategies for linkage to case reports can be found in the Record Linkage chapter.

Existing ELR Resources for Implementing HIV/AIDS Reporting

Background

Over the past several years clinical laboratories, hospitals, and government have recognized the need to develop standard ways of communicating electronic data. More recently a number of standards have evolved, and an increasing number of partners have adopted them.

The Department of Health and Human Services has identified three standards relevant to the content and format of transferring electronic laboratory data for the purposes of public health reporting. These include HL7 (Health Level 7), LOINC (Logical Observation Identifiers Names and Codes), and SNOMED (Systematized Nomenclature of Medicine), which are each described further below and in Appendix A, Electronic Laboratory Reporting Resources. Any public health jurisdiction that is considering implementing an electronic laboratory reporting (ELR) system for HIV infection data should first identify whether other public health staff have already developed systems for transferring non-HIV data. In most cases the same systems can be employed, whether or not the data are contained within the same messages.

CDC has supported two major initiatives related to ELR. First proposed in 1999, the National Electronic Disease Surveillance System (NEDSS) incorporated a standards-based architecture that included processing electronic laboratory data. In 2003, NEDSS

became a component of the broader Public Health Information Network (PHIN) initiative. Both initiatives have brought together experts in a wide variety of fields to further develop and adopt these standards. Two Web sites, http://www.cdc.gov/nedss/ and http://www.cdc.gov/phin/, contain numerous resources necessary for implementing ELR.

NEDSS (National Electronic Disease Surveillance System)

ELR is an important component of NEDSS. The general goals are to

- detect outbreaks rapidly and provide national health data
- facilitate electronic transfer of data from clinical to public health settings
- reduce the burden of public health reporting
- enhance timeliness and quality of data

The NEDSS Base System is being developed for health departments to use and modify for surveillance and analysis of notifiable disease information. The base system elements include

- browser-based data entry and data management
- processing of electronic lab, clinical, and public health data in HL7 format
- an integrated data repository
- active data translation and exchange (integration broker) functionality
- use of contemporary application programming practices—component based, object oriented, and cross platform where possible
- data reporting and visualization capability
- a shareable directory of public health personnel
- a security system and appropriate security policies

PHIN (Public Health Information Network)

The NEDSS Base System and ELR are important components of PHIN. Its broader vision includes

- automated data exchange
- use of electronic clinical data for event detection
- manual data entry for event detection
- specimen and lab result information management and exchange
- management of case, contact, and threat data
- analysis and visualization
- directories of personnel
- public health information and alerting
- information technology (IT) security and critical infrastructure protection

ELR Requirements

For the system to work across the full range of partners, all transactions must be standards-based. Some of the standards that must be in place regardless of the content and format of data transfers include the security and messaging standards. These are well beyond the scope of this document, but information on secure file transfer methods, the PHIN messaging system, XML (eXtensible Markup Language), and other terminology is available at the PHIN Web site. The standards that *are* specific to ELR are HL7, SNOMED, and LOINC.

HL7 (Health Level 7)

Health Level 7 provides interoperability between health care information systems by specifying standards for the exchange, management, and integration of data that support clinical patient care and the management, delivery, and evaluation of health care services. More information about this American National Standards Institute-accredited standard is available at www.hl7.org.

HL7 defines the structure and syntax of an electronic message, so the sender knows where to place information and the receiver knows where to find it. Each defined segment contains several fields. Common segments for electronic laboratory reporting include the Patient Identifier, Next of Kin, Common Order (ordering facility and provider), Observation Request (requested tests), and Observation Result (the lab results).

Note: HL7 allows coding of the laboratory tests and results using SNOMED or LOINC.

SNOMED (Systematized Nomenclature of Medicine)

The College of American Pathologists has been developing this system since 1965. SNOMED CT is a dictionary of clinical terms that includes every clinically relevant organism that might be identified by a laboratory.

LOINC (Logical Observation Identifiers Names and Codes)

The Regenstrief Institute developed this dictionary of lab tests and results.

Note: See Appendix A for additional information about general ELR in ELR Resources.

Electronic Case Reporting

Electronic case reporting, as of June 2004, was being conducted by a small number of HIV/AIDS reporting areas[1], and most are in the early stages of implementation. Of the 55 sites that were not currently receiving case reports electronically, eight were planning to begin electronic case reporting within the next 12 months (1).

Because so few areas have begun implementing electronic case reporting, it is recommended that the steps listed below be implemented in stages—beginning with providers that report the largest number of cases—with the goal that all providers will eventually report all cases electronically. The growing use of electronic medical records may make electronic case reporting less burdensome for facilities and providers. Nevertheless, the need to conduct field follow-up of potential new cases will not be completely eliminated with electronic case reporting.

Review Reporting Laws

Review your state laws and work toward making additions or modification as needed. Consider these examples:

- If possible, implement mandatory electronic reporting of HIV/AIDS case data from facilities and providers, specifying the timeframe in which the HIV/AIDS surveillance program should receive the report. If this is not possible, the surveillance program should strongly encourage and facilitate voluntary reporting.
- Identify which specific case fields are required to be reported.
- Include provisions for dealing with facilities and providers within and outside of the state's boundaries.
- Clarify policies for out-of-state case reporting that occurs via paper or electronically.

Erroneous interpretations of HIPAA can be a serious hindrance to establishing case reporting. The state's reporting regulations should acknowledge the provisions in HIPAA that deal with public health reporting. HIPAA permits reporting requirements established under state law. Therefore, the methods and procedures mandated by state laws or regulations are not prohibited by HIPAA. (HIPAA policies are covered in Vol. III: Security and Confidentiality Guidelines.

1. Los Angeles, NC, PA, PR, San Francisco

Develop a System for Reporting

- Determine if your program can use or adapt an existing secure electronic system in your health department. In some instances, there may be a way to integrate HIV/AIDS surveillance data into an existing or planned state system.
- If your state does not have an existing secure electronic system, a universally available method, such as a database, that stores case information in a uniform, standard format should be available to all HIV/AIDS surveillance sites.
- eHARS or NEDSS should provide a platform for collecting case information securely and confidentially.

Identify All Reporting Facilities and Providers

Identify all facilities and providers that need to report to your program, using local HIV/AIDS surveillance data and other available data sources, and identify a contact person at each facility/provider; establish a relationship with this person. Simply knowing who to call saves time. Educating contact persons so that they have a clear understanding of reporting requirements and mechanisms promotes prompt, accurate, and complete reporting from providers of surveillance data.

- Explore programs at other agencies—explore other programs such as the Ryan White Comprehensive AIDS Resources Emergency Act administered by the Health Resources and Services Administration (HRSA) or the Medicaid program, which may maintain a list of providers that they use for administrative and/or reimbursement purposes.
- Consult with other HIV/AIDS surveillance activities—consider special, new, or existing projects—to identify reporting providers and established relationships. Activities to consider include HIV Incidence and Viral Resistance Surveillance and the Morbidity Monitoring Project.
- Have your facility and provider list reviewed by field surveillance staff for completeness and accuracy.
- Contact each large reporting provider, establish a relationship with someone who has a good overview of the entire operation, and ask that person to complete a survey of how many new HIV/AIDS cases the facility has annually.
- Identify those provider facilities that are ready and willing to participate in electronic case reporting and have the capacity to generate and submit electronic files.

Documentation for Reporting Facilities and Providers

Develop and disseminate reporting documentation for your state that includes data structure and layout.

- Provide facilities and providers with a list of reportable diseases/conditions reflecting your state reporting regulations. In some instances, reference to the actual regulations may also need to be provided.

- Provide facilities and providers with a list of fields that must be completed to report a case electronically. (Some common fields are probably recorded for use with electronic patient medical records. Means by which these data fields can be exported for electronic case reporting should be explored.)
  - ASCII or area-specific alternative format should be accepted.
  - Paper reports should continue to be accepted.
- Resolve confidentiality and data transfer/security issues with each provider facility depending upon its IT capability. Adherence to CDC as well as HIPAA standards for confidentiality and secure data transfer will be required.
- Establish a reporting schedule that meets the needs of both the surveillance program and local laws and regulations.
- Consider holding training session(s) for facility staff as a mechanism for explaining reporting details.
  - Because of the clinical- and surveillance-specific nature of many of the data fields required for reporting, it may be necessary to work closely with and educate facility or provider staff unfamiliar with surveillance. For example, the type of information necessary to complete risk factor fields may not be appreciated or understood by provider staff responsible for "abstracting" information into an electronic case report. This is especially true if IT staff are storing and extracting information from data warehouses or electronic medical records for surveillance purposes.
- Use CDC's paper case report form as the template for the electronic case report form.
- Develop a user-friendly software to facilitate electronic reporting.

Note: See Appendix B for examples of variable lists for electronic case reporting. Two historical examples—(A) and (B) of Documentation for Reporting Facilities and Providers—have been provided.

Alternative Reporting Methods

Offer alternative methods of reporting if provider facility cannot report electronically or cannot adapt to using your particular system.

- Be aware that facilities have limits in their own systems.
- Resolve confidentiality and data transfer/security issues with each provider facility depending upon its IT capability.
- Consider an HIV/AIDS-specific electronic method that small facilities or providers could use. Web-based, secure reporting to enter facility or provider data may be needed.

Data Processing of Electronic Case Reports

- All electronic provider data of disparate formats will need to be converted to a standardized format that is compatible with HARS or eHARS. All reported cases should be linked with the registry to determine if the same person has been previously reported. (See Record Linkage.)
- Linked case reports can be used to update HARS records in the following order:
  1) cases reclassified from HIV to AIDS
  2) risk factor information changed from undetermined to one of the valid risk factor groups
  3) cases whose date of HIV/AIDS diagnosis changed to an earlier year
  4) other types of updates that will enhance the quality of the data
- Nonlinked cases may represent new case reporting. These new case reports and the linked ones that will reclassify the case from HIV to AIDS should be validated. When working with a large volume of reports, select a random sample for a validation project. This can help identify recurring problems from reporting sites.
- To evaluate the quality of the information extrapolated from electronic medical records, sample electronic case reports that link with manually reported HARS records and compare with the information obtained from manual reporting.
- New case reports with high-quality, complete data may be uploaded to the surveillance database followed by routine data re-abstraction. (See Data Quality.)

Reporting Back to Reporting Facilities and Providers

Develop a procedure for reporting back to provider facilities or local area public health departments responsible for case follow-up on incomplete reports. Consider use of an "audit," a "quality control," or a "report card" as a formal means to report back.

- Provide local area public health departments with periodic, ongoing case report summaries or case line lists to ensure consistency in case counts and quality of aggregate data.
- Supply feedback to providers with the evaluation results. Systematic errors that are caused by programming should be resolved before accepting the first batch of cases. A good practice is to inform the provider of errors such as incorrect death date, diagnosis, and birth date to improve the consistency between provider data and HARS.

Electronic Laboratory Reporting

Electronic laboratory reporting (ELR), also referred to as electronic laboratory-based reporting, provides an opportunity for HIV/AIDS surveillance programs to enhance their ability to conduct timely and complete surveillance. Drawing on experience from sites using ELR, this section outlines steps surveillance areas should contemplate for initiating ELR.

A number of structural requirements are necessary for ELR. They include

- hardware to support ELR functions
- software that allows importing, linking, and retention of electronic reports
- trained personnel to manage ELR and ELR-related activities

In addition to this guide for getting started, areas lacking ELR experience should consult with and take advantage of the experience and knowledge from those already conducting ELR. Additionally, 28 areas[2] reported in 2004 that they were receiving laboratory data electronically (1).

Review Reporting Laws

Laboratory test results, in accordance with state reporting laws, can be categorized as one of two types: *required* or *optional reports*. Those HIV and HIV-associated test results that are legislatively mandated to be reported to surveillance programs are required reports; those HIV and HIV-associated tests results for which reporting is not mandatory but may be sent voluntarily to surveillance programs are optional reports.

A first step toward implementing ELR is to review your state laws and work toward making additions or modifications as needed. Consider these examples:

- If possible, make reporting of HIV and HIV-associated test results from laboratories mandatory, specifying timeframe in which the HIV/AIDS surveillance program should receive these required reports.
- If possible, make *electronic* reporting of HIV data from laboratories mandatory, specifying timeframe in which the HIV/AIDS surveillance program should receive these required reports.
- Identify specifically what test results are to be reported along with specimen identification code (i.e., accession number) and either a LOINC code or local code.
- Identify specifically what types of patient demographic data are reportable.
- Specify what type of information is required to identify the laboratory and provider, i.e., Clinical Laboratory Improvement Amendments (CLIA) number, name and address of laboratory, name and phone number of laboratory contact, name and address of facility that specimen originates from, and name of physician ordering test.
- Specify how to handle data that are from an intermediate or a pass-through laboratory. If the laboratory performing the test cannot provide the name and address of the true provider, it must provide contact information for the intermediate laboratory as well as the accession number of the test so that it is possible to work back to the point where the patient and the true provider can be identified.

2. AZ, CO, HI, Houston, IL, IN, KS, KY, Los Angeles, LA, MD, MN, MS, MT, NJ, NY, NC, OH, OK, PA, PR, San Francisco, SC, TX, UT, WA, WI, WY

- Specify provisions for dealing with laboratories within and outside of the state's boundaries.
- Clarify policies for out-of-state case reporting that occurs via paper or electronically.

Three-quarters of reporting areas surveyed indicated they communicate lab report information received for nonresident individuals to the individual's home state HIV/AIDS program. One area indicated that state regulations prohibit forwarding lab results. (1)

  - Areas have developed various methods of handling out-of-state reporting, depending on their reporting regulations and resources. Three examples are
    - Reports are simply forwarded to the appropriate state.
    - Lab reports identifying new cases are entered into the local HIV/AIDS surveillance database but attributed to the appropriate state.
    - Lab reports are entered into a separate database and communicated to the appropriate state.
  - Areas must also develop policies for receiving ongoing lab reports on individuals with a current residence locally but with an HIV or AIDS diagnosis in another state.
- In some cases it is valuable to have the clear authority to collect HIV-negative test results, such as for seroreverting perinatal exposures or undetectable viral loads. Many laboratories will only report positive test results, however.
- Extra planning in developing reporting laws can save a lot of time and frustration in the long run. Making assumptions that certain items will be negotiable or that they are implied in existing regulations can become very problematic. Large hospital and private laboratories may wish to conduct an internal legal review of the state's reporting regulations before implementing new reporting procedures. Also, each HIV program should consult its public health laboratory for assistance in writing laboratory reporting regulations that use appropriate language for securing reporting from a wide range of laboratories.
- Erroneous interpretations of HIPAA can be a serious hindrance to establishing laboratory reporting. The state's reporting regulations should acknowledge the provisions in HIPAA that deal with public health reporting. HIPAA permits disease reporting requirements established under state law. (HIPAA policies are covered in Vol. III: Security and Confidentiality Guidelines and in a 2003 CDC *MMWR* publication available at http://www.cdc.gov/mmwr/PDF/wk/mm52SU01.pdf.)
- Hospital laboratories may try to defer reporting to Infection Control. The reporting regulations should clearly state that all laboratories must additionally report to Infection Control.

Develop a Mechanism for Electronic Reporting

Determine if your health department has an existing secure system for electronic reporting that your program can use or adapt. In some instances, there may be a way to integrate ELR for HIV/AIDS surveillance into an existing or a planned state system. If your state does not have an existing system that can be used, your program will need to consider the resources necessary for designing and developing your own secure system for laboratories to use to transmit data. Some options to consider are to

- hire consultants
- borrow internal resources
- hire systems persons or developers/programmers
- research the possibility of using third party products—some laboratories may already be familiar with some of these

While implementing electronic reporting, the surveillance program may need to reorganize human resources by recruiting more staff with skills to process data, such as SAS programmers, or sending existing staff for training on data management and data analysis.

Identify All Reporting Laboratories

Identify laboratories that are conducting tests required to be reported, such as Western blot, CD4, and viral load tests. Identify a specific contact person at each laboratory and establish an ongoing relationship with this person.

Of 32 reporting areas not conducting electronic laboratory reporting, over 90% have identified which laboratories report which test results to them. Over 60% have begun discussions with their laboratories about reporting electronically to their program. (1)

To compile a list of laboratories needing to report, beginning with an inventory of laboratories in HARS, consider these available sources:

- Start by approaching the state agency responsible for conducting laboratory proficiency testing. For most states, with the exception of Washington and New York, the agency will be CMS (Centers for Medicare and Medicaid Services)-CLIA (Clinical Laboratory Improvement Amendments). CMS-CLIA regional offices can provide the same information. Contact information for the regional CLIA offices can be found at http://www.cms.hhs.gov/clia. The list of laboratories from these sources will be very extensive but can be easily filtered to extract hospital laboratories, independent laboratories, blood banks, public health laboratories, and reference laboratories to produce a shorter list of primary candidates doing HIV-related work. (See Appendix C for a CLIA list of types of laboratories.) The CLIA proficiency level given to laboratories may be helpful in

distinguishing between laboratories that conduct HIV testing in-house compared with laboratories that draw the specimen but send it to another laboratory for HIV testing.

- Be aware that states that conduct licensing or credentialing of laboratories within their state often have lists of laboratories and the type of tests they conduct.
- Explore other state programs such as the state Medicaid program, which maintains a list of licensed laboratories that they use for administrative reimbursement purposes.
- Consider other HIV/AIDS surveillance activities—special, new, or existing projects that have established relationships with laboratories and that have created a list of licensed laboratories. Activities to consider include HIV Incidence and Viral Resistance Surveillance and the Morbidity Monitoring Project, especially for identifying laboratories that are not reporting to the HIV/AIDS surveillance program.
- Have field surveillance staff and possibly reporting clinicians review laboratory lists for completeness and accuracy.
- Contact each laboratory and establish a relationship with someone who has a good overview of the entire operation (e.g., laboratory directors); and ask them to complete a survey of what tests they are doing, who serves as their reference laboratory, who are their referring labs, how do they manage their data, what type of IT support do they have, etc.
- Once surveillance staff know where testing is being done, they can start the process of making sure that the laboratories report in some manner, even if it is only on paper.
- The next step is to identify those laboratories that have the capacity to generate and submit electronic files for ELR.
- Be sure to monitor this list routinely as laboratories doing HIV testing may change frequently.

  See Appendix C for a list of types of CLIA laboratories and an example of a laboratory survey in Identifying All Reporting Laboratories.

Documentation for Reporting Laboratory

Develop and disseminate laboratory reporting documentation that includes data structure and layout to suit your state's specific needs.

- Because of the complex nature of reporting, make documentation as detailed as possible. It might be beneficial to consult with other states to aid in writing this documentation. Topics to consider include confidentiality and IT issues regarding how data must be formatted and what data elements are required. Alternatively, states may want to approach laboratories with the idea of creating memoranda of understanding (MOUs) itemizing issues on both sides.

- Provide laboratories with a list of tests and test results that are required for each reportable condition.
- Create a data dictionary. Identify and submit to the laboratory a complete listing of fields, data elements, and codes to be used for each lab test and result. The lab results data dictionary is covered in detail in <*Volume II, Data Dictionary*> of *Technical Guidance for HIV/AIDS Surveillance Programs*.
  - Because it is the national standard for laboratories, HL7 is the prototype for the file structure and format that should be encouraged for use in reporting. Some form of documentation on HL7 should be provided in the recommendations. (Keep in mind that HL7 is the standard, but many laboratories are not ready to adopt this standard.)
  - ASCII files should also be accepted with a goal toward transitioning to the HL7 format; the naming conventions, allowable values, and data structures should mimic HL7.
  - Paper reports should continue to be accepted, particularly for low volume laboratories. (These laboratories may not have the resources to convert to an electronic reporting system.)
  - Develop a procedure to follow when a lab transitions its method of reporting from paper to ASCII or HL7, or from ASCII to HL7. The process should involve dual reporting (paper and electronic) and a record-to-record comparison of the data. Do this until the lab has submitted data without error using the new format. All data elements should be compared, but, for higher volume labs, comparison of only a percentage of the records may be the most feasible. This comparison should be conducted over the course of several lab submissions.
- Resolve confidentiality and data transfer/security issues with each laboratory, depending upon its IT capability. Adherence to CDC as well as HIPAA standards for confidentiality and secure data transfer will be required; these standards include encryption before electronic transfer of any confidential surveillance data or information, and are described in Vol. III: Security and Confidentiality Guidelines.
- Establish a reporting schedule: daily, weekly, monthly, etc. Consider lab resources and volume of reports as well as field staff activities in the development of a schedule.
- Consider holding training session(s) for laboratory staff where reporting procedures can be clarified.

See Appendix D for examples of laboratory data elements that should be used for each lab test and result. Two examples of Documentation for Reporting Laboratory have been included.

Alternative Reporting Methods

Offer alternative methods of reporting if a laboratory cannot report electronically or cannot adapt to using your particular system.

- Be aware that laboratories have limits in their own systems.
- Consider an HIV/AIDS-specific electronic method that small laboratories unable to provide reporting through HL7 messaging could use. Some states have created secure Web-based reporting systems for laboratories to manually enter and transfer lab results to surveillance programs. Other states have customized MS-Access databases for laboratories to enter data, encrypt, and transfer these data.

Optional Reports

Depending on your state's regulations, your program may need to develop a method for handling optional reports (e.g., CD4 count ≥200 when state regulations require reporting of CD4 count <200). Unless your regulations prohibit receipt or maintenance of these optional reports, surveillance programs are best served by trying to accommodate all laboratory reporting.

For CD4 counts, it may be impossible for laboratories to exclude CD4 results that are unrelated to HIV/AIDS. In these instances, it will be important to be able to identify cancer treatment facilities in your state. This will allow your program to exclude or prioritize as low this type of optional report that originates from these facilities.

Resolving Referral Laboratory Problems

Determine a mechanism to resolve problems with incomplete data from referral laboratories. This happens most commonly when the testing laboratory indicates the ordering or referral laboratory, but does not include the name of the provider who ordered the test.

- Incomplete data from referral laboratories should be addressed in the writing of the reporting regulations and again when the health department approaches the laboratory about reporting. It is imperative that the accession number or some identifying number that connects the patient with the specimen is preserved. When a specimen has passed from laboratory to laboratory or facility to laboratory, the laboratory actually performing the test should report this identifier. (See Review Reporting Laws.)

Managing Increased Number of Lab Reports

ELR will increase the number of lab results that a surveillance program will need to manage and accurately link to an individual case. An automated system for record linkage may be needed to minimize the number of possible record links that require manual review for final disposition.

- Electronic lab data will likely arrive in a variety of formats. All data of disparate formats must be mapped to a standard format as is being proposed for eHARS. Linking to the registry will be done using the standard format. Procedures for

linking must be developed that include mechanisms for handling links, nonlinks, possible links, and follow-up on both false links and false nonlinks. Strategies for linking and eventual linkage to case reports can be found in the Record Linkage chapter.

- Lab data identifying potential new cases must be communicated with the field staff for completion of the case report form. A secure mechanism for supplying this information to the field staff will need to be implemented. Vol. III: Security and Confidentiality Guidelines includes standards for this kind of intra-program data sharing.

Laboratories First Reporting Electronically

As each laboratory starts a new procedure to report test results electronically (such as switching from paper reports to electronic reports, or shifting reporting responsibility from one group to another), the laboratory should continue to send paper reports until the procedure is well established. Your HIV/AIDS surveillance program should confirm that the compiled line list of paper test reports is the same as the line list of electronic case reports. This can be achieved by manually checking all or a random sample of accession numbers reported manually with the ones reported electronically from the same laboratory. When these two lists agree with each other, your surveillance program can then rely on the completeness of electronic reports from that laboratory and end the receipt of duplicate paper reports.

Reporting Back to Laboratories

Of the 28 areas conducting electronic laboratory reporting, 21 do not report back to laboratories on the quality of data received. (1)

- Develop a procedure for reporting back to laboratories on incomplete reports—an "audit," a "quality control," or a "report card" of sorts. Laboratory feedback should be a goal of the HIV/AIDS surveillance program, and should be modified as resources and experiences dictate. Ongoing communication with laboratories can help to improve accuracy and completeness of reporting the required and recommended data elements. In addition, careful monitoring of laboratory reporting patterns can result in detecting batches of missing reports that can be submitted by the lab in future submissions. Problems of this nature are often identified at the time of laboratory system modifications or staff turnover.

  - Data quality performance indicators that can be reported back to laboratories include completeness, timeliness, data validation, etc. A definition and method for calculating these indicators can be found in Data Quality.

  See Appendix E for two examples, (A) and (B), of quality control elements shared with reporting laboratories in Reporting Back to Laboratories.

- Develop and communicate to laboratories a procedure for reporting false-positive reports or lab errors so the surveillance team can update the previous records as appropriate.
  - Laboratories should be interested in false positives and incorrectly reported results. The surveillance program, through other activities, will be identifying inaccurate data and should report these back to the laboratory performing the test. This will involve comparing other surveillance data with laboratory data to obtain an evaluation or accuracy assessment.
    - Example: 'Do the lab data conflict with the data observed in the patient's chart?'
    - Examples of faulty data:
      - very high viral load with no positive HIV diagnostic test result
      - an incorrect HL7 message structure
      - a data entry error on the part of the laboratory
- Depending on the nature and implication of the lab error, the appropriate individuals, both within the surveillance program and laboratory, should be identified and notified.

HIV and HIV-associated Laboratory Tests

Diagnosing HIV Infection

The traditional sequence of laboratory testing for HIV infection relies on detecting antibodies to HIV in the infected person. The initial test (i.e., screening) technology is based on an immunoassay such as enzyme-linked immunosorbent assay (ELISA) or enzyme immunoassay (EIA). The test results are interpreted as either positive (reactive) or negative (nonreactive). When the screening test is repeatedly positive, a confirmatory test such as a Western blot (WB) assay or immunofluorescent antibody (IFA) is then performed. The confirmatory test results are interpreted as positive, negative, or indeterminate. These confirmatory tests use a different technology but again specifically detect antibodies to HIV.

Because the screening and confirmatory tests detect antibodies to HIV, there is a brief window period (usually weeks) in which a person may be HIV-infected but antibodies have not been produced to a level that can be detected by these tests. New testing technology or testing algorithms are being developed to narrow the window period.

Other Tests Indicative of HIV Infection

HIV-1 is the most common HIV type found in the United States. HIV-2 screening and diagnostic tests are also available and have been specified in the detailed test information below. When HIV/AIDS surveillance programs have evidence of an HIV-2 infection, they should contact CDC's HIV/AIDS surveillance branch directly and speak to the coordinator of Cases of Public Health Importance.

Other tests that have been used to diagnose HIV infection include HIV culture (which is unable to distinguish between HIV-1 and HIV-2), HIV-1 P-24 antigen, HIV-1 proviral DNA polymerase chain reaction (PCR), and HIV-1 RNA PCR.

Testing related to HIV incidence and viral resistance (e.g., serologic testing algorithm for recent HIV seroconversion [STARHS] and viral drug resistance) is not covered in this document. Further information about these laboratory tests should be directed to the CDC Incidence and Viral Resistance Team.

HIV Disease Progression and Markers for Access to Care

CD4+ T-lymphocyte tests are used to assess immunosuppression in HIV-infected patients to guide therapeutic decisions. Consequently, CD4 testing at diagnosis and every 3–6 months thereafter is recommended in the management of HIV disease (2). Recent treatment guidance recommends considering antiretroviral therapy when CD4 count is ≤350 cells/µl (2). CD4 results are also used in determining whether an HIV-infected person meets the AIDS immunologic case definition (i.e., CD4 count <200 cells/µl or CD4 percent <14% of total lymphocytes); in 1993, these immunologic criteria were added to the AIDS-defining diseases and conditions (3).

CD4 count, obtained at the time of initial HIV diagnosis, enables the staging of HIV disease. In general, it can be used as an indicator of elapsed time from initial infection to diagnosis: the greater the elapsed time or delay to HIV testing, generally the greater the immunosuppression and the lower the CD4 count. HIV/AIDS programs should make every attempt to collect initial CD4 count (i.e., within 3 months of HIV diagnosis) because the presence and interpretation of CD4 counts reflects whether early HIV testing occurred—from recognition of risky exposures, need for testing, and actual HIV testing—and linkage to care efforts by public health. This may be done by collecting all CD4 counts from laboratories or during case ascertainment field investigations, depending on the resources and structure of each HIV/AIDS surveillance program.

For CD4 reporting, it may be impossible for laboratories to exclude CD4 results that are unrelated to HIV/AIDS (e.g., those that may originate from oncology patients). In these instances, it will be important to be able to identify cancer treatment facilities or cancer treatment providers that may request CD4 testing. This will enable identification of those reports related to HIV disease monitoring.

Viral load monitoring is also important in managing HIV disease care and treatment. Viral load should be obtained at diagnosis and may be a factor in deciding when to initiate antiretroviral therapy (2). A goal of therapy is to reduce viral load levels below the level of test detection (i.e., undetectable viral load result).

Viral load and CD4 testing are also markers of a person's access to care. Furthermore, CD4 testing can be quite specific for HIV disease; >90% of CD4 testing was for HIV-infected persons, as reported by a state conducting HIV/AIDS surveillance (4). Presence of viral load and CD4 test results indicates the patients' clinical status is being monitored, and they are thus receiving some health care. Ongoing testing that meets the current treatment standards is a marker for established, ongoing care and provides some evidence

for the quality of care received. Lack of any viral load or CD4 test result following HIV diagnosis in surveillance data would be interpreted as an unmet health care need and that linkage to care has somehow failed.

Because CD4+ T-lymphocyte and HIV viral load testing are key components of monitoring and managing HIV disease, CDC recommends that all states require laboratory reporting of all levels of CD4 and both detectable and nondetectable viral load results to their state public health departments. Further justification can be found in a CDC-supported Council of State and Territorial Epidemiologists (CSTE) position statement (5).

Changes in Laboratory Reporting Initiated by ELR

Reporting laboratories are encouraged to use the standard HL7 format when transmitting lab result information. (See SNOMED, LOINC, and mapping in Appendix A for more information.) Additionally, because of the needs of public health, reporting laboratories have been encouraged to submit information used for surveillance purposes (e.g., race, date of birth, etc.) when available. Furthermore, because of performance characteristic differences between manufacturer-specific tests, HIV/AIDS surveillance programs should collect information that can help distinguish between tests. Viral load testing is an example for which the identification of manufacturer has been recommended as part of standardized reporting (6). Also, as future testing algorithms for HIV diagnosis are introduced, there will be a further need to collect manufacturer and specimen source.

LOINCs have the potential to uniquely identify a clinical test, test method, and specimen source. To avoid the propagation of additional, redundant LOINCs, HIV/AIDS surveillance programs are encouraged to use the LOINCs specified in this document and to propose that laboratories reporting to them also use these unique LOINCs.

See Appendix F for a list of LOINCs specific for HIV and HIV-associated testing.

Performance Standards

Performance standards will evolve over time. They are intended to provide a framework of goals and activities that all HIV/AIDS surveillance programs can and should accomplish. Some of the performance standards are goals that programs can use for future planning; others are process standards that programs should strive to meet to ensure local, state, and national data are of high quality.

As electronic lab reports become an increasingly important vehicle for initial case identification, the discriminating information that the lab report contains will be critical to link a lab result to an existing case or to conduct follow-up for a potentially new case.

Case Follow-up from a Lab Report

The flow chart below presents the general outline for receiving laboratory reports, completing investigations when appropriate, and entering data. The initial step is to determine if the laboratory information is sufficiently complete to conduct further steps, to reject the result, or return the information to the laboratory for clarification. The essential elements of any laboratory report include

- the name of the laboratory or CLIA code
- the date of the test (e.g., specimen collection date or date test run)
- the type of test that was conducted, LOINC, or local code
- the test result (which must be a test of interest, i.e., not an incorrectly reported result such as a hepatitis result)
- a patient identifier, such as name, Social Security Number, medical record number, lab accession number, insurance number, Department of Corrections number, coded-identifier or facility-specified identifier
- the name of the physician or clinic ordering the test

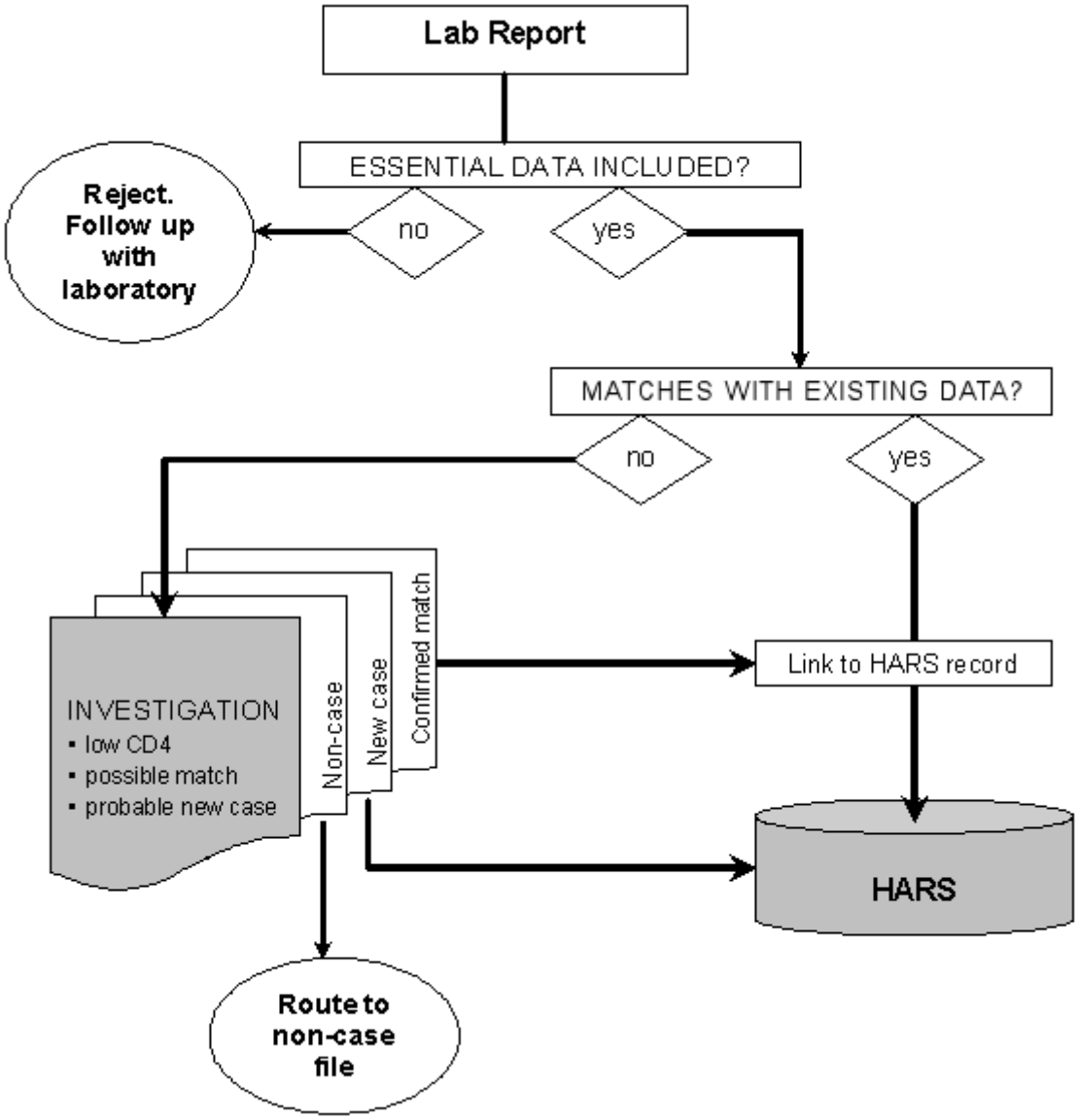
If any of these essential elements are missing, the laboratory needs to be contacted to clarify any missing information. If all of these elements are supplied, then continue and

- Reject certain records based upon the type of test or the test result. If these are results that were not requested, some quality control follow-up is required with the laboratory.
- Contact the provider to obtain information needed to conduct a linkage.
- Link records to determine if the lab result is for an existing case or a potential new case.

Figure 10 depicts the steps taken in follow-up of a lab result. Many details have been omitted for the purpose of generalization.

See Appendix G for examples of more detailed program-specific case follow-up flow charts. Three examples—(A), (B), and (C)—of Case Follow-up from a Lab Report have been included.

Figure 10 Lab report follow-up



The Ideal Complete Lab Report

While essential data elements are necessary to keep or reject each individual lab report (Figure 10), the data elements considered essential will differ for each HIV/AIDS surveillance program. In general, an extensive list of variables is recommended for conducting laboratory surveillance. Not all of these variables will be available from all laboratories, but this list may be used as a reference for the ideal laboratory report.

Table 5 includes a list of variables laboratories are strongly encouraged to report because of their value to surveillance. They have been grouped by subject elements, but the relative importance of one field element over another differs for reporting areas. As such, this list should **not** be distributed to laboratories reporting to your surveillance program but appears here for completeness. Your program may choose a subset of these data

elements, depending on your program's needs and priorities. Lists of lab result variables that laboratories should transmit to your program can be found in <*Volume II, Data Dictionary*> or see Appendix D for examples of lab data dictionaries in use by surveillance programs.

Table 5 Data elements in an ideal complete lab report

FIELD ELEMENT	Notes	ALTERNATIVE FIELD ELEMENT	Notes
Patient Identifier			
Patient Last Name		**Elements that comprise state-specific code**	
Patient First Name			
Patient Middle Name/Middle Initial			
		OR	
		Medical record number from referral facility	If no referral facility is involved in testing the specimen, this field will be blank.
		Medical record number from testing facility	
		Alternate Patient ID	If an outside laboratory has performed the test, and the referral facility is reporting the result, include the patient identifier from the outside laboratory in this field.
		OR	
		Dept. of Corrections ID	Inmate number
		Insurance or Billing Number	
		Patient Middle Initial	
Patient Social Security Number	Field may be valued as Social Security Number or Railroad Retirement Number.		

Table 5 Data elements in an ideal complete lab report

FIELD ELEMENT	Notes	ALTERNATIVE FIELD ELEMENT	Notes
Patient Demographics			
Patient Date of Birth	Data must be formatted as MM/DD/YYYY (e.g., 12/01/1952).	**Patient Age**	Age at time of first documented HIV+ diagnosis
Patient Gender	Allowable field values are as follows: • Female • Hermaphrodite/ Undetermined • Male • Other • Transsexual • Unknown		
Patient Race	Allowable field values are as follows: • Asian • Black • American Indian or Alaska Native • Multiracial • Other • Pacific Islander • White • Unknown		
Patient Ethnic Group	Allowable field values are as follows: • Hispanic • Non-Hispanic • Unknown		

Table 5 Data elements in an ideal complete lab report

FIELD ELEMENT	Notes	ALTERNATIVE FIELD ELEMENT	Notes
Patient Residence			
Patient Street Address			
Patient County of Residence	If valued, field must be a valid Federal Information Processing Standard (FIPS) code for the patient's county of residence. Valid FIPS county codes by state (as defined by the United States Environmental Protection Agency) can be found at the following Web site: http://www.epa.gov/enviro/html/codes/state.html		
Patient City of Residence		**Patient ZIP Code**	
Patient State of Residence	If valued, field must be a valid USPS state abbreviation code to identify the state of residence. Valid state codes (as defined by the United States Postal Service) can be found at the following Web site: http://www.usps.com/ncsc/lookups/abbr_state.txt		
Patient Telephone Number			

Table 5 Data elements in an ideal complete lab report

FIELD ELEMENT	Notes	ALTERNATIVE FIELD ELEMENT	Notes
Provider Identification			
Provider Last Name **Provider First Name** **Provider Middle Initial**		**Ordering Facility Name**	
Provider Street Address		**Ordering Facility Street Address**	
Provider City		**Ordering Facility City** ***OR*** **Ordering Facility ZIP**	
Provider State	If valued, field must be a valid USPS state abbreviation code to identify the state of residence. Valid state codes (as defined by the United States Postal Service) can be found at the following Web site: http://www.usps.com/ncsc/lookups/abbr_state.txt	**Ordering Facility State**	If valued, field must be a valid USPS state abbreviation code to identify the state of residence. Valid state codes (as defined by the United States Postal Service) can be found at the following Web site: http://www.usps.com/ncsc/lookups/abbr_state.txt
Provider ZIP Code			
Provider Phone Number		**Ordering Facility Phone Number**	
Originating Laboratory (when specimen has been sent to another laboratory for testing, e.g., a specialty laboratory)			
Sending Facility Name	Source of electronic lab report	**Referral Facility Name**	If no referral facility is involved in testing the specimen, this field will be blank.
Sending Facility CLIA	Unique Clinical Laboratory Improvement Amendments Identifier number.		

Table 5 Data elements in an ideal complete lab report

FIELD ELEMENT	Notes	ALTERNATIVE FIELD ELEMENT	Notes
Testing Facility Identification			
Testing Facility Name			
Testing Facility CLIA Code			
Specimen Identification			
Accession Number			
Specimen Received Date	Date specimen received by testing facility. Date field must be formatted as MM/DD/YYYY.		
Specimen Collection Date & Time	Date field must be formatted as MM/DD/YYYY. Time must be formatted as HH:MM (e.g., 12:07) in military time format. Valid values range from 00:00 through 23:59. A blank space should be included between the date and time components (e.g., 10/01/2000 17:42).		
Specimen Analysis Date	Date specimen tested. Date field must be formatted as MM/DD/YYYY.		
Specimen Result Report Date	Date specimen result reported. Data must be formatted as MM/DD/YYYY (e.g., 10/01/2000).		
Specimen Description		**Observation Specimen Source Code**	
		OR	
		Observation Specimen Source Text	

Table 5 Data elements in an ideal complete lab report

FIELD ELEMENT	Notes	ALTERNATIVE FIELD ELEMENT	Notes
Test Identification and Result			
Lab Code		**LOINC Code**	
Lab Code Description		**LOINC Description**	
Manufacturer's Name of Test Kit			
Test Result			
Observation Unit of Measure		**Reference Range**	
Electronic Transmission Identifiers			
Receiving Application	Example of allowable field values are • Bureau of HIV/AIDS • Cancer Registry • Bureau of Communicable Disease Control • Lead Program • Bureau of Sexually Transmitted Diseases • Bureau of Tuberculosis Control		
Lab Report Transmission Date and Time	Data must be formatted as MM/DD/YYYY HH:MM (e.g. 10/01/2000 17:39) with time submitted in military time format. Valid military time values range from 00:00 through 23:59.		
Record Termination Indicator	Field must be valued with 2 exclamation points (!!) to indicate the end of each record.		

Measurable Performance Standards

Because electronic reporting of cases and lab results are relatively new practices to enhance surveillance and not used by all reporting areas, only process standards have been provided in this guideline. Over time, with more experience and as these practices become universal activities to HIV/AIDS surveillance programs, outcome measures may be developed.

Quality of Laboratory Reporting

- HIV/AIDS surveillance programs should send a completeness and data quality "report card" of lab data elements regularly (no fewer than two times a year) to reporting laboratories, meeting a minimum reporting volume.

  For examples of report cards used by surveillance areas, see Appendix E.2.

- HIV/AIDS surveillance programs should receive batched electronic lab reporting at least monthly. Reporting occurring less frequently than monthly should be followed up by the surveillance program to determine reason for lapse.

  Use the report card to indicate

  - the total number of records reported by the laboratory
  - the number of records containing all of the essential elements outlined in Case Follow-up from a Lab Report
  - the number of records that were rejected because they were not appropriate HIV tests or results
  - the number of records that were sent back to the laboratory for clarification because not all of the essential elements were supplied

  Use the report card to indicate the completeness of the nonessential variables.

  - Indicate the completeness (presence) of each variable required by law or regulation.
  - Indicate the completeness (presence) of each variable requested by surveillance staff.

Assessing Laboratory Reporting of CD4 and Viral Load Testing

Laboratory testing (i.e., CD4 and viral load) can be used as a marker of entry to receipt of health care. Conversely, the lack of laboratory testing suggests lack of entry to medical care and is an unmet health care need for HIV-infected persons. The completeness of CD4 and viral load reporting will affect its usefulness as a marker of access to health care, and wide variation in area-to-area estimates of access to care will inevitably lead to an understanding of or explanation for the area-to-area variability. Over time and with reporting area participating, national standards may be established.

At least 50% of persons living with HIV/AIDS will have an initial CD4 and/or viral load test results (i.e., within 3 months of diagnosis) reported to the national surveillance system by the HIV/AIDS surveillance program.

Assessing Laboratory Reporting of CD4 Counts Following Initial HIV Diagnosis

Laboratory reporting of CD4 counts serves many surveillance purposes. One use of CD4 reporting is for CD4 count obtained at the time of initial HIV diagnosis to stage HIV disease. Population-based CD4 count at diagnosis has implications for prevention and outreach interventions. Although the actual obtainment of a CD4 count at diagnosis requires health care "intervention" for the ordering of the test, a minimum level of CD4 reporting should be expected for all reporting areas. Areas not able to achieve the minimum level of CD4 reporting may need to investigate surveillance practices and adjust those practices to ensure adequate collection and reporting of CD4 count results.

- Minimum standard:
  - For each calendar year, at least 50% of newly diagnosed persons living with HIV/AIDS, aged ≥13 years, will have an initial CD4 count (i.e., CD4 specimen collected within 3 months of HIV diagnosis) reported to the national HIV/AIDS surveillance system no later than 12 months following diagnosis.
- Optional, additional standards include
  - Determine median and mean CD4 count for persons with
    - HIV without AIDS for >12 months following initial HIV diagnosis
    - HIV that progresses to AIDS 1–12 months following initial HIV diagnosis
    - HIV with AIDS at diagnosis

References

1 2004 Surveillance Coordinators' Survey results. Available with password at http://www2a.cdc.gov/hicsb/. Accessed October 14, 2004.

2 Department of Health and Human Services. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents, October 29, 2004. Available at http://aidsinfo.nih.gov/ContentFiles/AdultandAdolescentGL10292004002.pdf.

3 Centers for Disease Control and Prevention. 1993 revised classification system for HIV infection and expanded surveillance case definitions for AIDS among adolescents and adults. *MMWR* 1992; 41: 1–19.

4 Centers for Disease Control and Prevention. Assessment of laboratory reporting to supplement active AIDS surveillance — Colorado. *MMWR* 1993; 42: 749–752.

5 Council of State and Territorial Epidemiologists. CSTE Position Statement (04-ID-07): Laboratory reporting of clinical test results indicative of HIV infection: new standards for a new era of surveillance and prevention. Available at http://www.cste.org/ps/2004pdf/04-ID-07-final.pdf.

6 Centers for Disease Control and Prevention. Guidelines for laboratory test result reporting of human immunodeficiency virus type 1 ribonucleic acid determination: recommendations from a CDC working group. *MMWR* 2001;50 (No. RR-20): 1–14. Available at http://www.cdc.gov/mmwr/PDF/RR/RR5020.pdf.

List of Appendixes

Appendix A 5-36
Electronic Laboratory Reporting Resources 5-36
Appendix B 5-39
Appendix B.1: Fields for HIV and AIDS Reporting by Position 5-39
Appendix B.2: Variables in Order 5-42
Appendix C 5-51
Identifying All Reporting Laboratories 5-51
Appendix D 5-56
Data Record Layout 5-56
Sample ASCII File Data Record 5-62
Appendix E 5-63
Reporting Back to Laboratories 5-63
Appendix E.1 5-63
Appendix E.2 5-71
Appendix F 5-75
Changes in Laboratory Reporting Initiated by Electronic Laboratory Reporting (ELR) 5-75
Screening Tests 5-75
Confirmatory Tests 5-76
Other HIV Detection Tests 5-77
CD4+ T-Lymphocyte Tests 5-79
Viral Load Tests 5-80
Viral Susceptibility Tests 5-81
Appendix G 5-82
Case Follow-up from a Lab Report 5-82
Processing Laboratory Test Reports 5-82
Overview of Laboratory Data Flow 5-83
HIV/AIDS Surveillance System Overview 5-84

Appendix A

Electronic Laboratory Reporting Resources

General Electronic Laboratory Reporting

1. Electronic reporting of laboratory information for public health (1999), 41 pages
http://0-www.cdc.gov.mill1.sjlibrary.org/nedss/ELR/ELR_LabInfo_1999.pdf
Outlines experiences in several states
Includes good descriptions of security issues, encryption, etc.
Main content covers pages 18–32

2. Electronic reporting of lab data for public health (1997), 146 pages
http://wwwn.cdc.gov/dls/pdf/elr1030.pdf
Summary is on pages 4–6
Background and recommendations are on pages 7–17
Appendixes pages 18–145

Health Level 7

3. HL7 specifications, 1997, 70 pages
http://www.cdc.gov/nedss/ELR/HL7Spec.pdf

4. Implementation guide for transmission of laboratory-based reporting of public health information using version 2.3.1 of the Health Level Seven (HL7) standard protocol
http://www.cdc.gov/PHIN/Architecture/Implementation_Guides/Laboratory/PHIN_Laboratory_Result_%20ELR_v231.pdf
This 86-page guide is dated May 2005.

5. Introduction to the Public Health Information Network (PHIN)
http://www.cdc.gov/phin/index.html
This is the CDC/PHIN website. Content ranges from a general overview of PHIN to Functional Requirements and Technical Specifications (Implementation Guides, KPMs, and Data Models).

SNOMED, LOINC, and mapping

6. The National Library of Medicine has developed a Unified Medical Language System metathesaurus which describes both SNOMED and LOINC, and allows you to download each free of charge.
http://www.nlm.nih.gov/research/umls/

7. Description of SNOMED
http://www.nlm.nih.gov/research/umls/Snomed/snomed_main.html

8. SNOMED organization home page
http://www.snomed.org

9. Description of LOINC
http://www.nlm.nih.gov/research/umls/loinc_main.html

10. LOINC, a universal standard for identifying laboratory observations: a 5-year update
http://www.clinchem.org/cgi/content/full/49/4/624?ijkey=oUwNonbF33rao&keytype=ref&siteid=clinchem
This article provides an overview, background, and description of LOINC codes.

11. Regenstrief organization home page
http://www.regenstrief.org/loinc

12. Download of LOINC version 2.13
http://www.regenstrief.org/loinc/download (71 megabyte)
RELMA version 3.13 is also available at this site. RELMA is mapping software for relating local codes to the LOINC database.

13. Introduction to PHIN notifiable condition mapping tables (NCMT, May 2004)
http://www.cdc.gov/PHIN/Vocabulary/Introduction_to_the_PHIN_Notifiable_Condition_Mapping_Tables.doc

13a. Description of LOINC to condition mappings in the PHIN NCMT
http://www.cdc.gov/PHIN/Vocabulary/LOINC_Introduction.doc

13b. Excel table of LOINC to condition mappings, style A
http://www.cdc.gov/phin/vocabulary/LOINC_to_Condition_StyleA.xls
Note there are two options for mapping LOINC codes to conditions. Both styles produce incorrect mappings, but will be updated in 2005. The existing codes are not specific enough for pediatric vs. adult HIV infection vs. AIDS. Style A maps each HIV test to one of the three HIV conditions.

13c. Excel table of LOINC to condition mappings, style B
http://www.cdc.gov/phin/vocabulary/LOINC_to_Condition_StyleB.xls
Style B maps each HIV test to all three HIV conditions.

13d. Description of the SNOMED organism tables in the PHIN NCMT
http://www.cdc.gov/phin/vocabulary/SNOMED_Organism_Lists.doc

13e. SNOMED organism tables
http://www.cdc.gov/phin/vocabulary/SNOMED_Organism_Lists.xls

14. Dwyer tables for mapping organisms to SNOMED codes
http://www.cdc.gov/nedss/DataModels/DWYERIII_SNOMED.pdf

Related Topics

Information about ebXML is available at http://www.ebxml.org/
ebXML (Electronic Business using eXtensible Markup Language) is a modular suite of specifications that enables enterprises of any size and in any geographical location to conduct business over the Internet. Therefore, ebXML is the mechanism or processes by which XML documents can be moved from one business partner to another. Using ebXML, companies now have a standard method to exchange business messages, conduct trading relationships, communicate data in common terms, and define and register business processes.

Published references

McDonald CJ, Huff SM, Suico JG, Hill G, Leavelle D, Aller R, et al. LOINC, a universal standard for identifying laboratory observations: a 5-year update. *Clin Chem* 2003; 49(4): 624–633.

Wurtz R and Cameron BJ. Electronic laboratory reporting for the infectious disease physician and clinical microbiologist. *Clin Infect Dis* 2005; 40(1 June): 1638–1643.

Appendix B

Documentation for Reporting Facilities and Providers

(Appendixes B.1 and B.2)

Appendix B.1: Fields for HIV and AIDS Reporting by Position

Obs	name	var_type	length	label
1	**ID**	**Character**	**10**	**Patient's ID (at health facility)**
2	**COMPLTED**	**Character**	**10**	**Date initial AIDS/HIV form completed**
3	**SOURCER**	**Character**	**2**	**Source of AIDS report**
4	**HSOURCE**	**Character**	**2**	**Source of HIV report**
5	NAME	Character	45	Name of AIDS case (can be longer to include all text)
6	CPHONE	Character	12	Patient's current phone number (AIDS cases only)
7	SNDX	Character	4	Soundex code for last name
8	BIRTH	Character	10	Date of patient's birth
9	SEX	Character	1	Sex of patient
10	SOCSEC	Character	11	Full Social Security Number for AIDS (xxx xx xxxx); last 4 digits for HIV (xxxx)
11	LAB_NO	Character	15	Lab-generated Accession Number/Lab Report No
12	CT_NO	Character	15	Confidential Counseling &Testing Number
13	**DIAGSTAT**	**Character**	**1**	**Diagnostic Status at Report; 1= HIV; 2=AIDS**
14	**HAGE_YRS**	**Character**	**2**	**Age in years when first found HIV+**
15	**AGE_YRS**	**Character**	**2**	**Age in yrs at dx of AIDS**
16	**STAT**	**Character**	**1**	**Current mortality status**
17	**DEATH**	**Character**	**10**	**Date of patient death**
18	**DEATH_ST**	**Character**	**2**	**State of death**
19	**HISP**	**Character**	**1**	**Patient's Ethnicity (1=Hispanic; 2=non-Hispanic; 9=Unknown)**
20	**RACE_A**	**Character**	**1**	**Race for Asian; 1=yes**
21	**RACE_B**	**Character**	**1**	**Race for Black; 1=yes**
22	**RACE_I**	**Character**	**1**	**Race for American Indian/Alaska Native; 1=yes**
23	**RACE_P**	**Character**	**1**	**Race for Hawaiian/Pacific Islander; 1=yes**
24	**RACE_W**	**Character**	**1**	**Race for White; 1=yes**
25	**RACE_U**	**Character**	**1**	**Race for Unknown; 1=yes**
26	**XRACE**	**Character**	**2**	**Expanded race (see hard copy for codes)**

EXAMPLE

Obs	name	var_type	length	label
27	**ORIGIN**	**Character**	**1**	**Country or territory of birth (1=US born; 2=foreign born; 9=unk)**
28	**ORIG_OTH**	**Character**	**3**	**Fill this field with XRACE code if not US born**
29	**US_DEPND**	**Character**	**20**	**US dependency where born**
30	**HOMELESS**	**Character**	**1**	**Is the patient homeless at time of AIDS/HIV dx? 1=yes**
31	**RCITY**	**Character**	**27**	**City of residence when AIDS/HIV dx**
32	**ADDRESS**	**Character**	**45**	**Patient's address (AIDS patient only)**
33	**CURR_ST**	**Character**	**2**	**Patient's current state**
34	**CURRCITY**	**Character**	**27**	**Patient's current city**
35	**CURR_ZIP**	**Character**	**9**	**Patient's current ZIP code**
36	**CURRCNTY**	**Character**	**27**	**Patient's current county**
37	**ST**	**Character**	**2**	**State of residence when AIDS dx or when HIV+**
38	**ZIP_CODE**	**Character**	**9**	**ZIP of residence when AIDS dx or when HIV+**
39	**RCOUNTY**	**Character**	**27**	**County of residence at AIDS dx or when HIV+**
40	**RCNTRY**	**Character**	**3**	**Country of residence at AIDS dx or when HIV+**
41	HOSP_DX	Character	27	Facility where AIDS dx
42	HOSP_ST	Character	2	State of facility where AIDS dx
43	HOSP_CTY	Character	27	City of facility where AIDS dx
44	FEDSET	Character	1	Facility setting for AIDS dx
45	FAC_TYPE	Character	2	Type of facility where AIDS dx
46	HHOSP_DX	Character	27	Facility where HIV+
47	HHOSP_ST	Character	2	State of facility where HIV+
48	HHSP_CTY	Character	27	City of facility where HIV+
49	HFEDSET	Character	1	Facility setting for HIV dx
50	HFAC_TYP	Character	2	Type of facility where HIV dx
51	**SEX_MALE**	**Character**	**1**	**Sexual relations with male**
52	**SEX_FMLE**	**Character**	**1**	**Sexual relations with female**
53	**IV**	**Character**	**1**	**IV drug user**
54	**S_IV**	**Character**	**1**	**Sex with IV drug user**
55	**S_BI**	**Character**	**1**	**Sex with bisexual man**
56	**S_HEMO**	**Character**	**1**	**Sex with hemophiliac**
57	**S_TX**	**Character**	**1**	**Sex with transfusion recipient**
58	**S_TRNPLT**	**Character**	**1**	**Sex with transplant recipient**
59	**S_HIV**	**Character**	**1**	**Sex with person with AIDS/HIV**

EXAMPLE

Obs	name	var_type	length	label
60	**BLDPRD**	**Character**	**1**	**Rec'd blood prod(clotting fac)**
61	**TYP_HEMO**	**Character**	**1**	**Type coagulation disorder**
62	**OTH_HEMO**	**Character**	**2**	**Other coagulation disorder**
63	**TRANSFUS**	**Character**	**1**	**Rec'd blood or blood components**
64	**TRANDTE1**	**Character**	**7**	**Date of first transfusion (mm/yyyy)**
65	**TRANDTE2**	**Character**	**7**	**Date of second transfusion (mm/yyyy)**
66	**TRANPLNT**	**Character**	**1**	**Tissue/transplant recipient**
67	**HCW**	**Character**	**1**	**Patient a health care worker**
68	**OCCUP**	**Character**	**3**	**Patient's hcw occupation**
69	EIA	Character	1	First EIA test
70	EIA_MOYR	Character	7	Date 1st EIA test (mm/yyyy)
71	COMBI	Character	1	First combi test
72	COMBMOYR	Character	7	Date of 1st combi test (mm/yyyy)
73	WBIFA	Character	1	First Western blot test
74	WBMOYR	Character	7	Date of 1st Western blot test (mm/yyyy)
75	HIVAB	Character	2	First other AB test
76	HIVABMOYR	Character	7	Date of 1st other AB test (mm/yyyy)
77	HIVABRES	Character	1	First other AB result
78	HIVDTEST	Character	2	First HIV detection test
79	HIVDMOYR	Character	7	Date 1st HIV detection test (mm/yyyy)
80	HIVDET	Character	1	Result of first HIV detection test
81	HVLOAD	Character	8	Viral load 1 (copies per ml)
82	HVDMOYR	Character	7	DATE FIRST VIRAL LOAD TEST (mm/yyyy)
83	LNEGMOYR	Character	7	Date of last documented negative HIV test (mm/yyyy)
84	LNEGTYPE	Character	2	Specify type of last documented negative HIV test
85	DOC_DIAG	Character	1	If HIV lab test not documented, is HIV diagnosis documented by a physician? 1=yes
86	DOC_MOYR	Character	7	Date of physician's diagnosis of HIV (mm/yyyy)
87	**HIVPMOYR**	**Character**	**7**	**First HIV positive date (mm/yyyy)**
88	TH1CNT	Character	4	Most current CD4 count
89	TH1PCT	Character	2	Most current CD4 percent
90	TH1MOYR	Character	7	Date of most current CD4 test (mm/yyyy)
91	CD4CNT	Character	4	First low CD4 count (<200)
92	CD4PCT	Character	2	First low CD4 percent (<14)
93	CD4MOYR	Character	7	Date of first low CD4 test (mm/yyyy)
94	**REVIEWED**	**Character**	**1**	**Clinical record reviewed**

EXAMPLE

Obs	name	var_type	length	label
95	**ASYMMOYR**	**Character**	**7**	**Date diagnosed Asymptomatic (mm/yyyy)**
96	**SYMPMOYR**	**Character**	**7**	**Date diagnosed Symptomatic (mm/yyyy)**
97	**CANDLUNG**	**Character**	**1**	**Candidiasis bronchi/lungs/trachea**
98	**CLNGMOYR**	**Character**	**7**	**Dx date for candida bronchi/lungs/trachea (mm/yyyy)**
99	**CANDESOP**	**Character**	**1**	**Candidiasis esophageal**
100	**CESOMOYR**	**Character**	**7**	**Dx date for esophageal candidiasis (mm/yyyy)**
101	**CERVDIS**	**Character**	**1**	**Carcinoma, invasive cervical**
102	**CDISMOYR**	**Character**	**7**	**Dx date for carcinoma cervical (mm/yyyy)**
103	**COCCI**	**Character**	**1**	**Coccidioidomycosis**
104	**CCMOYR**	**Character**	**7**	**Dx date for coccidioidomycosis (mm/yyyy)**

Appendix B.2: Variables in Order

Order	Variable	Typ	Len	lenbegin	lenend	Label	Format
1	NAME	Char	40	1	40	Name of patient	(Last, First, Middle Initial)
2	ADDRESS	Char	40	41	80	Patient's current street address	Street and Number
3	CURRCITY	Char	27	81	107	Patient's current city	For city name, use Table 5
4	CURR_ST	Char	2	108	109	Patient's current state	Abbreviate (such as PA)
5	CURR_ZIP	Char	9	110	118	Patient's current ZIP code	5 digit
6	SSN	Char	11	119	129	Patient's Social Security Number	xxx xx xxxx
7	HCOMPLTD	Char	10	130	139	Date initial HIV+ form compltd	mm/dd/yyyy
8	COMPLTD	Char	10	140	149	Date initial AIDS+ form compltd	mm/dd/yyyy
9	HSOURCE	Char	2	150	151	Source of HIV+ infection report	see Table 1 for code of source
10	SOURCER	Char	2	152	153	Source of AIDS report	see Table 1 for code of source
11	REP_ST	Char	2	154	155	State of report	Abbreviate (such as PA)
12	REP_CITY	Char	27	156	182	City of report	For city name, use Table 5
13	STATENO	Char	10	183	192	State patient number	If available

EXAMPLE

Order	Variable	Typ	Len	lenbegin	lenend	Label	Format
14	CITYNO	Char	10	193	202	Reporting city patient id num	If available
15	DIAGSTAT	Char	1	203	203	Diagnostic Status at Report	1=adult HIV, 2=Adult AIDS
16	BIRTH	Char	10	204	213	Date of patient birth	mm/dd/yyyy
17	STAT	Char	1	214	214	Current mortality status	1=Alive, 2=Dead, 3=Moved, 4=Lost to flwup, 9=Unknown
18	DEATH	Char	10	215	224	Date of patient death	mm/dd/yyyy
19	DEATH_ST	Char	2	225	226	State of death	Abbreviate (such as PA)
20	SEX	Char	1	227	227	Patient's Sex	1=male, 2=female
21	RACE	Char	1	228	228	Patient's Race/Ethnicity	1=White, 2=Black, 4=Asian, 5=Am. Indian, 9=Not specified
22	ETHNIC	Char	1	229	229	Patient's Ethnicity	1=Hispanic, 0=non-Hispanic
23	ORIGIN	Char	1	230	230	Country or territory of birth	1=USA, 2=Canada, 3=Dominican Republic, 4=Haiti, 5=Mexico, 7=US dependency, 8=Other, 9=Unknown
24	US_DEPND	Char	2	231	232	US dependency where born	If you report 7 in Origin, see Table 2 for US Dependency
25	ORIG_OTH	Char	3	233	235	If not US country of birth	If you report 8 in Origin, see Table 3 for other counties
26	HCITY	Char	27	236	262	City of residence when HIV positive	For city name, use Table 5
27	HCOUNTY	Char	27	263	289	County of residence when HIV positive	For county name, use Table 6
28	HST	Char	2	290	291	State of residence when HIV positive	Abbreviate (such as PA)
29	HZIP	Char	9	292	300	ZIP code of residence when HIV positive	5 digit
30	RCITY	Char	27	301	327	City of residence when AIDS dx	For city name, use Table 5

EXAMPLE

Order	Variable	Typ	Len	lenbegin	lenend	Label	Format
31	RCOUNTY	Char	27	328	354	County of residence at AIDS dx	For county name, use Table 6
32	ST	Char	2	355	356	State of residence when AIDS dx	Abbreviate (such as PA)
33	ZIP_CODE	Char	9	357	365	ZIP of residence when AIDS dx	5 digit
34	HFEDSET	Char	1	366	366	Facility setting for HIV dx	1=public, 2=private, 3=federal
35	HFAC_TYP	Char	2	367	368	Type of facility where HIV dx	See Table 1 for type of facility
36	HHOSP_DX	Char	27	369	395	Facility where HIV dx	See Table 4 for facility names
37	HHSP_CTY	Char	27	396	422	City of facility where HIV dx	For city name, use Table 5
38	HHOSP_ST	Char	2	423	424	State of facility where HIV dx	Abbreviate (such as PA)
39	FEDSET	Char	1	425	425	Facility setting for AIDS dx	1=public, 2=private, 3=federal
40	FAC_TYPE	Char	2	426	427	Type of facility where AIDS dx	See Table 1 for type of facility
41	HOSP_DX	Char	27	428	454	Facility where AIDS dx	See Table 4 for facility names
42	HOSP_CTY	Char	27	455	481	City of facility where AIDS dx	For city name, use Table 5
43	HOSP_ST	Char	2	482	483	State of facility where AIDS dx	Abbreviate (such as PA)
44	SEX_MALE	Char	1	484	484	Sexual relations with male	0=No, 1=Yes, 2=CDC Confirmed, 9=Unknown
45	SEX_FMLE	Char	1	485	485	Sexual relations with female	0=No, 1=Yes, 2=CDC Confirmed, 9=Unknown
46	IV	Char	1	486	486	IV drug user	0=No, 1=Yes, 2=CDC Confirmed, 9=Unknown
47	BLDPRD	Char	1	487	487	Rec'd blood prod(clotting fac)	0=No, 1=Yes, 9=Unknown
48	TYP_HEMO	Char	1	488	488	Type coagulation disorder	1=Hemo A, 2=Hemo B
49	OTH_HEMO	Char	2	489	490	Other coagulation disorder	0=No, 1=Yes, 9=Unknown

EXAMPLE

Order	Variable	Typ	Len	lenbegin	lenend	Label	Format
50	S_IV	Char	1	491	491	Sex with IV drug user	0=No, 1=Yes, 9=Unknown
51	S_BI	Char	1	492	492	Sex with bisexual man	0=No, 1=Yes, 9=Unknown
52	S_HEMO	Char	1	493	493	Sex with hemophiliac	0=No, 1=Yes, 9=Unknown
53	S_TX	Char	1	494	494	Sex with transfusion recipient	0=No, 1=Yes, 9=Unknown
54	S_TRNPLT	Char	1	495	495	Sex with transplant recipient	0=No, 1=Yes, 9=Unknown
55	S_HIV	Char	1	496	496	Sex with person with AIDS/HIV	0=No, 1=Yes, 9=Unknown
56	TRANSFUS	Char	1	497	497	Rec'd bld or bld components	0=No, 1=Yes, 2=CDC Confirmed, 9=Unknown
57	TRANDTE1	Char	10	498	507	Date of first transfusion	mm/dd/yyyy
58	TRANDTE2	Char	10	508	517	Date of last transfusion	mm/dd/yyyy
59	TRANPLNT	Char	1	518	518	Tissue/transplant recipient	0=No, 1=Yes, 2=CDC Confirmed, 9=Unknown
60	HCW	Char	1	519	519	Patient a health care worker	0=No, 1=Yes, 9=Unknown
61	EIA	Char	1	520	520	Result of EIA Test	1=Pos, 0=Neg, 9=Not done
62	EIA_MOYR	Char	10	521	530	Date of EIA test	mm/dd/yyyy
63	COMB	Char	1	531	531	Result of HIV1/HIV2 combination test	1=Pos, 0=Neg, 9=Not done
64	COMBMOYR	Char	10	532	541	Date of HIV1/HIV2 combination test	mm/dd/yyyy
65	WBIFA	Char	1	542	542	Result of Western blot	1=Pos, 0=Neg, 9=Not done
66	WBMOYR	Char	10	543	552	Date of Western blot test	mm/dd/yyyy

EXAMPLE

Order	Variable	Typ	Len	lenbegin	lenend	Label	Format
67	HIVAB	Char	2	553	554	Type of other HIV antibody test	01=RIPA, 02=Latex Ag, 03=Peptide, 04=Rapid 11=IgA, 12=IVAP, 88=Other, 99=Unspec
68	HIVABRES	Char	1	555	555	Result of other HIV antibody test	1=Pos, 0=Neg, 9=Not done
69	HVABMOYR	Char	10	556	565	Date of other HIV antibody test	mm/dd/yyyy
70	HIVDTEST	Char	2	566	567	Type of HIV detection test	01=Culture, 02=Antigen, 03=DNA PCR 11=NASBA, 12=RT-PCR, 13=bDNA 14=RNA PCR, 18=Other viral, 88=Other
71	HIVDET	Char	1	568	568	Result of HIV detection test	1=Pos, 0=Neg, 8=Indeterminate
72	HVLOAD	Char	10	569	578	Viral load of HIV detection test (if available)	n=copies/ml
73	HIVDMOYR	Char	10	579	588	Date of HIV detection test	mm/dd/yyyy
74	DOC_DIAG	Char	1	589	589	Physician diagnosis	0=No, 1=Yes, 9=Unknown
75	DOC_MOYR	Char	10	590	599	Date of physician's diagnosis	mm/dd/yyyy
76	CD4CNT	Char	4	600	603	Low CD4 count	number
77	CD4PCT	Char	2	604	605	Low CD4 percent	N=N%
78	CD4MOYR	Char	10	606	615	Date of low CD4 percent	mm/dd/yyyy
79	PHYSNAME	Char	45	616	660	Physician's name	(Last, First, Middle Initial)
80	PPHONE	Char	12	661	672	Physician's phone number	xxx-xxx-xxxx
81	MEDRECNO	Char	14	673	686	Medical record number	
82	PERS_COM	Char	24	687	710	Person completing form	(Last, First, Middle Initial)
83	COMPHONE	Char	12	711	722	Interviewer's phone number	xxx-xxx-xxxx

EXAMPLE

Order	Variable	Typ	Len	lenbegin	lenend	Label	Format
84	REVIEWED	Char	1	723	723	Clinical record reviewed	0=No, 1=Yes
85	ASYMMOYR	Char	10	724	733	Date diagnosed Asymptomatic	mm/dd/yyyy
86	SYMPMOYR	Char	10	734	743	Date diagnosed Symptomatic	mm/dd/yyyy
87	CANDLUNG	Char	1	744	744	Candidiasis lungs	1=definitive, 0=not diagnosis
88	CLNGMOYR	Char	10	745	754	Dx date for candida lungs/trac	mm/dd/yyyy
89	CANDESOP	Char	1	755	755	Candidiasis esophageal	1=definitive, 2=presumptive, 0=not diagnosis
90	CESOMOYR	Char	10	756	765	Dx date for esophageal candida	mm/dd/yyyy
91	CERVDIS	Char	1	766	766	Carcinoma cervical	1=definitive, 0=not diagnosis
92	CDISMOYR	Char	10	767	776	Dx date for carcinoma cervical	mm/dd/yyyy
93	COCCI	Char	1	777	777	Coccidioidomycosis	1=definitive, 0=not diagnosis
94	CCMOYR	Char	10	778	787	Dx date for coccidioidomycosis	mm/dd/yyyy
95	CRYPTOCO	Char	1	788	788	Cryptococcosis	1=definitive, 0=not diagnosis
96	CTCCMOYR	Char	10	789	798	Dx date for cryptococcosis	mm/dd/yyyy
97	CRYPTOSP	Char	1	799	799	Cryptosporidiosis	1=definitive, 0=not diagnosis
98	CRYPMOYR	Char	10	800	809	Dx date for cryptosporidiosis	mm/dd/yyyy
99	CMV	Char	1	810	810	Cytomegalovirus disease	1=definitive, 0=not diagnosis
100	CMVMOYR	Char	10	811	820	Dx date for cytomegalovirus	mm/dd/yyyy
101	CMVRET	Char	1	821	821	Cytomegalovirus retinitis	1=definitive, 0=not diagnosis
102	CMVRMOYR	Char	10	822	831	Dx date for CMV retinitis	mm/dd/yyyy
103	DEMENTIA	Char	1	832	832	HIV encephalopathy	1=definitive, 0=not diagnosis
104	DEMMOYR	Char	10	833	842	Dx date for HIV encephalopathy	mm/dd/yyyy
105	HS	Char	1	843	843	Chronic mucocutaneous herpes	1=definitive, 0=not diagnosis

EXAMPLE

Order	Variable	Typ	Len	lenbegin	lenend	Label	Format
106	HSMOYR	Char	10	844	853	Dx date for chronic herpes	mm/dd/yyyy
107	HISTO	Char	1	854	854	Histoplasmosis	1=definitive, 0=not diagnosis
108	HISTMOYR	Char	10	855	864	Dx date for histoplasmosis	mm/dd/yyyy
109	ISO	Char	1	865	865	Isosporiasis	1=definitive, 0=not diagnosis
110	ISOMOYR	Char	10	866	875	Dx date for isosporiasis	mm/dd/yyyy
111	KS	Char	1	876	876	Kaposi's sarcoma	1=definitive, 2=presumptive, 0=not diagnosed
112	KSMOYR	Char	10	877	886	Dx date for Kaposi's sarcoma	mm/dd/yyyy
113	BURKL	Char	1	887	887	Burkitt's lymphoma	1=definitive, 0=not diagnosis
114	BURKMOYR	Char	10	888	897	Dx date for Burkitt's lymphoma	mm/dd/yyyy
115	IBL	Char	1	898	898	Immunoblastic lymphoma	1=definitive, 2=presumptive, 0=not diagnosed
116	IBLMOYR	Char	10	899	908	Dx date for immunoblastic lymp	mm/dd/yyyy
117	PLB	Char	1	909	909	Primary lymphoma of brain	1=definitve, 0=not diagnosed
118	PLBMOYR	Char	10	910	919	Dx date for lymphoma of brain	mm/dd/yyyy
119	MAVIUM	Char	1	920	920	Mycobacterium avium complex	1=definitive, 2=presumptive, 0=not diagnosed
120	MAVMOYR	Char	10	921	930	Dx date for M. avium complex	mm/dd/yyyy
121	PULM_TB	Char	1	931	931	Pulmonary TB	1=definitive, 2=presumptive, 3=Pre 93 Result, 0=not diagnosed
122	PTBMOYR	Char	10	932	941	Date of pulmonary TB diagnosis	mm/dd/yyyy
123	TB	Char	1	942	942	M. tuberculosis	1=definitive, 2=presumptive, 0=not diagnosed
124	TBMOYR	Char	10	943	952	Dx date for M. tuberculosis	mm/dd/yyyy
125	MYCO	Char	1	953	953	Atypical mycobact diagnosed	1=definitve, 2=presumptive, 0=not diagnosed

EXAMPLE

Order	Variable	Typ	Len	lenbegin	lenend	Label	Format
126	MYCOMOYR	Char	10	954	963	Dx date for atypical mycobact.	mm/dd/yyyy
127	PC	Char	1	964	964	Pneumocystis carinii pneumonia	1=definitive, 2=presumptive, 9=unknown
128	PCMOYR	Char	10	965	974	Dx date for pneumocystis pneu.	mm/dd/yyyy
129	RP	Char	1	975	975	Pneumonia recurrent	1=definitive, 2=presumptive, 0=not diagnosed
130	RPMOYR	Char	10	976	985	Dx date for pneu. recurrent	mm/dd/yyyy
131	PML	Char	1	986	986	Progress multifoc leukoenceph	1=definitive, 0=not diagnosed
132	PMLMOYR	Char	10	987	996	Dx date for multifoc. leuko.	mm/dd/yyyy
133	SALS	Char	1	997	997	Salmonella septicemia	1=definitive, 0=not diagnosed
134	SALSMOYR	Char	10	998	1007	Dx date for salmonella sept.	mm/dd/yyyy
135	TP	Char	1	1008	1008	Toxoplasmosis of brain	1=definitive, 2=presumptive, 0=not diagnosed
136	TPMOYR	Char	10	1009	1018	Dx date for toxoplasmosis	mm/dd/yyyy
137	WASTING	Char	1	1019	1019	Wasting syndrome	1=definitive, 0=not diagnosis
138	WASTMOYR	Char	10	1020	1029	Dx date for wasting syndrome	mm/dd/yyyy
139	RVCTNO	Char	9	1030	1038	RVCT Case Number	
140	OTH_IMM	Char	1	1039	1039	Other immunodeficiency would disqualify AIDS	0=No, 1=Yes, 9=Unknown
141	INFORMED	Char	1	1040	1040	Patient informed of HIV	0=No, 1=Yes, 9=Unknown
142	NOTIFIED	Char	1	1041	1041	Patient's partners notified by	1=Health Dept., 2=Provider, 3=Patient, 9=Unknown
143	REF_MS	Char	1	1042	1042	Referred for medical service	0=No, 1=Yes, 9=Unknown
144	REF_SATS	Char	1	1043	1043	Refer for substance abuse Rx	1=Yes, 0=No, 8=N/A, 9=Unknown

EXAMPLE

Order	Variable	Typ	Len	lenbegin	lenend	Label	Format
145	ANTIRETV	Char	1	1044	1044	Rec'd antiretroviral Rx	0=No, 1=Yes, 9=Unknown
146	PCPPROPH	Char	1	1045	1045	Receive PCP prophylaxis?	0=No, 1=Yes, 9=Unknown
147	TRIAL	Char	1	1046	1046	Patient enrolled in trial	1=NIH sponsored, 2=Other, 3=None, 9=Unknown
148	CLINIC	Char	1	1047	1047	Patient enrolled in clinic	1=HRSA sponsored, 2=Other, 3=None, 9=Unknown
149	INSURNCE	Char	1	1048	1048	Patient's insurance type	1=Medicaid, 2=Private Coverage, 3=No Coverage, 4=Other public fund, 7=Gov't program 9=Unknown
150	PRENATAL	Char	1	1049	1049	Patient-received OB/GYN service	0=No, 1=Yes, 9=Unknown
151	PREGNANT	Char	1	1050	1050	Patient currently pregnant	0=No, 1=Yes, 9=Unknown
152	LIVE_INF	Char	1	1051	1051	Patient delivered live infant	1=Yes, 0=No, 9=Unknown
153	CBDATE	Char	10	1052	1061	Child's birthdate	mm/dd/yyyy
154	CHOSP	Char	27	1062	1088	Hospital of child birth	See Table 4 for facility names
155	CHCITY	Char	27	1089	1115	City of hosp of child's birth	For city name, use Table 5
156	CHOSP_ST	Char	2	1116	1117	State of hosp of child's birth	Abbreviate (such as PA)
157	CSTATENO	Char	10	1118	1127	Child's state patient no.	If available
158	COMMENT1	Char	65	1128	1192	First line of comments	Any comments for DOH

Note:
1. Variable names in the list are names used in the CDC database.
2. We hope that providers submit the data to us in the same order of variables and in the same length and format for each variable.
3. If an observation does not have any value for a variable, please just leave it empty.
4. All dates should be in the format of mm/dd/yyyy, such as 10/18/2002.
5. Please send an ASCII, HL7, or Excel file, or use PA EPI_INFO database.

EXAMPLE

Appendix C

Identifying All Reporting Laboratories

Provider Type Information

CLIA Laboratories

01 Ambulatory Surgery Center

02 Community Clinic

03 Comprehensive Outpatient Rehab

04 Ancillary Test Site

05 ESRD (End Stage Renal Disease Dialysis)

06 Health Fair

07 HMO

08 Home Health Agency

09 Hospice

10 Hospital

11 Independent Laboratory

12 Industrial

13 Insurance

14 Intermediate Care Facility—Mentally Retarded

15 Mobile Unit

16 Pharmacy

17 School/Student Health Service

18 Skilled Nursing/Nursing Facility

19 Physician Office

20 Other Practitioner

21 Tissue Bank/Repositories

22 Blood Banks

23 Rural Health Clinic/Federally Qualified Health Center

24 Ambulance

25 Other

Laboratory Surveillance Survey Questions

Laboratory Name_____

Laboratory Contact_____

Address_____

_____**ZIP Code**_____

Phone_____

Fax_____

E-mail_____

(Circle the Most Appropriate)	**Type of Laboratory**	
Public	**Private**	**Federal**
Hospital	Hospital	Hospital
Clinic	Blood Bank	Clinic
Other: Specify Below>	Physician – POL	Other: Specify Below>
	Reference Laboratory	

EXAMPLE

I. Data Management

1. Please indicate which of the following your laboratory has available for each lab procedure performed.

Data Element	Yes	No
Patient's Name		
Patient's Address		
Date of Birth		
Sex		
Race		
Provider's Name		
Provider's Phone Number		
Doctor's Name		
Medical Record Number		
Social Security Number		

2. Indicate your method of data maintenance. If you are using more than one method, indicate all that apply.

Data Maintenance	Yes	No
Computer File		
Log Book		
Paper File of Lab Slips		
Other _____		

3. If you are using a Laboratory Information System, provide the name of the vendor as well as the name of the person(s) within your lab/institution who is responsible for maintaining the system.

Vendor_____

Information System Manager_____

4. Are you familiar with the HL7 data format for the electronic exchange of health care related information? Yes No

5. Would your laboratory be interested in serving as a pilot for the testing of software to automate the reporting of communicable disease data?

Yes No

EXAMPLE

II. Human Immunodeficiency Virus Testing

6. Does your laboratory perform any rapid test for HIV detection?

Yes No

If you are doing rapid tests, please give the name of the test you are using_____.

7. Does your laboratory perform any of the following HIV tests in house? Check all that you are currently performing.

❏ Human Immunodeficiency Virus (HIV) EIA Screen
❏ HIV Western Blot
❏ HIV DNA by ❏bDNA or ❏PCR
❏ HIV RNA by ❏bDNA or ❏PCR
❏ HIV Resistance Testing: ❏ Genotype, ❏ Phenotype
❏ HIV P24 Antigen
❏ HIV Culture
❏ CD4+ Counts

EXAMPLE

8. Reference Laboratories: Do you use a Reference Laboratory(s) to assist your facility with testing for human immunodeficiency virus? If so please indicate the specific test, the lab's name and phone number, and the name of the contact person. If you are using multiple reference laboratories, please provide the information requested below for each.

Reference Lab #1_____
Test(s)_____
Name of Contact Person_____
Phone Number_____

Reference Lab#2_____
Test(s)_____
Name of Contact Person_____
Phone Number_____

Reference Lab#3_____
Test(s)_____
Name of Contact Person_____
Phone Number_____

EXAMPLE

Appendix D

Documentation for Reporting Laboratory

Data Record Layout

ID No	Field Description	Length	Optional/ REQUIRED	Variable Type	Notes
1.	Medical record number from referral facility	16	Optional	Character	If no referral facility is involved in testing the specimen, this field will be blank.
2.	Referral Facility Name	30	Optional	Character	If no referral facility is involved in testing the specimen, this field will be blank.
3.	Medical record number from testing facility	20	REQUIRED	Character	
4.	Testing Facility CLIA Code	10	Optional	Character	
5.	Alternate Patient ID	16	Optional	Character	If an outside laboratory has performed the test, and the referral facility is reporting the result, include the patient identifier from the outside laboratory in this field.
6.	Patient Last Name	48	REQUIRED	Character	
7.	Patient First Name	48	REQUIRED	Character	
8.	Patient Middle Name	48	Optional	Character	
9.	Patient Name Suffix	48	Optional	Character	
10.	Patient Date of Birth	10	REQUIRED	Character	Data must be formatted as MM/DD/YYYY (e.g., 12/01/1952).
11.	Patient's Age	20	Optional	Numeric	
12.	Patient's Place of Birth	1	Optional	Character	Allowable field values are as follows: 0 = Patient was Foreign Born 1 = Patient was born in the US
13.	Patient's Country of Origin	30	Optional	Character	
14.	Patient's Length of Residency in the US	6	Optional	Character	Submit as the number of years and months the patient has been a resident of the United States. Field should be formatted as follows: YY MM (e.g. 12 08 = A resident of the US for 12 years and 8 months)

EXAMPLE

ID No	Field Description	Length	Optional/ REQUIRED	Variable Type	Notes
15.	Patient's Race	1	Optional	Character	Allowable field values are as follows: A = Asian or Pacific Islander B = Black I = American Indian or Alaska Native M = Multiracial O = Other U = Unknown W = White
16.	Patient's Ethnic Group	1	Optional	Character	Allowable field values are as follows: H = Hispanic N = Non-Hispanic U = Unknown
17.	Patient's Gender	1	Optional	Character	Allowable field values are as follows: F = Female H = Hermaphrodite/Undetermined M = Male O = Other T = Transsexual U = Unknown
18.	Patient's Street Address	62	REQUIRED	Character	
19.	Patient's City	30	REQUIRED	Character	
20.	Patient's State	2	REQUIRED	Character	If valued, field must be a valid USPS state abbreviation code to identify the state of residence. Valid state codes (as defined by the United States Postal Service) can be found at the following Web site: http://www.usps.com/ncsc/lookups/abbr_state.txt
21.	Patient's Country of Residence	3	Optional	Character	
22.	Patient's ZIP Code	9	REQUIRED	Character	
23.	Patient's County of Residence	20	Optional	Character	If valued, field must be a valid Federal Information Processing Standard (FIPS) code for the patient's county of residence. Valid FIPS county codes by state (as defined by the United States Environmental Protection Agency) can be found at the following Web site: http://www.epa.gov/enviro/html/codes/state.html

EXAMPLE

ID No	Field Description	Length	Optional/ REQUIRED	Variable Type	Notes
24.	Patient's Telephone Number	40	Optional	Character	
25.	Patient's Social Security Number	16	Optional	Character	Field may be valued as Social Security Number or Railroad Retirement Number. Social Security Number should be reported for cancer results only.
26.	Parent's Last Name	25	Optional	Character	
27.	Parent's First Name	20	Optional	Character	
28.	Parent's Middle Initial	1	Optional	Character	
29.	Parent's Name Suffix	2	Optional	Character	
30.	Parent's Street Address	62	Optional	Character	
31.	Parent's City	30	Optional	Character	
32.	Parent's State	2	Optional	Character	If valued, field must be a valid USPS state abbreviation code to identify the state of residence. Valid state codes (as defined by the United States Postal Service) can be found at the following Web site: http://www.usps.com/ncsc/lookups/abbr_state.txt
33.	Parent's Country	3	Optional	Character	If valued, recommend using valid Universal Postal Union country abbreviations code, ISO3166 Alpha-3 standard as detailed at the following Web site: http://www.upu.int/upu/AN/Pays_membres.html
34.	Parent's ZIP Code	9	Optional	Character	
35.	Parent's Phone Number	40	Optional	Character	
36.	Pregnancy Flag	1	Optional	Character	Allowable field values are as follows: Y = Patient is Pregnant N = Patient is not Pregnant
37.	Dept. of Corrections ID	7	Optional	Character	Inmate number
38.	Insurance or Billing Number	20	Optional	Character	

EXAMPLE

ID No	Field Description	Length	Optional/ REQUIRED	Variable Type	Notes
39.	Type of Insurance	1	Optional	Character	Allowable field values are as follows: 1 = Patient is Self-Insured 2 = Blue Cross 3 = Commercial Insurer 4 = Medicaid 5 = Medicare 6 = CHAMPUS 7 = Managed Care Plan
40.	Vaccine Trial Participation	1	Optional	Character	Allowable field values are as follows: 0 = Participated in vaccine trial 1 = Did not participate in vaccine trial
41.	Provider Last Name	25	Optional	Character	
42.	Provider First name	20	Optional	Character	
43.	Provider Middle Initial	1	Optional	Character	
44.	Provider Name Suffix	2	Optional	Character	
45.	Provider Street Address	62	Optional	Character	
46.	Provider City	30	Optional	Character	
47.	Provider State	2	Optional	Character	If valued, field must be a valid USPS state abbreviation code to identify the state of residence. Valid state codes (as defined by the United States Postal Service) can be found at the following Web site: http://www.usps.com/ncsc/lookups/abbr_state.txt
48.	Provider ZIP Code	9	Optional	Character	
49.	Provider Phone Number	40	Optional	Character	
50.	Ordering Facility Name	40	Optional	Character	
51.	Ordering Facility Street	62	Optional	Character	
52.	Ordering Facility City	30	Optional	Character	
53.	Ordering Facility State	2	Optional	Character	If valued, field must be a valid USPS state abbreviation code to identify the state of residence. Valid state codes (as defined by the United States Postal Service) can be found at the following Web site: http://www.usps.com/ncsc/lookups/abbr_state.txt
54.	Ordering Facility ZIP	9	Optional	Character	

EXAMPLE

ID No	Field Description	Length	Optional/ REQUIRED	Variable Type	Notes
55.	Ordering Facility Phone Number	40	Optional	Character	
56.	Pathologist Name	178	Optional	Character	
57.	Pathologist License Number	20	Optional	Character	
58.	Pathologist State of License	2	Optional	Character	
59.	Accession Number	75	Optional	Character	
60.	SNOMED Code	200	Optional	Character	Valued with type of analysis performed.
61.	Observation Date & Time	16	Optional	Character	Date field must be formatted as MM/DD/YYYY. Time must be formatted as HH:MM (e.g. 12:07) in military time format. Valid values range from 00:00 through 23:59. A blank space should be included between the date and time components (e.g., 10/01/2000 17:42).
62.	Observation Specimen Source Code	8	Optional	Character	
63.	Observation Specimen Source Text	40	Optional	Character	
64.	Result Report Date	10	Optional	Character	Data must be formatted as MM/DD/YYYY (e.g., 10/01/2000).
65.	Result Status Code	1	Optional	Character	Allowable field values are as follows: C = Correction to results F = Final result
66.	Data Type	2	Optional	Character	
67.	LOINC Code	10	Optional	Character	
68.	LOINC Description	50	Optional	Character	
69.	Lab Code	10	REQUIRED	Character	
70.	Lab Code Description	50	REQUIRED	Character	
71.	Observation Unit of Measure	60	Optional	Character	
72.	Reference Range	10	Optional	Character	
73.	Observation Result Status	1	Optional	Character	
74.	Observation Method	60	Optional	Character	
75.	Test Result	100	REQUIRED	Character	

EXAMPLE

ID No	Field Description	Length	Optional/ REQUIRED	Variable Type	Notes
76.	Path Report	48000	Optional	Character	
77.	SNOMED Code	10	Optional	Character	
78.	Specimen Description	4000	Optional	Character	
79.	ICD Code	10	Optional	Character	
80.	ICD Rev No	10	Optional	Character	
81.	Clinical History	4000	Optional	Character	
82.	Nature of Specimen	4000	Optional	Character	
83.	Gross Pathology	4000	Optional	Character	
84.	Microscopic Pathology	4000	Optional	Character	
85.	Final DX	4000	Optional	Character	
86.	Comment	4000	Optional	Character	
87.	Supplemental Reports	4000	Optional	Character	
88.	Staging Parameters	4000	Optional	Character	
89.	Sending Facility Name	170	REQUIRED	Character	
90.	Sending Facility CLIA	10	REQUIRED	Character	Unique Clinical Laboratory Improvement Amendment Identifier number.
91.	Message Date and Time	26	REQUIRED	Character	Data must be formatted as MM/DD/YYYY HH:MM (e.g., 10/01/2000 17:39) with time submitted in military time format. Valid military time values range from 00:00 through 23:59.
92.	Receiving Application	60	Optional	Character	Allowable field values are as follows: AIDS = Bureau of HIV/AIDS CANCER = Cancer Registry CD = Bureau of Communicable Disease Control LEAD = Lead Program STD = Bureau of Sexually Transmitted Diseases TB = Bureau of Tuberculosis Control
93.	Record Termination Indicator	2	REQUIRED	Character	Field must be valued with 2 exclamation points (!!) to indicate the end of each record.

EXAMPLE

Sample ASCII File Data Record

||MR40155|33D1234567||Duck|Daffy||Jr.|10/10/1966||1|USA||O|U|T|211 LaLa Land|Binghamton|NY|USA|13901|7|607.754.5454|122-33-4556||||||||||||||ProviderLastName|ProviderFirstName|||212 Central Avenue|Albany|NY|12201|518.432.3209 X247|General Hospital|9 East Overshoot Drive Street|Gloversville|NY|12078|518.999.9999|Dr. Shiny Diamond||NY|Accession#10117||10/09/2000 12:07|Throat|Throat|10/09/2000 12:07| F||6819-7|Streptococcus AB|Strep AB|Streptococcus AB|CELLS/UL|Reference Range||| Positive for Strep Culture. Variant AB.||SNOMED Code||ICD Code|ICD Rev Code| Clinical History|Nature of Specimen|Gross Pathology|Microscopic Pathology|Final DX|Comments|Supplemental Reports|Staging Parameters|General Hospital| 33D9999999|10/09/2000 11:01|CD|!!

EXAMPLE

Appendix E

Reporting Back to Laboratories

Appendix E.1

Standardized ASCII Format for Electronic Laboratory Reporting

This document provides a standardized ASCII format for morbidity records reported by laboratories participating in Electronic Laboratory Reporting (ELR). A standardized ASCII format will facilitate importation of morbidity data into specific database applications maintained by the Texas Department of Health's Bureau of HIV/STD Prevention and will allow for dissemination of these data to local health authorities on a daily basis.

A. General Requirements

1. Date fields may be represented in the following format: MM/DD/YYYY.

2. Telephone numbers may be represented in either XXX-XXX-XXXX or XXXXXXXXXX format.

3. Social Security Number may be represented in either XXX-XX-XXXX or XXXXXXXXX format.

B. Data Variables

Patient ID **Must be defined in the file, but may be missing in the records that are provided**.
Up to 20 characters.

Last Name **Must be defined in the file and must be present in the records that are provided**.
Up to 20 characters.

First Name **Must be defined in the file and must be present in the records that are provided**.
Up to 20 characters.

MI **Must be defined in the file, but may be missing in the records that are provided**.
One (1) character.

DOB **Must be defined in the file, but may be missing in the records that are provided**.
Ten (10) characters.

Age **Must be defined in the file, but may be missing in the records that are provided**.
Up to 3 numbers.

SSN **Must be defined in the file, but may be missing in the records that are provided**.
Nine (9) or eleven (11) characters, depending on the SSN format.

Sex **Must be defined in the file, but may be missing in the records that are provided**.
Seven (7) characters.

Race **Must be defined in the file, but may be missing in the records that are provided**.
One (1) character.

EXAMPLE

Ethnicity **Must be defined in the file, but may be missing in the records that are provided**.
One (1) character.

Street **Must be defined in the file, but may be missing in the records that are provided**.
Up to 30 characters.

City **Must be defined in the file, but may be missing in the records that are provided**.
Up to 20 characters.

County **Must be defined in the file, but may be missing in the records that are provided**.
Up to 20 characters.

State **Must be defined in the file, but may be missing in the records that are provided**.
Two (2) characters.

ZIP **Must be defined in the file, but may be missing in the records that are provided**.
A 5-character field.

Phone **Must be defined in the file, but may be missing in the records that are provided**.
Ten (10) or 12 characters, depending on format.

Provider ID **Must be defined in the file, but may be missing in the records that are provided**.
Up to 20 characters.

Provider **Must be defined in the file and must be present in the records that are provided**.
Up to 30 characters.

Accession **Must be defined in the file and must be present in the records that are provided**.
Up to 15 characters.

PrCity **Must be defined in the file and must be present in the records that are provided**.
Up to 20 characters.

PrCounty **Must be defined in the file, but may be missing in the records that are provided**.
Up to 20 characters.

PrPhone **Must be defined in the file, but may be missing in the records that are provided**.
Ten (10) or 12 characters, depending on format.

PrState **Must be defined in the file, but may be missing in the records that are provided**.
Two (2) characters.

PrZIP **Must be defined in the file and must be present in the records that are provided**.
A 5-character field.

Collected **Must be defined in the file, but may be missing in the records that are provided**.
Ten (10) characters.

Received **Must be defined in the file, but may be missing in the records that are provided**.
Ten (10) characters.

EXAMPLE

Reported **Must be defined in the file and must be present in the records that are provided**. Ten (10) characters.

Specimen **Must be defined in the file, but may be missing in the records that are provided**. Up to 20 characters.

Lab ID **Must be defined in the file and must be present in the records that are provided**. Up to 10 characters.

Analysis **Must be defined in the file, but may be missing in the records that are provided**. Ten (10) characters.

Test Type **Must be defined in the file and must be present in the records that are provided.** Up to 30 characters.

Qual Resul **Must be defined in the file and must be present in the records that are provided. (if qualitative results are appropriate for the test type)**. Up to 15 characters.

Quan Resul **Must be defined in the file and must be present in the records that are provided. (if quantitative results are appropriate for the test type)**. Up to 15 characters.

Units **Must be defined in the file, but may be missing in the records that are provided**. Up to 15 characters.

Ref Range **Must be defined in the file, but may be missing in the records that are provided**. Up to 15 characters.

EXAMPLE

C. Record Layout

The following record layout is recommended for all ASCII files received from ELR participating laboratories:

VARIABLE	POSITION	LENGTH	RECORD TYPE
PATIENT ID	1 - 20	20	PATIENT
LAST NAME	21 - 40	20	PATIENT
FIRST NAME	41 - 60	20	PATIENT
MI	61	1	PATIENT
DOB	62 - 71	10	PATIENT
AGE	72 - 74	3	PATIENT
SSN	75 - 85	11	PATIENT
SEX	86 - 92	7	PATIENT
RACE	93	1	PATIENT
ETHNICITY	94	1	PATIENT
STREET	95 - 125	30	PATIENT
CITY	126 -146	20	PATIENT
STATE	147 - 148	2	PATIENT
ZIP	149 - 153	5	PATIENT
PHONE	154 - 165	12	PATIENT
PROVIDER ID	166 - 186	20	SPECIMEN
PROVIDER	187 - 217	30	SPECIMEN
ACCESSION	218 - 233	15	SPECIMEN
PRCITY	234 - 254	20	SPECIMEN
PRCOUNTY	255 - 275	20	SPECIMEN
PRPHONE	276 - 287	12	SPECIMEN
PRSTATE	288 - 289	2	SPECIMEN
PRZIP	290 - 294	5	SPECIMEN

EXAMPLE

VARIABLE	POSITION	LENGTH	RECORD TYPE
COLLECTED	295 - 305	10	SPECIMEN
RECEIVED	306 - 315	10	SPECIMEN
REPORTED	316 - 325	10	SPECIMEN
SPECIMEN	326 - 346	20	SPECIMEN
LAB ID	347 - 357	10	TEST
ANALYSIS	358 - 368	10	TEST
TEST TYPE	369 - 399	30	TEST
QUAL RESUL	400 - 414	15	TEST
QUAN RESUL	415 - 429	15	TEST
UNITS	430 - 444	15	TEST
REF RANGE	445 - 459	15	TEST

EXAMPLE

D. Data Dictionary

FIELD NAME	LENGTH	DESCRIPTION/CODED VALUES	REQUIRED
Patient ID	20	Laboratory system-generated patient ID	Y
Last Name	20	Patient last name	Y
First Name	20	Patient first name	
MI	1	Patient middle initial	
DOB	10	Patient date-of-birth	
Age	3	Patient age at testing	
SSN	9	Patient Social Security Number	
Sex	7	Patient gender: **Male** **Female** **Unknown**	
Race	1	Patient race: (race codes) **1 = American Indian/Alaska Native** **2 = Asian/Pacific Islander** **3 = Black** **4 = White** **5 = Other** **9 = Unknown**	
Ethnicity	1	Patient ethnicity: (ethnicity codes) **1 = Hispanic** **2 = Non-Hispanic** **9 = Unknown**	
Street	30	Patient street address	
City	20	Patient city of residence	
County	20	Patient county of residence	
State	2	Patient state of residence	
ZIP	5	Patient ZIP code	
Phone	10	Patient telephone number	
Provider ID	20	Laboratory system-generated provider ID	
Provider	30	Provider name	Y

EXAMPLE

FIELD NAME	LENGTH	DESCRIPTION/CODED VALUES	REQUIRED
Accession	15	Laboratory system-generated order of aquisition	Y
PrCity	20	Provider city	Y
PrCounty	20	Provider county	
PrPhone	10	Provider telephone number	
PrState	2	Provider state	
PrZIP	5	Provider ZIP code	Y
Collected	10	Date of specimen collection	
Received	10	Date specimen received by laboratory	
Reported	10	Date results reported to Public Health	Y
Specimen	20	Anatomical site where specimen obtained	
Lab ID	10	Name of reporting laboratory	Y
Analysis	10	Date of specimen analysis	
Test Type	30	Type of test	Y
Qual Resul	15	Qualitative result: (i.e.) **Positive** **Reactive** **Negative** **Nonreactive** **Indeterminate** **Contaminated**	Y
Quan Resul	15	Quantitative result	Y
Units	15	Qualifier for quantitative result	
Ref Range	15	Reference range for test performed	

EXAMPLE

HIV Laboratory Reporting Summary

For the period January – June 2003

	Your Lab		All Labs (N=20)	
Volume of Test Data				
CD4	131		3,368	
CD4 %	132		3,440	
Western Blot	111		461	
Viral Load	167		6,317	
Completeness of Information				
Name*	541	100% Complete	11,578	100% complete
Date of Birth*	534	99% Complete	10,681	92% complete
Age*	534	99% Complete	1,588	14% complete
Gender*	538	99% Complete	11,202	96% complete
Last 4 digits of Social Security Number*	0	0% Complete	5.215	45% complete
Patient ID#	495	91% Complete	8,114	70% complete
Specimen Collection Date*	541	100% Complete	11,575	100% complete

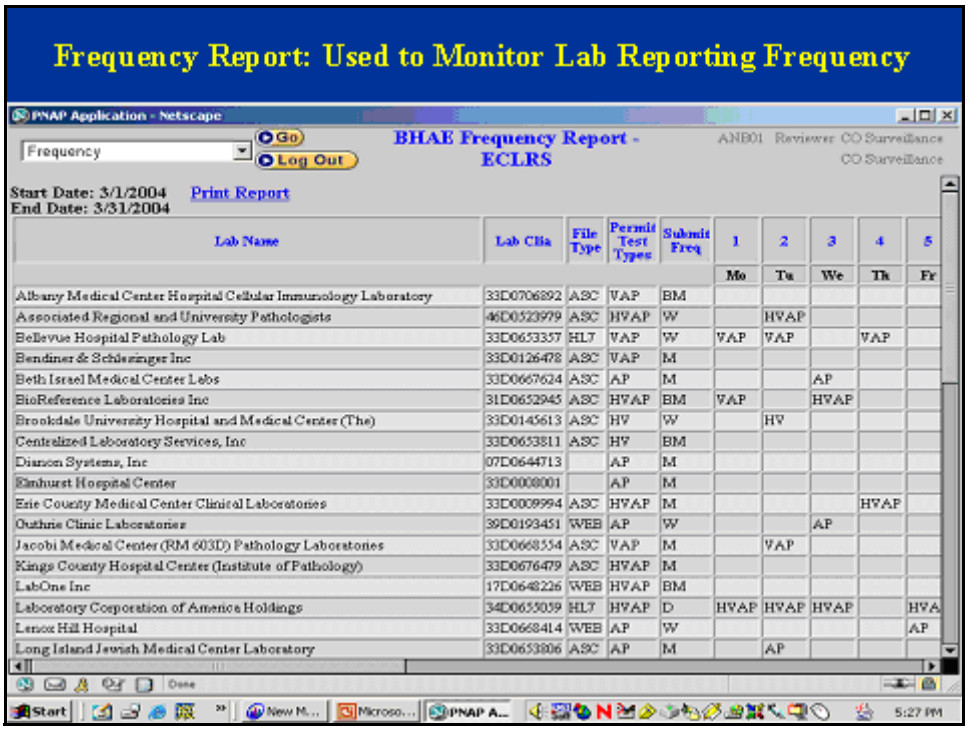
* Required component of Washington HIV reports

Data Transmission Details		
Method of submitting data (N=20 including nonelectronic methods)		
Virtual Private Network - HL7 (preferred)		0
Electronic file on diskette (acceptable)	Yes	7
U.S. Mail of paper records (acceptable)		13
Manual pick up of paper records (not preferred)		0
File protection (N=7)		
PGP Encrypted (preferred method)		2
Password protected (acceptable method)	Yes	4
No protection (poor method)		1
File formatting for electronic records (N=7)		
Excel (acceptable)	Yes	4
Tab or character-delimited (acceptable)		3
Space delimited (acceptable)		0

AIDS has been reportable since 1984. As of September 1, 1999, the Washington Administrative Codes (WAC 246-101) were expanded to require reporting, by name, of all cases of HIV infection. Laboratories are required to report—to their local or state health officer—AIDS-indicative CD4 counts (<200 cells per microliter or <14% of total lymphocytes), positive HIV viral load tests, confirmed HIV antibody tests, and other tests diagnostic of HIV infection.

EXAMPLE

Appendix E.2



Frequency Report: Used to Monitor Lab Reporting Frequency

PNAP Application - Netscape

Frequency | Go | Log Out

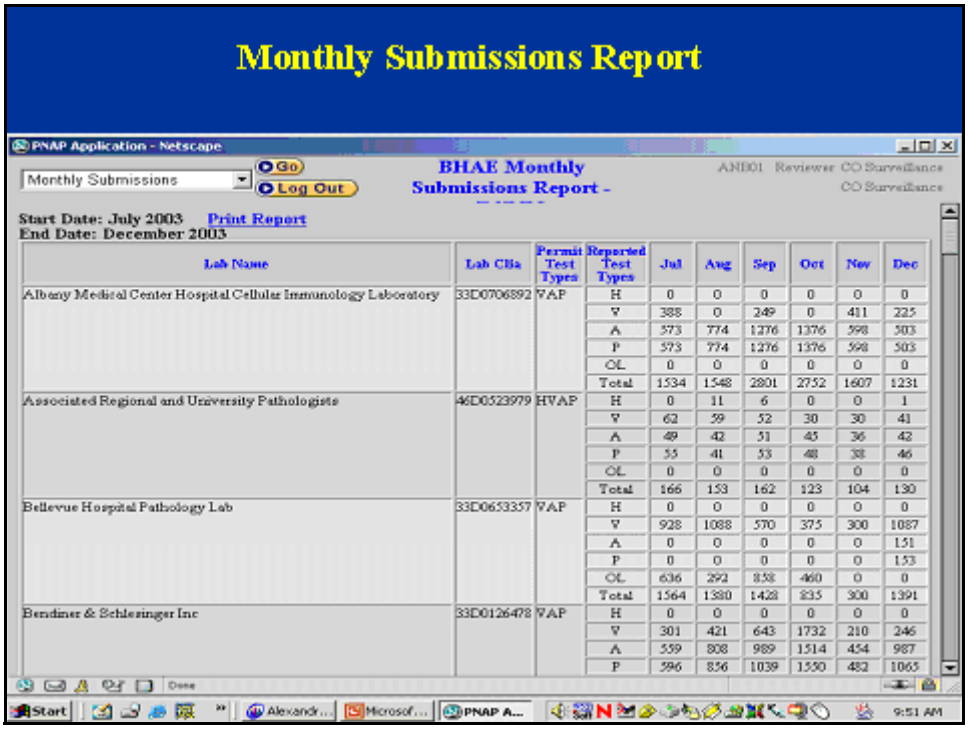
BHAE Frequency Report - ECLRS

ANB01 Reviewer CO Surveillance
CO Surveillance

Start Date: 3/1/2004 Print Report
End Date: 3/31/2004

Lab Name	Lab Clia	File Type	Permit Test Types	Submit Freq	1	2	3	4	5
					Mo	Tu	We	Th	Fr
Albany Medical Center Hospital Cellular Immunology Laboratory	33D0706892	ASC	VAP	BM					
Associated Regional and University Pathologists	46D0523979	ASC	HVAP	W		HVAP			
Bellevue Hospital Pathology Lab	33D0653357	HL7	VAP	W	VAP	VAP		VAP	
Bendiner & Schlesinger Inc	33D0126478	ASC	VAP	M					
Beth Israel Medical Center Labs	33D0667624	ASC	AP	M			AP		
BioReference Laboratories Inc	31D0652945	ASC	HVAP	BM	VAP		HVAP		
Brookdale University Hospital and Medical Center (The)	33D0145613	ASC	HV	W		HV			
Centralized Laboratory Services, Inc	33D0653811	ASC	HV	BM					
Dianon Systems, Inc	07D0644713		AP	M					
Elmhurst Hospital Center	33D0008001		AP	M					
Erie County Medical Center Clinical Laboratories	33D0009994	ASC	HVAP	M				HVAP	
Guthrie Clinic Laboratories	39D0193451	WEB	AP	W			AP		
Jacobi Medical Center (RM 603D) Pathology Laboratories	33D0668554	ASC	VAP	M		VAP			
Kings County Hospital Center (Institute of Pathology)	33D0676479	ASC	HVAP	M					
LabOne Inc	17D0648226	WEB	HVAP	BM					
Laboratory Corporation of America Holdings	34D0655039	HL7	HVAP	D	HVAP	HVAP	HVAP		HVA
Lenox Hill Hospital	33D0668414	WEB	AP	W					AP
Long Island Jewish Medical Center Laboratory	33D0653806	ASC	AP	M		AP			

Example: **Frequency Report** — Used to compare the licenses or permits held by each lab and their designated reporting schedule with the reports that are actually being submitted. This can be monitored by test type for any given month/year.



Monthly Submissions Report

PNAP Application - Netscape

Monthly Submissions | Go | Log Out

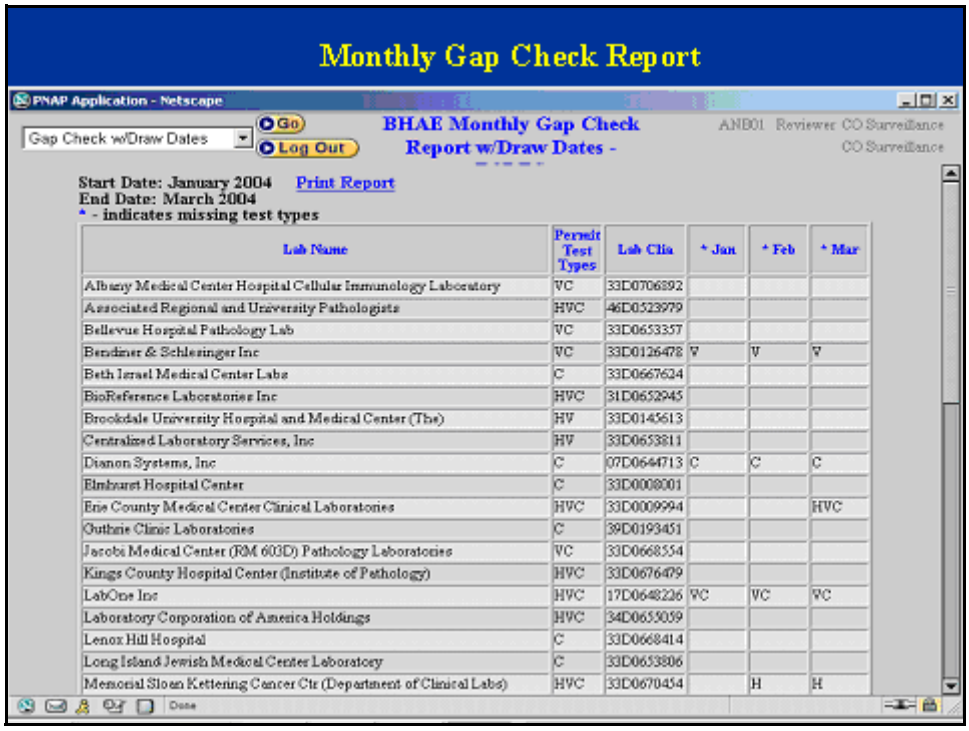
BHAE Monthly Submissions Report -

ANB01 Reviewer CO Surveillance
CO Surveillance

Start Date: July 2003 Print Report
End Date: December 2003

Lab Name	Lab Clia	Permit Test Types	Reported Test Types	Jul	Aug	Sep	Oct	Nov	Dec
Albany Medical Center Hospital Cellular Immunology Laboratory	33D0706892	VAP	H	0	0	0	0	0	0
			V	388	0	249	0	411	225
			A	573	774	1276	1376	598	503
			P	573	774	1276	1376	598	503
			OL	0	0	0	0	0	0
			Total	1534	1548	2801	2752	1607	1231
Associated Regional and University Pathologists	46D0523979	HVAP	H	0	11	6	0	0	1
			V	62	59	52	30	30	41
			A	49	42	51	45	36	42
			P	55	41	53	48	38	46
			OL	0	0	0	0	0	0
			Total	166	153	162	123	104	130
Bellevue Hospital Pathology Lab	33D0653357	VAP	H	0	0	0	0	0	0
			V	928	1088	570	375	300	1087
			A	0	0	0	0	0	151
			P	0	0	0	0	0	153
			OL	636	292	858	460	0	0
			Total	1564	1380	1428	835	300	1391
Bendiner & Schlesinger Inc	33D0126478	VAP	H	0	0	0	0	0	0
			V	301	421	643	1732	210	246
			A	559	808	989	1514	454	987
			P	596	856	1039	1550	482	1065

Example: **Monthly Submissions Report** — This report is used to monitor labs for any unusual change in volume of reporting by test type for each laboratory.



Monthly Gap Check Report

PNAP Application - Netscape

Gap Check w/Draw Dates | Go | Log Out

BHAE Monthly Gap Check Report w/Draw Dates -

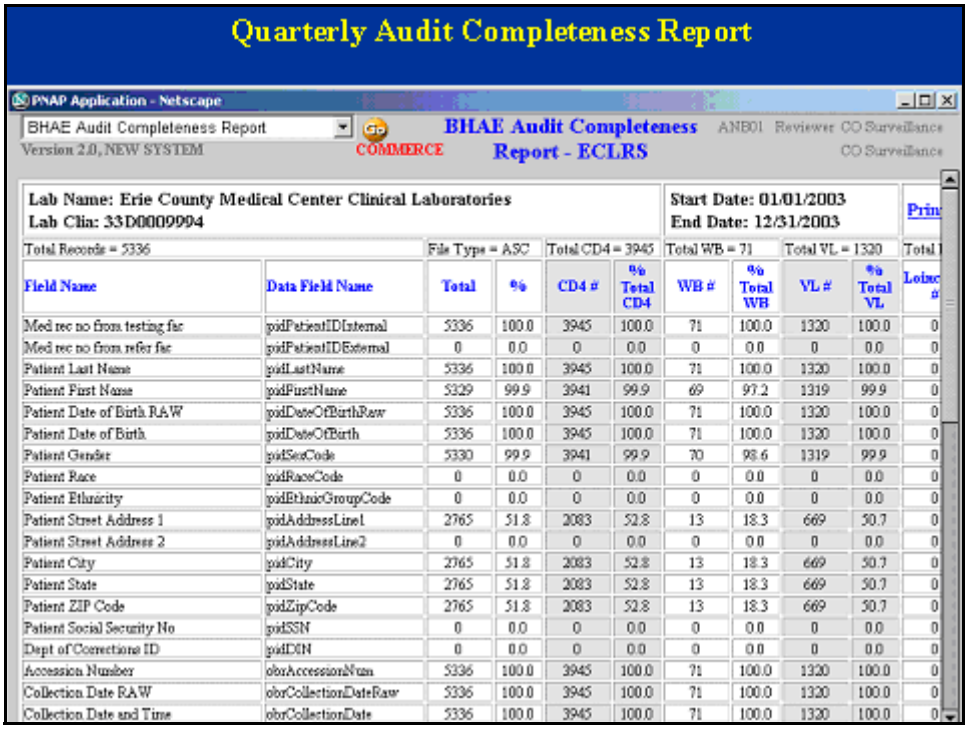
ANB01 Reviewer CO Surveillance
CO Surveillance

Start Date: January 2004 Print Report
End Date: March 2004
* - indicates missing test types

Lab Name	Permit Test Types	Lab Clia	* Jan	* Feb	* Mar
Albany Medical Center Hospital Cellular Immunology Laboratory	VC	33D0706892			
Associated Regional and University Pathologists	HVC	46D0523979			
Bellevue Hospital Pathology Lab	VC	33D0653357			
Bendiner & Schlesinger Inc	VC	33D0126478	V	V	V
Beth Israel Medical Center Labs	C	33D0667624			
BioReference Laboratories Inc	HVC	31D0652945			
Brookdale University Hospital and Medical Center (The)	HV	33D0145613			
Centralized Laboratory Services, Inc	HV	33D0653811			
Dianon Systems, Inc	C	07D0644713	C	C	C
Elmhurst Hospital Center	C	33D0008001			
Erie County Medical Center Clinical Laboratories	HVC	33D0009994			HVC
Guthrie Clinic Laboratories	C	39D0193451			
Jacobi Medical Center (RM 603D) Pathology Laboratories	VC	33D0668554			
Kings County Hospital Center (Institute of Pathology)	HVC	33D0676479			
LabOne Inc	HVC	17D0648226	VC	VC	VC
Laboratory Corporation of America Holdings	HVC	34D0655059			
Lenox Hill Hospital	C	33D0668414			
Long Island Jewish Medical Center Laboratory	C	33D0653806			
Memorial Sloan Kettering Cancer Ctr (Department of Clinical Labs)	HVC	33D0670454		H	H

Done

Example: **Monthly Gap Check Report** — This report is used to monitor gaps in reporting by specimen collection date. The Monthly Gap Check Report is shown above, but weekly and daily reports can also be generated. For each lab, for each month requested, a code in the column indicates that there was NO specimen collect date for that month. This report can identify a problem with the manner in which the lab has selected their data for submission.



Quarterly Audit Completeness Report

PNAP Application - Netscape

BHAE Audit Completeness Report

Version 2.0, NEW SYSTEM

COMMERCE

BHAE Audit Completeness Report - ECLRS

ANB01 Reviewer CO Surveillance

CO Surveillance

Lab Name: Erie County Medical Center Clinical Laboratories
Lab Clia: 33D0009994

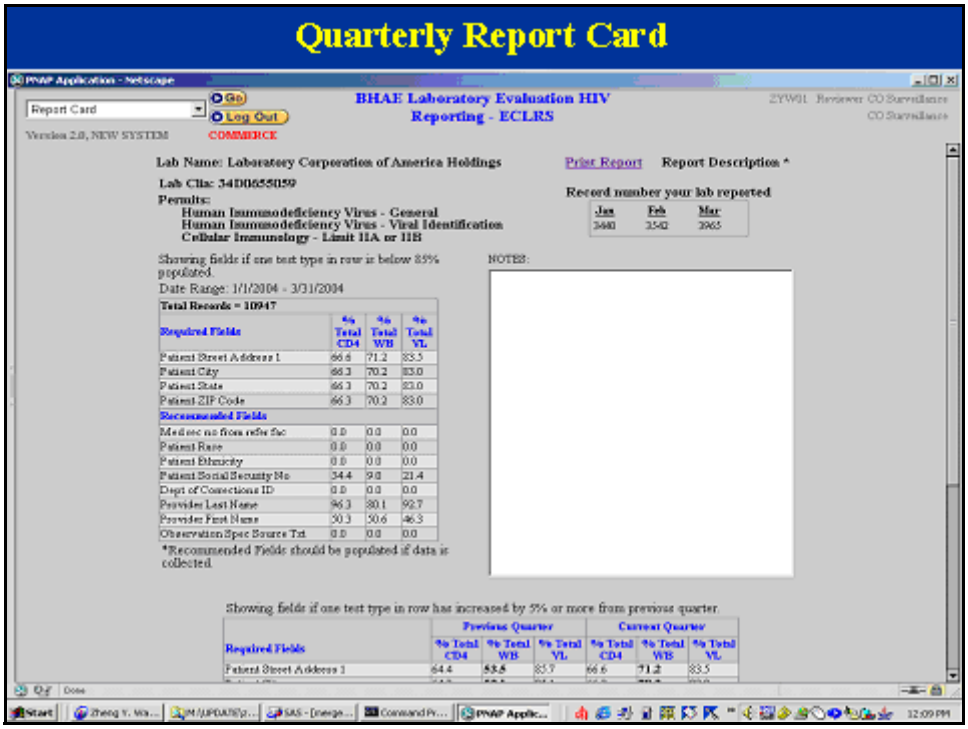
Start Date: 01/01/2003
End Date: 12/31/2003

Prin

Total Records = 5336 | File Type = ASC | Total CD4 = 3945 | Total WB = 71 | Total VL = 1320 | Total

Field Name	Data Field Name	Total	%	CD4 #	% Total CD4	WB #	% Total WB	VL #	% Total VL	Loinc #
Med rec no from testing fac	pidPatientIDInternal	5336	100.0	3945	100.0	71	100.0	1320	100.0	0
Med rec no from refer fac	pidPatientIDExternal	0	0.0	0	0.0	0	0.0	0	0.0	0
Patient Last Name	pidLastName	5336	100.0	3945	100.0	71	100.0	1320	100.0	0
Patient First Name	pidFirstName	5329	99.9	3941	99.9	69	97.2	1319	99.9	0
Patient Date of Birth RAW	pidDateOfBirthRaw	5336	100.0	3945	100.0	71	100.0	1320	100.0	0
Patient Date of Birth	pidDateOfBirth	5336	100.0	3945	100.0	71	100.0	1320	100.0	0
Patient Gender	pidSexCode	5330	99.9	3941	99.9	70	98.6	1319	99.9	0
Patient Race	pidRaceCode	0	0.0	0	0.0	0	0.0	0	0.0	0
Patient Ethnicity	pidEthnicGroupCode	0	0.0	0	0.0	0	0.0	0	0.0	0
Patient Street Address 1	pidAddressLine1	2765	51.8	2083	52.8	13	18.3	669	50.7	0
Patient Street Address 2	pidAddressLine2	0	0.0	0	0.0	0	0.0	0	0.0	0
Patient City	pidCity	2765	51.8	2083	52.8	13	18.3	669	50.7	0
Patient State	pidState	2765	51.8	2083	52.8	13	18.3	669	50.7	0
Patient ZIP Code	pidZipCode	2765	51.8	2083	52.8	13	18.3	669	50.7	0
Patient Social Security No	pidSSN	0	0.0	0	0.0	0	0.0	0	0.0	0
Dept of Corrections ID	pidDIN	0	0.0	0	0.0	0	0.0	0	0.0	0
Accession Number	obrAccessionNum	5336	100.0	3945	100.0	71	100.0	1320	100.0	0
Collection Date RAW	obrCollectionDateRaw	5336	100.0	3945	100.0	71	100.0	1320	100.0	0
Collection Date and Time	obrCollectionDate	5336	100.0	3945	100.0	71	100.0	1320	100.0	0

Example: **Quarterly Audit Completeness Report** — This report displays the completeness rate for each of the required and recommended fields that get submitted. They are separated by type of test.



Quarterly Report Card

PNAP Application - Netscape

Report Card

Go

Log Out

Version 2.0, NEW SYSTEM

COMMERCE

BHAE Laboratory Evaluation HIV Reporting - ECLRS

ZYW01 Reviewer CO Surveillance

CO Surveillance

Lab Name: Laboratory Corporation of America Holdings

Lab Clia: 34D0655059

Permits:
Human Immunodeficiency Virus - General
Human Immunodeficiency Virus - Viral Identification
Cellular Immunology - Limit IIA or IIB

Print Report

Report Description *

Record number your lab reported

Jan	Feb	Mar
3440	3542	3965

Showing fields if one test type in row is below 85% populated.

Date Range: 1/1/2004 - 3/31/2004

NOTES:

Total Records = 10947

Required Fields	% Total CD4	% Total WB	% Total VL
Patient Street Address 1	66.6	71.2	83.5
Patient City	66.3	70.2	83.0
Patient State	66.3	70.2	83.0
Patient ZIP Code	66.3	70.2	83.0
Recommended Fields			
Med rec no from refer fac	0.0	0.0	0.0
Patient Race	0.0	0.0	0.0
Patient Ethnicity	0.0	0.0	0.0
Patient Social Security No	34.4	9.0	21.4
Dept of Corrections ID	0.0	0.0	0.0
Provider Last Name	96.3	80.1	92.7
Provider First Name	50.3	50.6	46.3
Observation Spec Source Txt	0.0	0.0	0.0

*Recommended Fields should be populated if data is collected.

Showing fields if one test type in row has increased by 5% or more from previous quarter.

Required Fields	Previous Quarter % Total CD4	Previous Quarter % Total WB	Previous Quarter % Total VL	Current Quarter % Total CD4	Current Quarter % Total WB	Current Quarter % Total VL
Patient Street Address 1	64.4	53.6	83.7	66.6	71.2	83.5

Example: **Report Card** — A report card is sent to laboratories quarterly to help them monitor how well they are doing in their data submissions. Each lab can see where completion rates for required and recommended fields fall below 85%. A note field is

included and populated with comments tailored to each individual lab. These are usually based on the problems noted in the other reports, or problems seen with their data upon a manual review.

Example: **Report Card (cont)** — At the bottom of the report card, improvements in reporting are noted. Any field that was less than 85% complete on the previous audit appears here if there is a greater than 5 percent improvement in reporting this data element since the last report. It is helpful to give positive feedback to labs when their reporting improves.

Appendix F

Changes in Laboratory Reporting Initiated by Electronic Laboratory Reporting (ELR)

LOINC (Logical Observation Identifiers Names and Codes)

PLEASE NOTE: The difference in the LOINC codes in this document and LOINC codes used in eHARS is temporary. The codes in eHARS will eventually be changed to match those used in this document.

Screening Tests

HIV-1 EIA

Test Name	Unit	Other	Specimen Source	LOINC
HIV 1 Antibody EIA	positive/negative	—	serum	29893-5 HIV1 Ab Ser Ql EIA
HIV 1 Antibody EIA	positive/negative	—	serum/donor	21007-0 HIV1 Ab Ser Donr Ql
HIV 1 Antibody EIA	positive/negative	—	unspecified	29327-4 HIV1 Ab Fld Ql

HIV-1/2 EIA

Test Name	Unit	Other	Specimen Source	LOINC
HIV 1&2 Antibody EIA	positive/negative	—	serum	31201-7 HIV1+2 Ab Ser Ql EIA
HIV 1&2 Antibody EIA	positive/negative	—	CSF	32602-5 HIV1+2 Ab CSF Ql

HIV-2 EIA

Test Name	Unit	Other	Specimen Source	LOINC
HIV 2 Antibody EIA	positive/negative	—	serum	30361-0 HIV2 Ab Ser Ql EIA

WARNING: The LOINC short name (i.e., name listed under the LOINC number) may change. It has been included because some laboratories find the short name useful. The LOINC number, however, will not change.

HIV-1 Rapid Test

Test Name	Unit	Other	Specimen Source	LOINC
OraQuick Rapid HIV-1/2 Antibody Test	positive/negative	only HIV-1 in oral fluid	oral fluid	35437-3 HIV1 Ab Sal QL EIA
OraQuick Rapid HIV-1 Antibody Test	positive/negative	—	blood	hold
Reveal Rapid HIV-1 Antibody Test	positive/negative	—	plasma/serum	hold
Uni-Gold Recombigen HIV	positive/negative	—	plasma/serum	hold
Uni-Gold Recombigen HIV	positive/negative	—	blood	hold
HIV-1 Rapid EIA	positive/negative	used only in delivery room	cord blood	33866-5 HIV1 Ab BldC Ql EIA

HIV-1/2 Rapid Test

Test Name	Unit	Other	Specimen Source	LOINC
Multispot HIV-1/HIV-2 Rapid Test	positive/negative	—	plasma/serum	hold
OraQuick ADVANCE Rapid HIV-1/2 Antibody Test	positive/negative	—	oral fluid	hold
OraQuick Rapid HIV-1/2 Antibody Test	positive/negative	—	plasma/serum	hold

Confirmatory Tests

HIV-1 IFA

Test Name	Unit	Other	Specimen Source	LOINC
HIV 1 IFA	positive/negative	—	serum	14092-1 HIV1 Ab Ser Ql IF

WARNING: The LOINC short name (i.e., name listed under the LOINC number) may change. It has been included because some laboratories find the short name useful. The LOINC number, however, will not change.

HIV-1 Western Blot

Test Name	Unit	Other	Specimen Source	LOINC
HIV-1 WB	positive/negative	—	serum	5221-7 HIV1 Ab Ser Ql IB
HIV-1 WB	positive/negative	band pattern	plasma/serum	not used by CDC
HIV-1 WB	bands appearing	band-specific LOINC	serum	not used by CDC
Cambridge Biotech HIV-1 Urine WB	positive/negative	—	urine	32571-2 HIV1 Ab Ur Ql IB
HIV-1 WB	positive/negative	—	CSF	16977-1 HIV1 Ag CSF Ql
HIV-1 WB	positive/negative	• OraSure HIV-1WB	oral fluid	35439-9 HIV1 Ab Sal Ql IB
HIV-1 WB	positive/negative	—	unspecified	34592-6 HIV1 Ab Fld Ql IB

*HIV-2 Western Blot**

Test Name	Unit	Other	Specimen Source	LOINC
HIV-2 WB	positive/negative	—	serum	5225-8 HIV2 Ab Ser Ql IB
HIV-2 WB	bands appearing	band pattern	serum	not used by CDC

* ***All HIV-2 confirmatory testing should be conducted through CDC. Contact HICSB's Cases of Public Health Importance coordinator.***

Other HIV Detection Tests

HIV-1 P24 Antigen

Test Name	Unit	Other	Specimen Source	LOINC
HIV-1 P24 Antigen	positive/negative	EIA method	serum	18396-2 HIV1 p24 Ag Ser Ql EIA
HIV-1 P24 Antigen	positive/negative	unspecified method	serum	9821-0 HIV1 p24 Ag Ser Ql
HIV-1 P24 Antigen	positive/negative	neutralization confirm	serum	33660-2 HIV1 p24 Ag Ser Ql Nt
HIV-1 P24 Antigen	quantitative	unspecified method	CSF	16979-7 HIV1 p24 Ag CSF-aCnc

WARNING: The LOINC short name (i.e., name listed under the LOINC number) may change. It has been included because some laboratories find the short name useful. The LOINC number, however, will not change.

HIV Culture (HIV 1 or 2)

Test Name	Unit	Other	Specimen Source	LOINC
HIV Culture	positive/negative	—	blood	6429-5 HIV Bld Cult
HIV Culture	positive/negative	—	semen	6430-3 HIV Smn Cult
HIV Culture	positive/negative	—	unknown	6431-1 HIV XXX Cult

HIV-1 proviral DNA (qualitative detection of deoxynucleic acid)

Test Name	Unit	Other	Specimen Source	LOINC
HIV-1 Proviral DNA	detectable/non-detectable	probe and/or PCR	plasma/serum	30245-5 HIV1 DNA SerPl Ql PCR
HIV-1 Proviral DNA	detectable/non-detectable	probe and/or PCR	blood	9837-6 HIV1 DNA Bld Ql Amp Prb
HIV-1 Proviral DNA	detectable/non-detectable	probe and/or PCR	tissue	hold

HIV-2 proviral DNA (qualitative detection of deoxynucleic acid)

Test Name	Unit	Other	Specimen Source	LOINC
HIV-2 Proviral	detectable/non-detectable	probe and/or PCR	plasma/serum	34699-9 HIV2 SerPl Ql PCR
HIV-2 Proviral	detectable/non-detectable	probe and/or PCR	blood	25841-8 HIV2 DNA Bld Ql PCR
HIV-2 Proviral	detectable/non-detectable	probe and/or PCR	unknown	25842-6 HIV2 DNA XXX Ql PCR

* ***All HIV-2 detection testing should be conducted through CDC. Contact HICSB's Cases of Public Health Importance coordinator.***

WARNING: The LOINC short name (i.e., name listed under the LOINC number) may change. It has been included because some laboratories find the short name useful. The LOINC number, however, will not change.

HIV-1 RNA (qualitative detection of ribonucleic acid)

Test Name	Unit	Other	Specimen Source	LOINC
HIV-1 RNA	detectable/non-detectable	probe and/or PCR	plasma/serum	25835-0 HIV1 RNA SerPl Ql PCR
HIV-1 RNA	detectable/non-detectable	probe and/or PCR	blood	5017-9 HIV1 RNA Bld Ql PCR
HIV-1 RNA	detectable/non-detectable	probe and/or PCR	unknown	5018-7 HIV1 RNA XXX Ql PCR

HIV-2 RNA (qualitative detection of ribonucleic acid)

Test Name	Unit	Other	Specimen Source	LOINC
HIV-2 RNA	detectable/non-detectable	probe and/or PCR		hold

HIV-1 DNA/RNA (qualitative detection of deoxyribonucleic acid a/o ribonucleic acid)

Test Name	Unit	Other	Specimen Source	LOINC
HIV-1 DNA/RNA	detectable/non-detectable		plasma/serum	hold

CD4+ T-Lymphocyte Tests

CD4 Count

Test Name	Unit	Other	Specimen Source	LOINC
CD4 T-lymphocyte count	cells/µl	—	blood	8127-3 CD4 Cells # Bld
CD4 T-lymphocyte count	cells/µl	—	unspecified	20605-2 CD4 Cells # XXX

WARNING: The LOINC short name (i.e., name listed under the LOINC number) may change. It has been included because some laboratories find the short name useful. The LOINC number, however, will not change.

CD4 Percent

Test Name	Unit	Other	Specimen Source	LOINC
CD4 T-lymphocyte percent	%	—	blood	8128-1 CD4 Cells % Bld
CD4 T-lymphocyte percent	%	—	unspecified	20606-0 CD4 Cells % XXX

Viral Load Tests

HIV-1 RNA Viral load—grouped by minimum detection limit, various methods

Test Name	Unit	Other	Specimen Source	LOINC
HIV-1 viral load, min det ≤400 c/ml	copies/ml	• Amplicor Monitor (**standard**): 400–750,000	plasma/serum	41513-3 HIV1 RNA #SerPl Amp Prb DL 400 copies
HIV-1 viral load, min det ≤2.6 logc/ml	log copies/ml	• Amplicor Monitor (**standard**): 2.6–5.9	plasma/serum	41514-1 HIV1RNA SerPl Amp Prb DL 2.6 log copies-log count
HIV-1 viral load, min det ≤75 c/ml	copies/ml	• Amplicor Monitor (**ultrasensitive**): 50–75,000 • NucliSens QT: 40–15,000,000 • Versant bDNA: 75–500,000	plasma/serum	41515-8 HIV1 RNA #SerPl Amp Prb DL 75 copies
HIV-1 viral load, min det ≤1.9 logc/ml	log copies/ml	• Amplicor Monitor (**ultrasensitive**): 1.7–4.7 • NucliSens QT: 1.6–7.2 • Versant bDNA: 1.9–5.7	plasma/serum	41516-6 HIV1 RNA SerPl Amp Prb DL 1.9 log copies-log count

WARNING: The LOINC short name (i.e., name listed under the LOINC number) may change. It has been included because some laboratories find the short name useful. The LOINC number, however, will not change.

HIV-1 Viral load—unspecified or unknown method & unknown minimum detection limit

Test Name	Unit	Other	Specimen Source	LOINC
HIV-1 RNA	copies/ml	any method	serum/plasma	20447-9 HIV1 Viral Load SerPl PCR
HIV-1 RNA	log copies/ml	any method	serum/plasma	29541-0 HIV1 Log Viral Load Plas PCR

HIV-2 Viral load—unspecified or unknown type

Test Name	Unit	Other	Specimen Source	LOINC
HIV-2 RNA	copies/ml	—	serum/plasma	hold
HIV-2 RNA	log copies/ml	—	serum/plasma	hold

Viral Susceptibility Tests

HIV-1 Genotype (viral resistance/susceptibility)

Test Name	Unit	Other/misc	Specimen Source	LOINC
HIV-1	Protease and RT nucleotide sequence; mutations associated with HIV drug resistance det/nondetectable	ELR to states should consist of a text file (“FASTA”) representing the nucleotide sequence	plasma/serum	hold
HIV-1			blood	hold
HIV-1			dried bloodspot	hold
HIV-1			dried plasma/serum spot	hold

HIV-1 Phenotype (measured viral drug resistance/susceptibility)

Test Name	Unit	Other/misc	Specimen Source	LOINC
HIV-1	IC50	antiretroviral agents	plasma/serum	hold
HIV-1	IC50	antiretroviral agents	dried plasma spot	hold

WARNING: The LOINC short name (i.e., name listed under the LOINC number) may change. It has been included because some laboratories find the short name useful. The LOINC number, however, will not change.

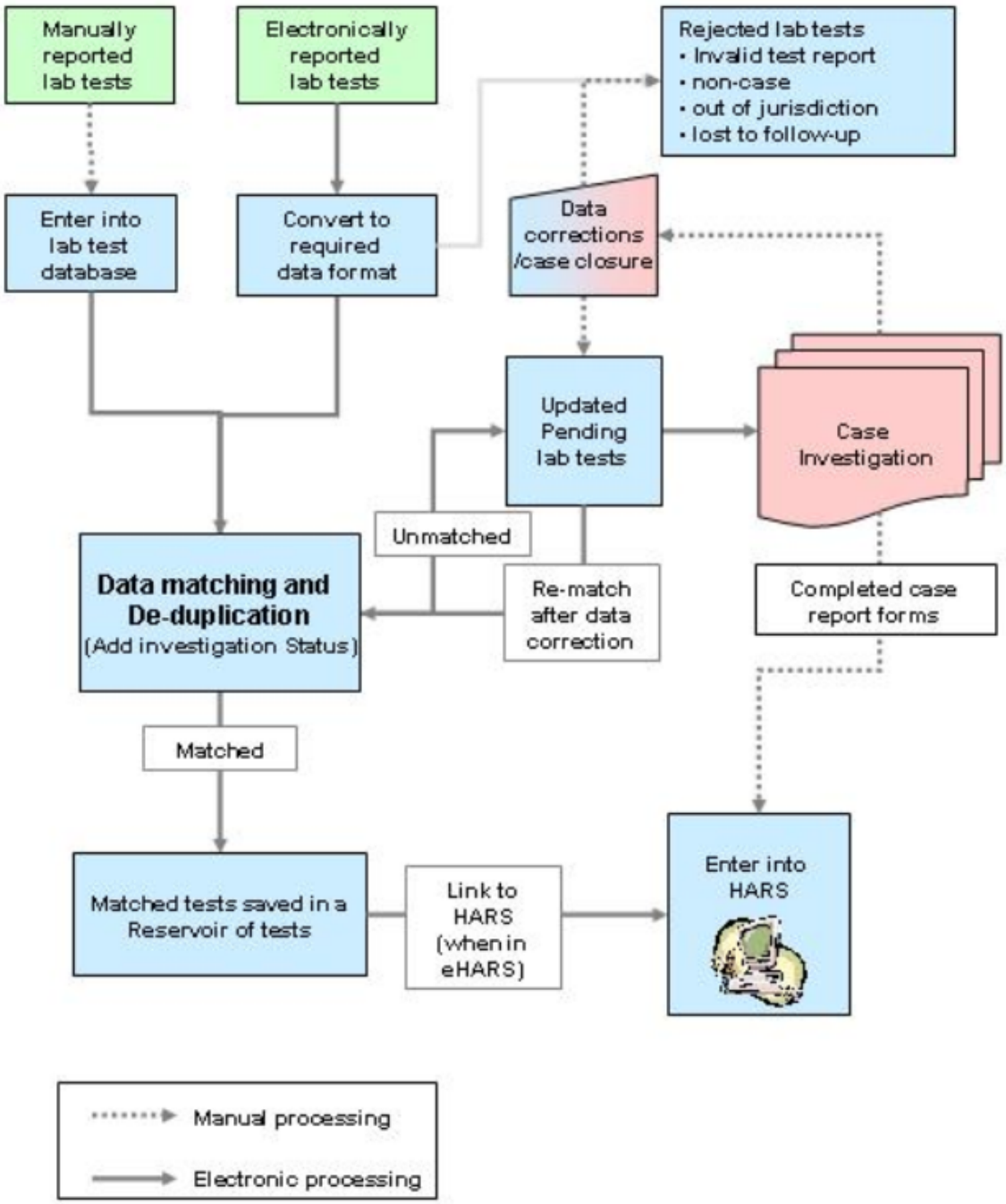
Appendix G

Case Follow-up from a Lab Report

Appendixes G.1 – G.3

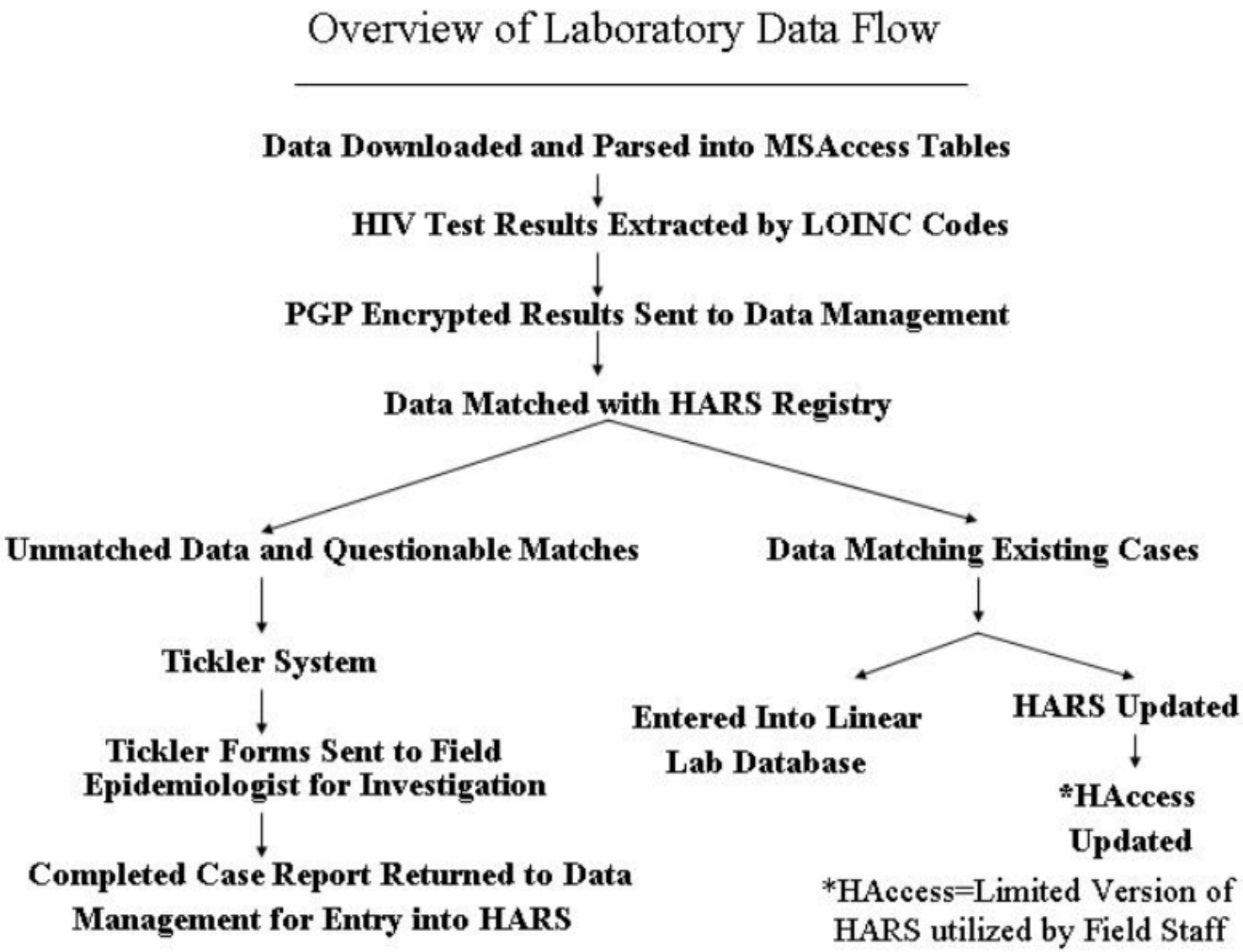
Appendix G.1

Processing Laboratory Test Reports



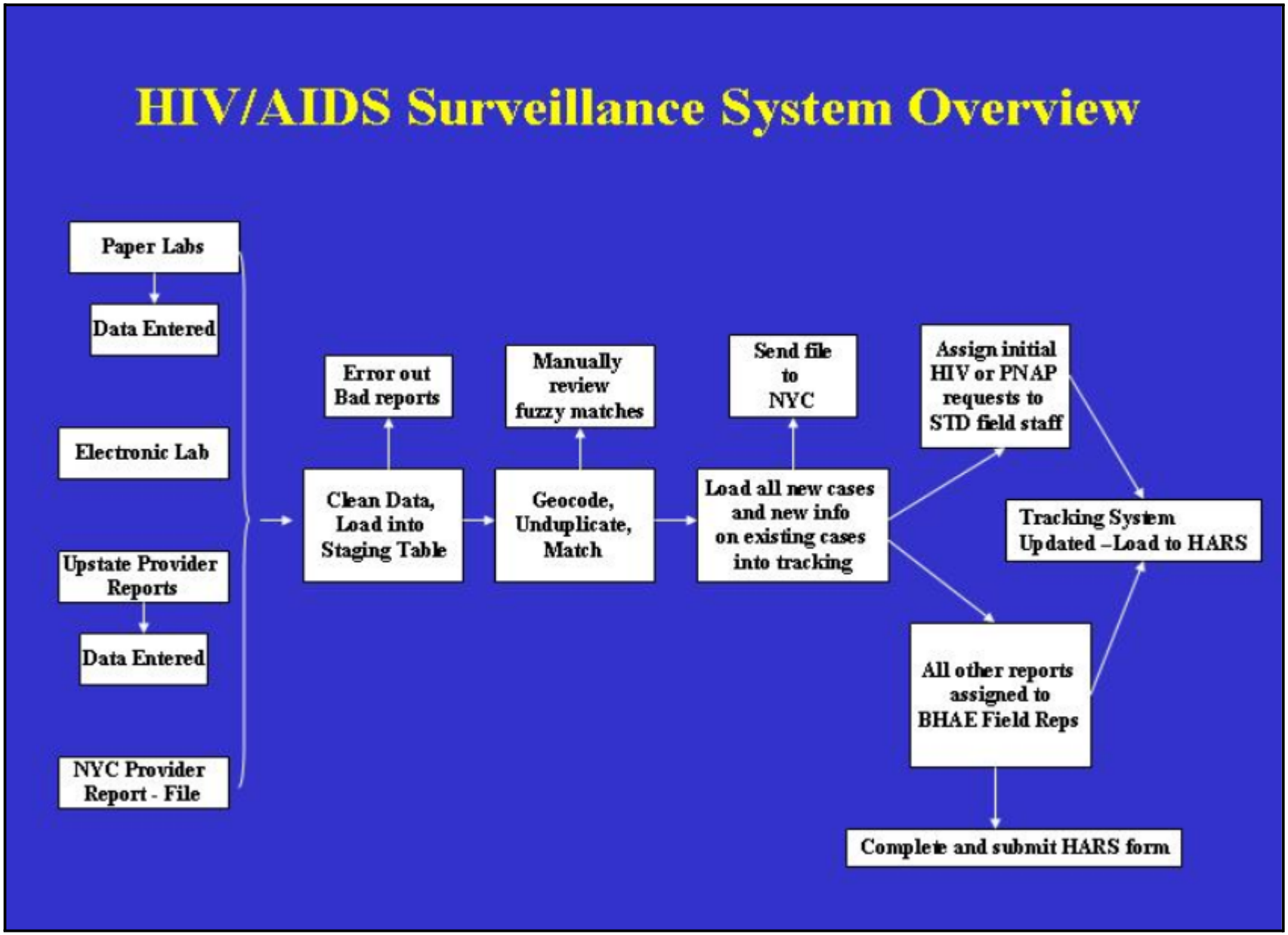
Appendix G.2

Overview of Laboratory Data Flow



Appendix G.3

HIV/AIDS Surveillance System Overview



Technical Guidance for HIV/AIDS Surveillance Programs

Case Residency

HIV Incidence and Case Surveillance Branch
Atlanta, Georgia

Contents — Case Residency

Case Residency Assignment 6-3
HIV Diagnosis Date 6-3
Structural Requirements to Establish HIV Diagnosis Date 6-3
Process Standards 6-4
Determination of Case Residency 6-6
Structural Requirements to Determine Case Residency 6-6
Process Standards 6-7
Residency Assignment of HIV Cases in States with Newly Implemented HIV Reporting 6-9
Appendix A 6-11
Appendix B 6-12
Appendix C 6-14

Technical Guidance for HIV/AIDS Surveillance Programs — Case Residency

Case Residency Assignment

Cases of HIV and AIDS are reported based on the place of residence at diagnosis of HIV and/or AIDS regardless of where exposure may have occurred. Unless different residence information is reported or provided from other sources, the address given by the person at time of diagnosis, recorded in the medical chart, and subsequently reported by physicians, laboratories, or surveillance staff for case surveillance purposes can be assumed to be the "usual residence." When there is only one report for a case, and only one jurisdiction, this residency determination is not an issue. The complexity is introduced when a case is reported by more than one jurisdiction as often occurs in the setting of HIV/AIDS surveillance.

HIV Diagnosis Date

HIV diagnosis date is the earliest date at which HIV infection was diagnosed from either a positive confirmatory laboratory test result or, in the absence of laboratory documentation, a documented physician diagnosis date. This date is used for epidemiologic monitoring in estimates of incidence and prevalence of HIV in the population. Although AIDS diagnosis dates have been clearly defined as either the date of the earliest CD4 T-lymphocyte count <200 cells/µL or the date of the first opportunistic infection diagnosis, the interpretation of HIV diagnosis date has been ambiguous and has included anonymous test dates, patient self-report dates, physician diagnosis dates, and laboratory report dates. So, although an HIV infection can have only one 'true' diagnosis date, multiple dates can be collected that are believed to be the diagnosis date. Collection of the most accurate date and an understanding of the calculated variables in the data management system directly associated with it are necessary to ensure correct use and interpretation.

Structural Requirements to Establish HIV Diagnosis Date

1. **Case Report and/or Laboratory Report**
2. **HIV/AIDS Surveillance System Software**
3. **HIV/AIDS Surveillance System Software User Instructions**
4. **Variables Used to Calculate HIV Diagnosis Date**
   - The Earliest of the Following Dates:
     - First positive HIV antibody test,
     - First detectable viral load,
     - Positive antigen or culture, or

- In the absence of laboratory documentation, date documented by the physician.

- Date of the Earliest AIDS Diagnosis—Based on the earlier of two dates: the date of the first CD4 T-lymphocyte count <200 cells/µL or CD4 T-lymphocyte count <14% or the date of the first opportunistic infection based on the current case definition.
- HIV Disease Date—A newly created variable that takes the earliest of diagnosis dates for HIV or AIDS. This date can be used to establish the earliest date of HIV infection diagnosis regardless of whether the method of diagnosis was through a laboratory test, a physician diagnosis, or an AIDS-defining condition.
- Laboratory Tests Used for Establishing HIV Diagnosis Date—First positive/detectable/reactive laboratory test using the specimen collection date for the first positive of any of the following types of tests: (see software user instructions for variable names)
  - HIV-1 EIA, Rapid, HIV-1/2 EIA
  - Latex Ag, IVAP, IgA, RIPA, Peptide
  - HIV-1 IFA
  - HIV-1 Western blot
  - HIV Antibody test—Other and unspecified
  - Viral load: HIV-1 RNA NASBA, Viral load: HIV-1 RNA RT-PCR, Viral load: HIV-1 RNA bDNA, Viral load: HIV-1 RNA, Other and unspecified viral load assay test HIV-1
  - Proviral DNA (qualitative), HIV-1 RNA PCR (qualitative)
  - HIV-1 Culture
  - HIV-1 P24 Antigen, HIV Antigen/detection/viral load—other and unspecified that are positive, reactive, or detectable
- Diagnosis Date for HIV-2—The HIV-2 Western blot is not approved by the Food and Drug Administration (FDA). Laboratory results for HIV-2 can be entered to meet the HIV case definition in the surveillance software, but the confirmatory HIV diagnosis must be documented by the physician with an associated date. Once the physician diagnosis date is entered, the calculated diagnosis date will be the date of the first positive EIA (as with HIV-1).

Process Standards

Because new generation EIA tests can identify new infections before a positive Western blot and with the use of rapid tests, a new level of complexity has been added to the testing algorithm for confirming HIV infection. The surveillance system software calculates the HIV diagnosis date based on the earliest date of all positive screening HIV tests once a positive confirmatory test result has been entered. The dates of the screening and

confirmatory tests are not compared. **If a person has a positive screening test with a negative confirmatory test but at a much later date has a positive screening test with a positive confirmatory test, the system will calculate the HIV diagnosis date based on the earliest screening test result. This could result in an HIV diagnosis date being erroneously calculated earlier than the true date.**

To eliminate this problem, any positive screening test without a positive confirmatory antibody test or viral detection test (therefore, the patient is considered negative and not an HIV case) should be deleted as soon as the negative confirmatory result is received. If the confirmatory result is 'indeterminate,' the tests should be retained until a second confirmatory result is received (usually performed within 1 month of the first result). If the second confirmatory result is also negative or indeterminate and viral detection tests (if any) are negative, then the initial false-positive tests should be deleted.

Current clinical HIV practice requires a positive EIA with a positive Western blot to confirm HIV infection. With varying methods and antigens used in the EIA test, the practice of confirming HIV infection with a Western blot may be discontinued and replaced by multiple EIA or other diagnostic tests. When this occurs, this document will be revised and the surveillance software system updated to reflect the change in practice.

HIV Diagnosis Date and AIDS Diagnosis Date Relationship

In the HIV/AIDS Reporting System (HARS), if the AIDS date of diagnosis is earlier than the HIV diagnosis date, the HIV date does not change to reflect the earlier date and, therefore, will not correctly reflect the earliest diagnosis date. In the HARS replacement software, you can use the HIV disease date for the earliest HIV/AIDS diagnosis date. In HARS, a comparison of the two dates should be made and the earliest date taken. For purposes of surveillance, the earlier of the two should be considered the HIV diagnosis date.

Anonymous Test Results

These are not reportable or to be included in cases sent to the national surveillance database because the anonymous nature of the tests results in an inability to ascertain duplicates or verify results. States should follow their own guidelines for the use of anonymous test results.

Patient Self-Reported Date of Prior Positive HIV Test

This type of self-reported test result is not allowed as a laboratory test date.

Physician-Documented Diagnosis Date

This is the date of the patient visit to the health care provider who documented the diagnosis of HIV. This date is used in the absence of HIV laboratory test results.

Note: Date should include a month and a year.

Laboratory Tests

For any two tests within the same year, a date that includes month and year vs. just the year will take precedence.

Pediatric Cases

These cases should follow the current pediatric case definition for HIV diagnosis date and use the same rules for physician-documented diagnosis date as outlined previously.

Determination of Case Residency

These policies are intended for use in routine surveillance, as well as in circumstances where ambiguous information may be reported from various sources and a determination needs to be made. They are adapted for HIV surveillance from revised guidelines for determining residence for nationally reportable diseases presented in position statement 03-ID-10 of the Council of State and Territorial Epidemiologists (CSTE) (Appendix C).

Structural Requirements to Determine Case Residency

1. **Case Report**
2. **HIV/AIDS Surveillance System Software**
3. **HIV/AIDS Surveillance System Software User Instructions**
4. **CSTE Position Statement 03-ID-10 (2003)** (Appendix C)
5. **CSTE Position Statement 17 (1986): AIDS Case Reporting: Reciprocal Notification** (Appendix A)
6. **CSTE Position Statement 01-ID-04 (2002): Reciprocal (Interstate) Notification of HIV Cases** (Appendix B)
7. **Description of Variables**
   - State of Residence at HIV Diagnosis
   - State of Residence at AIDS Diagnosis
   - County of Residence at HIV Diagnosis
   - County of Residence at AIDS Diagnosis
   - City of Residence at HIV Diagnosis
   - City of Residence at AIDS Diagnosis
   - ZIP Code of Residence at HIV Diagnosis
   - ZIP Code of Residence at AIDS Diagnosis

Process Standards

Residence Information

This information is entered at the local jurisdiction, and should reflect the residence at the time of HIV and/or AIDS diagnosis. The home address given by the patients at the time of diagnosis is considered their usual residence.

Circumstances in Which Questions May Arise Concerning Residence Assignment:

- Multiple Residences—Residency for people who move regularly between residences should be assigned by using the address where they live most of the time. If their time is equally divided, residence assignment should be based on where they were living at the time of diagnosis. If an individual assumes a new residence for an indefinite period without intending to return to the previous residence, the address of the new residence should be considered as the usual residence, even if this change occurred shortly before HIV or AIDS diagnosis.
- Homeless Persons—People without a usual residence should be reported by the jurisdiction where they were staying at the time of diagnosis.
- Vacation—Residence for people diagnosed while on vacation should be assigned by using the usual residence, not the vacation residence.
- Students
  - Residency for college or boarding school students on a typical yearly academic cycle should be assigned by using the address where they live most of the time. Intermittent or part-time students without a regular schedule for moving between parental and school residences should be assigned by using the address where they were living at the time of diagnosis.
- Live-ins—For foster children, residency should be assigned by using the address where they were living at the time of diagnosis.
- Military or Merchant Marine Personnel in the United States
  - For people in the military residing in the United States, residency should be assigned by using the address where they were living at the time of diagnosis, either on the base or off the base. For crews of military or U.S. flag merchant vessels, residency should be assigned by using the address where they live most of the time when they are onshore, if available. If the only available address is the temporary port in which the person resided at diagnosis, then that should be used.

- Institutionalized Persons
  - For persons who are incarcerated in state or federal correctional facilities at the time of diagnosis, residence of diagnosis should be defined as the residence of the correctional facility.
  - For persons who are incarcerated in city or county jails for short-term stays (a year or less), place of residency should be assigned by using the home address. Facility address should only be used if home address is not available.
  - For persons who are institutionalized for indefinite or long-term stays, residency should be assigned by using the address of the facility where they are staying at the time of diagnosis. Examples of facilities include
    - Chronic or long-term hospitals
    - Hospices
    - Nursing or convalescent homes
    - Inpatient drug/alcohol recovery facilities
    - Homes, schools, hospitals, or wards for the physically or mentally disabled
    - Orphanages and residential care facilities for neglected/abused children
- Foreign Citizens—Individuals, regardless of citizenship, who are diagnosed in the United States with HIV or AIDS while residing in the United States, should be reported to CDC. The CDC does not collect surveillance information or forward case reports to other countries for persons whose country of residence is not the United States. Foreign citizens who are visiting in the United States and are not considered residents should not be reported.
  - Residency for foreign citizens who have established a household or *are part of an established household in the United States, including those here for work or study,* should be assigned by using the address of their usual residence in the United States.
  - Residency for foreign citizens who live on diplomatic compounds (e.g., embassies, consulates) should be assigned by using the address of the facility.
- U.S. Residents Diagnosed Abroad—HIV in U.S. residents diagnosed abroad is only notifiable in the United States if treatment and care occur in the United States, reportable diagnostic tests are performed, and/or AIDS-related conditions are diagnosed by a physician in the United States. The residence at diagnosis should be based on the location of the usual residence at the time of treatment or care in the United States.
- If the usual place of residence for a U.S. citizen at the time of diagnosis is another country but diagnosis and/or care was in the United States, the case is reportable by the jurisdiction of usual residence in the United States. If there is no usual residence in the United States, it should be reported by the jurisdiction where the person was staying at the time of diagnosis.

- Assignment of residency is based as closely as possible on U.S. Census guidelines. If there are further questions, contact the CDC Routine Interstate Duplicate Review (RIDR) coordinator for clarification.

Designation of State or Territory

For any given case reported to CDC, only one state or territory can be listed under *HIV Residence at Diagnosis* and only one state or territory under *AIDS Residence at Diagnosis*. Because residency at diagnosis of HIV and AIDS can have implications for funding, every effort should be made to obtain accurate information. On the basis of CSTE position statements 17 and 01-ID-04, state health departments should follow reciprocal notification guidelines and communicate with each other regarding possible out-of-jurisdiction (OOJ) cases. Case information for individuals whose home is outside the reporting state should be forwarded to the appropriate state. State health departments should provide updated information to one another (such as updates on vital status) on an "as needed" basis.

Using Physician-Documented Diagnosis Date to Assign Residency

If a case is reported in multiple jurisdictions with a diagnosis date in the same year and one jurisdiction has only the year of diagnosis but the other jurisdiction has the month and year of diagnosis, residency will be assigned to the jurisdiction with the month/year date.

Maintaining Out-of-Jurisdiction Cases

Using surveillance software or hard copies, surveillance units should maintain limited information on OOJ cases for future reference. Maintaining OOJ cases using surveillance software eliminates the need to create additional space for a hard copy file. Routine computerized record searches can match OOJ cases along with resident case reports so that they will not be investigated and/or reported again. States should use their surveillance software to track the date that the information was given to the state of residence. Out-of-jurisdiction cases can be easily excluded from local data analysis.

Residency Assignment of HIV Cases in States with Newly Implemented HIV Reporting

1. **Follow Guidelines for Assignment of Residency**

2. **Reporting of Cases**

   2.1. States should follow state-legislated mandates for reporting.

   2.2. HIV reporting can be incident or retrospective as per a particular date. No matter what schema is used, the state of residency at diagnosis of HIV is the state with the earliest HIV diagnosis date, even if this results in a transfer of residency assignment from a state that had reported the case previously.

2.3. Where reporting of cases that were diagnosed by confidential testing before HIV name-based reporting was implemented, the state of residency at earliest diagnosis of HIV is assigned residency.

2.4. Residency is assigned to the state with the earliest diagnosis of HIV infection, regardless of jurisdiction of first report. If a case is first reported by a state with earlier implementation of HIV reporting and residency is assigned to that state, but a state with newly implemented HIV reporting finds an earlier diagnosis date in their state, residency assignment at HIV diagnosis is transferred to the state with newly implemented reporting. The case should be retained as an out-of-jurisdiction case in the state reporting the later date of diagnosis.

2.5. If both states have the same HIV diagnosis date, the state reporting the case first is assigned residency based on the date the record is entered in the surveillance software system.

2.6. Anonymous test results are not allowed.

2.7. Self-reported HIV diagnosis date is not allowed.

2.8. Follow policies and procedures for HIV diagnosis date.

Appendix A

CSTE POSITION STATEMENT 1986-17

COMMITTEE: Surveillance

TITLE: AIDS Case Reporting: Reciprocal Notification

ISSUE: All states and territories now conduct AIDS surveillance. Currently, case counting by states and territories is delayed by sending AIDS case reports on individuals now residing in other states and territories to the Centers for Disease Control rather than the state or territory of residence at time of onset of illness. Occasionally, duplicate counting of cases reported from different states or territories has occurred and some problems have arisen in exchange of information between states.

POSITION TO BE ADOPTED: The CDC should encourage states and territories to use the same state to state reciprocal notification process for AIDS case reports that is used for all other reportable diseases, including providing appropriate identifying information. In addition, states, and territories should be encouraged to notify by telephone the state or territory of residence of cases reported when some public health action might be appropriate. Also, CDC should institute a process to remove duplicate reported cases of AIDS based on soundex-birthdate matching, if such process does not presently exist.

BACKGROUND/JUSTIFICATION: None

CONSULTATION WITH OTHER ORGANIZATIONS: CDC

CONTACT: Dr. Hall

Appendix B

01-ID-04

Committee: Infectious Disease

Title: Reciprocal (Inter-state) Notification of HIV cases

Statement of the Problem:

In 1986, CSTE adopted a resolution calling for reciprocal notification for AIDS cases which are reportable in all states. Reciprocal notification involves communication between state surveillance programs regarding out-of-jurisdiction cases that may have been previously reported in another state to resolve duplicates. With the advent of HIV reporting in an increasing number of states, the reciprocal notification system needs to be extended to include HIV cases. This includes cases who have been diagnoses with HIV in different jurisdictions, or who have tested HIV positive in one state, moved to another state, and subsequently been reported with AIDS. Although CDC currently estimates that duplication in the national surveillance database is low, it is anticipated that duplication will increase as more states adopt HIV reporting policies and more prevalent HIV cases are reported.

Statement of desired action(s) to be taken:

CSTE recommends that states should follow the same reciprocal notification process for cases of HIV infection as has been used for many years for AIDS. The same systems should be used for reciprocal notification of perinatal HIV exposure and deaths among persons with HIV infection, including providing appropriate identifying information (e.g., name or other identifier). States should ensure that reciprocal notification is carried out by appropriately trained and authorized surveillance staff, in a confidential manner consistent with local security, confidentiality and reporting policies and procedures. CDC should provide states with lists of potential duplicate reports based on soundex and birth date matching in the national database to assist in the identification and removal of duplicate cases. In addition, CDC should provide software tools and technical assistance to facilitate interstate communication regardless of whether name, code, or name-to-code reporting is in effect.

Public Health Impact:

This recommendation will help states maintain and improve the quality of local and national HIV and AIDS surveillance data. In an effort to achieve minimum performance standards set forth in CDC guidelines for national HIV case surveillance, all states will be evaluating the performance of their HIV reporting systems, including the accuracy of case

counts. Therefore, this resolution will support states efforts to unduplicate HIV/AIDS case registries. In addition, because federal funding for care programs will be based on HIV as well as AIDS case counts in the future, improving data accuracy will result in more equitable distribution of care dollars.

Coordination:

Agencies for Response:

Dr. Jeffrey P. Koplan, MD, MPH

Director

Centers for Disease Control and Prevention

1600 Clifton Road, NE

Mailstop D-14

Atlanta, GA 30333

Author:

Gus Birkhead, MD, MPH

Director, AIDS Institute

New York State Department of Health

Room 1483 Corning Tower, ESP

New York, New York 12237

United States

Telephone: (518) 402-5382

Fax: (518) 486-1455

Email: gsb02@health.state.ny.us

References:

None

Appendix C

Council of State and Territorial Epidemiologists
Position Statement

03-ID-10

Committee: **Infectious Disease**

Title: **Revised Guidelines for Determining Residency for Disease Reporting Purposes**

Statement of the Problem:
In general, cases of nationally notifiable diseases are reported based on the case's place of residence, regardless of where exposure may have occurred. While usually obvious, in a small proportion of cases (e.g., children in split parental custody, travelers becoming ill away from home, persons with no fixed address) there can be ambiguity about how to determine residence for reporting purposes. With input from several CSTE members, CDC has drafted updated guidelines for determining residence for disease reporting purposes [Attachment -- Updated guidelines for determining jurisdiction responsible for reporting notifiable diseases to CDC under the National Notifiable Disease Surveillance System (NNDSS)].

Statement of the desired action(s) to be taken:
CSTE should review, amend where appropriate, and adopt these updated guidelines for determining residence for disease reporting (Attachment).

CDC programs participating in NNDSS will follow the updated guidelines for determining residence for notifiable disease reporting. CDC's Epidemiology Program Office, which manages NNDSS, will offer technical assistance as needed regarding interpretation of the guidelines.

Public Health Impact:
Adapting these guidelines is not expected to significantly alter broad geographic patterns of disease incidence. However adapting these uniform guidelines is expected to help clarify and streamline local decision-making about a small number of cases each year. In some cases there is no one “right” answer to the question of residency. The revised guidelines will help ensure complete reporting and avoid duplicate reporting of cases with ambiguous residency—achieving consistency without any need to agonize over each determination.

Agencies for Information:

(1) Sam Groseclose, DVM, MPH
Centers for Disease Control and Prevention
Chief, Surveillance Systems Branch
Epidemiology Program Office (EPO)
4770 Buford Highway, Mailstop K74
Atlanta, GA 30341
Telephone: (770) 488-8403
Fax: (770) 488-8445
Email: sgroseclose@cdc.gov

(2) George E. Hardy, MD, MPH
Executive Director
The Association of State and Territorial Health Officials
1275 K Street NW Suite. 800
Washington, DC 20005-4006
Telephone: (202) 371-9090
Fax: (202) 371-9797
Email: ghardy@astho.org

Council of State and Territorial Epidemiologists
Position Statement

Agencies for Response:

(1) Steve Thacker, MD, MPH
Centers for Disease Control and Prevention
Director, Epidemiology Program Office
1600 Clifton Road NE, Mailstop C-08
Atlanta, GA 30333
Telephone: (404) 639-3661
Fax: (404) 639-4088
Email: Sthacker@cdc.gov

(2) James Hughes, MD
Director, National Center for Infectious Diseases
Centers for Disease Control and Prevention
1600 Clifton Road NE, Mailstop C-12
Atlanta, GA 30333
Telephone: (404) 639-3401
Fax: (404) 639-3039
Email: jhughes@cdc.gov

(3) Walter A. Orenstein, MD
Director, National Immunization Program
Centers for Disease Control and Prevention
1600 Clifton Road, NE Mailstop E05
Atlanta, GA 30333
Telephone: (404) 639-8200
Fax: (404) 639-8626
Email: worenstein@cdc.gov

(4) Harold Jaffe, MD
Director, National Center for HIV, STD, & TB Prevention
Centers for Disease Control and Prevention
1600 Clifton Road, NE Mailstop E07
Atlanta, GA 30333
Telephone: (404) 639-8000
Fax: (404) 639-8600
Email: hjaffe@cdc.gov

Authors:

(1) William E Keene
Epidemiologist
Oregon Dept of Human Services
Acute & Communicable Disease Program
800 NE Oregon St, Ste 772
Portland OR 97232
Telephone: (503) 731-4024
Fax: (503) 731-4798
Email: william.e.keene@state.or.us

Council of State and Territorial Epidemiologists
Position Statement

ATTACHMENT

Updated guidelines for determining jurisdiction responsible for reporting notifiable diseases to CDC under the National Notifiable Disease Surveillance System (NNDSS)

Effective Date: _____

Summary of updated guidelines

For purposes of notifiable disease reporting to CDC, cases should be reported by the jurisdiction of the person's "usual residence" at the time of disease onset. For most people, usual residence is obvious and unambiguous. However, situations do arise for many people in which usual residence is less clear. The following guidelines are intended to provide uniform standards for determining usual residence for disease reporting purposes. The guidelines are modeled after provisions developed for the U.S. Census. The overarching aim of these guidelines is that all cases should be reported, but no case should be reported by multiple jurisdictions. It is important to note that following these guidelines may result in cases being reported by a jurisdiction other than where the infection was acquired. In such instances, other variables can be used to reflect "imported" infections acquired outside the jurisdiction reporting the case.

For instances in which usual residence remains ambiguous, the public health jurisdictions involved should discuss the situation and come to agreement on which jurisdiction will report the case, based on the principles contained in these guidelines. When jurisdictions cannot agree, the Division of Public Health Surveillance and Informatics (DPHSI), Epidemiology Program Office (EPO), Centers for Disease Control and Prevention (CDC) is willing to arbitrate the disagreement and recommend a reporting jurisdiction.

I. Rationale for basing disease reporting guidelines on U.S. Census residency rules.

Although not developed specifically for disease surveillance purposes, residency rules used by the U.S. Census have been developed over many years to account for most circumstances of ambiguous residence. In addition, since notifiable disease data are often combined with population data, disease reporting guidelines based on census residence rules will contribute toward greater consistency in the methods used to collect numerator and denominator data used in disease rates.

II. Concept of usual residence

Usual residence is defined as the place where the person lives and sleeps most of the time, which is not necessarily the same as the person's voting residence, legal residence, or the place where they became infected with a reportable disease. Determining usual residence for most people is easy and unambiguous. However, the usual residence for some people is not obvious. A few examples are people without housing, commuter workers, retirees who spend the winter months in warmer climates (i.e. "snowbirds"), college students, military personnel, and migrant workers.

III. Parameters for disease reporting

It is important to note that disease reporting is not intended to capture the location of exposure per se. If the patient is known to have acquired their infection outside the reporting jurisdiction, the IMPORTED variable should be used in the National Electronic Telecommunications System for Surveillance (NETSS) or the nationally notifiable disease (NND) report message to reflect acquisition of infection outside the reporting jurisdiction. Additional guidance on use of the IMPORTED variable is provided in Appendix I.

To determine usual residence, it is necessary to define a fixed *reference point* in time, analogous to the "census day" used for the census. In addition, for persons who regularly move between residences, it may be necessary to consider a *reference period* preceding the fixed reference point.

Council of State and Territorial Epidemiologists
Position Statement

A. Reference Point

Date of symptom onset is selected as the reference point for establishing "usual residence." If date of symptom onset is not available, the date of lab culture, diagnosis, or the first case report to the health department are recommended, in that order, as the reference point. This is consistent with the use of the EVENTDATE field in NETSS, which gives priority to the "earliest known date associated with this incidence of disease." The advantages to using symptom onset as the reference point rather than diagnosis date are that onset is a more meaningful date from an epidemiological point of view (i.e. more proximal to the date of exposure). In addition, date of diagnosis is frequently unavailable or even non-existent, particularly for cases that are not lab-confirmed or physician-diagnosed (e.g., epi-linked cases identified during an outbreak investigation).

B. Reference Period

If the person is on a regular schedule or cycle for moving between two or more residences, a reference period preceding onset date may be necessary to determine usual residence. Ideally this reference period might coincide with the incubation period of the disease being reported. However, given the variability and uncertainty of incubation periods for the range of notifiable diseases, basing the reference period on disease specific incubation periods would be unnecessarily complicated and impractical. Therefore, we propose defining the reference period consistent with Census Bureau rules, based on the cycle that an individual has for moving between residences. This cycle could be weekly, monthly, yearly, or some other interval. Again, reference period is only relevant for determining usual residence for individuals with a regular cycle for moving back and forth between two or more residences. When the individual takes up a new residence for an indefinite period without intending to return to the previous residence, the jurisdiction of the new residence will be the recommended reporting authority, even if this change of residence occurred shortly before disease onset.

IV. Specific guidelines for determining usual residence at time of symptom onset.

A. People away on vacation or business

People temporarily away on vacation or a business trip at the time of disease onset should be reported by the jurisdiction of their usual residence.

B. People without housing

People without a usual residence should be reported by the jurisdiction where they were staying on the day of disease onset.

C. People with multiple residences

1. Commuter workers living away part of the week while working (on a weekly cycle) should be reported by the jurisdiction where they stay most of the week.

2. People who live in one state most of the year but who regularly spend part of the year in another state (e.g., snowbirds) can be said to have an annual cycle and should be reported by the jurisdiction of the residence where they live most of the year.

3. Children in joint custody should be reported by the jurisdiction of the residence where they live most of the time. If the time is equally divided, they are reported by the jurisdiction where they were staying at the time of disease onset.

4. People who move between residences without any regular cycle should be reported by the jurisdiction of the residence where they live most of the time. If their time is equally divided, report based on where they were staying at the time of disease onset.

Council of State and Territorial Epidemiologists
Position Statement

D. Students

1. College or boarding school students on a typical yearly academic cycle should be reported by the jurisdiction of the residence where they live most of the year.

2. If the individual is an intermittent or part-time student without a regular cycle for moving between parental and school residences, then report by the jurisdiction where they were living at the time of disease onset.

E. Live-ins

Foster children should be reported by where they are living at the time of disease onset.

F. Military or merchant marine personnel in the U.S.

1. People in the military residing in the United States should be reported by the jurisdiction at their usual residence, either on- or off-base.

2. Crews of military vessels with a U.S. homeport should be reported by the jurisdiction at their usual onshore residence if they report one (the place where they live and sleep most of the time when they are onshore); otherwise, at their vessel's homeport.

3. Crews of U.S. flag merchant vessels engaged in inland waterway transportation should be reported at their usual onshore residence (the place where they live and sleep most of the time when they are onshore).

4. Crews of U.S. flag merchant vessels docked in a U.S. port or sailing from one U.S. port to another U.S. port should be counted at their usual onshore residence if they report one (the place where they live and sleep most of the time when they are onshore). If they have no onshore residence, follow rule IV.B and report from nearest jurisdiction at the time of illness onset.

G. Institutionalized persons

1. Patients in general hospitals or wards at the time of symptom onset should be reported by the jurisdiction of their usual residence (the place where they live and sleep most of the time when they are not hospitalized). Newborn babies who have not yet been discharged following delivery should be reported by the mother's usual residence.

2. In general, persons who are institutionalized for indefinite or long-term stays should be reported by the jurisdiction of the facility where they are staying at the time of disease onset. Examples of such facilities include:

> chronic or long-term disease hospitals; hospices; nursing or convalescent homes; inpatient drug/alcohol recovery facilities; homes, schools, hospitals, or wards for the physically handicapped, mentally retarded, or mentally ill; federal and state prisons, jails, detention centers, and halfway houses; orphanages; residential care facilities for neglected or abused children.

3. Staff members living in hospitals, nursing homes, prisons, or other institutions should be reported by the jurisdiction of their usual residence (the place where they live and sleep most of the time); otherwise by the jurisdiction where the institution is located.

H. Foreign citizens (Individuals, regardless of citizenship, who are diagnosed in the U.S. with a notifiable disease, should be reported to CDC.)

1. Foreign citizens who have established a household or are part of an established household in the U.S., including those here for work or study, should be reported by the jurisdiction of their usual residence in the U.S.

Council of State and Territorial Epidemiologists
Position Statement

2. Foreign citizens who live on diplomatic compounds (e.g., embassies, consulates) should be reported by the jurisdiction where the facility is located.

I. Individuals diagnosed in the U.S. but with disease onset outside the U.S. or its overseas territories.

If the person's usual residence at the time of symptom onset was outside the U.S., then the jurisdiction of usual residence at the time of *diagnosis* should report the case, using the same guiding principles contained in this document to determine usual residence. If the person has no usual residence in the U.S. (as per section B or H.3) then the case should be reported by the jurisdiction where the person was staying at the time of diagnosis.

J. U.S. residents diagnosed abroad

When disease onset and diagnosis occurs overseas, illness among U.S. residents is only notifiable in the U.S. if treatment or care occur in the U.S. Health care providers in the U.S. treating patients diagnosed with a notifiable disease while traveling or temporarily living outside the U.S. should notify their local or state health department of the continued treatment or care of a notifiable condition, along with information regarding the location where the disease was likely acquired. The case jurisdiction should be based on location of "usual residence" at the time of treatment or care, and the case should be classified as "imported" as defined in Appendix I.

K. Reporting involving U.S. territories and possessions outside the fifty states and D.C.

See Appendix I.

V. Resolution of disagreements between states

When there is disagreement between states regarding who should report a case, states are encouraged to resolve their disagreement amongst themselves based on the underlying principles contained in these guidelines. If states are unable to come to agreement, the Branch Chief, Surveillance Systems Branch, CDC/EPO/DPHSI is available to arbitrate the disagreement, and recommend a reporting jurisdiction.

Contact Information:
Chief, Surveillance Systems Branch
DPHSI/EPO/CDC
Mailstop K-74
4770 Buford Highway
Atlanta, GA 30341-3717

Telephone: 770-488-8359
Fax: 770-488-8445

Council of State and Territorial Epidemiologists Position Statement

Appendix I. Clarification of response categories for *Imported* variable, NETSS core record

The imported variable should be used in instances when the case is believed to have acquired their infection outside the reporting jurisdiction, based on the usual incubation period for the disease. Below are the current response categories for the *Imported* variable, taken from page 4-12 of the Manual of Procedures for the Reporting of Nationally Notifiable Diseases to CDC:

Coding: Indicates if the case was locally acquired or imported into the state or the US.

Values:

1 = Indigenous (acquired in U.S. in reporting state)
2 = International (acquired outside U.S.)
3 = Out of State (acquired in U.S. but outside the reporting state)
9 = Unknown

Questions have arisen regarding how to categorize cases acquired in U.S. Territories. The following clarification is proposed, which is derived from language in the Code of Federal Regulations related to Foreign Quarantine. U.S. territories include only Puerto Rico, the U.S. Virgin Islands, Guam, American Samoa, the Commonwealth of Northern Mariana Islands, Palau, Marshall Islands, and Federated States of Micronesia. For the purposes of categorizing cases regarding the *Imported* variable, only Puerto Rico and the U.S. Virgin Islands should be considered inside the U.S. The rationale for this is based on regulations for foreign quarantine. All other territories should be considered outside the U.S. Therefore, while the response categories remain the same, the description of values for the *Imported* variable should be amended to read:

1 = Indigenous (acquired in state or territory reporting the case)

2 = International (acquired outside U.S. [i.e. outside 50 states, District of Columbia, Puerto Rico, and the U.S. Virgin Islands]). This includes cases imported to the U.S. from the U.S. overseas territories of Guam, American Samoa, Commonwealth of Northern Mariana Islands, Palau, Marshall Islands, and Federated States of Micronesia.

3 = Out of State (acquired in U.S. [i.e. in 50 states, District of Columbia, Puerto Rico, or the U.S. Virgin Islands] but outside the reporting state).

9 = Unknown

Note: Citizenship or immigration status of the patient has no bearing on the coding of the Imported variable.

The following chart is intended to further assist in classifying individuals with respect to the imported variable.

Location Infection Acquired	Location Reporting Case	Value for *Imported* variable
In any state (including District of Columbia, Puerto Rico, or the U.S. Virgin Islands)	In the same state as infection acquired (including District of Columbia, Puerto Rico, or the U.S. Virgin Islands)	1 = Indigenous
In any state (including District of Columbia, Puerto Rico, or the U.S. Virgin Islands)	In a different state as infection acquired (including District of Columbia, Puerto Rico, or the U.S. Virgin Islands)	3 = Out of State

Council of State and Territorial Epidemiologists
Position Statement

In a U.S. overseas territory (including Guam, American Samoa, the Commonwealth of Northern Mariana Islands, Palau, Marshall Islands, and Federated States of Micronesia.)	In the same U.S. overseas territory as infection acquired (including Guam, American Samoa, the Commonwealth of Northern Mariana Islands, Palau, Marshall Islands, and Federated States of Micronesia.)	1 = Indigenous
In a U.S. overseas territory (including Guam, American Samoa, the Commonwealth of Northern Mariana Islands, Palau, Marshall Islands, and Federated States of Micronesia.)	In a different U.S. overseas territory as infection acquired (including Guam, American Samoa, the Commonwealth of Northern Mariana Islands, Palau, Marshall Islands, and Federated States of Micronesia.)	2 = International
In a U.S. overseas territory (including Guam, American Samoa, the Commonwealth of Northern Mariana Islands, Palau, Marshall Islands, and Federated States of Micronesia.)	In any state (including District of Columbia, Puerto Rico, or the U.S. Virgin Islands)	2 = International
Outside of any state, District of Columbia, Puerto Rico, the U.S. Virgin Islands, Guam, American Samoa, the Commonwealth of Northern Mariana Islands, Palau, Marshall Islands, and Federated States of Micronesia	In any state, District of Columbia, Puerto Rico, the U.S. Virgin Islands, Guam, American Samoa, the Commonwealth of Northern Mariana Islands, Palau, Marshall Islands, and Federated States of Micronesia	2 = International

Notes

Technical Guidance for HIV/AIDS Surveillance Programs

Duplicate Review

HIV Incidence and Case Surveillance Branch
Atlanta, Georgia

Contents — Duplicate Review

Intrastate 7-3
- Structural Requirements 7-3
- Process Standards 7-4
- Outcome Measures 7-4

Interstate 7-4
- Routine Interstate Duplicate Review (RIDR) 7-4
- Structural Requirements 7-5
- Process Standards 7-5
- Outcome Measures 7-7

Appendix A 7-8

Appendix B 7-9

Technical Guidance for HIV/AIDS Surveillance Programs — Duplicate Review

The ability of a surveillance system to correctly distinguish newly reported persons from persons previously reported should be measured to determine accuracy of case counts. Failing to properly link case reports to existing cases leads to overcounting; incorrectly linking reports to existing cases can lead to undercounting and contamination of existing records with data from another case. At the local level, surveillance software and other routine surveillance practices are used to eliminate duplicate reports. At the national level, CDC does not receive patient names or other specific patient identifiers such as Social Security number, so duplicate reports cannot be eliminated with the same degree of accuracy as at the state/local level where the patient identifiers and other details are maintained.

Because doctors, hospitals, laboratories, and other sources may be required to report all HIV diagnoses whether or not they are newly identified, duplicate case reporting within a state may not be identified during routine case entry into the surveillance database (i.e., duplicate case reports may not be readily identified as such). Eliminating duplicate reports on a regular basis is an important component of maintaining a reliable surveillance database and ensures accurate case counts at the state and national levels. Using the following procedures, all jurisdictions should routinely perform data quality checks to verify that a person in their surveillance database has a single state-assigned number.

Intrastate

Structural Requirements

1. **Case Report, Laboratory, and Other Reports Received on an Individual**
2. **HIV/AIDS Surveillance System Software**
3. **Data Processing Procedures and Tools for Record Linkage**
4. **Procedures for Evaluating Accuracy of Integrated HIV/AIDS Surveillance Systems** (Data Quality)
5. **Surveillance Policies and Procedures for Record Linkage**
6. **Variables Used for Ascertaining Potential Duplicates**
   - Exact variable match for determining that a duplicate case is the same person (if present): first name, last name, middle name, Soundex, date of birth, sex, full Social Security number, death date
   - Decision maker variables useful when they are identical: medical record number, inmate identification number, diagnosis date, others

Process Standards

1. **Frequency of Procedure**
   - Jurisdictions with centralized data management should perform this duplicate review process monthly.
   - Jurisdictions with disseminated data management systems should perform duplicate review monthly at the local level and quarterly at the statewide level.
   - Jurisdictions should perform exact and 'fuzzy matching' using tools provided by data management.
2. **Valid Duplicates**
   - Valid duplicates are those cases where Soundex, date of birth, sex, and state of residency at diagnosis are the same, but the cases are, in fact, different persons.
   - HARS help desk should be notified in order to flag a 'valid duplicate.'
   - HARS replacement software contains a screen that allows the jurisdiction to enter information on the outcome of intrastate and interstate duplicate review.

Outcome Measures

1. **Evaluation Standard**: less than 5% duplicates. See methods in Data Quality.
2. **Results Reporting to CDC**
   - Although CDC does not receive results of routine intrastate duplicate review procedures, jurisdictions should perform these procedures no less than quarterly and be able to answer in the affirmative if queried about performance.

Interstate

Routine Interstate Duplicate Review (RIDR)

Interstate duplicate case reporting can result from persons moving to different states over time; as persons receive care for HIV infection in their new state of residence, they are reported to that state health department in accordance with local reporting requirements. Because interstate migration can occur at any time and at multiple times throughout the course of illness and because case information is reported at several sentinel events (HIV diagnosis, AIDS diagnosis, and death), there are multiple opportunities for duplicate reporting to occur. It is anticipated that the potential for duplicate reporting of cases in the national database will increase as more areas adopt name-based HIV reporting, as persons with HIV remain healthier longer because of advances in therapy, and as laboratory-based surveillance increases. Therefore, routine procedures to ensure that state and national data sets do not 'overcount' cases are necessary.

At the national level, cases reported to CDC are linked by Soundex, date of birth, sex, and state of residence at AIDS diagnosis or HIV diagnosis if no AIDS diagnosis has been made. Conversely, valid duplicates for different persons (cases with matching Soundex, date of birth, sex, and same state of residence at AIDS diagnosis that are actually two different individuals) should be flagged to prevent inappropriate linkages. As data management procedures progress, increased software ability to determine and manage exact and valid duplicates is being developed.

In 1986, CSTE passed a resolution for a state-to-state reciprocal notification process for AIDS case reporting to encourage resolution of duplicate case counting. [Most states have adopted this policy for HIV case reports as well.] A new resolution that includes HIV as well as AIDS case reports was approved at the 2001 CSTE meeting. HIV/AIDS surveillance program staff should communicate with other states to resolve potential duplicates using procedures outlined next.

Structural Requirements

1. **CSTE Position Statement 17 (1986): AIDS Case Reporting: Reciprocal Notification** (Appendix A)

2. **CSTE 01-ID-04 Reciprocal (Interstate) Notification of HIV Cases** (Appendix B)

3. **HIV/AIDS Surveillance System Software**

4. **CDC Procedures for Linking Cases in the National HIV/AIDS Database**

5. **Variables Used for National Case Linking**: Soundex, date of birth, sex, state of residence at diagnosis

6. **Surveillance Policies and Procedures for Record Linkage**

7. **Case Residency Assignment Policies and Procedures**

8. **Performance of Intrastate Duplicate Review**

9. **Access to Secure Data Network; Current Digital Certificate**

Process Standards

1. **Frequency of Procedure**—Semiannually or quarterly [to be determined by CDC after implementation of new surveillance system software]

2. **Out-of-Jurisdiction Case Handling**

   States should maintain information on out-of-jurisdiction cases. Often, persons diagnosed in one state may be receiving care and services in another jurisdiction. Use policies and procedures for state residency assignment to ensure that cases are counted appropriately.

3. Resolution of Potential Duplicates

- Contact the surveillance coordinator or designee from the appropriate state, and use additional case information (such as patient name) to determine whether a potential duplicate is the same or a different person. Persons who are approved to release information about HIV/AIDS cases to other jurisdictions can be found on the CSTE HIV/AIDS Contact Board Web site. Please contact the HIV/AIDS surveillance support staff at CSTE for information on obtaining sign-on identifications and passwords to access the Web site. Other questions that can be asked are these:
  - Do the two states share a border?
  - Does the Social Security number prefix come from the other state?
  - Are there any comments that reference the other state?
  - Is there a death date match?
  - Is there a current residence match?
  - Is there an unusual mode of exposure?
- If the cases are deemed to be different persons, indicate this in the appropriate area in your data management system. In HARS, this could be a local field or supplemental database. In eHARS, a duplicate review screen captures this information.
- If the cases are deemed to be the same person, assign residency at diagnosis of HIV and residency at diagnosis of AIDS (as necessary) using policies and procedures for residency assignment. Results of this review should be entered in the data management system.
- Use communication with the other state surveillance coordinator (or designee) to collect additional information (such as risk factor, death date, HIV diagnosis information, AIDS-defining conditions, etc.) to complete your case record and to support the diagnosis of HIV and/or AIDS.
- Contact the CDC Routine Interstate Duplicate Review (RIDR) coordinator for resolution of any case residency assignment disputes.

4. Valid Duplicates (Different Persons)

- Valid duplicates are those cases where Soundex, date of birth, sex, and state of residence at diagnosis are the same, but the cases are, in fact, different persons.
- HARS help desk should be notified to flag a ‘valid duplicate’ for intrastate and interstate duplicates.
- HARS replacement software contains a screen that allows the jurisdiction to enter information on the outcome of intrastate and interstate duplicate review.

Outcome Measures

1. **Standard**: less than 5% duplicates in the national database
2. **Procedures for Evaluation** [Pending]

Appendix A

CSTE POSITION STATEMENT 1986-17

COMMITTEE: Surveillance

TITLE: AIDS Case Reporting: Reciprocal Notification

ISSUE: All states and territories now conduct AIDS surveillance. Currently, case counting by states and territories is delayed by sending AIDS case reports on individuals now residing in other states and territories to the Centers for Disease Control rather than the state or territory of residence at time of onset of illness. Occasionally, duplicate counting of cases reported from different states or territories has occurred and some problems have arisen in exchange of information between states.

POSITION TO BE ADOPTED: The CDC should encourage states and territories to use the same state to state reciprocal notification process for AIDS case reports that is used for all other reportable diseases, including providing appropriate identifying information. In addition, states, and territories should be encouraged to notify by telephone the state or territory of residence of cases reported when some public health action might be appropriate. Also, CDC should institute a process to remove duplicate reported cases of AIDS based on soundex-birthdate matching, if such process does not presently exist.

BACKGROUND/JUSTIFICATION: None

CONSULTATION WITH OTHER ORGANIZATIONS: CDC

CONTACT: Dr. Hall

Appendix B

01-ID-04

Committee: Infectious Disease

Title: Reciprocal (Inter-state) Notification of HIV cases

Statement of the Problem:

In 1986, CSTE adopted a resolution calling for reciprocal notification for AIDS cases which are reportable in all states. Reciprocal notification involves communication between state surveillance programs regarding out-of-jurisdiction cases that may have been previously reported in another state to resolve duplicates. With the advent of HIV reporting in an increasing number of states, the reciprocal notification system needs to be extended to include HIV cases. This includes cases who have been diagnoses with HIV in different jurisdictions, or who have tested HIV positive in one state, moved to another state, and subsequently been reported with AIDS. Although CDC currently estimates that duplication in the national surveillance database is low, it is anticipated that duplication will increase as more states adopt HIV reporting policies and more prevalent HIV cases are reported.

Statement of desired action(s) to be taken:

CSTE recommends that states should follow the same reciprocal notification process for cases of HIV infection as has been used for many years for AIDS. The same systems should be used for reciprocal notification of perinatal HIV exposure and deaths among persons with HIV infection, including providing appropriate identifying information (e.g., name or other identifier). States should ensure that reciprocal notification is carried out by appropriately trained and authorized surveillance staff, in a confidential manner consistent with local security, confidentiality and reporting policies and procedures. CDC should provide states with lists of potential duplicate reports based on soundex and birth date matching in the national database to assist in the identification and removal of duplicate cases. In addition, CDC should provide software tools and technical assistance to facilitate interstate communication regardless of whether name, code, or name-to-code reporting is in effect.

Public Health Impact:

This recommendation will help states maintain and improve the quality of local and national HIV and AIDS surveillance data. In an effort to achieve minimum performance standards set forth in CDC guidelines for national HIV case surveillance, all states will be evaluating the performance of their HIV reporting systems, including the accuracy of case

counts. Therefore, this resolution will support states efforts to unduplicate HIV/AIDS case registries. In addition, because federal funding for care programs will be based on HIV as well as AIDS case counts in the future, improving data accuracy will result in more equitable distribution of care dollars.

Coordination:

Agencies for Response:

Dr. Jeffrey P. Koplan, MD, MPH

Director

Centers for Disease Control and Prevention

1600 Clifton Road, NE

Mailstop D-14

Atlanta, GA 30333

Author:

Gus Birkhead, MD, MPH

Director, AIDS Institute

New York State Department of Health

Room 1483 Corning Tower, ESP

New York, New York 12237

United States

Telephone: (518) 402-5382

Fax: (518) 486-1455

Email: gsb02@health.state.ny.us

References:

None

Technical Guidance for HIV/AIDS Surveillance Programs

Data Management

HIV Incidence and Case Surveillance Branch
Atlanta, Georgia

Contents — Data Management

Data Management 8-3
Structural Requirements 8-3
Process Standards 8-3

Technical Guidance for HIV/AIDS Surveillance Programs — Data Management

Data Management

The surveillance practices described in this chapter include the management of paper and electronic documents, record linkage, and documentation of procedures. These practices are supplemented by the technical manual for the HIV/AIDS surveillance software, *eHARS Technical Reference Guide*.

Structural Requirements

The surveillance program must be prepared to receive and process different kinds of documents because case information will be received on different media (paper forms, electronic reports, records from other databases) and from different report sources. The structural requirements for data management include hardware and software that allow searching for existing cases, entry and retention of multiple documents per case (relational database), importing of documents and linkage to other databases, transmission of data, data quality functions (e.g., data edits), and report generation in a secure environment (see Vol. III: Security and Confidentiality Guidelines). Further, staff must receive adequate training (initial training and continuing education) in data processing to assure consistency.

Process Standards

The processes for data management must be documented in procedures manuals and kept up-to-date to maintain an adequately functioning system and to train staff. The documentation may include manuals for the data entry software as well as documentation of processes specific to the local health department.

Data Entry

Paper documents received at the health department should be marked with the date of receipt and, after visual editing (proofreading), data items from these documents should be entered into the HIV/AIDS surveillance software. Local programs should develop written procedures for handling and storage of paper and electronic documents, including an acceptable time frame from document receipt to data entry. Computerized data collection, where implemented, combines abstracting, coding, data entry, and editing into one process. This manner of collecting data may be implemented by providing software to the reporting facilities or laptops to health department field staff. Some verification of keyed data should be in place regardless of the methods used, such as duplicate data entry to compare original and duplicate documents for discrepancies (see Data Quality). A date of receipt should also be entered for any data received by electronic reporting.

The HIV/AIDS surveillance software is designed to facilitate case investigations. Documents for potential cases of HIV infection and those for cases with established HIV infection but other missing information should be entered, and progress in obtaining information should be tracked (see Processing Follow-ups, Corrections, and Deletions for further details).

Record Linkage

See Record Linkage.

Rules for Managing Multiple Documents per Case

Information received from multiple documents needs to be consolidated to yield one case record for analysis. With a system that manages multiple documents per case, most of this consolidation will occur through an automated process; specifications (hierarchy rules) for this process are outlined in the *eHARS Technical Reference Guide, Person View Hierarchy.* However, edit checks or other quality control procedures may discover errors in documents, and the surveillance program's investigators may want to override the hierarchy rules to ensure that a correct value is selected for a given data field.

Edits

Data edits are automatically applied when information is entered into the system and are applied to all data on a routine schedule (see Data Quality). The edits are described in the eHARS software documentation. For a system in which multiple documents are entered for each case, each document should be retained as originally received. If data edits indicate incorrect information in a particular document, surveillance staff should verify that the information was correct as indicated by the source records. However, information in a document should not be changed based on information from other documents. The rules for managing multiple documents should be applied to select the best value for a data item, or to override information from a document that is proven to be incorrect.

Processing Follow-ups, Corrections, and Deletions

In many instances, follow-up is required to obtain all necessary information on a case. Information obtained from a source during follow-up for missing information on a case may range from one particular data item (e.g., risk factor information) to many data items; this information is recorded on a new case report form if the follow-up information is obtained from a new data source. If the follow-up is a re-visit or re-contact of a source from which data were already collected, and the information is on a paper form that has not been entered into the surveillance software, the additional information may be recorded on the original form. When information is gathered on a new case report form during a follow-up investigation, this form must be marked as a follow-up report; this information is also entered into the surveillance software.

In some instances, information is found to be incorrect (e.g., a data edit revealed an invalid value and the report source is re-contacted or re-visited and the correct information is obtained). Then the information should be corrected on paper and electronic forms. Changes to original documents should only occur if verification is obtained from the original source. If information is incorrect on a document but known to be correct in another document for the same person, staff can assure the correct value is used for the person's record for analysis by reviewing the person record in the software and, if necessary, setting the software so the correct value is used (see *eHARS Technical Reference Guide, Person View Hierarchy*). Any errors created during data entry should be corrected on the appropriate document.

Forms or entire cases may also need to be deleted after they have been entered into the surveillance software (for example, person is found not to be HIV infected). After a document or record for a person is marked for deletion and the updated information is sent to CDC, final deletion occurs after feedback from CDC is received that the document/person record deletion was noted and was also deleted in the national database.

Documentation

Data management processes—data collection, additions, changes, deletions, and the like—should be documented, including the staff performing the action, the date, and the action (e.g., changed sex from male to female). For the most part, these processes will be automatically documented in the surveillance software. In addition, data management reports should be routinely created and shared with the appropriate parties. Such reports include

- Reports of database and surveillance operations (e.g., staff access to data, changes/deletions, workflow [e.g., cases by process completed such as visual editing, follow-up investigations], and listing of cases missing critical data items)
- Reports of cases reported by report source
- Reports of timeliness and completeness of reporting by report source to provide feedback to reporting sources
- Reports on quality control operations (see Data Quality)

Notes

Technical Guidance for HIV/AIDS Surveillance Programs

Record Linkage

HIV Incidence and Case Surveillance Branch
Atlanta, Georgia

Contents — Record Linkage

Introduction 9-3
Definition of Record Linkage 9-3
Purposes of Record Linkage 9-4
Structural Requirements and Organizational Barriers to Record Linkage 9-4
Memorandum of Understanding 9-5
Human Subjects Approval 9-5
Information Technology Resources 9-5
Data Files That May Be Useful for Record Linkage to HIV/AIDS Cases 9-5
Identification of HIV-specific Data Files 9-7
Variables of Interest When Updating Existing Cases by Record Linkage 9-9
Variables on Which to Search for Agreement 9-10
Fixing Problems in Data Files before Record Linkage 9-13
Record Linkage Strategies 9-16
Considerations for Selecting a Computer Software Program for Record Linkage 9-23
Visual Review of Indeterminate Records 9-25
Tracking Log 9-26
Measures or Standards for Record Linkage 9-27

Appendixes
Appendix A 9-30
Appendix B 9-32
Appendix C 9-34
Appendix D 9-37
Appendix D.1 9-38

Technical Guidance for HIV/AIDS Surveillance Programs — Record Linkage

Introduction

The following are policies and procedures for HIV/AIDS surveillance programs to conduct record linkage between the HIV/AIDS surveillance case data file and other ("external") data files, such as the state death certificate data file or laboratory test data files (e.g., results of tests for HIV antibodies or CD4 T-lymphocyte counts). These policies and procedures are intended mainly for linkage with electronic data files, but some of them may also apply to linkage with paper files.

Definition of Record Linkage

Record linkage is a comparison of records to assess the extent of agreement on the values of selected variables; on the basis of such agreement, records are classified as "links" (referring to the same [a single] person) or "nonlinks" (referring to different [separate] persons). Records that a linkage could not classify as links or nonlinks are classified as "indeterminate." If the classification of records into links and nonlinks does not perfectly correspond to the true status of the records, same-person records (those pertaining to the same person) and different-person records (those pertaining to different persons), the links and nonlinks may be subdivided into those that are true and those that are false (Table 1). The values of all variables may not need to agree completely for records to be classified as links, particularly if agreement between some variables is given more weight than agreement between other variables. If the records compared are in two separate data files, one data file may have more than one record linked to a single record in the other file (as may be true for laboratory test results)—a linkage of many to one instead of one to one. Record linkage can also be applied within a single file to find multiple records that refer to a single case (which are called "duplicates" if they are redundant).

Table 1

Results of Record Linkage Process	**Records Truly Refer to:**	
	Same Person	**Different Person**
Links	True link	False link
Nonlinks	False nonlink	True nonlink
Indeterminate	Indeterminate, same person	Indeterminate, different person
Total	Total same-person records	Total different-person records

Purposes of Record Linkage

Record linkage may be used for several purposes in a case registry, such as in HIV/AIDS case surveillance:

1. **To Find Previously Unreported (New) Cases**. Records from the external file classified as nonlinks can be used to find new cases if the records are known to represent persons who qualify as cases because the records are either exclusively for HIV-infected persons or have a data field in which a diagnosis of HIV infection has been entered (Figures 1 and 2). Follow-up investigations must be done to confirm these cases and to collect all data required for new cases.
2. **To Find New Information (Updating) on Previously Reported (Existing) Cases**. Records from the external file classified as links may provide new information, including
   - Date of death
   - Laboratory test results, such as CD4 counts
   - Risk factors for HIV infection
   - AIDS-defining diseases, especially communicable diseases (e.g., tuberculosis) of public health importance or cancers that should be reported to their own registries
3. **To Collect Data on Additional Variables for Epidemiologic Research** (as distinct from routine surveillance).
4. **To Find Duplicates for the Purpose of Eliminating Redundancy or Correcting Errors of Counting a Single Case as Multiple Cases**.

Structural Requirements and Organizational Barriers to Record Linkage

State laws or departmental policies may limit how record linkage can be done. These vary among states. In some states, policies and laws may allow the HIV/AIDS surveillance program to entrust its confidential data to another office or agency, such as the state office of vital records, so that the latter can search the death registry for links to HIV/AIDS records. In other states, the HIV/AIDS surveillance program may be prohibited from disclosing personally identifying information on cases to the staff of other offices or agencies—even those in the same department. Consequently, the HIV/AIDS surveillance programs may have to obtain personally identifying information from the records of the other data files to conduct record linkages.

Memorandum of Understanding

For one office/agency to disclose such confidential information to another, the one that owns the data to be disclosed may require the other to sign an official "Memorandum of Agreement" or "Memorandum of Understanding" that describes how the recipient of the data promises to protect the confidentiality of the information. An example of such a memorandum is in Appendix A.

Human Subjects Approval

Record linkage is recommended as a routine surveillance activity and does not usually require Institutional Review Board oversight. However, requirements for such review will depend on the purpose of the linkage, the funding source for the activity, the sources of the data systems, and local rules and regulations.

Information Technology Resources

Technical support and resources may affect which office/agency must disclose confidential data to the other. Record linkage will require knowledgeable and well-trained computer programming staff, record linkage software, and computers with enough speed, memory, and hard-drive capacity to handle large data sets. During linkage between two data files, a temporary data file may result that contains different combinations of pairs of records, with a maximum number of pairs equal to the product of the number of records in each file. For example, if the HIV/AIDS case file has 500 records and the death certificate data file has 100,000 records, then 500 × 100,000 = up to 50 million pairs of records could result; this would require 10–15 gigabytes of hard-drive space, which may exceed the capabilities of the computers of most HIV/AIDS surveillance programs. Servers on networks might be able to handle the load better than a stand-alone microcomputer. Doing the linkage on a mainframe computer instead of a microcomputer may be more practical in terms of hardware. Use of a mainframe or server may be more common in other offices (e.g., vital records, Medicaid) where the data files are generally much larger than the HIV/AIDS data files.

Data Files That May Be Useful for Record Linkage to HIV/AIDS Cases

The following list of types of data files may be useful for finding new information on existing cases or for finding new cases. The possibility, costs, and benefits of linkage to these files may vary among states. The yield of new data from some sources may be greater if data from other sources have not been collected.

1. **Death Registries** (see Death Ascertainment for more information.)
   - State or Local Death Certificate Data Files
   - National Death Index
   - Social Security Administration's Death Master File (DMF) or "Social Security Death Index" Files Derived from the DMF
2. **Birth Registry** (linkage of HIV/AIDS case records of women with birth records of their children may reveal potential new HIV/AIDS cases among the children; linkage of HIV/AIDS case records of children with their own birth records may reveal potential new HIV/AIDS cases among their mothers).
3. **TB Registry**
4. **Cancer Registry** (particularly for AIDS-defining cancers: lymphoma, Kaposi's sarcoma, and cervical carcinoma)
5. **Laboratory Test Data Files**
   - HIV Antibody Tests
   - HIV Antigen Tests
   - HIV Nucleic Acid Tests (including HIV viral loads)
   - CD4 T-lymphocyte Counts

   Viral loads and CD4 counts may be useful not only to find new cases of HIV infection or AIDS, but also to identify those receiving care, which may be information required to apply for Ryan White Comprehensive AIDS Resources Emergency (CARE) Act funding.
6. **Data Files from Clinics or Services That Are Specifically for HIV-infected Patients**
   - AIDS/HIV Drug Assistance Program (ADAP)/Ryan White CARE Act Clinics
   - HIV Early Intervention Clinics
   - HIV Counseling and Testing Programs
   - Research Projects Limited to HIV-infected Persons (but issues of informed consent to sharing of data with the surveillance program may need to be addressed, particularly if research data are protected by a Certificate of Confidentiality)

7. **Data Files from High-risk Clinics**
   - Sexually Transmitted Disease Clinics (e.g., STD*MIS)
   - Drug Addiction Treatment Clinics
   - Hepatitis C Clinic or Registry
   - Department of Corrections (may have a clinic file or simply a list of HIV/AIDS cases with minimal information)
8. **Administrative or Clinical Data Files That May Have ICD-9-CM Codes for HIV Disease or Drug Codes Specific for Antiretroviral Drugs**
   - Hospital Discharge Abstracts
   - Medicaid Claims Data Files
   - Medicare Claims Data Files
   - HMO (e.g., Kaiser Permanente) or Other Outpatient Data Files

Identification of HIV-specific Data Files

To enable users to find potential new cases among records found not to be linked to existing cases (nonlinks) (Figures 1 and 2), the data file that contains the records must either be limited to data on HIV-infected persons or must include a variable that allows a subset of the data to be created that is limited to HIV-infected persons. Data on positive HIV laboratory test results and data from ADAP and Ryan White CARE Act clinics are, by definition, limited to HIV-infected persons. Data from disease registries and administrative data files of hospitals and outpatient clinics are generally not limited to HIV-infected patients, but may be reduced to an HIV-specific subset if they include a variable that either is a flag for HIV infection or is a code for whatever disease the patient has. A commonly used list of such codes is the International Classification of Diseases (ICD) (e.g., ICD-10) for causes of death and the ICD-Clinical Modification (e.g., ICD-9-CM) codes for conditions associated with hospitalizations or outpatient encounters.

1. **HIV-specific Disease Codes**
   - ICD Codes (for Death Certificate Data) Specific for HIV Infection

     ICD-9 codes (effective 1987–1998):
     - 042.0–044.9 (AIDS or HIV disease)
     - 795.8 (positive HIV laboratory test)

     ICD-10 codes (effective 1999 to present):
     - B20.0–B24 (symptomatic [ever])
     - R75 (positive HIV laboratory test)

- ICD-CM Codes (for Hospital or Outpatient Data) Specific for HIV Infection

  ICD-9-CM codes effective 1987–1994:

  - 042.0–044.9 (AIDS or HIV disease)
  - 795.8 (positive HIV laboratory test)

  ICD-9-CM codes effective 1995 to present (2005):

  - 042 (symptomatic [ever])
  - V08 (asymptomatic [never previously symptomatic])

  ICD-10-CM codes (implementation date not yet known):

  - B20.0–B24 (symptomatic [ever])
  - R75 (positive HIV laboratory test)
  - Z21 (asymptomatic [never previously symptomatic])

2. Potential Leads to New Cases That Require Follow-up for Confirmation

The following sources of records may be useful to update previously reported cases or to find new (previously unreported) cases. If used to find new cases, the records must be followed up for confirmation of HIV infection because they may not necessarily represent cases of HIV infection.

- Undetectable Viral Loads

  Some health departments consider all records of HIV viral load tests to represent potential HIV/AIDS cases and attempt record linkage even to those with undetectable viral loads. However, on follow-up, some of these persons with undetectable viral loads have been found not to be infected (HIV antibody test negative).

- Low CD4 Counts or Percentages

  A low CD4 count (<200 per microliter or <14% of total white blood cells) can be used as an indicator of possible AIDS. However, on follow-up, the person may be found not to be HIV-infected, and the low CD4 count may have been due to a disease other than HIV infection (e.g., lymphoma, leukemia).

- AIDS-defining Opportunistic Illnesses

  Some health departments have used an opportunistic infection, such as Pneumocystosis pneumonia (ICD-9 code 136.3, ICD-10 code B59), as an indicator of possible AIDS. However, on follow-up, the person may be found not to be HIV-infected, and the opportunistic disease may have been unrelated to HIV infection. Previous studies have shown that pneumocystosis is the opportunistic disease most predictive of underlying HIV infection. Other opportunistic diseases are more likely to have been due to other underlying conditions or to have arisen spontaneously.

Variables of Interest When Updating Existing Cases by Record Linkage

1. **From Death Registries** (see also Death Ascertainment)
   - Vital status (if dead)
   - Date of death
   - State where death occurred
   - State of residence at death
   - Cause(s) of death
2. **From Laboratory Data Files and Clinical Data Files** (see also Electronic Reporting)
   - CD4 counts (for earliest known value, lowest value, most recent value, value <200 per microliter to indicate AIDS, or any value as an indicator of receipt of patient care and overall level of immunosuppression among infected to identify need for drug therapy).
   - HIV test results, including viral load tests (HIV tests can confirm "physician diagnosis" of HIV infection and confirm that AIDS-defining opportunistic illnesses represent AIDS. Viral load tests may also be an indicator of medical care or infectiousness.)
3. **From Data Files from Counseling and Testing Programs, STD Clinics, and Drug Addiction Treatment Clinics** (see also Risk Factor Ascertainment)
   - HIV risk factors (e.g., male-to-male sexual contact [MSM], injection drug use [IDU])
4. **From Cancer Registries: AIDS-indicative Cancers**
   - Non-Hodgkin's lymphoma (including primary brain lymphoma)
   - Kaposi's sarcoma
   - Cervical carcinoma

5. **From Hospital and Clinic Data Files**:

   Collection of data on most AIDS-defining diseases is not worthwhile for studying the spectrum of HIV/AIDS disease because it is done so incompletely (most cases being reported with only a low CD4 count). Such data are better collected in special studies (e.g., Morbidity Monitoring Project) rather than in routine surveillance. However, preventable or transmissible opportunistic infections are of special importance:

   - Tuberculosis cases should be reported to the proper unit of the health department as permitted by local law or policy.
   - Preventable infections, such as pneumocystosis, toxoplasmosis, and non-tuberculous mycobacteriosis, are a measure of inadequate medical care among previously reported cases. Hepatitis A/B/C may be indicators of behavioral risk factors (IDU/MSM).

6. **From STD Clinic Data Files**

   - STDs acquired after patients knew they had HIV infection (indicators of unsafe sexual behavior despite knowledge of HIV infection)

7. **From Hepatitis Registries**

Variables on Which to Search for Agreement

To select the variables to be used for record linkage ("linkage variables"—i.e., the variables on which to search for agreement [having the same values]), assess the proportion of records with missing or invalid values for those variables in each of the two data files to be compared. Only variables that have a low proportion of records with missing or invalid values in both data files should be used for linkage.

1. **Variables Generally Available in HIV/AIDS Surveillance Data**:

   - Soundex code of surname (generally derived from the last name, rather than collected separately). See special considerations in "Derivation of Variables with Less Discriminating Power for More Sensitivity" and a warning about Soundex in Appendix 2: "Examples of Standardization of Incompatible Variable Coding Schemes."
   - Last name (but it may change for adopted children and for women who marry or divorce).

     Note: In the HIV/AIDS Reporting System (HARS), "name" is a single variable from which the last name must be extracted by means of a comma separating it from the first name. In eHARS, first and last names are separate variables.]

- First name
- Date of birth
  - Year
  - Month
  - Day
- Sex
- Race (but it may not always be reliable; may change to multiracial)
- Ethnicity (Hispanic/Non-Hispanic)
- Country of birth
- Current residence (informative if it agrees, but not if it disagrees)
  - State
  - City
  - Street address
  - ZIP code

2. **Variables Less Commonly Maintained in Either HIV/AIDS Data Files or Other Data Files, but Which Could Be Very Helpful for Linkage if They Were Collected**:

- Social Security Number (SSN)
  - If available and accurate, the SSN can be extremely helpful because it uniquely identifies an individual, making it a powerful tool for record linkage.
  - However, the SSN is often unavailable in many data files, such as electronic data files for laboratory test results, from which many cases are now discovered. Also, because an SSN has nine digits, various errors in data collection and data entry are possible.
  - The SSN may erroneously be attributed to a person to whom it does not belong (e.g., spouse, child, or sibling).
  - Programs may be reluctant to collect the SSN because U.S. Public Law 93-759, entitled the Privacy Act of 1974, requires that all individuals be informed of the purposes and uses to be made of such information by the person who collects it from them, and states that health services may not be withheld from a person because he/she refuses to disclose his/her SSN. However, federal law does not prohibit state/local public health surveillance programs from copying SSNs from forms or data files into which they have already been recorded, and using them as identifiers in another data file. In fact, the use of the SSN as an identifier when conducting record linkage to their databases for research is encouraged by

both the Social Security Administration (http://www.ntis.gov/products/ssa-dmf.asp) and the CDC (through the National Death Index) (http://www.cdc.gov/nchs/data/NDICriteria_Front.pdf). Nonetheless, each HIV/AIDS surveillance program should consult the attorney for its health department about whether any state/local laws, regulations, or policies are more restrictive than federal law on their permission to collect SSNs from records and databases and use them as identifiers for the purpose of public health surveillance. Privacy laws are frequently ambiguous and open to differing interpretations. Court decisions, changes in related laws, and the background of specific attorneys providing support to health departments can all influence such interpretations. Therefore, advice from legal counsel concerning the use of SSNs for legitimate public health purposes, such as record linkage, should not be considered immutable. Review of such decisions should be requested on a recurring basis to ensure that procedures and practices are consistent with the most current laws, precedents, and perspectives.

  - The last 4 digits of the SSN are used by some surveillance programs as a component of a coded identifier, but they are not unique.

- Middle name or middle initial (sometimes confused with or switched with the first name on some records)
  - Father's surname (important for children)
  - Maiden name
  - State of birth. Although the country of birth is collected in most HIV/AIDS case reports, both the state and country of birth are usually missing in other data files.

    Middle name, father's surname, and state of birth are not on the case report form, but linkage may be improved if their collection became routine (use local data fields or designated fields in eHARS). However, they may not be present in other data files.

- Patient or medical record identification number (MRN) at a hospital or clinic (to uniquely identify a patient, the name of the health care facility must be combined with the medical record number).
  - This number may allow linkage to find multiple records from the same facility that can be linked to the HIV/AIDS case if the MRN has been previously entered in the case record.
  - It may also be used on laboratory data sent to that facility, allowing linkage to laboratory data.
  - In some locations, the MRN is the SSN, allowing the linkage of the MRN in HARS to another database.

- Department of Corrections Identification Number
  - This number may help discovery of links of laboratory results to HIV/AIDS cases if the ID number has previously been recorded in the HIV/AIDS case data file.

Fixing Problems in Data Files before Record Linkage

1. Discovery of Missing Values, Errors, Inconsistencies, and Incompatibilities

The quality of the data within each file and incompatibilities between supposedly equivalent variables in different files should be examined by doing the following:

- Tabulating the Frequencies of Values of All Important Variables:

  This can reveal missing, unknown, invalid, or erroneous values. Deviation of the frequencies from expected patterns can be a clue that something is wrong.

- Studying the Meanings of Variable Values in the Data Documentation (Data Dictionaries and File Layouts):

  In addition to comparing the documentation of the two data files that will undergo linkage, the documentation of each data file from different years should be examined to detect changes over time. For example, a major change occurred in death certificate data files in 1999 when the ICD-10 codes began to be used instead of the ICD-9 codes. Changes also have occurred from year to year in ICD-9-CM values for hospital discharge data files. Other variables, such as race and ethnicity, have undergone changes in the way they are coded. To the extent possible, such differences should be eliminated before using the variables for linkage or analysis. See "Standardization" on how to standardize codes.

- Running Edit Checks:

  Each data file should be screened for invalid or obviously erroneous values. Ideally, edit checks should be done when data are entered, rather than waiting until record linkage is planned. For HIV/AIDS surveillance data, edit checks should already have been done by following the Data Quality policies and procedures; but they may also need to be done for an external data file that is to undergo linkage with HIV/AIDS surveillance data. To recognize invalid values, the data documentation can be reviewed to learn about valid values. The following are examples of invalid or erroneous values:

  - Numeric values for month outside the range from 1 through 12
  - Numeric values for day outside the range from 1 through 31
  - Numeric values for year outside the range of what is possible (e.g., birth before 1890, death after the current date)

- Dates with impossible relationships to one another (e.g., birth date after death date)
- Age inconsistent with year of birth (e.g., age differing by more than 2 years from the difference between the year of the event [for example, death] and the year of birth)
- Social Security Numbers that include characters other than numerical digits or that have more or less than 9 digits
- Values for sex other than the valid ones
- Values for race/ethnicity other than the valid ones
- Values for state or county federal information processing standards (FIPS) codes other than the valid ones
- Names that include numerical digits instead of letters

If a year is represented by only 2 digits, Y2K-type errors could result. Low values (e.g., '02') are ambiguous and could be misinterpreted as belonging to the wrong century (e.g., '1902' or '2002').

When discovered, erroneous values should be corrected if possible, at least in a temporary data file that will be used only for linkage (since it may be impractical to change the permanent data file soon enough to be available for linkage). If an erroneous value cannot be corrected, then it should be changed to an unknown value if the value consists of only a single digit or a single letter. However, if the value is more complex, consisting of multiple digits (e.g., Social Security Number) or multiple letters (e.g., name)—and not all of those digits or letters are known to be erroneous—then only the portion known to be erroneous should be changed to an unknown value, so that the portion that may be correct will remain for matching.

- Distinguishing Missing and Unknown Values:

  For most variables on which agreement will be assessed, an unknown value is equivalent to a missing value and should be neutral, being neither for nor against linkage to a record in which the value is not missing. For other variables, a missing value could characterize the true status of the individual. For example, some persons truly have no middle name or, if they are still alive, have no date of death. A missing value for such variables should not be neutral but should contribute weight for linkage (if the other record also has a missing value) or against linkage (if the other record has a nonmissing value). For such variables in which a missing value can make a difference, a missing value should be distinguished from an unknown value, which should be neutral. The appropriate way to distinguish missing and unknown values from one another may depend on the linkage computer program. For example, a missing value could be represented by '(blank)', '0', or 'M', whereas an unknown value could be represented by a '.', '9', or 'U'.

2. Standardization

The following applies both to standardization of different components of a single data file (e.g., for different years) and to different data files.

- Standardization of File Types

  To undergo linkage, data files must be of the same type (e.g., ASCII text, SAS, or MS Access). Otherwise, the type of one file must be converted to the type of the other. Record linkage could be done within SAS, wherein it would first read in or import data files of different types, perform the linkage, then save the data as either a temporary or a permanent SAS data set, and could output the final results in a variety of file types (e.g., ASCII, Excel, MS Access, Comma Delimited). Conversely, record linkage could be done outside of SAS if a SAS case file is converted to a file of another type (e.g., ASCII) for linkage with another file using an external linkage program.

- Standardization of Variable Attributes

  The attributes of supposedly equivalent variables must be the same across data files. This includes length in bytes and variable type (character or numeric). In SAS, for example, when two data sets are to undergo linkage, by merger—and there is a variable with the same name in each data set on which the merger will be based—if there are differences in the variable length and/or type between the data sets, SAS may fail to recognize the variable as being the same in each data set, resulting in a failed merger and some false nonlinks.

- Standardization of Complex Variable Component Sequences

  The components of a complex variable must be arranged in the same sequential order in both data files if the components are not explicitly considered separate variables themselves. For example, if the sequence for the components of a date consists of 1) month, 2) day, and 3) year in one data file, then it should be the same, and not 1) year, 2) month, and 3) day in the other data file. The easiest way to deal with different sequences across data files may be to define each variable component as a separate variable. Some linkage software packages (e.g., DataFlux's dfPower Studio, MatchWare's AutoStan, Ascential's Integrity) may have special utilities for standardization.

- Standardization of Variable Coding Schemes of Valid Values

  Each variable to be used for linkage or analysis should be coded in a single standardized way, both across records within each data file and across data files, so that a given value on one record will mean the same thing on another record. If variables are coded in different ways to represent the same information (or the same way to represent different information), results of analysis will be erroneous, and if the variables are used for linkage, the results of the linkage will be erroneous. For accurate and successful linkage, the coding scheme in one or both data files must be changed so that they agree. For examples, see "Appendix B."

3. Deduplication

Before linkage of one data file with another is attempted, each data file should be deduplicated. This means that records erroneously classified as being for two or more separate persons, when they actually refer to the same person, should be combined into a single record or grouped by being given a common identification number. They may remain as separate records (as may be appropriate for multiple laboratory reports) provided that they are recognized as being for a single person who is counted only once in case tabulations. Deduplication is equivalent to linkage of records within a single file, rather than between two files. The procedures listed previously (correction and standardization within the data file) should be done before deduplication, as they should be before any record linkage.

Record Linkage Strategies

1. Obstacles to Agreement

The fundamental principle of record linkage is that classification of records as linked depends on whether they agree on the values of one or more variables. Agreement on a uniquely identifying variable, such as a Social Security Number, a medical record number from a particular facility, or a death certificate identification number, may be sufficient to classify the records as linked. However, many important data files (including the HIV/AIDS case data file) may not contain such identification numbers as variables, or, if they do, the values for these unique identifiers may be missing or erroneous because of data-entry errors or misuse (e.g., by someone other than the person to whom they belong, such as the spouse). Therefore, whether records are classified as links is generally more complex, relying on several variables instead of a single highly specific variable. Some of the variables used, such as sex or race, may have a small number of possible values or categories that are common in the data files, and consequently these variables may be relatively nonspecific. Agreement on such nonspecific variables could easily occur by chance and would not necessarily mean that the records were links. Such variables are said to have a low discriminating power.

Other variables, such as name and date of birth, have a large number of possible values, each of which would be relatively uncommon, and, as a result, such variables have a high discriminating power. These variables may have data-entry errors, such as misspellings or transposition of the components of a date. All of these variables may occasionally have missing values, which should not necessarily be considered evidence against the records being linked.

2. Partial and Inexact Agreement

To reduce the number of false nonlinks (missed links), the criteria for agreement between a pair of records may be broadened to include partial or inexact agreements. In partial agreement, a minimum number (not necessarily all) of the linkage variables agree, or there is agreement on any of several alternative combinations of variables. Closely related is inexact or "fuzzy" agreement, which can pertain to a single complex variable, such as a name or multidigit number. The "algorithm" (set of logical rules for classification of records as links) may include inexact agreement on a variable in which there is agreement on a component of the variable (e.g., the first 3 letters of a name, or 6 of the 9 digits of the SSN) or on a less specific derivative of the variable (e.g., the Soundex code of a name). Another type of inexact agreement is one defined in terms of the "spelling distance," a quantification of the extent to which the spellings of a name differ. This can be calculated by computer programs such as the "SPEDIS" function in SAS. Other computer programs can recognize when dates differ by only a small amount (e.g., plus or minus a single day, plus or minus a year), or by a transposition of two of the three components of a date (day, month, year). Another type of inexact agreement is that achieved by interchanging variables, particularly first name and last name, first name and middle name, or maiden name and last name. Linkage based on inexact or partial agreements is generally more sensitive and less specific (resulting in acceptance of more links, both true and false) than linkage based on exact agreement, but the results depend also on the discriminating power of the linkage variables. For example, exact agreement on only sex, birth-year, and birth-month would not be very specific. Both deterministic and probabilistic linkage algorithms (see sections "Probabilistic Linkage" and "Deterministic Linkage") can accept inexact agreements. To find inexact agreements on complex variables such as names, two tactics that may be used are the creation of extra records with alternative but related names (e.g., alias, nickname, maiden name) and the derivation of new variables (e.g., Soundex codes) that are less specific than the actual names.

- Creation of "Extra" Nonduplicate Records

  If some records have alternative valid values for variables that cannot be standardized (e.g., aliases, nicknames, maiden names, Hispanic equivalents to English first names, multiple medical record numbers from different health care providers), or have erroneous reversals between first and last names, or have reversals between the two parts of a hyphenated surname, then extra records with predictable alternative values may be created for the same person for the purpose of record linkage. The extra records will not be duplicates (erroneously assumed to be for separate persons) because they will be recognized to be for a single person by having the same identification number (e.g., STATENO, eHARS ID). Submission of such extra records to search for links in the National Death Index is allowed without additional charge. Appendix 4A contains a list to help identify given names that may be equivalent to nicknames. Some record linkage software programs may allow

alternative values to be used, including transpositions of month and day of birth, without the need for creation of extra records.

- Derivation of Variables with Less Discriminating Power for More Sensitivity

  The values of some variables, such as names, may have high discriminatory power and specificity, but may not be sufficiently sensitive to detect links if errors in spelling are likely. To increase the chance of finding a link, alternative variables for linkage may be derived that have less discriminatory power and greater sensitivity than the original variables. Five examples follow:

  - Soundex code of the last name
  - Soundex code of the first name
  - Substring of the last name, consisting of the first 3 letters
  - Substring of the first name, consisting of the first 3 letters
  - Last 4 digits of the Social Security Number

3. Probabilistic Linkage

Linkage is probabilistic if the linkage algorithm takes into consideration the discriminating power of the linkage variables and the rarity of the variable values. In probabilistic linkage, agreement on the value of a variable such as name, with a high discriminating power (because it has many possible values), would count more than agreement on the value of a variable with a low discriminating power, such as sex (because it generally has only 2 possible values); and agreement on a rare value for a variable (e.g., a rare name, such as “Brockovitch”) would count more than agreement on a common value (e.g., the name “Smith”). The mathematical basis for probabilistic linkage, as it is generally understood today, involves estimations of the expected frequency in which particular variables agree among pairs of same-person records, as compared with the expected frequency of those agreements among pairs of different-person records. Frequencies of agreement estimated without regard to the particular values that variables may have are called “global” or “general” frequencies, and frequencies of agreement defined by the specific values of the variables are called “value-specific” frequencies. For a pair of records with a particular combination of agreements and disagreements of their variables, the algorithm calculates the ratio of the expected frequency of that combination among pairs of same-person records to the expected frequency of that combination among pairs of different-person records. This frequency ratio is correlated with the probability that the records refer to the same person. Until recently, estimating these expected frequencies required a preliminary linkage (perhaps done by a nonprobabilistic, or deterministic, method) supplemented by visual review [Section 12]) on a “calibration” sample of records before the probabilistic linkage could be done by computer. The preliminary linkage serves as a “gold standard,” assumed to be highly accurate and reliable, and its resulting links and nonlinks are substituted for same-person records and different-person records because the true status of the records cannot otherwise be

known. Newer computer programs use an iterative, self-correcting algorithm that allows probabilistic linkage to be done without a preliminary linkage to calibrate the expected frequencies. The frequency ratios then undergo a logarithmic transformation into "weights." The output from the probabilistic linkage is a ranking of pairs of records according to these weights. The user usually selects two threshold measures of weights. The upper threshold is to distinguish between pairs of records that will be classified as links without further investigation and those that will be considered indeterminate (pending additional investigation by visual review). The lower threshold is to distinguish between the indeterminate pairs of records and those that will be classified as nonlinks without further investigation. The selection of the thresholds depends on how sensitive or specific the user wants the linkage to be (how many false links or false nonlinks are tolerable).

4. Deterministic Linkage

Linkage is deterministic to the extent that it is not probabilistic. In deterministic linkage, the classification of records as links is based simply on whether they agree on the values of the variables used for record linkage, without considering the relative frequency of agreements among same-person records as compared with different-person records, and without considering the rarity of particular values. A refinement of deterministic linkage may include a procedure that has a superficial resemblance to probabilistic linkage, in which pairs of records are ranked with scores according to the relative probability that they are linked based on subjective estimates of the discriminating power of the variables on which they agree. Agreements on one set of variables may be assigned a higher score than agreements on an alternative set of variables. Record pairs can also be scored according to their extent of partial or inexact agreement, from a little agreement to complete or exact agreement. The scores are not derived from mathematical estimates of agreement frequency ratios. Nonetheless, threshold values for such scores may be selected to identify pairs of records that require further investigation by visual review, similar to how threshold values for probabilities are used in probabilistic linkage.

5. Pros and Cons of Probabilistic Linkage vs. Deterministic Linkage

- Deterministic linkage is easier to understand and to program than probabilistic linkage.
- Deterministic linkage can be done with a greater variety of software packages, including in-house programs written in SAS or PRODAS, and a linkage program that is built into HARS (utility "M" or "N").
- Deterministic linkage can be done within eHARS (e.g., using SAS or dfPower Studio), whereas probabilistic linkage will require exporting a file of cases with the linkage variables from eHARS to allow the linkage to be conducted externally, after which the results can be imported back into eHARS.

- More technical support may be available for deterministic linkage than for probabilistic linkage.
- Deterministic linkage may require fewer computer resources (e.g., memory, hard drive space) than probabilistic linkage.
- Probabilistic linkage has become the "state-of-the-art" or "gold standard" of modern methods of record linkage.
- Probabilistic linkage may reduce the number of indeterminate records that require visual review to be classified as either links or nonlinks.
- Both deterministic and probabilistic linkage can be used with small data files.
- Both can accept inexact ("fuzzy") agreements.
- Both can be done with software that can be made available to health departments at no cost (see "Considerations for Selecting a Computer Software Program for Record Linkage.")
- Both can automatically triage records into three classes (Table 1): 1) links, 2) nonlinks, 3) indeterminate.

6. "Blocking"

Blocking is the restriction of linkage to records that agree on selected variables (the "blocking variables"). It is generally used as a preliminary step before probabilistic linkage. It reduces the computer resources required for probabilistic linkage by eliminating consideration of possible links between records that do not agree on the values of the blocking variables. Since this elimination does not involve estimation of the expected frequencies of agreements on the blocking variables among linked and nonlinked pairs, blocking is a deterministic element within probabilistic linkage. An effective blocking variable should have very few erroneous values, very few missing values, and a discriminating power (number of possible values) that is high, but not so high that the linkage would be insensitive (resulting in too few true links). Two or three blocking variables are usually used in combination, such as sex in combination with the Soundex code of the last name. To increase the sensitivity of the linkage, several alternative blocking algorithms (or "passes") may be used. If a pair of records do not agree on the values of the combination of blocking variables used in the first pass, they are given another opportunity to find agreement on the combination of blocking variables used in the next pass. Agreement on any of the blocking combinations allows the pair of records to be considered further as possible links with respect to agreement on other variables. The blocking variables used in one pass should not include those used in another pass. For example, sex and the Soundex code of the last name might be used on the first pass, followed by the Soundex code of the first name in combination with the month and year of birth on the second pass. Usually only two or three passes are used. Using a larger number of alternative blocking passes may defeat the purpose of blocking.

7. Selecting the Best Compromise between Sensitivity and Specificity

- Definitions of Sensitivity and Specificity:

  The sensitivity of the linkage algorithm is its ability to detect same-person records, which can be expressed mathematically as the proportion of true links among all same-person records (Table 1) (assuming all same-person records could be known independently of the record linkage). The specificity of the linkage algorithm is its ability to detect different-person records, which can be expressed mathematically as the proportion of true nonlinks among all different-person records. The false link rate would be the number of false links among all links, and the false nonlink rate would be the number of false nonlinks among all nonlinks. These concepts are similar to those of the sensitivity, specificity, false-positive rate, and false-negative rate of diagnostic laboratory tests.

- Conflict between Sensitivity and Specificity:

  Ideally, both the sensitivity and specificity should be made as high as possible, but there is often a conflict between them so that something that increases the sensitivity is likely to reduce the specificity, and vice versa. It may be necessary to decide whether sensitivity or specificity is the more important of the two. This situation arises when deciding whether to accept partial/inexact agreements or only complete and exact agreements as criteria for linkage. Accepting partial or inexact agreements will increase the sensitivity but reduce the specificity. Another situation in which a choice must be made between sensitivity and specificity is during selection of threshold weights or scores. If there were no indeterminate records, raising or lowering the threshold separating links from nonlinks would obviously reduce sensitivity whenever it increased specificity, and vice versa. With indeterminate records in the middle, however, the situation is more complicated because there are two thresholds: the upper one separating links from indeterminate records and the lower one separating indeterminate records from nonlinks. Raising the upper threshold will reduce the sensitivity and the false link rate by reducing the number of links, and will increase the number of indeterminate records that may need to be visually reviewed. Dropping the lower threshold will reduce the specificity and the false nonlink rate by reducing the number of nonlinks and will increase the number of indeterminate records.

- When Avoiding False Links May Be More Important Than Sensitivity:

  When linkage is done to update (collect new information on) existing cases, avoiding false links may be more important than sensitivity (finding as many links as possible) because it is important that ascertainment of the accepted information is accurate, particularly for critical events such as death. The threshold for acceptance as a link should be high in this circumstance.

- When Avoiding False Nonlinks May Be More Important Than Specificity:

  Depending on the purpose of the linkage, avoiding false nonlinks may be more important than specificity (finding as many nonlinks as possible). At first glance, it may not seem particularly desirable to find nonlinks, but they actually are an important way to find new cases. For example, some external data files (e.g., HIV antibody test results) are useful primarily to find new cases among the records classified as nonlinks (Figures 1 and 2). In this circumstance, it may be more important to be sure that an antibody test record is a different-person record because nonlinks will require the expenditure of resources to conduct a follow-up investigation to collect all the data needed on new cases. In contrast, the antibody test records classified as links or indeterminate may be ignored because they provide no new information. False nonlinks either will be discovered actually to be links as a result of the follow-up investigation or will become duplicates (erroneously counted as new separate cases).

  A situation in which sensitivity may be more important than specificity is a linkage done for a special study in which the ability to examine general patterns and overall trends may be more important than the accuracy of every record. This view would be based on the assumption that including a small proportion of false links would not change the analytical results significantly, and that increasing the sensitivity to capture more links would be necessary to obtain a sufficient number of observations for the results to be significant.

- When Both Sensitivity and Specificity Are Important:

  If a single external data file is to be used for the dual purposes of both finding new cases and updating existing cases (Figures 1 and 2), both the specificity and the sensitivity are important. A reasonable compromise may be reached between sensitivity and specificity by appropriately adjusting the linkage criteria (either the algorithms or the thresholds of the scores or weights used to classify records into links, nonlinks, and indeterminate records). The upper threshold should be lowered to increase sensitivity, and the lower threshold should be raised to increase specificity. An alternative to compromising on a single algorithm or threshold setting would be to use different linkage criteria for each purpose. This could be achieved by adjusting the thresholds differently to reclassify the results from a single linkage; or, if adjusting thresholds two different ways is not possible for a single linkage, the linkage may be conducted twice on the same data files, each time using different linkage criteria.

8. External Files Containing Some Records That Mention HIV and Others That Do Not

When a purpose of the linkage is to find new information about existing cases, and the external data file consists of a mixture of records that mention HIV and other records that do not mention HIV (e.g., death certificate records, hospital discharge records), all records in the external data file—not just those that mention HIV—should be included in the linkage to allow the records on which the values of the linkage variables most closely agree to be found, regardless of whether they mention HIV. Thus, a hospital record or death certificate linked to a case could be found even if it does not mention HIV. However, a linkage may have already been conducted with the subset of the external data file limited to records that mention HIV. This would be likely to happen when death certificate records that mention HIV are brought to the attention of HIV/AIDS surveillance staff before the complete death certificate data file for the year is available. In this circumstance, ideally the HIV/AIDS cases for which links were found in the HIV-specific external file should be submitted for linkage again with the entire external file to provide an opportunity for better (more closely agreeing) links to be found among the records that did not mention HIV. If links that agree more closely are found in the larger, mixed external data file used in the second linkage, then they should replace the links previously found in the smaller, HIV-specific external data file (which is a subset of the larger file). If a previous linkage with the subset of records that mention HIV has not already been done to allow follow-up of the nonlinks as potential new cases, then such a follow-up should be applied to the HIV-specific subset of nonlinks found in the entire external file (Figures 1 and 2).

Considerations for Selecting a Computer Software Program for Record Linkage

The following should be considered when choosing a computer program for record linkage:

- How much would it cost? Is it already owned by the health department, or is it provided by CDC without cost to the health department?
- How easy is it to use? What kind of training is available?
- Is the vendor reliable? Can the vendor provide adequate technical support for the life of the software, or does CDC offer technical assistance?
- What operating system and computer resources (memory, hard drive space) are required?
- What is the maximum number of records in the combined files that the software could handle?

  For example, if there were 10,000 HIV/AIDS case records and 2,000,000 Medicaid records, could the software process that many records? How long would it take—hours or weeks?

- Can it accept partial and inexact agreements ("fuzzy matching")? This is a definite advantage.
- What is the maximum number of linkage variables?
- How does the software handle missing variables?
- Can it do both automatic (batch, many-at-a-time) linkage and interactive (one-at-a-time) linkage?

  These policies and procedures focus mainly on automatic linkage of one electronic data file with another, in which the linkage is done for all records at once. An alternative to this is interactive linkage, in which the user enters the data for one record at a time, perhaps key-punching them in at that moment from hard copy records. Interactive linkage may be practical for only a small number of records, which could be a single record, in the file initiating the linkage. This could be the situation, for example, when doing linkage with copies of death certificate records received weekly from the office of vital records to search for links in the HIV/AIDS case surveillance database. Linkage software should allow for both automatic and interactive linkage to be done.

- Is it deterministic or probabilistic?

  See "Pros and Cons of Probabilistic Linkage vs. Deterministic Linkage" for the pros and cons between deterministic and probabilistic linkage. Theoretically, because probabilistic linkage would be more accurate, it is preferred if it can be done easily and inexpensively.

In addition to the record linkage programs that are available at no charge from CDC, many record linkage programs can be purchased. Particular programs that may be considered include, but are not limited to, the following:

- Deterministic Programs:
  - Utilities "M" and "N" in HARS.

    This program relies heavily on names, so it cannot be used easily where names are not retained (e.g., surveillance programs that use coded identifiers instead of names). This program will be phased out as HARS is replaced by eHARS. CDC's Division of HIV/AIDS Prevention (DHAP) provides technical support for using this program.

  - DataFlux's dfPower Studio:

    This program will be included in eHARS, but it is not probabilistic. The 2004 version uses the term "sensitivity" in a sense opposite to the way it is used in this document (setting the "sensitivity" high would make its algorithm less sensitive). DHAP will provide or support training in the use of this program.

- Probabilistic Programs:
  - LinkPlus:

    This program was developed with CDC funds for use by cancer registries, but can be used also by any other registry, such as the HIV/AIDS surveillance case data file. It can be downloaded for free from http://www.cdc.gov/cancer/npcr/tools/registryplus/lp_tech_info.htm.

    This program should not be confused with another program with the same name ("LinkPlus") made by a commercial company for another purpose. To search for the correct LinkPlus on the Internet, the exact phrase "Registry Plus" should be included. Technical support may be obtained from CancerInfo@cdc.gov or possibly from your state cancer registry.
  - QualityStage or its antecedents (AutoMatch, SuperMatch, Integrity):

    AutoMatch became famous in the 1980s and 1990s, but the manufacturer was bought out at least twice, and the program underwent a name change with each buyout. Currently, the manufacturer is Ascential, Inc., which produces QualityStage, but formerly called it Integrity. Several health departments already own versions of this program.

Visual Review of Indeterminate Records

Visual review (also called "manual," "clerical," or "interactive" review) involves individual, one-at-a-time inspection of either electronic records on a computer screen or hard copy paper records. The alternative to visual review is automatic linkage done by a computer, which can batch-process many records at the same time. Visual review used to be the only way that partial or inexact agreements could be recognized. Now sophisticated linkage programs have reduced the need for visual review by including algorithms based on partial or inexact agreements. However, some records are still classified as indeterminate. Although mathematical formulas are not used, consideration of the rarity of values during visual comparison of records is conceptually probabilistic, even if it is used after deterministic linkage to evaluate records of indeterminate status. The human user may also be able to resolve some indeterminate records by recognizing equivalent values that a computer program may miss. For example, the Soundex coding system used in HIV/AIDS surveillance provides a way to detect inexact agreement on surnames by detecting similarly sounding names, but it fails to recognize the similarity of "Rodgers" and "Rogers," "Pitcher" and "Picher," or "Carlson" and "Karlson." One useful capability of some record linkage computer programs that is applicable especially to interactive visual review is the ability to search for a selected fragment of a name. Familiarity with names may allow the user to select the fragment least likely to be misspelled. Visual review may be necessary to recognize equivalents between nicknames and proper given names, particularly if they differ in their first 3 letters (e.g., Peggy = Margaret, Trudy = Gertrude). To assist in visual review, Appendixes 4 and 4-A provide advice and a table of English proper-name equivalents to common nicknames.

Multiple medical record numbers from different health care providers could also be taken into account at this stage, allowing agreement on one of them to be accepted as sufficient for classification as links. Humans can also recognize correlations between what the computer may consider independent variables. For example, the location of the hospital where the person received care may be recognized as being geographically close to the person's residence, despite having a different ZIP code. The criteria for classification as links or nonlinks used in visual review may be somewhat subjective. Therefore, it is suggested that linkage by visual review be repeated by a different staff member, if practical. Records on which the other staff member reaches a different conclusion should either remain classified as indeterminate or be classified as nonlinks if the linkage is intended to have a very low false link rate, or be classified as links if the linkage is intended to err on the side of sensitivity.

Tracking Log

The log could be a spreadsheet or an MS Access database, or there may be another way to flag and track records in eHARS or HARS, perhaps using local data fields. The following purposes show the utility of a tracking log:

- It may be helpful to use a log to track the follow-up (or follow-back) investigations of potential new cases, regardless of whether they are identified as records in an HIV-specific external data file for which no matching existing cases were found or they were simply incompletely reported by a health care provider.
- A tracking log may also be useful for keeping track of linkages of new records (e.g., death certificate data) to previously reported cases. The reasons for classifying a new record as a link or for changing a previous classification from link to nonlink may be recorded. For example, a classification of a death certificate as a link may be reversed if a later record (e.g., laboratory test result) is classified as a link and accepted as evidence that the previous linkage was mistaken (because the person must still have been alive to have had the laboratory test done).
- Another possible use would be to track details of the outcome of record linkage, such as which variables or variable components of incoming records were the ones that agreed with those of existing records. This information could then be considered in fine-tuning the criteria used for classifying records as links or nonlinks.

Measures or Standards for Record Linkage

The quality of record linkage conducted by an HIV/AIDS surveillance program may be evaluated or measured in several ways.

1. **Process Measures**:

   - Frequency:

     The optimal frequency will depend on program need and how frequently an external data file becomes available or how frequently it is updated. Data on the underlying cause of death in the state death certificate data file, for example, generally become available annually, but data on laboratory test results probably will become available at least monthly.

   - Choice of External Data Files:

     Some external files may be considered essential for HIV/AIDS surveillance programs to include among those they choose for record linkage, whereas others may be considered optional. Among the essential ones would be the state death certificate data file. If the state has a state-wide hospital discharge data file that includes patient identifiers, that too should probably be considered essential. Depending on availability, electronic laboratory data files may also be considered essential. Other data files, such as the cancer registry or the tuberculosis registry, are likely to be less useful.

   - Use HIV-specific Records for Two Purposes:

     If an external data file or a subset of it is HIV-specific, it should be used for two purposes: 1) to find new cases among the HIV-specific nonlinks and 2) to find new information about existing cases in the links.

   - Also Use Records That Do Not Mention HIV to Find New Information on Existing Cases:

     If the first linkage with an external file was limited to an HIV-specific subset of a larger file that included records that did not mention HIV, it must be followed by a linkage with the entire external file or its HIV-nonspecific subset to allow the cases for which links were not found in the first linkage to have an opportunity for links to be found among the records that do not mention HIV. If the first linkage could be with the entire external file, the second linkage would be unnecessary.

   - Repeat Linkage on the Entire File if Previously Done Only on an HIV-specific Subset:

     To find the most closely agreeing links in the external file, repeat the linkage using both the entire external file and all cases involved in a previous linkage limited to an HIV-specific subset (including cases for which links were found in the first linkage). This may result in reclassification of an HIV-specific record from link to nonlink if a more closely agreeing link is found among the records that do not mention HIV. Repetition of linkage with cases for which

links were found in the first linkage may be omitted, however, if prior experience shows that the number of reclassifications resulting from it would be very small.

2. **Outcome Measures**:

Appropriate outcome measures may depend on the frequency of linkage and the number of years that linkage has been conducted.

- Yield of New Information on Existing Cases:

  The success or usefulness of record linkage may be measured in terms of the number and proportion of existing cases for which essential new information (e.g., progression to AIDS, death) was found by the linkage.

- Yield of New Cases:

  Another measure of success is the number of new cases found by linkage with an HIV-specific data file. The yield could be expressed also as the proportion of new cases found by the linkage among total new cases found by all methods for a given year. However, a low proportion may not necessarily mean that the linkage is working poorly; it could instead mean that the other methods of finding cases are working very well.

- Evaluate Automatic Linkage by Comparison with Visual-Review Linkage:

  Because a human can recognize equivalences that a computer may not, linkage by visual review is likely to be more sensitive in detecting inexact agreements than automatic linkage. Therefore, until the quality of the automatic linkage is well established, a reasonable sample of records should be processed both automatically and visually to assess the sensitivity of the automatic linkage. As a standard, the automatic linkage should classify as links at least 95% of the records classified as links by visual review.

Appendixes

Appendix A Prototype Example of a Memorandum of Understanding for Obtaining Access to a Data File

Appendix B Examples of Standardization of Incompatible Variable Coding Schemes

Appendix C Example of How to Standardize Incompatible Coding Systems for Race/Ethnicity

Appendix D Proper Name Equivalents to Nicknames

Appendix A

Prototype Example of a Memorandum of Understanding for Obtaining Access to a Data File

[*This is a form, usually supplied by the data owner to the data recipient, on which the recipient states his/her understanding of the conditions for his/her use of the data. The phrases enclosed by "<" and ">" symbols ("<for example>") are to be replaced by whatever is appropriate for the particular data file, data owner, data recipient, and specific circumstances.*]

The conditions for release of the <Data File Name> data files for <the year> by <the Data Owner or Custodian> are as follows:

A. **List of Conditions**: I, <name of data recipient>, representing <the name of the state or local HIV/AIDS surveillance program>, agree to observe the following conditions of use of the <name of the data file> data file released to me or information derived from this file:

   1. The file will be used only for the following purposes: For public health surveillance of cases of human immunodeficiency virus (HIV) infection and/or acquired immunodeficiency syndrome (AIDS) by the <state and/or local department of public health> as <authorized or mandated> by <public health law or regulation>. The information on HIV/AIDS-related disease and death in <the jurisdiction of the health department> that will result from these data is needed to set public health policy and to allocate funding for programs for prevention of HIV infection and AIDS and for the care of HIV-infected persons. These data may also be used to evaluate and improve the completeness and quality of HIV/AIDS surveillance conducted by <the health department>.
   2. <Describe here the state/local laws/regulations under which the confidentiality of the data will be protected. Also describe, if applicable, how data subsets or printouts listing individuals would be destroyed following their use. >
   3. <Name of data recipient> will be designated as custodian of this data file, and will be responsible for observance of all conditions of use, and for establishment and maintenance of security arrangements to prevent unauthorized use. If the custodianship is transferred within the organization, <the data owner> will be notified promptly. This individual must have the legal authority to keep the information confidential and maintain the confidentiality.
   4. No listing of information from individual records and no statistical tabulation will be published or otherwise publicly released in a manner that would allow individual persons or their health care providers to be identified. Subject to this limitation, statistical summaries of data from this file may be published or released.

B. Understanding of Penalties for Violation of the Previously Listed Conditions:

I understand that any violation of the previously listed conditions may result in prosecution under all relevant state and federal laws.

<Signature of data recipient or representative of the health department> <Date>
Representative of <the state/local HIV/AIDS surveillance program> Date

<Signature of representative of data owner or custodian> <Date>
Representative of <data owner or custodian> Date

Appendix B

Examples of Standardization of Incompatible Variable Coding Schemes

Soundex Code

The algorithm for determining the Soundex code used for HIV/AIDS cases may differ from that used for a variable with a similar name or a similar function in other databases. Therefore, the Soundex code for records in the external database should be newly derived from the surname by applying the same Soundex program used for HIV/AIDS cases.

Sex

For “sex,” valid values in one data file may be '1' for male and '2' for female, whereas in the other data file the corresponding values may be '1' and '0' or 'M' and 'F.' The values in one file must be changed to conform with those in the other file.

Race/Ethnicity

For race/ethnicity, valid values in one data file may differ from those in another. They may be able to be mapped directly to one another, or it may be necessary to create a new variable to which values from both files may be mapped. Appendix 3 has an example of how this can be done.

Dates

- If a month can be recorded as the full name (e.g., “November”), an abbreviation (e.g., “Nov.”), or a number (e.g., “11”), then only one of these styles should be used.
- If a year consists of 4 digits in one data file but only 2 digits in the other data file, one must be converted to the same form as the other. Conversion to 4 digits must avoid Y2K-type errors (mistaking a 20th century year for a 21st century year, or vice versa). If the 2-digit year has a value greater than what would correspond to the current year, then it probably refers to the previous century. If it has a smaller value (e.g., '03'), independent information (e.g., on age or pediatric information) may be needed to decide the correct century.

Names of Persons and Places

- If a person's name may or may not include titles (e.g., Mr., Dr.), suffixes (e.g., Jr., III, MD, Ph.D., Esq.), or punctuation (e.g., an apostrophe, a hyphen, or a space),

these either should always be included or (preferably) always omitted. Special software programs for standardizing names generally remove titles, prefixes, suffixes, and punctuation marks.

- If the name of a state, city, county, street, or health care facility can be recorded either using full words or abbreviations, then only one of these styles should be used. If abbreviations are used, it should be the same standard abbreviation across records and across files. Examples of inconsistent abbreviations:
  - "Michigan" may be abbreviated "MI" on one record and "Mich." on another.
  - "Broadway" may be recorded as "B'way" on one record and "Bdwy" on another.

  Where the names of a county, borough, and city refer to the same geographic area and are used interchangeably, a single standard way should be used to record the city or county. For example, "Brooklyn," "Kings County," and "New York City" should not all be used as the value for the same variable. Use of FIPS codes may be a way to avoid such inconsistencies.

Appendix C

Example of How to Standardize Incompatible Coding Systems for Race/Ethnicity

Within HIV/AIDS case data, race/ethnicity is currently summarized into these categories:

- NEW_RACE
  1) Hispanic (all races)
  2) non-Hispanic American Indian/Alaska Native
  3) non-Hispanic Asian
  4) non-Hispanic black or African American
  5) non-Hispanic Native Hawaiian/Pacific Islander
  6) non-Hispanic white
  7) non-Hispanic legacy Asian/Pacific Islander
  8) non-Hispanic multiple race
  9) Unknown

Another data file (e.g., death certificate data) may have a race variable with these values:

- RACE
  1) white
  2) black
  3) other

and an ethnicity variable with these values:

- ETHNICITY
  0) non-Hispanic
  1) Mexican
  2) Puerto Rican
  3) Cuban
  4) Central or South American Hispanic
  5) Other or unknown type of Hispanic
  9) Unknown if Hispanic

From these two variables (race and ethnicity) in the other data file, a new variable (race/ethnicity) that combines both race and ethnicity must be created that will be the same in both data files. In the HIV/AIDS case data file, categories 1 (Hispanic), 4 (non-Hispanic black), and 6 (non-Hispanic white) may remain as they are, and categories 2, 3, 5, 7, and 8 may be combined to form the category of "other non-Hispanic," except that they would be renumbered as follows:

RACE_ETH

1) Hispanic
2) non-Hispanic white
3) non-Hispanic black
4) Other non-Hispanic
9) Unknown

The SAS code for doing this in the HIV/AIDS case data file could be as follows:

```
If NEW_RACE = 1 then RACE_ETH = 1;

If NEW_RACE = 6 then RACE_ETH = 2;

If NEW_RACE = 4 then RACE_ETH = 3;

If NEW_RACE in (2, 3, 5, 7, 8) then RACE_ETH = 4;

If NEW_RACE = 9 then RACE_ETH = 9.
```

In the other data file, a compatible race/ethnicity variable can be created by defining its values using the following SAS code:

```
If 1 = ETHNICITY = 5 then RACE_ETH = 1; *Hispanic;

If RACE = 1 and ETHNICITY = 0 then RACE_ETH = 2; *non-Hispanic white;

If RACE = 2 and ETHNICITY = 0 then RACE_ETH = 3; *non-Hispanic black;

If RACE = 3 and ETHNICITY = 0 then RACE_ETH = 4; *other non-Hispanic;

If ETHNICITY = 9 then RACE_ETH = 9; *unknown if Hispanic.
```

Appendix D

Proper Name Equivalents to Nicknames

The following information, provided by Christine S. Cox of the National Center for Health Statistics (NCHS), was based on names in the National Death Index (NDI) for persons who died during 1979–2000.

The accompanying spreadsheet (Appendix D.1) is the table for converting nicknames to proper names, which has been used at NCHS when conducting record linkage involving nonproper names collected in survey data.

The variable NCOUNT is the number of times that a name listed as a "nickname" actually appears in the NDI data between 1979 and 2000, and PCOUNT is the frequency for the proper names. In cases where the nickname count exceeds the proper name count, a user of this table might decide that the nickname is actually a proper name and not create a second submission record for linkage. For example, "Annie" has a higher frequency count than either the more proper "Ann" or "Anne" (i.e., more decedents named Annie than Ann or Anne). In this case, the user may decide to simply submit "Annie" as a proper name.

NICKNAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
ANNIE	F	ANN	135373	87136
ANNIE	F	ANNE	135373	72562

Additionally, the user may choose only to substitute the most common proper name for a given nickname, which can be found by examining PCOUNT within a given nickname group. For example, "Albert" is the most common proper name available for converting "Al" to a proper name. Because the NCOUNT for "Al" (n= 2423) is less than all the proper names that "Al" is a known nickname for, the user may choose to substitute all listed proper names for a subject named "Al."

NICKNAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
AL	M	ALBERT	2423	201565
AL	M	ALFRED	2423	102688
AL	M	ALEXANDER	2423	42055
AL	M	ALAN	2423	17340

Appendix D.1

List of proper given name equivalents to nicknames

This list was provided by Christine S. Cox of the National Center for Health Statistics, based on names in the National Death Index for decedents who died during 1979–2000. Numbers for names found among fewer than 500 decedents (e.g., Allison) are not shown.

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
ABBIE	F	ABIGAIL	4047	1958
ABBY	F	ABIGAIL	710	1958
ABE	M	ABRAHAM	4700	21421
ABE	M	ABEL	4700	2896
ABE	M	ABRAM	4700	2483
ABRAM	M	ABRAHAM	2483	21421
ADA	F	ADELINE	49440	17423
ADA	F	ADELAIDE	49440	9608
ADA	F	ADALINE	49440	1200
ADDIE	F	ADELAIDE	20432	9608
ADELE	F	ADELAIDE	15589	9608
ADOLPH	M	ADOLPHUS	14951	1711
AGGIE	F	AGATHA	518	4576
AGGIE	F	AGNES	518	105135
AL	M	ALBERT	2423	201565
AL	M	ALFRED	2423	102688
AL	M	ALEXANDER	2423	42055
AL	M	ALAN	2423	17340
ALBERT	M	ELBERT	201565	14656
ALBERT	M	ADELBERT	201565	1558
ALEC	M	ALEXANDER	1065	42055
ALEX	M	ALEXANDER	20599	42055

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
ALF	M	ALFRED	941	102688
ALLIE	F	ALBERTA	10842	28926
ALLIE	F	ALICIA	10842	6811
ALLIE	F	ALLISON	10842	
ANA	F	ANASTASIA	12870	4528
ANDY	M	ANDREW	4848	98506
ANGIE	F	ANGELA	3452	28195
ANN	F	SUSANNA	87136	3075
ANN	F	ANNE	87136	72562
ANNA	F	SUSANNA	365939	3075
ANNA	F	HANNAH	365939	17138
ANNIE	F	ANN	135373	87136
ANNIE	F	ANNE	135373	72562
ARCHIE	M	ARCHIBALD	22186	2587
ARNIE	M	ARNOLD	596	35249
ART	M	ARTHUR	528	215065
AUGUST	M	AUGUSTUS	19435	4266
AUGUST	M	AUGUSTINE	19435	3042
BARNEY	M	BERNARD	6592	69999
BART	M	BARTHOLOMEW	1103	972
BECKIE	F	REBECCA	506	35077
BECKY	F	REBECCA	1429	35077
BELL	F	ISABEL	672	24765
BELL	F	ISABELLA	672	7242
BELL	F	BELINDA	672	1730
BELLA	F	ISABEL	5024	24765
BELLA	F	ISABELLE	5024	16930
BELLA	F	ISABELLA	5024	7242
BELLE	F	ISABEL	9130	24765

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
BELLE	F	ISABELLE	9130	16930
BELLE	F	ISABELLA	9130	7242
BEN	M	BENJAMIN	21469	72739
BEN	M	BENEDICT	21469	2729
BENNIE	M	BENJAMIN	14845	72739
BENNIE	M	BENEDICT	14845	2729
BERNIE	M	BERNARD	1812	69999
BERT	M	ALBERT	12983	201565
BERT	M	HERBERT	12983	104783
BERT	M	GILBERT	12983	34484
BERTIE	F	BERTHA	8426	130903
BERTIE	F	ALBERTA	8426	28926
BERTIE	F	ROBERTA	8426	23401
BESS	F	ELIZABETH	6075	305833
BESSIE	F	ELIZABETH	108497	305833
BETH	F	ELIZABETH	4278	305833
BETSY	F	ELIZABETH	2794	305833
BETTE	F	ELIZABETH	5194	305833
BETTY	F	ELIZABETH	159289	305833
BILL	M	WILLIAM	11189	1028647
BILL	M	WILLIS	11189	21183
BILLIE	F	WILHELMINA	13465	7539
BILLY	M	WILLIAM	47028	1028647
BIRDIE	F	BERTHA	7106	130903
BIRDIE	F	ROBERTA	7106	23401
BOB	M	ROBERT	4362	830576
BOBBIE	F	BARBARA	8032	125194
BOBBIE	F	ROBERTA	8032	23401
BOBBY	M	ROBERT	31721	830576

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
BONNIE	F	BARBARA	29505	125194
BRAD	M	BRADFORD	1351	2184
BRYAN	M	BRIAN	8001	24209
BUCK	M	CHARLES	929	603031
CARL	M	CHARLES	162018	603031
CAROL	F	CAROLINE	43517	33486
CARRIE	F	CAROLINE	59136	33486
CASSIE	F	CASSANDRA	4245	2176
CATHY	F	CATHERINE	2958	156439
CATHY	F	CATHLEEN	2958	1243
CELIA	F	CECILIA	19800	16597
CELIA	F	CELESTE	19800	3495
CHAD	M	CHARLES	3680	603031
CHARLIE	M	CHARLES	48014	603031
CHRIS	M	CHRISTOPHER	5759	37619
CHRIS	M	CHRISTIAN	5759	8889
CHRIS	M	KRISTOPHER	5759	708
CHRIS	F	CHRISTINE		40151
CHRISTY	F	CHRISTINE	1282	40151
CINDY	F	CYNTHIA	3019	19638
CINDY	F	LUCINDA	3019	5785
CLAIR	M	CLARENCE	5930	161844
CLAIRE	F	CLARA	17895	139402
CLARA	F	CLARISSA	139402	2101
CLARE	M	CLARENCE	2217	161844
CLARE	F	CLARISSA	4497	2101
CLEM	M	CLEMENT	1904	6860
CLIFF	M	CLIFFORD	853	63287
CLIFF	M	CLIFTON	853	17657

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
CONNIE	F	CONSTANCE	11439	19673
CORA	F	CORINNE	53142	7394
CURT	M	CURTIS	1096	34582
DAISY	F	MARGARET	35113	436797
DAISY	F	MARGARETTA	35113	1177
DAN	M	DANIEL	10806	120224
DANNY	M	DANIEL	12432	120224
DAVE	M	DAVID	5223	258920
DEB	F	DEBORAH		18713
DEBBIE	F	DEBORAH	1999	18713
DEBBIE	F	DEBORA	1999	575
DEE	F	DOROTHY	777	385229
DEE	F	AUDREY	777	33825
DEE	F	DELORES	777	17016
DELIA	F	CORDELIA	10964	3304
DELIA	F	ADELIA	10964	2227
DELL	F	ADELINE	693	17423
DELL	F	DELORES	693	17016
DELLA	F	DELORES	29716	17016
DELLA	F	ADELAIDE	29716	9608
DENNY	M	DENNIS	941	43985
DICK	M	RICHARD	2776	339756
DINA	F	GERALDINE	1680	41958
DOLLY	F	DOROTHY	5417	385229
DOM	M	DOMINIC		9987
DOM	M	DOMENIC		
DON	M	DONALD	16846	253754
DONNIE	M	DONALD		253754
DORA	F	DOROTHY	48803	385229

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
DORA	F	DORIS	48803	118677
DREW	M	ANDREW	1216	98506
ED	M	EDWARD	4568	364607
ED	M	EDWIN	4568	71736
ED	M	EDGAR	4568	42997
ED	M	EDMUND	4568	20169
EDDIE	M	EDWARD	32280	364607
EDDIE	M	EDWIN	32280	71736
EDDIE	M	EDGAR	32280	42997
EDDIE	M	EDMUND	32280	20169
EDDIE	F	EDWINA	4625	3013
EDDY	M	EDWARD	789	364607
EDDY	M	EDWIN	789	71736
EDDY	M	EDGAR	789	42997
EDDY	M	EDMUND	789	20169
EDIE	F	EDITH		139553
EDYE	F	EDITH		139553
EFFIE	F	EUPHEMIA	28277	713
ELAINE	F	ELEANOR	23991	92619
ELI	M	ELIJAH	6093	7792
ELI	M	ELIAS	6093	5863
ELI	M	ELISHA	6093	2010
ELIZA	F	ELIZABETH	10840	305833
ELLA	F	HELEN	78908	489417
ELLA	F	ELEANOR	78908	92619
ELLA	F	LUELLA	78908	14972
ELLA	F	LUELLA	78908	14972
ELLA	F	HELENE	78908	11856
ELLA	F	GABRIELLE	78908	1692

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
ELLA	F	ELENORA	78908	1451
ELLA	F	GABRIELLA	78908	753
ELLEN	F	ELEANOR	60201	92619
ELLIE	F	ELEANOR	2421	92619
ELLIE	F	LUELLA	2421	14972
ELLIE	F	ELENORA	2421	1451
ELLIS	M	ELISHA	13177	2010
ELLY	F	ELMIRA	687	1973
ELSIE	F	ELIZABETH	108795	305833
ELSIE	F	ALICE	108795	213847
ELSIE	F	ALICIA	108795	6811
ELVIE	F	ELVIRA	1328	12222
EMANUEL	M	MANUEL	8587	47320
EMELINE	F	EMILY	995	46545
EMILY	F	AMELIA	46545	27518
EMILY	F	EMELINE	46545	995
EMMA	F	EMILY	138554	46545
EMMA	F	EMELINE	138554	995
EMMIE	F	EMMA	1589	138554
EMMY	F	EMILY	1086	46545
EMMY	F	EMELINE	1086	995
ERIC	M	DERRICK	23958	4063
ERMA	F	EMMA	23907	138554
ERNA	F	ERNESTINE	9873	12396
ERNIE	M	ERNEST	1686	128779
ERNIE	M	EARNEST	1686	13162
ESSIE	F	ESTHER	19261	126512
ESTHER	F	HESTER	126512	8027
ETTA	F	LORETTA	20582	37613

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
ETTA	F	HENRIETTA	20582	26406
EVA	F	EVALINE	106182	608
EVE	F	EVELYN	2868	155563
EVE	F	GENEVIEVE	2868	39727
FANNIE	F	FRANCES	48811	220769
FANNY	F	FRANCES	5604	220769
FATE	M	LAFAYETTE	508	1497
FAY	F	FAITH	10785	4034
FLO	F	FLORENCE	840	210225
FLORA	F	FLORENCE	35700	210225
FLORRIE	F	FLORENCE	995	210225
FLOSSIE	F	FLORENCE	14923	210225
FORD	M	CLIFFORD	1659	63287
FORD	M	BRADFORD	1659	2184
FRAN	F	FRANCES		220769
FRAN	F	FRANCIS		113028
FRAN	F	FRANCINE		
FRANCIE	F	FRANCINE		
FRANK	M	FRANCIS	334255	113028
FRANK	M	FRANKLIN	334255	29851
FRANKIE	F	FRANCES	7238	220769
FRANKIE	M	FRANCIS	2546	113028
FRANKIE	M	FRANKLIN	2546	29851
FRANNIE	F	FRANCES		220769
FRANNIE	F	FRANCINE		
FRED	M	ALFRED	128645	102688
FRED	M	FREDERICK	128645	87355
FRED	M	FERDINAND	128645	5981
FREDA	F	ALFREDA	19661	2797

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
FREDDIE	M	FREDERICK	11048	87355
FREDDIE	F	WINIFRED	2326	23711
FREDDY	M	FREDERICK	1242	87355
FREDRIC	M	FREDERICK	1135	87355
FRIEDA	F	ALFREDA	22111	2797
FRITZ	M	FREDERICK	3468	87355
FRONA	F	SOPHRONIA	588	757
GAIL	F	ABIGAIL	8339	1958
GERRY	M	GERALD	811	77241
GERTIE	F	GERTRUDE	4825	130609
GIL	M	GILBERT	671	34484
GINA	F	REGINA	2680	22504
GREG	M	GREGORY	1014	27406
GRETA	F	MARGARET	3214	436797
GRETA	F	MARGARETHA	3214	727
GRETTA	F	MARGARET	740	436797
GRETTA	F	MARGARETTA	740	1177
GUS	M	AUGUSTUS	6375	4266
GUS	M	AUGUSTINE	6375	3042
GUSSIE	F	AUGUSTA	9001	12633
GUSSIE	F	AUGUSTINA	9001	891
GWEN	F	GWENDOLYN	1606	12582
HAL	M	HENRY	3148	246278
HAL	M	HAROLD	3148	229116
HAL	M	HOWARD	3148	127815
HATTIE	F	HARRIET	49386	35155
HATTIE	F	HANNAH	49386	17138
HATTIE	F	HARRIETT	49386	9409
HETTIE	F	ESTHER	4626	126512

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
HETTIE	F	HENRIETTA	4626	26406
HETTIE	F	HESTER	4626	8027
HOSIE	M	HOSEA	520	1499
HUGH	M	HUBERT	31190	27596
HUGO	M	HUBERT	6360	27596
IAN	M	JOHN	2215	1204477
IKE	M	ISAAC	1632	18984
INA	F	CHRISTINE	16447	40151
INA	F	AUGUSTA	16447	12633
INA	F	LAVINA	16447	2619
INA	F	LAVINA	16447	2619
INA	F	LAVINIA	16447	1775
INA	F	AUGUSTINA	16447	891
INEZ	F	AGNES	33963	105135
JACK	M	JOHN	123217	1204477
JACK	M	JACKSON	123217	4531
JACKIE	M	JOHN	7385	1204477
JACKIE	F	JACQUELINE	2699	18805
JAKE	M	JACOB	6127	39054
JAMIE	M	JAMES	2611	956086
JAMIE	M	BENJAMIN	2611	72739
JAN	F	JANET	1263	37597
JAN	F	JENNIFER	1263	14232
JANE	F	JEANNE	54166	22087
JANE	F	JOANNA	54166	4906
JANET	F	JEANETTE	37597	27643
JANIE	F	JANE	14611	54166
JAY	M	JACOB	16443	39054
JEAN	F	GENEVIEVE	82645	39727

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
JEAN	F	JEANETTE	82645	27643
JEAN	F	JOANNA	82645	4906
JEANNE	F	JEANETTE	22087	27643
JEANNIE	F	JEAN	1619	82645
JEANNIE	F	JEANNE	1619	22087
JEFF	M	JEFFREY	4888	25365
JEFF	M	JEFFERSON	4888	3555
JEFF	M	GEOFFREY	4888	1795
JEFFREY	M	GEOFFREY	25365	1795
JENNIE	F	VIRGINIA	63024	162308
JENNIE	F	JEAN	63024	82645
JENNIE	F	JENNIFER	63024	14232
JENNY	F	VIRGINIA	4596	162308
JENNY	F	GENEVIEVE	4596	39727
JENNY	F	JANET	4596	37597
JENNY	F	JEANETTE	4596	27643
JENNY	F	JENNIFER	4596	14232
JEREMY	M	JEREMIAH	6377	5272
JERRY	M	GERALD	53293	77241
JERRY	F	GERALDINE	1290	41958
JERRY	M	JEREMIAH	53293	5272
JESSE	F	JESSICA	1230	11086
JESSIE	F	JANET	73245	37597
JESSIE	F	JEANETTE	73245	27643
JESSIE	F	JESSICA	73245	11086
JESSIE	F	JESSICA	73245	11086
JILL	F	JULIA	2582	108499
JIM	M	JAMES	11504	956086
JIMMIE	M	JAMES		956086

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
JIMMY	M	JAMES	21949	956086
JO	F	JOSEPHINE	6282	129990
JOAN	F	JOANNA	47012	4906
JODY	M	JOSEPH	1005	588810
JODY	F	JOSEPHINE	810	129990
JODY	F	JOANNA	810	4906
JOE	M	JOSEPH	72786	588810
JOEY	M	JOSEPH	1287	588810
JOHN	M	JONATHAN	1204477	12683
JOHNNIE	M	JOHN	26431	1204477
JOHNNY	M	JOHN	30115	1204477
JON	M	JONATHAN	5867	12683
JORGE	M	GEORGE	10052	547639
JOSH	M	JOSHUA	1151	16128
JOY	F	JOYCE	7374	42923
JUDY	F	JUDITH	11866	26156
JULIA	F	JULIET	108499	1594
JULIE	F	JULIA	9664	108499
JUSTUS	M	JUSTIN	627	10739
KAREN	F	CATHERINE	21352	156439
KATE	F	KATHERINE	10521	92434
KATHARINE	F	CATHERINE	6558	156439
KATHLEEN	F	CATHERINE	53512	156439
KATHY	F	CATHERINE	5607	156439
KATHY	F	KATHERINE	5607	92434
KATHY	F	KATHLEEN	5607	53512
KATHY	F	CATHLEEN	5607	1243
KATIE	F	CATHERINE	25990	156439
KATIE	F	KATHERINE	25990	92434

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
KATY	F	KATHERINE	929	92434
KAY	F	CATHERINE	5434	156439
KAY	F	KATHERINE	5434	92434
KAY	F	KENDRA	5434	720
KEN	M	KENNETH	787	173749
KENDALL	M	KENNETH	1559	173749
KENDRICK	M	KENNETH	723	173749
KENNY	M	KENNETH	1264	173749
KIM	F	KIMBERLY	2690	9031
KITTIE	F	CATHERINE	1112	156439
KITTY	F	KATHERINE	2171	92434
LANNY	M	ROLAND	600	27412
LARRY	M	LAWRENCE	55489	122631
LARRY	M	LAURENCE	55489	9361
LARS	M	LAWRENCE	1053	122631
LAURA	F	LORETTA	99465	37613
LAURIE	F	LAUREN	2809	2263
LEE	F	SHIRLEY	8418	81045
LEE	M	LEROY	46491	61688
LEE	F	LENORA	8418	9955
LEE	M	LEVI	46491	7126
LEE	M	ELIAS	46491	5863
LEM	M	LEMUEL	560	3270
LEN	M	LEONARD	665	94676
LENA	F	ANGELINA	68540	18932
LENA	F	HELENE	68540	11856
LENA	F	HELENA	68540	10712
LENORA	F	ELEANOR	9955	92619
LEO	M	LEONARD	74317	94676

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
LEON	M	LEONARD	47600	94676
LEON	M	LIONEL	47600	6518
LEON	M	NAPOLEON	47600	2781
LEON	M	LEONIDAS	47600	591
LESLIE	F	LESTER	37237	59052
LESSIE	F	CELESTE	7105	3495
LETTIE	F	LETITIA	4685	1867
LETTY	F	CHARLOTTE	653	55762
LIBBY	F	ELIZABETH	1744	305833
LILA	F	DELILAH	14794	1545
LILLY	F	LILLIAN	6876	185329
LINA	F	ANGELINA	4287	18932
LINA	F	HELENA	4287	10712
LINA	F	MAGDALENA	4287	5138
LINA	F	PAULINA	4287	1682
LINA	F	SELENA	4287	1270
LINDA	F	MELINDA	51443	3069
LINDA	F	BELINDA	51443	1730
LINDA	F	ROSALYN	51443	1696
LISA	F	MELISSA	12002	9506
LISA	F	MELISSA	12002	9506
LISA	F	ALICIA	12002	6811
LIVIA	F	OLIVE	533	35225
LIVIA	F	OLIVIA	533	10363
LIZ	F	ELIZABETH		305833
LIZA	F	ELIZABETH	1074	305833
LIZZIE	F	ELIZABETH	14933	305833
LOIS	F	LOUISA	80954	6282
LOLA	F	DELORES	30866	17016

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
LON	M	LAWRENCE	1661	122631
LON	M	LAURENCE	1661	9361
LON	M	ALONZO	1661	8846
LONNIE	M	ALONZO	18575	8846
LONZO	M	ALONZO	1229	8846
LOREN	M	LORENZO	10006	8748
LORNE	M	LAWRENCE	828	122631
LORNE	M	LAURENCE	828	9361
LOTTIE	F	CHARLOTTE	27521	55762
LOTTIE	F	CARLOTTA	27521	1025
LOU	M	LOUIS	718	155906
LOU	F	LUCILLE	7057	96725
LOU	F	LOUISA	7057	6282
LOU	F	LUCINDA	7057	5785
LOUIE	M	LOUIS	7225	155906
LOUISE	F	ELOISE	134957	12582
LOUISE	F	ELOUISE	134957	2120
LUCY	F	LUCILLE	61627	96725
LUCY	F	LUCIA	61627	7470
LUCY	F	LUCINDA	61627	5785
LUKE	M	LUTHER	4959	30689
LUKE	M	LUCAS	4959	2350
LULA	F	LUELLA	41974	14972
LYNN	F	LINDA	4334	51443
LYNN	F	CAROL	4334	43517
LYNN	F	CAROLINE	4334	33486
LYNN	F	MADELINE	4334	29150
LYNN	M	LYNDON	8475	894
MAC	M	MALCOLM	1147	11344

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
MAE	F	MARY	51391	1199322
MAGDA	F	MADELINE	926	29150
MAGGIE	F	MARGARET	31097	436797
MAGGIE	F	MARGARETTA	31097	1177
MANDA	F	AMANDA	619	23029
MANDY	F	AMANDA	2060	23029
MANDY	F	MIRANDA	2060	1029
MANUEL	M	EMANUEL	47320	8587
MARGE	F	MARGARET	635	436797
MARGE	F	MARJORIE	635	83108
MARGIE	F	MARJORIE	18012	83108
MARGO	F	MARGARET	950	436797
MARIAN	F	MARIANNA	40638	2711
MARK	M	MARCUS	49717	10837
MARTY	M	MARTIN	1043	59621
MARY	F	MARION	1199322	85605
MARY	F	MARILYN	1199322	26848
MARY	F	MIRIAM	1199322	22625
MARY	F	MARIETTA	1199322	4112
MARY	F	MARIAH	1199322	1346
MATT	M	MATTHEW	2343	36898
MATT	M	MATTHIAS	2343	521
MAUD	F	MADELINE	6113	29150
MAUD	F	MATILDA	6113	17325
MAY	F	MARY	23072	1199322
MERCY	F	MERCEDES	554	9837
MICKEY	M	MICHAEL	1358	234607
MIKE	M	MICHAEL	9188	234607
MILLIE	F	MILDRED	9260	234780

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
MILLIE	F	EMILY	9260	46545
MILLIE	F	AMELIA	9260	27518
MILLIE	F	CAMILLE	9260	3706
MILLIE	F	EMELINE	9260	995
MINA	F	WILHELMINA	5415	7539
MINA	F	MINERVA	5415	5023
MINNIE	F	WILHELMINA	97523	7539
MINNIE	F	MINERVA	97523	5023
MONA	F	RAMONA	6401	12871
MONNIE	F	MONICA	847	7521
MONTY	M	LAMONT	874	1442
MYRA	F	ALMIRA	10305	849
NAT	M	NATHANIEL	894	17738
NAT	M	JONATHAN	894	12683
NATHAN	M	NATHANIEL	24670	17738
NATHAN	M	JONATHAN	24670	12683
NATHAN	M	JONATHAN	24670	12683
NATHANIEL	M	NATHAN	17738	24670
NEIL	M	CORNELIUS	10389	11360
NETTIE	F	JEANETTE	28450	27643
NETTIE	F	ANTOINETTE	28450	18750
NETTIE	F	ANNETTE	28450	14758
NICK	M	NICHOLAS	8027	43390
NICK	M	DOMINIC	8027	9987
NIKKI	F	NICOLE	551	4904
NITA	F	JUANITA	2358	48618
NONIE	F	JOAN	555	47012
NONIE	F	NORA	555	40878
NORA	F	ELEANOR	40878	92619

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
NORA	F	LENORA	40878	9955
NORA	F	ELNORA	40878	7879
NORA	F	LEONORA	40878	2704
NORA	F	ELENORA	40878	1451
NORA	F	LEONORE	40878	1006
NORA	F	HONORA	40878	705
NORA	F	ELENORE	40878	629
NORAH	F	HONORA	1174	705
OLIVE	F	OLIVIA	35225	10363
OLLIE	F	OLIVE	21492	35225
OLLIE	M	OLIVER	7625	25584
OLLIE	F	OLIVIA	21492	10363
ONA	F	ARIZONA	4504	814
ONIE	F	ARIZONA	1170	814
ONNIE	F	IONA	557	5036
ORA	F	CORINNE	21569	7394
ORA	F	AURELIA	21569	4963
OSWALD	M	WALDO	2679	3255
PAT	F	PATRICIA	589	99319
PAT	M	PATRICK	2538	43388
PATSY	F	PATRICIA	9305	99319
PATTY	F	PATRICIA	1958	99319
PEGGY	F	MARGARET		436797
PENNY	F	PENELOPE	2177	1383
PERCY	M	PERCIVAL	14885	727
PETE	M	PETER	6676	95533
PHIL	M	PHILIP	1906	53493
PHIL	M	PHILLIP	1906	30603
POLLY	F	PAULINE	4832	96947

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
POLLY	F	PAULA	4832	16139
POLLY	F	PAULINA	4832	1682
RAE	F	RACHEL	5391	30534
RALPH	M	RAPHAEL	155422	2778
RANDY	M	RANDOLPH	9826	7026
RAY	M	RAYMOND	47307	222296
REBA	F	REBECCA	11124	35077
REGGIE	M	REGINALD	558	12887
RENA	F	IRENE	11290	145546
RENA	F	SERENA	11290	1478
RETTA	F	LORETTA	1277	37613
RETTA	F	HENRIETTA	1277	26406
RICK	M	RICHARD	1797	339756
RICK	M	DERRICK	1797	4063
RICK	M	CEDRIC	1797	2342
RICKY	M	RICHARD	8967	339756
RICKY	M	DERRICK	8967	4063
RICKY	M	CEDRIC	8967	2342
RITA	F	MARGARITA	44535	10216
ROBBIE	M	ROBERT	1031	830576
ROBBIE	F	ROBERTA	4247	23401
ROBIN	F	ROBERTA	4938	23401
ROLAND	M	ORLANDO	27412	5763
RON	M	RONALD	662	93407
RON	M	AARON	662	21135
RON	M	CAMERON	662	2178
RONNIE	M	RONALD	8182	93407
RONNIE	M	AARON	8182	21135
ROSA	F	ROSALYN	56450	1696

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
ROSE	F	ROSALYN	195950	1696
ROSE	F	ROSEANN	195950	1046
ROSE	F	ROSEANNA	195950	834
ROSIE	F	ROSEANN	19823	1046
ROSIE	F	ROSEANNA	19823	834
ROSS	M	ROSCOE	11916	10944
ROXIE	F	ROXANNE	7076	1229
ROXIE	F	ROXANNA	7076	679
ROY	M	LEROY	132985	61688
RUBY	M	REUBEN	891	10364
RUDY	M	RUDOLPH	3503	25953
SADIE	F	MERCEDES	37083	9837
SAM	M	SAMUEL	35937	130315
SAM	M	SAMPSON	35937	949
SAM	M	SAMSON	35937	528
SAMMY	M	SAMUEL	3048	130315
SAMMY	M	SAMSON	3048	528
SANDRA	F	CASSANDRA	29516	2176
SANDRA	F	ALEXANDRIA	29516	1335
SANDY	F	SANDRA	707	29516
SANDY	M	SANFORD	1659	4601
SANDY	F	CASSANDRA	707	2176
SAUL	M	SOLOMON	4884	7896
SHELLY	F	RACHEL	1147	30534
SHELLY	F	MICHELLE	1147	9574
SHELLY	M	SHELDON	526	4104
SHELLY	M	SHELTON	526	1750
SID	M	SIDNEY	503	31133
SIM	M	SIMEON	735	1279

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
SIMON	M	SIMEON	11080	1279
SINA	F	ROSINA	571	3359
SOL	M	SOLOMON	5371	7896
SONNY	M	JEFFERSON	752	3555
SONNY	M	ANDERSON	752	3112
SONNY	M	JUDSON	752	2356
SOPHIA	F	SOPHRONIA	17056	757
STACY	F	ANASTASIA	2004	4528
STELLA	F	ESTELLE	65716	28649
STELLA	F	ESTELLA	65716	14707
STEVE	M	STEPHEN	20609	66393
STEVE	M	STEVEN	20609	44926
STEVIE	M	STEPHEN	506	66393
SUE	F	SUSAN	13165	47954
SULA	F	URSULA	774	4984
SUSIE	F	SUSAN	29489	47954
TED	M	THEODORE	8779	76431
TEDDY	M	THEODORE	3298	76431
TERRY	F	THERESA	2732	61538
TERRY	F	TERESA	2732	25817
TERRY	M	TERENCE	21458	1832
THAD	M	THADDEUS	1032	5210
THEO	M	THEODORE	1716	76431
THEO	F	THEODOSIA	1549	819
TILDA	F	MATILDA	835	17325
TILLIE	F	MATILDA	10152	17325
TILLIE	F	MATHILDA	10152	6083
TIM	M	TIMOTHY	1546	37098
TIMMY	M	TIMOTHY	673	37098

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
TINA	F	CHRISTINE	7259	40151
TINA	F	CHRISTINA	7259	15149
TINA	F	ERNESTINE	7259	12396
TINA	F	AUGUSTINA	7259	891
TOBY	M	TOBIAS	1092	785
TOM	M	THOMAS	13568	355230
TOMMY	M	THOMAS	13041	355230
TONY	M	ANTHONY	15050	133472
TORRIE	F	VICTORIA		32474
TORY	F	VICTORIA		32474
TRACY	F	THERESA	3338	61538
TRUDY	F	GERTRUDE	1325	130609
VAL	F	VALERIE		
VAL	M	VALENTINE	825	3154
VALLIE	F	VALENTINA	1848	1556
VERNA	F	LAVERNE	29782	9096
VIC	M	VICTOR		
VIC	M	VINCENT		51450
VICKI	F	VICTORIA	3120	32474
VICKY	F	VICTORIA	1030	32474
VINA	F	LAVINA	2342	2619
VINA	F	MELVINA	2342	2406
VINA	F	LAVINIA	2342	1775
VINA	F	LOUVENIA	2342	1280
VINCE	M	VINCENT	568	51450
VIRGIE	F	VIRGINIA	11338	162308
VONNIE	F	VERONICA	722	21323
WALDO	M	OSWALD	3255	2679
WENDY	F	GWENDOLYN	3416	12582

NICK_NAME	SEX	PROPER_NAME	NCOUNT	PCOUNT
WILL	M	WILLIAM	8029	1028647
WILL	M	WILBUR	8029	30254
WILL	M	WILBER	8029	2001
WILLIE	M	WILLIAM	128015	1028647
WILLIE	F	WILHELMINA	58965	7539
WILLIE	F	WILDA	58965	4343
WILLY	M	WILLIS	853	21183
WILMA	F	WILHELMINA	41607	7539
WINNIE	F	WINIFRED	11142	23711

Technical Guidance for HIV/AIDS Surveillance Programs

Data Quality

HIV Incidence and Case Surveillance Branch
Atlanta, Georgia

Contents — Data Quality

Introduction 10-3

Structural Requirements 10-3

- Procedure Manuals, Coding Manuals, and Other Documentation 10-3
- Edits and Data Processing 10-3

Process Standards 10-4

- Adherence to Data Standards 10-4
- Training 10-4
- Quality Control Activities 10-4
- Re-abstraction Studies 10-5

Outcome Measures 10-8

Appendixes

- Appendix A 10-10
- Appendix B 10-17
- Appendix C 10-18
- Appendix D 10-20
- Appendix E 10-22
- Appendix F 10-27
- Appendix G 10-37
- Appendix H 10-38
- Appendix I 10-40

Technical Guidance for HIV/AIDS Surveillance Programs — Data Quality

Introduction

To monitor the HIV epidemic, state and local HIV/AIDS surveillance programs attempt to collect extensive information from multiple sources over time on all persons receiving a diagnosis of HIV infection or AIDS. The basic functions of such surveillance programs include quality control activities, which focus on assessing and improving the quality of the data. The surveillance program must have a quality assurance program with designated staff and a schedule of defined activities. The data quality evaluations and the outcome measures should be recorded and used to improve registry procedures if target performance levels are not met.

Structural Requirements

Procedure Manuals, Coding Manuals, and Other Documentation

The surveillance program must document all procedures, coding of information, and data edits, and this documentation must be available to all staff. The documentation must be aligned with the policies and standards for CDC's HIV/AIDS Reporting System. This documentation should also include the registry's quality assurance plan (checks on case ascertainment and checks on data quality), staff and report source training, and quality control activities.

Edits and Data Processing

Computerized data edits that check for item validity and internal consistency based on logic rules are generally part of the registry software. At a minimum, the standard data edit set must be available (the edits are described in the eHARS software documentation). Information from edit procedures should be retained and analyzed on a regular basis to detect and correct problems with specific data sources, disease investigators, or instruction manuals.

The software system should allow duplicate data entry and the preparation of reports comparing the original and duplicate entry.

The registry software system or some other software program should allow a sample of records to be selected for audits or special studies. In addition, the system should allow original data to be compared with data from re-abstraction studies. Either through built-in functions of the registry software, or through other analytic packages, quality control personnel must be able to produce reports to measure data quality standards. Standard methods must be used to produce the outcome measures.

Process Standards

Adherence to Data Standards

Data items should be collected according to standard codes as defined in the Instructions for Completing the Data Collection Form, Vol. II of Technical Guidance for HIV/AIDS Surveillance Programs. Using standard codes allows data exchange between local sites and aggregation of data on the national level. Adherence to standard codes is assessed with standard data edits.

Training

Training is an essential component to assuring accurate, consistent, and complete data collection. Training for registry staff and report sources should include

- Reporting requirements, including frequency of reporting, mechanism of reporting, and required data items
- Data collection, including reportable events, case-finding procedures, coding, and follow-up procedures
- Quality control, including visual and computer edits and feedback regarding edit results
- Data processing, including data entry and linkage

Quality Control Activities

The quality control activities described in this section primarily address the accuracy of the data collected. Checks on case ascertainment (accuracy of case counts, completeness of reporting) and timeliness of case reporting are described in Access to Source Data and Completeness of Reporting.

Data quality control activities include

- Visual editing (proofreading) of hard copy case report forms (all forms, all data items, and all comments) before data entry, if possible by a person other than the person completing the form. This includes checking readability, consistency, and coding, and verifying any inconsistent or unclear responses.
- Duplicate data entry (all or at least 10% of hard copy forms). An alternative option is to select 10%–20% of cases and cross-reference them with the hard copy. Any discrepancies identified in reports of the results from duplicate data entry (i.e., the comparisons of the original and duplicate documents) should be resolved, and the results should guide training efforts regarding data entry.
- Electronic edit checks of individual electronic documents (all standard data edits are applied to all documents upon data entry or electronic import).
- Electronic edit checks of the consolidated case-based records (all standard data edits [see <*Standard Data Edits*>] are applied to all records before data transfers to CDC, but at least quarterly), and resolution of errors.
- Duplicate abstracting (see Re-abstraction Studies below).

- Data analysis and use: Inconsistencies in the data are often discovered during data analyses. These problems should be communicated to the quality control staff for follow-up and improvements of procedures.

Re-abstraction Studies

Re-abstraction studies assess the agreement (accuracy) between information recorded in the surveillance system and information recorded in source records (e.g., hospital medical records). Re-abstraction studies are conducted to assess the quality of the data from

1. staff abstracting (active surveillance) and
2. provider reporting (passive surveillance, where it has been implemented).

At least once a year, all programs should routinely re-abstract demographic, risk factor, laboratory, and clinical data from a representative sample of records to assess the quality and validity of information collected as it existed in the source file when the record was initially abstracted.

Eligibility Criteria and Study Population

To assess the quality of information collected by program staff and determine data quality for a diagnosis year, programs will use cases with documents (case report forms) obtained through active surveillance during a diagnosis year to form the study population. Only cases diagnosed and reported will be available for sampling. Sampling may also be based on report year, but it may be difficult to obtain medical records for cases diagnosed several years earlier. Generally, cases that do not meet the HIV or AIDS case definitions should be excluded unless there is reason to believe that case definition criteria need to be reviewed.

To assess the quality of information reported to the program by providers, programs will include cases with documents received through passive surveillance. Generally, this will entail the re-abstracting of case report forms; data quality checks for laboratory reporting are discussed in Reporting Back to Laboratories in Electronic Reporting. Information received from electronic medical records, and transmitted and imported electronically, may not need to be re-abstracted unless information has been manually transcribed onto case report forms somewhere in the process (re-abstracting of such records should then be considered). The accuracy of the information provided in electronic documents, however, needs to be assessed as well (e.g., coding of values or accuracy of data entry at remote site).

Selection of cases for re-abstracting also depends on the date of the original abstraction. Re-abstracting on the same day as the original abstraction should be avoided because bias may be introduced when staff members know re-abstracting is immediately to follow. However, because of the changing nature of medical records and the potential for archiving of some files, re-abstracting should not be done too distant in the future. The time frame for re-abstracting should be 1 to 6 months after the initial case report (but may be later depending on the program's resources and logistic considerations). This time frame also allows name-to-code states to re-abstract. For code-based states, re-abstracting for HIV (not AIDS) cases might have to

be done on the same day. To re-abstract data that were available at the time of initial abstraction, the program needs to ensure proper tracking of the dates when initial case report forms are completed and the sources of the information; this allows re-abstracting staff to abstract from the date the form was completed backwards at the original source.

Sampling Strategy

Sampling strategies for simple and stratified random sampling are outlined in Appendix A. While the sample size is calculated once before the beginning of the re-abstraction study using the prior year's reported case count as a proxy for the expected reported case count, sampling of cases may occur throughout the year (e.g., quarterly or monthly) to accommodate the intended sampling frame and stay within the re-abstracting period of 1 to 6 months after original abstraction.

Frequency of Re-abstraction Studies

Re-abstraction studies should be performed annually.

Data Items Included in Re-abstraction Studies

The data items for re-abstracting include, at a minimum, the information required to report a case of HIV infection or AIDS to CDC's HIV/AIDS Reporting System. These data items are the alpha-numeric (Soundex) code of the patient's name; state-assigned patient identifier number; HIV/AIDS diagnosis information, including date(s) of diagnosis; and the patient's date of birth, race/ethnicity, and sex. In addition, risk factor information and key laboratory data (e.g., CD4 counts at or near diagnosis) should be collected as well as any other data items of interest to the program. See Appendix G for a detailed description of the data items.

Data Collection

Staff responsible for re-abstracting cases collect the information on hard copy or electronic case report forms that indicate the data elements to be abstracted. The form used for re-abstraction should be clearly marked as ***duplicate***. The date when the initial case report form was completed is the date from which to re-abstract the information backwards in time. If information is available that was added to the medical record after the date of the initial abstract, a separate case report form should be used to collect any additional information of interest. This form is not a duplicate form but is handled like any other document added to the document-based surveillance system. Any automated process of calculating agreement rates will be based on the original and the duplicate forms; therefore, information from re-abstracting and new information should not be collected on the same form.

During the study, re-abstracting staff should not have access to already available information on the data elements for the cases they are re-abstracting. Completed forms are entered into the software system marked as duplicate documents.

Analysis Plan and Use of Results

Statistical software or the built-in function of the data collection software is used to compare information from the original abstract with that from the duplicate abstracts. The errors are stratified by major and minor error levels (see Appendix G, Variables for Re-abstraction), and agreement rates are calculated (see Appendix I, Re-abstraction Report Table). Standards for agreement rates from re-abstraction studies have not been established to date. Ultimately, the goal is to achieve perfect agreement.

Any discrepancies observed should be reviewed to determine their possible causes. In addition, a mediator should reconcile the differences to assure the data are updated appropriately. Such reconciliation may necessitate contacting report sources for verification or comparing the discrepant information with additional documents. Rules and procedures should be updated based on relevant findings, and feedback should be provided to staff and reporting sources (for passive surveillance).

From the analyses the program may learn of problems with

- information recorded overall or for specific data items based on the agreement rates,
- specific data sources (e.g., medical records) or provider reporting,
- certain types of cases,
- staff and provider training, and
- interpretation of the data (limitations of the data).

Case Finding

Completeness of HIV/AIDS case ascertainment can be calculated by comparing the number of cases diagnosed and reported to the surveillance program for a diagnosis year with the number of cases expected to be diagnosed during that year. Capture-recapture methods are used to estimate the expected number of cases. The overlap of reporting of cases from different sources is determined, as well as the number of cases reported solely by each source. The total number of cases that were not reported by any source is estimated to calculate the total number of cases expected to be diagnosed in the population.

Where document-based surveillance is not performed so that capture-recapture methods can be used to determine completeness of reporting, a case-finding component should be added to re-abstraction studies. As staff visit facilities to re-abstract cases, facility records should be reviewed for the past diagnosis year to determine whether any cases were missed (see Access to Source Data and Completeness of Reporting for case-finding methods). A comparison between the total number of cases (reported and previously missed cases) and the reported cases should indicate the completeness of reporting (see Access to Source Data and Completeness of Reporting).

Outcome Measures

The validity and accuracy of estimates derived from individual data elements may be limited if these data elements contain errors or a large proportion of information is missing or unknown for them. Standards for individual data elements are measured by edits (proper values, internal consistency), the percentage of missing information, and re-abstraction studies.

Outcome standards are set to indicate the minimum at which data can be reliably used for analyses and should be assessed for each diagnosis year at the specified time for all cases meeting the HIV/AIDS case definitions. The results should be used to improve surveillance processes.

- Edits: The minimum standard for passing data edits is that ≥97% of case records pass all standard data edits (see <*Standard Data Edits*>). The target standard is that 100% of case records pass all standard data edits. The edits for the standard include those variables most important for data analysis (sex, date of birth, date of diagnosis, race/ethnicity, state of residence at diagnosis, initial CD4 count at HIV and/or AIDS diagnosis, vital status, and risk factors). The standard is assessed for the most recent diagnosis year at 12 months after that diagnosis year.
- Missing/Unknown Information: The proportion of case records missing information is assessed for Soundex, sex, date of birth, date of diagnosis, race/ethnicity, state of residence at diagnosis, initial CD4 count at HIV diagnosis and/or AIDS diagnosis, vital status, and date of death (for those known to be dead). The target is 0% missing information.

  Note: Percentage of missing information is measured for each data item at 12 months after the diagnosis year. While a code of "unknown" is not the same as a code for missing value, the percentage of "unknown" is also calculated. A separate standard addresses missing information for transmission category.

- Re-abstraction Studies: Agreement rates (see previous discussion at Re-abstraction Studies) are calculated from re-abstraction studies as an indicator of data quality. No standards have been set for agreement rates.
- Other Data Quality Indicators: Other data quality indicators include the percentage of cases reported from death certificates only (see Death Ascertainment), the percentage of cases with no information on risk factors (HIV/AIDS Risk Factor Ascertainment), and the number of duplicate reports (Case Residency Assignment), which are described in the respective chapters of *Technical Guidance for HIV/AIDS Surveillance Programs*. A summary of outcome standards is provided in the Introduction to Policies and Procedures chapter.

Appendixes

Appendix A Sampling Methodology for Re-Abstraction

Appendix B Taking a Simple Random Sample (SRS) Using the Random Number Table

Appendix C Desired Number of Strata

Appendix D Point Estimates and Sample Size Calculation for Simple Random Sampling (SRS)

Appendix E Point Estimates and Sample Size Calculation for Stratified Random Sampling

Appendix F Examples of Sampling for Re-abstraction Studies

Appendix G Variables for Re-Abstraction

Appendix H Re-abstraction Error Report Form, for Pairs of Reports for One Case

Appendix I Re-abstraction Report Table, for Diagnosis Year 20XX

Appendix A

Sampling Methodology for Re-Abstraction

Objective—The purpose is to provide a statistical methodology for all sites, which can be used to create a sample of cases for re-abstraction.

Background—Re-abstracting information on a case has been shown to be a good way to measure data quality in a data collection system. The method is simple. At some pre-designated time after initially collecting information on a case (via medical charts), filling out a case report form, and then entering the data into the site's registry, staff go back to the original source of the case report form and re-collect the information without knowing what was entered on the initial report.

The re-abstraction should only be done for important variables, including those that define a case (HIV or AIDS case definitions) or are required for reporting eligibility. These variables are listed in Appendix G.

The variables will be compared between the initial report and the re-abstracted report for agreement (assuming quality control procedures for data entry have been implemented to assure discrepancies are not due to data entry errors). Agreement is assessed for pairs of documents and for the surveillance program as a whole (see Appendix I). The results can be used to determine the level of data quality for the program, individual staff, and/or reporting sources (passive surveillance). This document will address how to sample cases to be re-abstracted.

Note: The map of county federal information processing standards (FIPS) codes for your state (see Web site below) may be of use in determining regions:

http://www.census.gov/geo/www/fips/fips65/

Sampling for Re-Abstraction

Feasibility of Sampling within a Site—It is strongly recommended that all state and local surveillance programs carefully consider the strategy for sampling HIV/AIDS cases on which to perform re-abstraction studies. The scientific and practical factors to consider in sampling include

- Number of persons receiving a diagnosis of HIV/AIDS and number of cases reported per year
- Number of providers/facilities within the surveillance program area
- Number of surveillance staff responsible for completing case report forms
- Geographic regions
- Resources available to design and conduct a sampling protocol

Note: Selection of providers, surveillance staff, and geographic regions may be weighted by the number of cases in each stratum.

A site with limited resources may not be able to re-abstract the number of cases suggested by sample size calculations (discussed below). Also, if no attempt is made to lessen the effect of geography and the locations of providers, even if the site had the resources, it may not have the staff needed to travel the state to re-abstract.

Consultation with your CDC core epidemiologic technical assistant (Epi TA) (see CDC Web site to identify responsible Epi TA) is strongly recommended for this evaluation. However, before consulting with your Epi TA, you need to determine the level of detail that you wish to achieve. For example, do you wish to have precise knowledge of error rates at the provider, facility, or surveillance staff levels? If so, you should run the following frequency tables using your HIV/AIDS surveillance data to determine the

- Number of reported HIV/AIDS cases in the past year
- Number of cases reported by each facility/provider
- Number of case reports completed by each surveillance staff member
- Number of cases reported by region

Note: Separate frequency tables may be useful to distinguish HIV and AIDS cases.

These numbers will not only help to determine if sampling is feasible, but also assist in deciding the appropriate sampling design and associated sample sizes. Otherwise, you should consult with your Epi TA to properly assess the feasibility of sampling within your site.

Choosing a Sampling Frame—The sampling frame is the list of cases from which a sample is drawn. For example, if the target population is all newly reported HIV/AIDS cases in a given calendar year in State X, the frame will be constructed from all HIV/AIDS cases reported to the HIV/AIDS surveillance system during that specific year within State X. Also, sampling may be conducted separately for HIV and AIDS cases (the latter may be first reported as or updated to an AIDS diagnosis).

When choosing a sampling frame for re-abstraction, the site may want to focus only on those facilities/providers and/or surveillance staff that report the majority of cases, and exclude infrequent reporters.

Consult with your Epi TA to help you choose which sampling frame is best for your site.

Other Sampling Frame Considerations

Time Frame for Sampling—The time frame for re-abstraction should be 1 to 6 months after the date of initial case report. This is especially important for name-to-code HIV reporting states (re-abstraction would have to occur on the same day as initial abstraction for HIV cases in code-based reporting systems). Information on the case may be unattainable if records are not re-abstracted within the appropriated time of keeping patient information available for surveillance activities.

Assuming the total number of cases to be sampled for re-abstraction in the coming year is determined to be *n*, the site may sample *n/4* cases every quarter (3 months) or *n/12* cases every month. To meet the desired sample size, *n*, by the end of the year, the site may need to sample more cases to guard against nonresponse (e.g., records may not be found for some cases).

Seasonal fluctuations in case reporting may exist; therefore it may be prudent to factor seasonality into the time frame for sampling. Though examples given in this document do not consider seasonality, suggested solutions may include over sampling in later months to compensate for earlier months not having enough reports to sample from, or allocating the sample according to past seasonal trends of case reporting.

Simple Random Sampling (SRS)—SRS assumes that the population to be sampled from is homogeneous and that each case in the sampling frame has an equal chance of being selected for the sample. It should be noted that throughout this document, we assume that sampling will be done without replacement (i.e., once a case is selected for the sample, it cannot be sampled again, except cases first reported as HIV and later updated to AIDS may be sampled in each category).

The advantage of SRS is that it is simple and easy to apply, particularly in lower morbidity areas. However, the fact that every case in the population has an equal chance of being selected into the sample may not be good for those surveillance programs that would like to focus their re-abstraction efforts on certain subgroups. For example, State X knows that 70% of its cases are reported within two large urban areas and the remaining 30% are reported throughout the rural areas of the state. An SRS, just by chance, may include only urban cases or include more rural cases than urban—a geographical distribution that would not be representative of the HIV/AIDS population in State X. On the other hand, if it is more imperative to sample from all cases reported in the time frame, then SRS should be done, with the acknowledgement that the sample may not be representative and that any analytic results based on the sample may not be the same as the results based on the true population.

Sample Size

Sample size refers to the target number of cases to be sampled, based on the decisions made by the site and its CDC core Epi TA. Sample size should be large enough to produce reliable population estimates. In general, the smaller the sample, the less reliable the resulting estimates will be of the population parameters. CDC recommends that a minimum of 30 cases be sampled per year.

In determining sample size, include the statistical criteria for the specified level of confidence and the desired level of precision (e.g., being 95% sure that the sample estimate falls within 3 percentage points of the true value). Appendix B illustrates an example of sample size calculation for SRS.

High nonresponse rates can compromise estimates of the population parameters with respect to the pre-specified level of confidence and precision at a specified confidence level. Nonresponse in relation to re-abstraction means attempts were made by field investigation staff, but they were unable to re-abstract the original information in the

medical record(s) or the original medical record could not be found. To guard against nonresponse, state and local surveillance programs should estimate response rates and inflate sample sizes by the reciprocals of the response rates. For instance, if one estimates the response rate to be 80%, then one would inflate the sample size by a factor of 1/0.8 = 1.25.

The statistical criteria may recommend a sample size or sampling fraction that a site may not be able to meet logistically, given time restraints, staffing, total number of cases, and resource limitations. These limitations should therefore be considered in selecting a sampling frame. In other words, if a site is not able to sample cases reported by every provider throughout the state, the site should limit the sampling frame to only those providers that report the majority of cases. It should be noted that this limiting of the sampling frame violates the assumption that the sample is representative of the true population. For the sake of enhanced data quality, this violation can be ignored.

Selecting the Sample

A surveillance program can use several approaches to take an SRS of size *n* from a sampling frame that contains *N* cases. One approach is to first assign a number from 1 to *N* to each case in the sampling frame, then pick a sample of *n* of these numbers by using random numbers generated by a computer or a random number table. Once the numbers are chosen, cases in the sampling frame that correspond to these numbers are selected to form the sample. Appendix C demonstrates how to take an SRS using the random number table.

Surveillance programs may also consult with their CDC core Epi TA in developing SAS programs to generate an SRS. The SURVEYSELECT procedure, available starting in SAS version 8, offers a variety of methods for selecting probability-based random samples, including selecting an SRS.

Stratified Random Sampling (StRS)

If the entire population of interest can be partitioned into epidemiologically meaningful, mutually exclusive and exhaustive homogeneous subpopulations (called strata) and precise estimates for the subpopulations are desired, then it is advantageous to generate the sample by selecting an SRS within each stratum independently. This powerful and commonly used technique is called Stratified Random Sampling (StRS). In essence, if there are *L* separate, nonoverlapping subgroups of a population to be sampled, then StRS is akin to performing *L* separate SRSs simultaneously. The benefits include combining the simplicity of SRS with increased reliability in estimating rates. Unlike SRS, StRS allows inferences to be made about the error rates at a specified level of reliability and precision for subpopulations (e.g., providers, facilities, or surveillance staff) as well as for the whole population. It must be emphasized that special analytic techniques are required in estimating the true overall population parameters when StRS is implemented.

Stratification Variables

The choice of stratification variables to be used should be guided by the site's analytic objectives. If the site does not need to make precise estimates for particular subgroups, then it does not need to stratify on that variable. When selecting the variable(s) to use for stratification, the site should consider those (categorical) variables that may have a strong association with data quality. It is important to create subgroups that are likely to be homogeneous (less variability) along the lines of the other variables. Surveillance programs may consider the following as potential stratification variables:

Facility/Provider—Each facility/provider may report any number of cases. Some report a large percentage of an area's HIV/AIDS cases, while others report a few cases. To determine data quality for a site, the program may need to consider the distributions of case reporting by facility. This assessment is particularly relevant for information obtained through passive surveillance. Because of the possible heterogeneous nature of case reporting by facility, special considerations may need to be made to standardize the distribution. These may include collapsing facilities that report few cases into one stratum; not using facility as a stratification variable but combining it with geographical region and stratifying on region (see below); or performing independent SRS separately for frequent and infrequent reporters (i.e., select two sampling frames for the site, one for frequent reporters and one for infrequent reporters), and instead of creating the samples simultaneously, create the samples separately.

Surveillance Staff—Surveillance staff who abstract medical record information onto Case Report Forms may complete forms with differing levels of frequency. Also, varying levels of experience in reporting information may exist among staff. These factors may need to be considered in the sampling design. **Note**: For training purposes, new staff should always have several of their cases re-abstracted to evaluate training efficacy.

Geographical Reporting Areas or Case Residence at Diagnosis—Facilities may be far apart, which may make re-abstraction time-consuming. Re-abstraction could be done at the time of routine surveillance visits for new cases. By stratifying on this variable, plans can be made to limit excess travel by staff. For some states, it may be desirable to choose one high morbidity county or a few such counties with certainty (with a probability of 1) because they are considered counties that must be represented to get an overall picture of the state. For example, each of the high morbidity counties may be considered as separate strata. **Note**: The correlation between geographic region, facility, staffing, and surveillance practices (i.e., passive versus active surveillance) may be strong in any one site. For that reason, the suggestion to combine facility and staff into region (as stratification variables), and stratify on region may be the best method for re-abstraction in terms of a site's resource limitations. A simpler sampling design is preferred over an overly complex sampling design.

Number of Strata

State and local surveillance programs may wish to consult with CDC in defining the desired number of strata. Although using more strata ensures more complete coverage of the population of interest, a large number of strata may not be beneficial because they increase the possibility of strata with small sizes. Furthermore, rate estimates are less stable if there are relatively few cases per stratum (small denominators).

For some states, when setting up the sampling plan, it may be desirable to stratify counties by morbidity, so that those with the highest morbidity would be selected with certainty (assigned a probability of 1) because they would be considered counties that must be represented to get an accurate overall picture of what is happening in the state. In that case, each of the few high morbidity counties may be considered as a separate stratum.

In general, it is recommended that if only the overall estimate is desired, little is gained in the precision of the estimate from having more than five or six strata. If estimates are wanted for subpopulations, a larger number of strata may be necessary. Appendix D contains an example that illustrates the process of deciding the desired number of strata.

Sample Size Calculation

Statistical criteria for sample size determination in StRS are similar to those discussed in the section for SRS.

Once the strata are specified and the total number of cases, *n*, to be sampled is known, an important step in StRS is to determine how many cases should be taken from each stratum under the constraint that a total of *n* is to be selected from all strata. Two of the simplest and most commonly used methods are

1. equal allocation and
2. proportional allocation.

In equal allocation, the same number of cases is sampled from each stratum. Under proportional allocation, the sample is allocated to the strata based on the proportion of the sampling frame (i.e., population) that each stratum represents. The site should consult with the CDC core Epi TA to decide the most appropriate method of allocation for the site's surveillance program. Appendix E illustrates an example for sample size calculation in StRS with proportional allocation.

Similar to sample size determination for SRS, state and local surveillance programs should estimate response rates within each stratum and inflate sample sizes by the reciprocals of the response rates.

Selecting the Sample

CDC has developed SAS programs to generate StRS from cases reported to the HIV/AIDS surveillance system. The SAS program uses the SURVEYSELECT procedure. The program is available upon request from your CDC core Epi TA. CDC and state or local surveillance programs will review the program and make necessary modifications to fit

the specific needs of the states. The state or local program, in consultation with CDC, will generate the sample of cases for each stratum using the SAS program developed by CDC.

Examples of Different Sampling Strategies

Each site will need to determine, with the help of its core Epi TA, the sampling strategy best suited for that site. Low morbidity sites will not be able to stratify, whereas high morbidity sites—those reporting several thousand cases each year—may need to stratify on several levels. Some sites may want to concentrate on certain aspects of their surveillance system of data collection. For example, if only a few providers report the vast majority of cases, then the sampling strategy should focus more on allowing a higher probability of sampling from these providers.

Generally, a site will need to sample at least 1% of its total reported cases annually or select at least 30 cases if the total number of reported cases annually is less than 3,000. For those sites that report less than 50 cases annually, the sites and their Epi TA will have to decide how many cases should be re-abstracted to be able to calculate meaningful error rates. If possible, sites should sample more of their total reported cases, up to 5%.

See Appendix F for examples:

- Example #1: Low Morbidity Site
- Example #2: Medium Morbidity Site, Reporting Concentrated in 2 Regions
- Example #3: Medium Morbidity Site, Reporting Concentrated by Providers
- Example #4: High Morbidity Site

Note: If the site is interested in investigating re-abstraction based on staff, it should consider that other stratification variables, such as region and provider, may overlap with staff. If so, the sampling should focus only on one or the other, not both. Sampling at least 30 cases per staff member may be too difficult, but a minimum number of cases need to be set by the site and its Core Epi TA.

The site is also able to post-stratify by additional variables in the analysis. For example, if the site stratifies on geographical region, after the samples are collected, the site could further stratify each category of region by staff. Although this does not guarantee balanced sample sizes in the strata, it may nevertheless give insights into problem areas.

Bibliography

1. CDC. Guidelines for national human immunodeficiency virus case surveillance, including monitoring for human immunodeficiency virus infection and acquired immunodeficiency syndrome. *MMWR* 1999; 48(RR13):1–28.

2. Cochran WG. *Sampling techniques*. 3rd ed. New York: John Wiley & Sons; 1977.

3. Levy PS, Lemeshow S. *Sampling of populations: methods and applications*. 3rd edition. New York: John Wiley & Sons; 1999.

4. Kalton G, Kasprzyk D. The treatment of missing survey data. *Survey Methodology* 1986; 12:1–16.

Appendix B

Taking a Simple Random Sample (SRS) Using the Random Number Table

The following example illustrates how to take an SRS using a random number table. Suppose State X has 313 HIV/AIDS cases reported in 2003. They decided to randomly select 30 (9%) of the 313 cases to perform their re-abstraction study.

The first step in selecting an SRS is to assign a number from 1 to 313 to each of the 313 cases. This can be done by first sorting the 313 cases by their STATENO value and then starting with the first case, assign it a 1; the next case will be a 2, etc. until the last case, which gets assigned number 313. The next step is to pick a sample of 30 numbers using a random number table. To do this, we need to ensure that the numbers selected from the table are all different and that none are equal to 0 or greater than 313.

Using the random number table in Table A.1 in reference 3, starting at some randomly selected point, with (for the sake of this example) row three and column 1, we proceed down the column and take different 3-digit numbers between 001 and 313 until 30 such numbers are selected. In this case, we get 241, 289, 099, 103, 071, 023, 010, 070, 293, 024, 296, 007, 053, 005, 259, 097, 179, 145, 089, 156, 050, 069, etc. Note that when we come to the bottom of a particular column, we proceed to the top line of the very next column on the same page. Also, note a commonly used technique to randomly select a starting point in the random number table is to close one's eyes and point to the table.

To form the sample of 30 cases for re-abstraction, we take from the 313 cases the 241st, the 289th, the 99th, the 103rd case, and so on.

Note: For the purpose of re-abstraction, the sample size is calculated first, based on the prior year's reported case count as a proxy for the expected case count in the present year of study. The incremental selection of cases for the study (in 2004) will occur on a quarterly or monthly basis. For instance, for the 30 cases to be selected above, on a quarterly basis, 8 cases need to be sampled from the collection of cases reported in the first quarter of the study year, 8 cases from those reported in the second quarter, 8 from those reported in the third quarter, and only 6 from those reported in the fourth quarter. For dealing with seasonality effects or nonresponse, over sampling during any one quarter or month is suggested.

Appendix C

Desired Number of Strata

The following example illustrates construction of the number of strata. Suppose State X has decided to use the stratified random sampling approach to sample all HIV and AIDS cases reported in 2003. State X will consider county of residence at diagnosis, surveillance staff, and provider as possible stratification variables. Of the 5 counties in State X—A, B, C, D, and E—the proportion of cases in County E is much smaller relative to those in the remaining counties. Two providers report the vast majority of cases; therefore, State X has decided to include only those two in their stratification schemes.

Note: For this example, we are assuming all four staff report cases from both providers.

Stratification Variables	Number of Levels	Values
County	5	• A • B • C • D • E
Staff	4	• Joe • Sue • Bob • Jean
Provider	2	• Hospital • Private MD

Following are 5 possibilities for stratification schemes.

Scheme	Stratification	Number of Strata
1	(County) by (Staff) by (Provider)	40
2	(County) by (Joe, Sue) by (Provider) + (Bob/Jean)* by (Provider)	22
3	(County E, Other Counties) by (Joe, Sue, Bob/Jean)* by (Provider)	12
4	(County E, Other Counties) by (Joe, Sue) by (Provider) + (Bob/Jean)* by (Provider)	10
5	(Joe, Sue, Bob/Jean)* by (Provider)	6

* ‘Bob/Jean’ is the same as ‘Bob or Jean.’

Scheme 1 has a large number of strata. This increases the probability that a number of strata will have small sizes and consequently increases the potential for requiring all elements in those strata to be included in the sample instead of only a portion of them. Furthermore, estimates based on small cell sizes are unreliable, which is a major drawback to having a highly stratified design.

Scheme 2 reduces the number of strata somewhat, but the number is still considerable and probably unmanageable. Bob's and Jean's cases are collapsed into one category because of the small number of cases they abstracted.

The number of strata in Scheme 3 begins to approach a manageable number. This scheme collapses all counties except County E into one category, but allows for over sampling in the County E stratum to ensure adequate number of cases from County E for analysis purposes.

Scheme 4 refines Scheme 3 in the similar manner that Scheme 2 refines Scheme 1.

Scheme 5 offers a very manageable number of strata. This proposal would require that county information be collected on all the sampled cases. Post-stratification (stratification after sampling) by county would then allow estimation at the county level. However, this scheme does not address the problem of a county with a low number of cases (County E). Inference for that county may be difficult because of the small size of the sample in that stratum.

Appendix D

Point Estimates and Sample Size Calculation for Simple Random Sampling (SRS)

Let us suppose that State X uses an SRS of HIV/AIDS cases reported in calendar year 2004 to perform re-abstraction. The goal is to estimate the error rate, or the percentage of incorrect reports. In an 'incorrect report,' at least one of the selected fields from the re-abstracted report form does NOT correspond with the respective field from the initial report form.

Sampling Frame: All cases reported to the state health department between January 1, 2004, and December 31, 2004

Sampling Method: SRS

- N = the total number of reports from 2003 (proxy for 2004)
- n = the number of reports to be sampled from 2004
- δ = the number of incorrect reports from the sampled reports
- P = the true error rate

Then the point estimate for P is

$$p = \frac{d}{n} \qquad (1)$$

and

$$var(p) = \frac{N-n}{N-1} \times \frac{P(1-P)}{n} \qquad (2)$$

where $\frac{N-n}{N-1}$ is the finite population correction (fpc) for the variance. A 95% confidence interval for P, based on asymptotic normality, is

$$p \pm 1.96\sqrt{var(p)} \qquad (3)$$

Let $\delta = 1.96\sqrt{var(p)}$, δ is called the *precision* or *margin of error.* Using the formula for *var(p),*

$$\delta = 1.96\sqrt{\frac{N-n}{N-1} \times \frac{P(1-P)}{n}} \qquad (4)$$

To calculate sample size based on a specified precision, we simply solve the δ-formula for n.

$$n = \frac{1.96^2 NP(1-P)}{\delta^2(N-1) + 1.96^2 P(1-P)} \qquad (5)$$

Note that in the absence of knowledge about *P*, we assume *P*=0.50, which results in the most conservative (largest) sample size necessary.

Given the total number of cases reported in 2003 is *N* = 798, the following table shows sample sizes required for various designated levels of precision, δ, assuming 95% confidence limit. Again, the population error rate, *P*, is assumed to be 0.50. Sample sizes are rounded to the nearest integer. (Formula #5 was used to calculate sample size.)

Precision, δ	**Sample Size, n**
0.02*	599 (719)**
0.04	343 (411)
0.05	260 (311)
0.075	141 (169)
0.10	86 (103)

* Sample size estimate is within 2 percentage points of the true value.
**Numbers in parentheses are sample size inflated for 20% nonresponse.

Note: The finer the precision, the greater the sample size is needed.

If the sample size is predetermined without any statistical criteria, as may happen with low morbidity sites, the resulting precision can be calculated based on the predetermined sample size. Again, given the total number of cases reported last year is *N* = 798, and the population error rate, *P*, is assumed to be 0.50, the following table shows precision levels for given sample sizes. (Formula #4 was used to calculate level of precision.)

Sample Size Percent of Total, 100*(n/N)	**Sample Size, n**	**Level of Precision, δ**
1%	8* (10)**	0.35*
2%	16 (19)	0.24
3%	24 (29)	0.20
4%	32 (38)	0.17
5%	40 (48)	0.15

* Sample size estimate is within 35 percentage points of the true value.
**Numbers in parentheses are sample size inflated for 20% nonresponse.

Note: The smaller the sample size, the larger the amount of variability surrounding the estimated error rate.

Appendix E

Point Estimates and Sample Size Calculation for Stratified Random Sampling

Let us suppose that State X uses a stratified random sample of HIV/AIDS cases reported in calendar year 2004, stratified on provider, to re-abstract. The goal is to estimate the error rate of incorrect reports. In an 'incorrect report,' at least one of the selected fields from the re-abstracted report form does NOT correspond with the respective field from the initial report form. State X wishes to obtain separate estimates for males and females. However, they would like to maintain a specified level of precision at the population level only.

Sampling Frame:	Cases reported by State X between January 1, 2004, and December 31, 2004
Sampling Method:	StRS
Stratification Variable:	Provider (Providers B & C)
Precision:	Overall population level
Allocation:	Proportional (equal sampling weight in each stratum)

We shall use the following notations throughout this example and assume from this point on that we are referring to cases reported during the specified time period.

- N = the total number of reports
- N_1 = the number of reports from Provider B
- N_2 = the number of reports from Provider C
- n = the number of sampled reports
- n_1 = the number of sampled reports from Provider B
- n_2 = the number of sampled reports from Provider C
- d_1 = the number of incorrect reports sampled from Provider B
- d_2 = the number of incorrect reports sampled from Provider C
- P = the true error rate of incorrect reports from both Providers
- P_1 = the true error rate of incorrect reports from Provider B
- P_2 = the true error rate of incorrect reports from Provider C
- p_1 = the estimated error rate of incorrect reports from Provider B
- p_2 = the estimated error rate of incorrect reports from Provider C

Then the estimate for P is

$$p = \frac{(N_1 p_1) + (N_2 p_2)}{N} \qquad (6)$$

where

$$p_1 = \frac{d_1}{n_1} \quad \text{and} \quad p_2 = \frac{d_2}{n_2}$$

are estimates for P_1 and P_2, respectively, and

$$var(p) = \left(\frac{N_1}{N}\right)^2 \left(\frac{P_1(1-P_1)}{n_1}\right)\left(\frac{N_1 - n_1}{N_1 - 1}\right) + \left(\frac{N_2}{N}\right)^2 \left(\frac{P_2(1-P_2)}{n_2}\right)\left(\frac{N_2 - n_2}{N_2 - 1}\right) \qquad (7)$$

Assume the terms $1/N_1$ and $1/N_2$ are negligible, and with proportional allocation,

$$\frac{n_1}{N_1} = \frac{n_2}{N_2} = \frac{n}{N}$$

The formula for var(p) can then be simplified into

$$var(p) = \frac{N-n}{Nn}\left[\left(\frac{N_1}{N}\right)P_1(1-P_1) + \left(\frac{N_2}{N}\right)P_2(1-P_2)\right]$$

A 95% confidence interval for P, based on asymptotic normality, is

$$p \pm 1.96\sqrt{var(p)}$$

Let $\delta = 1.96\sqrt{var(p)}$, δ is called the *precision* or *margin of error*. Using the formula for *var(p)*,

$$\delta = 1.96\sqrt{\frac{N-n}{Nn}\left[\left(\frac{N_1}{N}\right)P_1(1-P_1) + \left(\frac{N_2}{N}\right)P_2(1-P_2)\right]} \qquad (8)$$

To calculate sample size based on a specified precision, we solve the δ-formula above for n.

$$n_1 = \frac{1.96^2 N[N_1 P_1(1-P_1) + N_2 P_2(1-P_2)]}{\delta^2 N^2 + 1.96^2[N_1 P_1(1-P_1) + N_2 P_2(1-P_2)]} \qquad (9)$$

Note that in the absence of knowledge about P_1 and P_2, we assume P_1=P_2=0.50, which results in the most conservative (largest) sample size necessary. But, for re-abstraction studies, we are estimating an error rate. The process standards suggest an error rate less than 0.05, so for the sake of consistency, we will assume that P_1=P_2=0.05.

To calculate sample size based on the above sample size, now allocated proportionally, we use

$$n_1 = N_1\left(\frac{n}{N}\right) \quad \text{and} \quad n_2 = N_2\left(\frac{n}{N}\right) \qquad (10)$$

Given the total number of cases reported last year is $N = 798$, where Provider B reported 605 cases and Provider C reported 193, the following table shows sample sizes required for various designated levels of precision, δ, assuming 95% confidence limit. The population proportions, P_1 and P_2, are assumed to be 0.05. Sample sizes are rounded to the nearest integer. (Formulas #9 and #10 were used to calculate sample sizes.)

Precision, δ	**Sample Size**		
	Overall (n)	Provider B (n_1)	Provider C (n_2)
0.02*	290 (348)	220 (264)**	70 (84)**
0.04	100 (120)	76 (91)	24 (29)
0.05	67 (80)	51 (61)	16 (19)
0.075	31 (37)	24 (29)	8 (10)
0.10	18 (22)	14 (17)	4 (5)

* Sample size estimate is within 2 percentage points of the true value.
**Numbers in parentheses are sample size inflated for 20% nonresponse.

If the overall sample size is selected and not calculated, as may happen with low morbidity sites, precision can be calculated. Again, given the total number of cases reported last year is $N = 798$, where Provider B reported 605 cases and Provider C reported 193, and the population error rates, P_1 and P_2, are assumed to be 0.05, the following table shows precision levels for given overall sample sizes. (Formula #8 was used to calculate level of precision.)

Sample Size Percent of Total, 100*(n/N)	**Overall Sample Size**			
	Overall (n)	**Provider B (n_1)**	**Provider C (n_2)**	**Level of Precision, δ, at Overall Level**
4%	32 (38)**	24 (29)**	8 (10)**	0.07398*
12%	96 (115)	73 (86)	23 (28)	0.04089
18%	144 (173)	109 (131)	35 (42)	0.03223
24%	192 (230)	146 (175)	46 (55)	0.02687
30%	240 (288)	182 (218)	58 (70)	0.02306

* Sample size estimate is within 7 percentage points of the true value.
**Numbers in parentheses are sample size inflated for 20% nonresponse.

Now let us suppose that State X again wants to estimate the error rate of incorrect reports as in the previous example, except with a specified level of precision at the provider level instead.

Sampling Frame:	Cases reported by State X between January 1, 2004, and December 31, 2004
Sampling Method:	StRS
Stratification Variable:	Provider (Providers B & C)
Precision:	Provider level
Allocation:	Proportional (equal sampling weight in each stratum)

Using the same notations as those specified in the previous example, we derive the point estimates and the required sample sizes.

The estimates for P1 and P2 are

$$p_1 = \frac{d_1}{n_1} \quad \text{and} \quad p_2 = \frac{d_2}{n_2}$$

and the variances for p_1 and p_2 are

$$var(p_1) = \frac{N_1 - n_1}{N_1 - 1} \times \frac{P(1-P)}{n_1} \quad \text{and} \quad var(p_2) = \frac{N_2 - n_2}{N_2 - 1} \times \frac{P(1-P)}{n_2}$$

The 95% confidence intervals for P_1 and P_2, based on asymptotic normality, are

$$p_1 \pm 1.96\sqrt{var(p_1)} \quad \text{and} \quad p_2 \pm 1.96\sqrt{var(p_2)}$$

Let $\delta_1 = 1.96\sqrt{var(p_1)}$ and $\delta_2 = 1.96\sqrt{var(p_2)}$ be the corresponding *precision* or *margin of error* for the estimate of males and females, respectively. Using the formula for $var(p_1)$ and $var(p_2)$,

$$\delta_1 = 1.96\sqrt{\frac{N_1 - n_1}{N_1 - 1} \times \frac{P_1(1-P_1)}{n_1}} \quad \text{and} \quad \delta_2 = 1.96\sqrt{\frac{N_2 - n_2}{N_2 - 1} \times \frac{P_2(1-P_2)}{n_2}}$$

To calculate sample size based on a specified precision, we simply solve the δ-formulas for n_1 and n_2.

$$n_1 = \frac{1.96^2 N_1 P_1(1-P_1)}{\delta^2(N_1 - 1) + 1.96^2 P_1(1-P_1)} \quad \text{and} \quad n_2 = \frac{1.96^2 N_2 P_2(1-P_2)}{\delta^2(N_2 - 1) + 1.96^2 P_2(1-P_2)}$$

Note that in the absence of knowledge about P_1 and P_2, we assume P_1=0.05 and P_2=0.50 following the same logic as mentioned previously.

With N_1 = 2,375 and N_2 = 1,472, the table below shows sample sizes for sex categories at varying levels of precision, δ, calculated at the sex level, assuming 95% confidence limit. The population proportions for males and females, P1 and P2, are assumed to be 0.50 to maximize sample sizes used to calculate sample size.)

Precision, δ	**Sample Size, n**		
	Provider B (n_1)	**Provider C (n_2)**	**Overall (n = n_1+n_2)**
0.02*	260 (312)**	136 (163)**	396 (475)**
0.04	96 (115)	72 (86)	168 (201)
0.05	65 (78)	53 (64)	118 (142)
0.075	31 (37)	28 (34)	59 (71)
0.10	18 (22)	17 (20)	35 (42)

* Sample size estimate is within 2 percentage points of the true value.
**Numbers in parentheses are sample size inflated for 20% nonresponse.

Comparing this table with that in the previous example, one notices that larger sample sizes are required if precision is desired at a lower level of the total population (at the provider level as compared with at the overall population level).

Appendix F

Examples of Sampling for Re-abstraction Studies

In the following examples, four possible situations are described with a suggested solution. Each site will need to discuss with their core Epi TA at CDC their specific situation and how sampling can best be used to benefit the re-abstraction study. These examples emphasize the problems and complexities inherent in any sampling methodology. Each site must plan accordingly and address each unique situation.

For each example, an introduction to the site's situation is described. Their decisions are noted, and calculations are made to show how their decisions will affect the sample methodology they have chosen to use. A summary of decisions is then shown. A table follows that attempts to show, over the course of a hypothetical year, how their sampling plan proceeds.

Note: We are NOT stratifying over time, either 4 quarters or 12 months, when we are calculating sample size and precision levels. We are assuming that the sample size calculated for a site will be for the year—but for practical purposes, it will be split over time to aid in re-abstraction. What this means is that the cells over time (as shown in the tables) will have small counts, less than the suggested minimum of 30. These small counts are okay—as long as the total number of sampled cases for the site or for each stratum, for the year, is at least 30.

Example #1: Low Morbidity Site

Site A had reported 120 HIV/AIDS cases in 2003. For their re-abstraction study, they decide (with their core Epi TA) that they should use SRS. Because they have reported less than 3,000 cases, they will randomly select a total of 30 cases for the re-abstraction, pulled at the end of each quarter of 2004.

Given that the total number of cases reported last year is $N = 120$, and the site will select the minimum number of cases, the level of precision, δ, assuming 95% confidence limit, can be calculated using Formula #4 from Appendix D. The population proportion, P, is assumed to be 0.05.

Sampling Frame:	All cases reported by State X
Sampling Method:	Simple Random Sampling (SRS)
Stratification Variable:	None
Level of Precision:	0.068, or 6.8%
Allocation:	Equal
Timing of Sample:	Quarterly
Number of Cases Reported Last Year:	120
Number of Cases to Sample:	30

		Number of Cases Proposed to Sample	Number of Cases Actually Sampled	**Number of Successful Re-Abstractions**
Quarter 1	At the end of March 2004, of those cases reported to the state health department, they randomly select 8 cases.	8	8	**7**
Quarter 2	By June 2004, 7 of the 8 cases had been successfully re-abstracted (the original source for each case had been found and the selected data fields were scored as to agree or not agree, and the cases were updated in the HIV/AIDS surveillance data collection system with any new information discovered). Another 8 cases are randomly selected from the number of cases reported since March. *The number of successful re-abstractions is the 8 plus 1 extra, for 9.	8	8	**9***
Quarter 3	By September 2004, all 16 cases were successfully re-abstracted. Another 8 cases are sampled from the 2004 workload.	8	8	**4**
Quarter 4	By December 2004, only 4 of the most recent 8 cases were successfully re-abstracted. To make the 30 case minimal sample size for the re-abstraction study for 2004, the site needs another 10 cases. The site decides to sample 12 cases to allow for the possibility of 2 cases failing re-abstraction. It's discovered at the end of January 2005 that all 12 were successfully re-abstracted.	10	12	**12**
Total		34	36	**32**

Example #2: Medium Morbidity Site, Reporting Concentrated in 2 Regions

Site B had reported 2,500 HIV/AIDS cases in 2003. About 1,800 cases (72%) were reported by facilities located in Regions 1 and 3 of the 30 total regions (counties) of the state. The surveillance team is located in Region 1, with Region 3 located only 100 miles away. Because of travel constraints, the site decides to sample cases only from the two high morbidity regions. (Limitations are being introduced here, mainly that 28% of cases being reported for this site will not be included in the re-abstraction study and so

agreement rates will be biased away from the facilities/providers and staff working in these areas.) Just like Site A, Site B decides to use SRS. However, Site B will only sample from cases being reported from the two regions. They will do this by excluding from their HIV/AIDS surveillance data system any case not reported from either region. They also decide to select 36 cases over the next year, monthly, which comes to 3 cases per month.

Given that the total number of cases reported last year is N = 2,500, and the site will randomly select 36 cases from the 1,800 reported from Regions 1 and 3 (due to limited resources), the level of precision, δ, assuming 95% confidence limit, can be calculated using Formula #4 from Appendix D. The population proportion, P, is assumed to be 0.05.

Sampling Frame:	Only cases reported by Regions 1 & 3
Sampling Method:	Simple Random Sampling (SRS)
Stratification Variable:	None
Level of Precision:	0.07, or 7%
Allocation:	Equal
Timing of Sample:	Monthly
Number of Cases Reported Last Year:	2,500 Total, 1,800 from Regions 1 & 3
Number of Cases to Sample:	36, from Regions 1 & 3

		Number of Cases Proposed to Sample	Number of Cases Actually Sampled	**Number of Successful Re-Abstractions**
Month 1	At the end of each month, the site randomly selects 3 cases reported from either Region 1 or Region 3.	3	3	**3**
Month 2	Over the following month, re-abstraction is completed for 2 of the 3 cases, with the outstanding case located in Region 3. A day is set aside for a staff member to drive to Region 3 and re-abstract the last case.	3	3	**2**
Month 3	Same as Month 1	3	3	**3**
Month 4	Same as Month 1	3	3	**3**
Month 5	Same as Month 1	3	3	**3**
Month 6	Same as Month 1	3	3	**3**
Month 7	Same as Month 1	3	3	**3**
Month 8	Same as Month 1	3	3	**3**
Month 9	Same as Month 1	3	3	**3**
Month 10	Same as Month 1	3	3	**3**

		Number of Cases Proposed to Sample	Number of Cases Actually Sampled	**Number of Successful Re-Abstractions**
Month 11	Same as Month 1. Re-abstraction is completed for 2 of the 3 cases.	3	3	**2**
Month 12	At the end of 2004, Site B has 34 successfully re-abstracted cases from Regions 1 and 3.	3	3	**3**
Total		36	36	**34**

Example #3: Medium Morbidity Site, Reporting Concentrated by Providers

Site C had reported 3,200 HIV/AIDS cases in 2003. Four providers in the state had reported 3,000 cases, whereas the remaining 200 cases were reported by several providers. Site C has decided to use SRS at each of the four main providers. (This is stratifying on provider, and so it could also be defined as StRS, but there is a reason it is not. The site will calculate a total sample size for all four providers and then proportionally allocate that sample size into the four smaller sample sizes. The calculations are not being performed simultaneously, as would be the case if StRS was being used.) In the diagnosis year of 2003,

- Provider A reported 1,500 cases,
- Provider B reported 750 cases,
- Provider C reported 500 cases, and
- Provider D reported 250 cases.

The site plans to do their re-abstracting once a quarter; they will sample 1% of the total number of cases reported by the four major providers (3,000), which is 30 cases. They also decide to proportionally allocate the reporting (apply the percentage breakdown) across the four providers (50%, 25%, 16.7%, and 8.3%, respectively) to determine the sample size for each provider. This calculates to selecting 15, 8, 5, and 3 cases from Providers A, B, C, and D, respectively (which will sum to 31 cases total).

Given that the total number of cases reported last year is N = 3,200, and the site will randomly select 30 cases from the 3,000 reported from the four major providers, the level of precision, δ, assuming 95% confidence limit, can be calculated using Formula #4 from Appendix D. The population proportion, P, is assumed to be 0.05.

Sampling Frame:	Cases reported by each of four providers
Sampling Method:	Simple Random Sampling (SRS) at each level of provider (A, B, C, D)
Level of Precision:	0.078, or 7.8%, at the overall level
Allocation:	Proportional (Equal sampling weight in each stratum)
Timing of Sample:	Quarterly
Number of Cases Reported Last Year:	3,200 Total, 3,000 from the four providers
Number of Cases to Sample:	1% (30) from the four providers

Provider A				
		Number of Cases Proposed to Sample	Number of Cases Actually Sampled	**Number of Successful Re-Abstractions**
Quarter 1	Successful sampling and re-abstraction for the first quarter of all 5 cases.	5	5	**5**
Quarter 2	Same as above.	5	5	**5**
Quarter 3	Same as above.	5	5	**5**
Quarter 4	Site decides to go ahead and sample an extra 5 cases.	0	5	**5**
Total		15	20	**20**

Provider B				
		Number of Cases Proposed to Sample	Number of Cases Actually Sampled	**Number of Successful Re-Abstractions**
Quarter 1	Successful sampling and re-abstraction for the first quarter of all 3 cases.	3	3	**3**
Quarter 2	Same as above.	3	3	**3**
Quarter 3	Site decides to sample one extra case.	2	3	**3**
Quarter 4	Site decides to go ahead and sample an extra 3 cases.	0	3	**3**
Total		8	12	**12**

Provider C				
		Number of Cases Proposed to Sample	Number of Cases Actually Sampled	**Number of Successful Re-Abstractions**
Quarter 1	Successful sampling and re-abstraction for the first quarter of their one sampled case.	1	1	**1**
Quarter 2	Same as above.	1	1	**1**
Quarter 3	Same as above.	1	1	**1**
Quarter 4	Site decides to only sample one case, and not two.	2	1	**1**
Total		5	4	**4**

Provider D				
		Number of Cases Proposed to Sample	Number of Cases Actually Sampled	**Number of Successful Re-Abstractions**
Quarter 1	Successful sampling and re-abstraction for the first quarter of their one sampled case.	1	1	**1**
Quarter 2	Same as above.	1	1	**1**
Quarter 3	Site was unable to re-abstract the one sampled case.	1	1	**0**
Quarter 4	Site decides to sample one extra case to replace the one lost last quarter.	0	1	**2**
Total		3	4	**4**

Total				
		Number of Cases Proposed to Sample	Number of Cases Actually Sampled	**Number of Successful Re-Abstractions**
Quarter 1	At the end of March 2004, they randomly select 5 cases reported from Provider A, 3 from Provider B, 1 from Provider C, and 1 from Provider D. The staff is able to successfully re-abstract the 10 cases.	10	10	**10**

		Total		
		Number of Cases Proposed to Sample	Number of Cases Actually Sampled	**Number of Successful Re-Abstractions**
Quarter 2	In June 2004, they decide to again select 5, 3, 1, and 1 and again they are successful in re-abstraction. This leaves only 10 more cases for the rest of the year. Since the re-abstraction has not been difficult, the site decides to go ahead and sample the same number of cases for the next two quarters.	10	10	**10**
Quarter 3	After the September 2004 sampling and re-abstraction, they had successfully re-abstracted 9 of the 10.	9	10	**9**
Quarter 4	In December 2004, the same sampling is done, and they are able to re-abstract all 10 cases, plus the one they missed.	2	10	**11**
Total		31	40	**40**

Example #4: High Morbidity Site, Stratified Random Sampling

Site D is primarily interested in re-abstracting stratified by provider and region, and in the feasibility of re-abstracting stratified by staff. The site had reported over 10,000 cases in 2003. The site has 5 major and 10 minor regions, 35 major and 130 minor providers, and 20 staff members that work at abstracting case information. The difference between major and minor regions or providers relates to the number of reports; if the number is over 100 for a year, then it is assumed to be a major region/provider; if less than 100, then it is a minor region/provider. The site acknowledges that it could not stratify by region, provider, and staff, for fear of small cell counts (that would be a total of 15x165x20 = 49,500 strata) without needing to collapse categories of region and provider.

Because 30 of the 35 major providers are located in the 5 major regions of the state, the site decides not to stratify on provider, but on region. The site will then collapse regions into 6 categories: the 5 major regions and all of the minor regions. That comes out to 6 strata. If the site wanted to stratify by staff also, that would entail 6x20 = 120 strata; with this large number of strata, it would be difficult to maintain adequate stratum size, plus the complexity would be resource draining. The site decides to do their re-abstraction study stratified by region only.

The site decides, after discussing with their core Epi TA, that they will sample 2% of their overall total number of cases (about 200). They use the frequency distribution of case reporting of the 6 regions to determine the sampling weights to apply. The 6 regions

follow this distribution: 26%, 21%, 20%, 16%, 12%, and 5%, respectively. They decide to sample quarterly. At the end of each quarter, a stratified random sample of cases will be selected to re-abstract the case information during the next quarter.

Two methods will work to calculate sample size and level of precision. The site could select a specific number of cases, and then calculate levels of precision; or decide on a level of precision, and calculate the sample size (see table).

	Number of Cases Reported, N_i	Distribution of Cases Reported, $N_i/N*100$ %	Sample Size Calculated for $\delta = 0.02$	Sample Size Calculated for $\delta = 0.01$	Selected Sample Sizes (based on the distribution of cases)	Calculated Level of Precision, δ
Region 1	2,600	26%	30	123	52	0.015
Region 2	2,100	21%	20	80	42	0.014
Region 3	2,000	20%	18	73	40	0.014
Region 4	1,600	16%	12	47	32	0.012
Region 5	1,200	12%	7	26	24	0.010
Region 6 (all other regions)	500	5%	1	5	10	0.007
Total	10,000	100%	88	354	200	—

As the above table shows, the level of precision can greatly affect the sample sizes calculated for each of the 6 categories, and chosen sample sizes can greatly affect the level of precision. The question to be posed to the site would be which of these methods would be best used at their site, given their resource level and staffing. Also, sampling will need to be done for each of the 20 staff members. Will the site be able to sample 88, 200, or 354 cases for each staff member? Are all staff members reporting cases from all 6 categories of region? These questions were discussed at length with the core Epi TA, and it was decided that each staff member would re-abstract 200 cases. If a staff member did not report from a region, then that region could be skipped for that staff member, thereby making it possible for the staff member to have less than 200 cases to re-abstract.

Given that the total number of cases reported last year is $N = 10{,}000$, and the site has chosen to randomly select a total of 200 cases for each staff member, the level of precision, δ, assuming 95% confidence limit, is shown in the previous table, in the far right column. Formulas #8, #9, and #10 from Appendix E were used to calculate the numbers. The population proportion, P, is assumed to be 0.05.

Sampling Frame:	Cases stratified by region, 6 levels, by staff, 20 levels
Sampling Method:	Stratified Random Sampling (StRS)
Stratification Variable:	Region (6 levels)

Level of Precision:	Range is 0.007 to 0.015, or 0.7% to 1.5% for each region
Allocation:	Proportional (Equal sampling weights ineach stratum)
Timing of Sample:	Quarterly
Number of Cases Reported Last Year:	10,000 plus
Number of Cases to Sample:	2% of all cases for each staff member (200): 52 cases in Region 1 42 cases in Region 2 40 cases in Region 3 32 cases in Region 4 24 cases in Region 5 10 cases in Region 6

The next table is a transposed version of the tables used in the three prior examples; that is, time in quarters now progresses over subsequent columns, instead of descending rows. Also, the table only shows the numbers for Staff Person A; it is assumed a similar table was created for each staff person. Comments are below the table.

	Quarter 1	Quarter 2	Quarter 3	Quarter 4	**Total**
Staff Person A					
Region 1 (26% of cases)					
Number of Cases Proposed to Sample	13	13	13	13	**52**
Number of Cases Actually Sampled	13	13	13	13	**52**
Number of Successful Re-Abstractions	13	13	12	12	**50**
Region 2 (21% of cases)					
Number of Cases Proposed to Sample	10	10	11	11	**42**
Number of Cases Actually Sampled	10	10	11	11	**42**
Number of Successful Re-Abstractions	9	8	8	9	**34**
Region 3 (20% of cases)					
Number of Cases Proposed to Sample	10	10	10	10	**40**
Number of Cases Actually Sampled	10	10	10	10	**40**
Number of Successful Re-Abstractions	10	10	10	10	**40**
Region 4 (16% of cases)					
Number of Cases Proposed to Sample	8	8	8	8	**32**
Number of Cases Actually Sampled	0	0	0	0	**0**
Number of Successful Re-Abstractions	0	0	0	0	**0**

	Quarter 1	Quarter 2	Quarter 3	Quarter 4	**Total**
Region 5 (12% of cases)					
Number of Cases Proposed to Sample	6	6	6	6	**24**
Number of Cases Actually Sampled	6	6	6	6	**24**
Number of Successful Re-Abstractions	6	6	6	6	**24**
Region 6 (5% of cases)					
Number of Cases Proposed to Sample	2	3	3	2	**10**
Number of Cases Actually Sampled	2	3	3	2	**10**
Number of Successful Re-Abstractions	2	3	3	2	**10**
Total of Successful Re-Abstractions over all Regions:	40	40	39	39	**158**
Staff Person B					
…ditto…					

Comments: Staff Person A reports cases from all regions except Region 4. Over the year, for each quarter, cases were sampled and re-abstracted for each of the five out of six regions. Because of administrative constraints, no over sampling was done to replace nonresponders (unsuccessfully re-abstracted cases). As the table shows, 2 cases were lost in quarters 3 and 4 in Region 1, and 8 cases were lost in all four quarters in Region 4. This loss was attributed to the providers in this region not keeping records available for re-abstraction; i.e., records were lost or stored after a certain period.

The "Number of Cases Proposed to Sample" is constant for each staff member across regions. It is assumed that each staff person has the same probability of reporting from each region. This is not the reality, but it was felt that the overall number of cases re-abstracted should be enough for adequate estimation of error rates for the site as a whole. The site may want to return to the resulting sample and analyze the patterns of reporting and error rates over staff and region.

Appendix G

Variables for Re-Abstraction

Variables for Re-Abstraction

Variables	Factors to Be Compared	Major Error	Minor Error
		including failure to obtain information*	
Required			
Name	legal, aliases	errors that would result in incorrect Soundex	Junior, numbers (II, III, etc.)
Sex		any error	
Date of Birth	day, month, year	month or year	day
Race	multiple races	any error	
Ethnicity		any error	
Vital Status	alive/dead	any error	
Date of Death	day, month, year	month or year	day
Current Residence	street, city, county, state, ZIP Code	county, state	street
Residence at Diagnosis	street, city, county, state, ZIP Code	county, state	street
Earliest HIV Test	type, result, month, year	type, result (pos/neg), month, year	
Other HIV Tests	type, result, month, year		any error
Earliest Viral Load Test	type, result, month, year	year	
Other Viral Load Tests	type, result, month, year	type, result, month, year	any error
Physician Diagnosis	yes/no, month, year	month, year	
First CD4 Count <200, or	count, month, year	count, month, year	
First CD4 Percent <14%	percent, month, year	percent, month, year	
CD4 Count Closest to HIV Diagnosis			
Opportunistic Illnesses	definitive/presumptive; month, year	any omission	
Risk Factors		any omission	
Recommended			
Facility of Diagnosis			
Recent CD4 Count	count, month, year		
Recent CD4 Percent	percent, month, year		
Recent Viral Load Test	type, result, month, year		
Treatment			

***NOTE: Failure to obtain information (missing information) is also coded as a minor or major error.**

Appendix H

Re-abstraction Error Report Form, for Pairs of Reports for One Case

Re-Abstraction Error Report Form

STATENO: **Date of Sampling:**

Initial Report:	Re-Abstracted Report:
Date:	Date:
Staff:	Staff:

Check boxes if the re-abstracted information differs from the initial report. Please distinguish between non-errors, minor errors, and major majors. If there were any errors, mark the last box in the row.

Required Variables to Compare:

Data Elements		Difference	Minor Error	Major Error	Any Error
Name, Legal					
Name, Aliases (if applicable)					
Sex					
Date of Birth					
Ethnicity					
Race					
Vital Status					
Date of Death					
Current Residence					
Street Address					
City					
County					
State					
ZIP Code					
Residence at Diagnoses					
Street Address					
City					
County					
State					
ZIP Code					
Earliest HIV Test					
Type					
Result					
Date (Month and Year)					
Other HIV Tests					
Type					
Result					
Date (Month and Year)					
Earliest Viral Load Test					
Type					
Result					
Date (Month and Year)					
Other Viral Load Tests					
Type					
Result					
Date (Month and Year)					

Required Variables to Compare, continued:

Data Elements	Difference	Minor Error	Major Error	Any Error
Physician Diagnosis				
Result				
Date (Month and Year)				
First CD4 Count <200 or <14%				
Count / Percent				
Date (Month and Year)				
CD4 Count closest to HIV diagnosis				
Count				
Date (Month and Year)				
Risk Factors				
Sex with male				
Sex with female				
Injection Drug User (IDU)				
Received clotting factor				
Heterosexual contact (HC) with IDU				
HC with bisexual male				
HC with person with hemophilia				
HC with HIV+ transfusion recipient				
HC with HIV+ transplant recipient				
HC with HIV+, risk factor unspecified				
Received blood or blood components				
Transplant recipient				
Worked in health-care/laboratory setting				
Other exposure to human blood/body fluids				
Opportunistic Illnesses				
Diagnosis				
Date (Month and Year)				

Recommended Variables to Compare:

Facility at Diagnoses				
Name of Facility				
City				
County				
State				
ZIP Code				
Recent CD4 Count / Percent				
Count / Percent				
Date (Month and Year)				
Recent Viral Load Test				
Type				
Result				
Date (Month and Year)				
Treatment				
Total Errors:				

Appendix I

Re-abstraction Report Table, for Diagnosis Year 20XX

Re-Abstraction Report Table, for Diagnosis Year 20XX:

Site:

Date:

Below, fill in the total number of cases reported for this year, the total number of cases sampled, and the number of cases that were successfully re-abstracted. To the right, fill in the summary numbers for each box.

Total number of cases reported for this diagnosis year:

Total number of cases sampled:

		Number of cases without an error	Number of cases with at least 1 error	Number of cases with at least 1 major error	Agreement Rate	Error Rate (at least 1 error)	Error Rate (at least 1 major error)
Total number of cases successfully re-abstracted:							

Below, fill in the numbers of cases where the data element was compared between the initial report and the re-abstracted report. Then, fill in the number of cases that were in complete agreement, that had errors, that had a major error for that data element.

Data Elements		Number of Cases		Number of cases without an error	Number of cases with At Least 1 Error	Number of cases with At Least 1 Major Error		Agreement Rate	Error Rate (at least 1 error)	Error Rate (at least 1 major error)
Required Variables to Compare:										
Name, Legal										
Name, Aliases (if Applicable)										
Sex										
Date of Birth										
Ethnicity										
Race										
Vital Status										
Date of Death										
Current Residence										
Street Address										
City										
County										
State										
ZIP Code										
Residence at Diagnoses										
Street Address										
City										
County										
State										
ZIP Code										
Earliest HIV Test										
Type										
Result										
Date (Month and Year)										
Other HIV Tests										
Type										
Result										
Date (Month and Year)										
Earliest Viral Load Test										
Type										
Result										
Date (Month and Year)										
Other Viral Load Tests										
Type										
Result										
Date (Month and Year)										

Data Elements	Number of Cases	Number of cases without an error	Number of cases with At Least 1 Error	Number of cases with At Least 1 Major Error	Agreement Rate	Error Rate (at least 1 error)	Error Rate (at least 1 major error)
Required Variables to Compare, continued:							
Physician Diagnosis							
Result							
Date (Month and Year)							
First CD4 Count <200 or <14%							
Count / Percent							
Date (Month and Year)							
CD4 Count closest to HIV diagnosis							
Count							
Date (Month and Year)							
Risk Factors							
Sex with male							
Sex with female							
Injected nonprescription drugs (IDU)							
Received clotting factor							
Heterosexual contact (HC) with IDU							
HC with bisexual male							
HC with person with hemophilia							
HC with HIV+ transfusion recipient							
HC with HIV+ transplant recipient							
HC with HIV+, risk factor unspecified							
Received blood or blood components							
Transplant recipient							
Worked in health-care/laboratory setting							
Other exposure to human blood/body fluids							
Opportunistic Illnesses							
Diagnosis							
Date (Month and Year)							
Recommended Variables to Compare:							
Facility at Diagnoses							
Name of Facility							
City							
County							
State							
ZIP Code							
Recent CD4 Count / Percent							
Count / Percent							
Date (Month and Year)							
Recent Viral Load Test							
Type							
Result							
Date (Month and Year)							
Treatment							

Notes

Technical Guidance for HIV/AIDS Surveillance Programs

Pediatric HIV/AIDS Surveillance

HIV Incidence and Case Surveillance Branch
Atlanta, Georgia

Contents — Pediatric HIV/AIDS Surveillance

Background 11-3
Introduction 11-3
Diagnosing HIV Infection in Children <13 Years of Age 11-4
Case Definition 11-5
The CDC Pediatric Case Report Form 11-5
Structural Requirements 11-6
Data Collection for Children <13 Years of Age 11-6
Structural Requirements 11-6
Data Collection for HIV Exposure Reporting 11-8
Structural Requirements 11-8
Process Standards for Identification and Follow-up of Pediatric HIV/AIDS Cases 11-8
Pediatric HIV/AIDS Case Reporting (Children <13 Years of Age) 11-9
Pediatric HIV Case Finding (Children <13 Years of Age) 11-9
Information Collection and Follow-up 11-9
Process Standards for Data Collection for Children <13 Years of Age 11-10
Pediatric Cases of Public Health Importance (CoPHI) 11-11
Perinatal Exposure Reporting for Infants <18 Months of Age (recommended) 11-11
Exposure Case Finding for Infants <18 Months of Age 11-11
Information Collection 11-11
Process Standards for Data Collection for HIV Exposure Reporting 11-12
Outcome Measures 11-12
Data Collection for Children <13 Years of Age 11-12
Data Collection for HIV Exposure Reporting 11-13
References 11-14

Technical Guidance for HIV/AIDS Surveillance Programs — Pediatric HIV/AIDS Surveillance

Background

Introduction

The Centers for Disease Control and Prevention (CDC), in conjunction with the Council of State and Territorial Epidemiologists (CSTE), recommends that all states and territories extend their pediatric AIDS case surveillance activities to include HIV case surveillance and the reporting of HIV-exposed infants[1–3]. In addition, to facilitate follow-up to assess infection status and access to care, the American Academy of Pediatrics issued a policy statement in 1998 in support of all states and territories conducting pediatric HIV surveillance to include reporting of all perinatally exposed infants[4]. Surveillance for perinatal HIV exposure, HIV infection, and AIDS enables timely and complete monitoring of the effectiveness of perinatal HIV prevention efforts[5]. These surveillance data are also used to monitor HIV incidence trends, identify groups in which prevention strategies are less successful, evaluate the impact of zidovudine (ZDV) and other antiretroviral therapy on perinatal HIV incidence, assess resources required to provide care for HIV-exposed children, examine the timeliness of receipt of HIV-related care, and evaluate potential short- or long-term adverse effects of in utero exposure to ZDV and other antiretroviral therapy[6].

Since the early 1980s, all states and territories of the United States have conducted routine case surveillance for pediatric AIDS for children <13 years of age. As of April 2006, pediatric HIV infection surveillance is also being conducted in 48 areas that use name-based HIV infection reporting and 8 areas that use coded patient identifiers. In addition, 29 areas also conduct population-based perinatal HIV exposure surveillance for infants born to HIV-infected mothers and follow the infants until serostatus of the child is determined, which is normally confirmed by 18 months of age. State and local health departments conduct follow-up of these exposure cases to determine the HIV infection status of the child and progression to AIDS. Most of the cumulative pediatric AIDS cases reported, and virtually all new pediatric HIV infections, are perinatally acquired.

In July 1995, the U.S. Public Health Service (PHS) published recommendations for HIV counseling and voluntary testing for all pregnant women[7]. These recommendations provide guidance to health care professionals on preventing mother-to-child transmission and educating women about the importance of knowing their HIV infection status. Revised recommendations for HIV screening of pregnant women that further emphasized HIV testing as a routine part of prenatal care, including rapid testing at labor and delivery, were published in 2001[8]. Advances in the early 1990s in preventing perinatal HIV infection and in diagnosing and treating HIV in infants also resulted in improved diagnostic methods for HIV-exposed infants, such as the polymerase chain reaction (PCR) and virus culture assays for detecting HIV infection in children born to infected mothers[9].

In 2005, the Public Health Service Task Force published updated guidelines for use of antiretroviral drugs in pregnant women[10]. The revisions include 1) updated information on nevirapine and hepatic/rash toxicity and recommendations regarding the use of nevirapine-based highly active antiretroviral therapy (HAART) for prevention of mother-to-child transmission of HIV; 2) updates to Clinical Scenario #1 regarding the use of efavirenz and nevirapine during pregnancy and change in efavirenz's Food and Drug Administration (FDA) pregnancy classification (reclassified to Pregnancy Class D—positive evidence of fetal harm); and 3) updates to the abacavir, lamivudine, stavudine, zidovudine, efavirenz, nevirapine, indinivir, and nelfinavir sections of the *Safety and Toxicity of Individual Antiretroviral Agents in Pregnancy* supplement. See http://aidsinfo.nih.gov.

Diagnosing HIV Infection in Children <13 Years of Age

In children ≥18 months of age, standard immunoglobulin G (IgG) antibody tests can be used to diagnose HIV infection. However, the diagnosis of HIV infection in children <18 months of age is complicated by the passive transfer of maternal HIV IgG antibody across the placenta to the fetus, which may persist as long as 18 months. Therefore, standard enzyme immunoassay (EIA) antibody tests cannot be used to determine HIV infection in this age group.

In recent years, significant advances have been made in diagnostic tests for HIV infection in infants that can reliably detect HIV infection in virtually all infected infants by 1–2 months of age. DNA polymerase chain reaction is the preferred method for diagnosis of HIV infection and has been recommended for use in perinatally exposed children from birth onwards. When the test is performed by experienced virology laboratories, its sensitivity can be as high as 38% by 48 hours of age, 93% by 14 days of age, and 96% by 4 weeks of age in the absence of breast feeding[11]. The standard tests for p24 antigen are specific for HIV infection; however, the sensitivity for p24 antigen is less in the first months of life. Immune-complex-dissociated, p24 antigen-capture assay has increased the sensitivity of this test, but data concerning sensitivity in early infancy are limited.

With HIV nucleic acid detection tests, HIV infection can be detected in nearly all infants over the age of 1 month. The timing of the HIV serologic and HIV nucleic acid detection tests and the number of HIV nucleic acid detection tests needed to determine HIV infection are based on evaluations of the performance of these tests for children in this age group (1 month – 18 months)[12–22].

In document-based data management, all HIV tests are recorded for children <13 years of age, including the first test, tests at each diagnostic status (i.e., perinatal HIV exposure, HIV infection, AIDS, or not infected with HIV), and the most recent tests. The software will classify perinatal cases (in infants <18 months of age) as perinatal HIV exposure, HIV-infected, not infected with HIV, or AIDS according to the 1999 pediatric HIV case definition[1], the 1994 pediatric HIV classification system[9], and the 1987 AIDS case definition[23]. For HIV cases among children aged ≥18 months but <13 years, the 1999 revised criteria for adults, adolescents, or children aged ≥18 months apply (Guidelines for national human immunodeficiency virus case surveillance, including monitoring for human immunodeficiency virus infection and acquired immunodeficiency syndrome)[1].

For perinatally exposed children who are not infected with HIV, all negative viral detection and negative antibody tests are included in the document-based system. Some children develop severe clinical conditions resulting from HIV infection before their infection status has been sufficiently established. For the purposes of classification, a child aged <18 months born to a woman with an HIV infection that was documented during pregnancy or at delivery and meeting the criteria for AIDS in the 1987 case definition should be considered HIV-infected—even in the absence of definitive laboratory assays results.

Case Definition

In 1999, in collaboration with CSTE, CDC revised the case definition for HIV infection in adults and children, which includes surveillance criteria for HIV infection and incorporates the surveillance criteria for AIDS [23–25]. These changes to the case definition and guidelines for case surveillance are described in the December 10, 1999, *MMWR* entitled, Guidelines for National Human Immunodeficiency Virus Case Surveillance, Including Monitoring for Human Immunodeficiency Virus Infection and Acquired Immunodeficiency Syndrome[1].

For adults and children aged ≥18 months, the HIV surveillance case definition includes laboratory and clinical evidence specifically indicative of HIV infection and severe HIV disease (AIDS). For children aged <18 months (except for those who acquired HIV infection other than by perinatal transmission), the HIV surveillance case definition updates the definition in the 1994 revised classification system. In addition, the new case definition is based on recent data regarding the sensitivity and specificity of HIV diagnostic tests in infants and clinical guidelines for Pneumocystis carinii pneumonia (PCP) prophylaxis for children and for use of antiretroviral agents for pediatric HIV infection. The revised surveillance case definitions for adults and children became effective January 1, 2000[1].

Importantly, the pediatric AIDS case definition does NOT include immunologic criteria (i.e., CD4 cell counts less than 200/µL or less than 14% as in adults/adolescents) for case-defining purposes. (Per the preceding paragraph, severe clinical conditions indicative of AIDS have been incorporated into the pediatric surveillance case definition for HIV infection[1].) However, CD4 test results are useful information to collect for children, and these laboratory data are included on the Pediatric HIV/AIDS Confidential Case Report Form. These data should be recorded on the report form whenever they are posted in the child's medical records or when laboratory documents are added to the database. The Pediatric CRF contains further instructions on completing the Pediatric HIV/AIDS Confidential Case Report Form for surveillance.

The CDC Pediatric Case Report Form

The CDC Pediatric HIV/AIDS Confidential Case Report is a double-sided, two-page foldout form designed to collect, in a confidential manner, information that will assist in understanding HIV infection and AIDS among children aged <13 years. For further information on the CDC Pediatric Case Report Form, see the Pediatric Case Report Form instructions.

Structural Requirements

Data Collection for Children <13 Years of Age

Structural Requirements

Personnel

- Epidemiologist/supervisor/program manager to oversee activities
- Field staff to conduct medical record reviews, quality assurance re-abstractions
- Staff to conduct quality assurance review of forms and data entry
- Staff to manage databases to create/maintain system for tracking cases that need to be investigated or that require follow-up, and to provide reminders for pending cases (recommended)
- Programmer to match the HIV/AIDS Reporting System (HARS)/eHARS database with public health registries (specifically, birth registry records) and other data sources that could help identify potential cases or potential adverse events related to treatments; create data quality control reports/check mechanisms
- Epidemiologist to analyze local data for local planning, implementation of interventions, and program evaluation

Resources

- Offices, telephones, storage, etc. for staff consistent with the requirements in the Security and Confidentiality Guidelines
- Information technology (IT) support for staff
- Appropriate data management, data analysis, and encryption software (with current licenses)
- Mechanisms to distribute cases to field for investigation (certified mail/courier, or ability to download from secure site) and for cases to be returned to the program for data review/data entry/transfer to CDC
- Secure networks for storing databases and related files, and a mechanism to securely back up databases
- Resources to train field and data entry staff on the information that is to be abstracted from medical records and entered into data system

Equipment

- Computers/printers for staff
- Secure filing cabinets
- Access to equipment for presenting data to prevention and services partners, planning groups, and personnel at staff and statewide meetings

Partnerships

- Relationships with reporting sites so that cases are actively reported and follow-up investigations can occur without barriers
- Relationships with laboratories to send notification of positive HIV/AIDS results to the health department
- Memoranda of understanding (MOUs) or other agreements with registries, public health programs, and facilities to be able to link with HARS/eHARS for case ascertainment
- Collaborative relationships with HIV prevention and services programs so that data can be shared and used for program planning and evaluation
- Relationships with partner counseling and referral services (PCRS) staff at local health departments to provide a referral resource for HIV-positive pregnant women and their exposed partners and infants

Data Management Systems (recommended)

- A data management system so that local data are accessible for timely local analysis for quality assurance (QA), planning, and evaluation efforts
- A data management system to track cases pending investigation or follow-up

Legislation/Regulations

- Required reporting of all test results possibly indicative of HIV infection
- Local institutional review board (IRB) approvals needed to conduct registry matches or linkages (necessary in areas where such practices are not considered routine components of public health practice)

Data Management

- HARS database in place at the local site
- Data entry staff to enter information on pediatric cases
- Staff to develop quality assurance reports and analyze data for local use

Data Analysis

- Biostatistician or epidemiologist to analyze local data and to develop standard pediatric surveillance reports/local profiles, program reports (annual updates, annual reports, regional annual reports, etc.)
- Policies and procedures for custom data requests and analyses

Data Dissemination

- Staff trained and available to present surveillance data at perinatal prevention meetings, community planning groups, annual meetings, and as requested by other organizations
- Stakeholders and partners identified for distribution of reports; established distribution list and methods (mail, electronic, etc.)

- Perinatal surveillance data posted on Internet website
- Database to document presentation, dissemination of data
- Software, laptop, and projection equipment for presentations

Data Collection for HIV Exposure Reporting

(in addition to requirements listed under section on Data Collection for Children <13 Years of Age)

Structural Requirements

Partnerships

- Relationships with reporting sites for active reporting of maternal HIV/AIDS cases and follow-up investigations
- Relationships with laboratories for notification of maternal HIV/AIDS results to health department for follow-up
- Relationships with PCRS staff for notification of partners who have had contact with an HIV-positive pregnant female to make appropriate referrals
- Relationships with Ryan White-funded programs that serve pregnant women and their children

Legislation/Regulations (recommended)

- State surveillance provisions should either require or allow for investigation of cases of perinatal exposure.
- HIV and AIDS cases are recommended to be reportable by name.
- All HIV test results for children under 3 years of age should be reported to resolve “exposed” cases that may have been determined as HIV-negative.

Data Management

- Develop mechanism for tracking estimated delivery dates of HIV-positive pregnant women to facilitate epidemiologic follow-up. Last menstrual period (LMP) and estimated date of confinement (EDC) are needed to do this type of tracking.
- Develop database to record the testing history of exposed infants to prompt longitudinal follow-up of children without definitive HIV status.

Process Standards for Identification and Follow-up of Pediatric HIV/AIDS Cases

The surveillance processes follow those described in Access to Source Data and Completeness of Reporting. However, there are additional considerations for pediatric case finding for children <13 years of age (including perinatal cases), perinatal exposure reporting (for infants <18 months of age), data collection, and follow-up.

Pediatric HIV/AIDS Case Reporting (Children <13 Years of Age)

All state and local programs should collect a standard set of surveillance data for all pediatric cases (children <13 years of age) that meet the reporting criteria for HIV infection and AIDS. The standard data set includes 1) patient identifier, 2) earliest date of diagnosis of HIV infection, 3) earliest date of diagnosis of an AIDS-defining condition, 4) demographic information (e.g., date of birth, race/ethnicity, and sex), and residence (i.e., city and state) at diagnosis of HIV infection and of AIDS, 5) HIV risk factor(s), 6) facility of diagnosis, and 7) date of death and state of residence at death. For perinatal cases, information on birth and maternal history is also collected on the pediatric case report form. In addition, the date(s) of HIV diagnostic testing, the results of these tests, and information on exposure to antiretroviral treatment for reducing perinatal HIV transmission should be collected for all children with perinatal exposure to HIV[1]. New information should also be updated in the birth history section of the maternal record in HARS/eHARS. For information on required, recommended, and optional variables, see the Pediatric Case Report Form instructions.

Information on children under the age of 13 years should be collected for four time points in the course of HIV disease: 1) perinatal HIV exposure, 2) diagnosis of HIV infection, 3) diagnosis of AIDS, and 4) death.

Pediatric HIV Case Finding (Children <13 Years of Age)

CDC recommends active case finding of HIV-infected children <13 years of age. Active case finding consists of health department staff regularly contacting reporting facilities (hospitals, clinics, physician offices, laboratories) to identify HIV/AIDS cases. Health department staff review medical records at provider sites or receive information from sites to identify an HIV/AIDS case and to elicit information for HIV/AIDS case report forms. In either active or passive surveillance, collection of HIV or AIDS case data is usually mandated by state law. Instructions for data collection can be found in Instructions for Completing the Data Collection Form and in the sections under Specific Data Collection Guidelines. Specific data collection activities for pediatric surveillance are outlined below.

Active case finding of perinatally exposed children often requires review of both maternal and pediatric records. These records include, but are not limited to, prenatal care records, maternal HIV clinic records, labor and delivery records, birth records, all pediatric medical records, birth certificate, death certificate, and health department records. In both name-based and code-based perinatal HIV reporting sites, patient identifiers should be used to conduct record linkages (see Record Linkage).

Information Collection and Follow-up

Patient information should be recorded, reported, and updated to reflect the earliest date that the child meets each reporting event (i.e., perinatal HIV exposure, HIV infection, AIDS, or not infected with HIV) or changes in diagnostic or vital status. When a child who was previously reported as HIV-infected progresses to AIDS or dies, the information should be obtained and the surveillance record updated.

Similarly, when an infant <18 months of age who was previously reported as perinatally exposed has been either diagnosed with HIV infection or has seroreverted, this information should be recorded.

Process Standards for Data Collection for Children <13 Years of Age

- Ongoing passive surveillance of pediatric HIV/AIDS cases; i.e., providers (obstetric clinics, etc.) who care for pregnant women with HIV and their newborns send case reports to health department staff; an annual review of reporting sites to assess compliance with reporting requirements as indicated in Access to Source Data and Completeness of Reporting
- Ongoing active surveillance of pediatric HIV/AIDS cases; i.e., health department staff regularly contact reporting facilities (hospitals, clinics, physician offices, laboratories) to identify infants/children with HIV or AIDS; at least an annual review of surveillance staff activities
- Matches between HIV/AIDS surveillance databases and vital records (birth/death) and birth defects registries to identify perinatal cases not reported to the surveillance system

Data Management

- Data fields in HARS/eHARS to track pediatric cases of HIV, AIDS, or perinatal HIV exposure
- All pediatric cases entered into HARS/eHARS database within one month of receipt of case reports
- Each additional laboratory entry to generate an automatic update of the HARS/eHARS record

Data Analysis

- Annual analysis of surveillance data to report on 1) demographics, risk factors, and clinical diagnosis of HIV/AIDS-infected pediatric cases (sex, race/ethnicity, age [date of birth], residence [ZIP code, city, county], risk/mode of HIV transmission); 2) trends (month and year of diagnosis, opportunistic infections, CD4 counts, and other conditions); and 3) rates of infection, incidence, prevalence, and survival
- Ad hoc analyses of pediatric HIV surveillance data in response to requests

Data Dissemination

- Participate in regular work group meetings to disseminate perinatal data
- Disseminate and present annual statewide and regional reports via e-mail, mail, and website
- Respond to special presentation requests as needed

Pediatric Cases of Public Health Importance (CoPHI)

Local reporting staff should continue to discuss certain priority cases directly with CDC surveillance staff. These include cases of HIV infection in a health care setting; infection with HIV-2; infection attributed to tissue or organ transplantation; suspected transmission due to sexual contact; mother-to-infant transmission due to breast feeding; infections due to blood transfusions after March 1985; or any unusual transmission circumstances. This direct communication with CDC will ensure timely technical support. For details, please see Cases of Public Heath Importance (COPHI) and Appendix C, Procedures for Investigations of HIV/AIDS Cases of Epidemiologic, Diagnostic, or Special Public Health Importance in the Risk Factor Ascertainment chapter.

Perinatal Exposure Reporting for Infants <18 Months of Age (recommended)

Perinatal exposure reporting for infants <18 months of age differs from reporting of pediatric cases of HIV infection among children <13 years of age, as HIV-exposed infants may or may not be infected with HIV. Antibody tests will be positive at birth because of transplacental transfer of maternal antibodies, regardless of the infant's true HIV status.

The types of information collected on the pediatric HIV/AIDS case report form include dates of maternal HIV tests, receipt of prenatal care, and maternal and neonatal use of ZDV and other antiretrovirals during pregnancy. Infants known to be HIV-exposed are monitored after birth up to 18 months of age to determine whether they have been infected with HIV. Common ways that mother-infant pairs are identified include birth registry linkages (e.g., programs matching HIV cases in women of childbearing age with the birth registry database), reports of HIV-infected pregnant women being sent to surveillance programs, and review of hospital discharge summaries.

Exposure Case Finding for Infants <18 Months of Age

As part of surveillance activities, health department staff should regularly contact delivery hospitals and clinics to identify potential/suspect HIV/AIDS perinatal exposure cases, as consistent with local/state reporting laws. Health department staff should review medical records at provider sites or receive notification from sites to obtain all relevant information for maternal HIV/AIDS cases. Active reporting systems should also be established with laboratories, pediatric clinics, and HIV clinics to receive testing results on exposed infants, including all positive and negative PCRs. Matching known HIV-positive women to birth certificate registries helps identify unreported perinatally exposed children. If an HIV-infected woman gave birth to an exposed infant who was not reported, the case should be sent to appropriate surveillance staff for follow-up.

Information Collection

A pediatric HIV/AIDS case report form should be completed for each child born to a diagnosed HIV-infected mother. This includes infants <18 months of age whose infection status has not yet been determined and those determined not to be infected with HIV; inclusion of such information is for public health surveillance purposes

only. State laws protect the confidentiality of local information on HIV-exposed infants; on the national level a federal assurance of confidentiality applies to information on infants exposed perinatally with or without consequent infection.

On the case report form surveillance staff should indicate "perinatally HIV exposed" if the patient is <18 months of age, was born to an HIV-infected mother, and does not meet the criteria for HIV infection or the criteria for "not infected with HIV."

More information on data collection can be found in Volume II: Data Collection Resources and Reporting. Additional specific data collection activities for perinatal exposure reporting for infants <18 months of age are outlined below.

Process Standards for Data Collection for HIV Exposure Reporting

(in addition to requirements listed under Data Collection for Children <13 Years of Age)

- Abstraction of maternal and pediatric records occurs within 2 months of the anticipated delivery date.

Data Management

- All infants with perinatal HIV exposure are followed periodically until serostatus of the child is determined for 18 months after birth or loss to follow-up.
- Estimated delivery dates of expectant mothers are routinely entered in local data fields; data are reviewed monthly and distributed to staff for follow-up (recommended); any HIV test results (positive, negative, indeterminate) are tracked.

Data Analysis

- Maternal/perinatal data are analyzed at least annually for trends in perinatal exposure and infection.

Data Dissemination

- Staff produce annual report on trends in the implementation of USPHS recommended prophylactic treatments and trends in transmission.

Outcome Measures

Data Collection for Children <13 Years of Age

- Completeness of case ascertainment as described in Access to Source Data and Completeness of Reporting.
  - Minimum performance standard of ≥85% of expected number of cases for a diagnosis year are reported by 12 months after the diagnosis year.
  - Target performance standard of ≥95% of expected number of cases for a diagnosis year are reported by 12 months after the diagnosis year.

- Timeliness standards as described in Access to Source Data and Completeness of Reporting.
  - Minimum performance standard of ≥66% of cases for a diagnosis year are reported within 6 months of diagnosis, and assessed at ≥85% completeness 12 months after the diagnosis year.
- Annually, the program monitors the source of pediatric HIV/AIDS cases as specified by the outcome standards. Results of these evaluations should be used to improve the system. Checks on the source of HIV/AIDS cases include these:
  - audits
  - measures to record and tally the number of HIV/AIDS cases identified through passive surveillance, active surveillance, or through a registry match, e.g., through creation of a tracking database
  - regular quarterly assessments of reported vs. expected number of cases received from a reporting source

Data Dissemination

- Use pediatric surveillance data for planning purposes for HIV/AIDS prevention and care services
- Use pediatric surveillance data for allocation of funds

Data Collection for HIV Exposure Reporting

(in addition to requirements listed under Data Collection for Children <13 Years of Age)

- Completeness of case ascertainment as described in Access to Source Data and Completeness of Reporting.
  - Minimum performance standard of ≥85% of expected number of cases for a diagnosis year are reported by 12 months after the diagnosis year.
  - Target performance standard of ≥95% of expected number of cases for a diagnosis year are reported by 12 months after the diagnosis year.
  - Minimum performance standard of ≥85% of cases of indeterminate serostatus are identified as being a positive or negative case by 3 years after the diagnosis year.
- Timeliness standards as described in the Access to Source Data and Completeness of Reporting.
  - Minimum performance standard of ≥66% of cases for a diagnosis year are reported within 6 months of diagnosis, and assessed at ≥85% completeness 12 months after the diagnosis year.

- Annually, the program monitors the source of mother/infant pairs as specified by the outcome standards. Results of these evaluations should be used to improve the system. Checks on the source of mother/infant pairs include these:
  - audits
  - measures to record and tally the number of mother/infant pairs identified through passive surveillance, active surveillance, or through a registry match, e.g., through creation of a tracking database
  - regular quarterly assessments of reported vs. expected number of mother/infant pairs received from a reporting source

Data Management

- Staff follow up on expected births in the pregnancy tracking database to determine outcome of pregnancy and initiate abstraction of pediatric records (if applicable).
- Staff review pediatric records until a definitive HIV infection status is found, or every 6 months for 18 months following birth or loss to follow-up.

Data Analysis

- Analyses are used to set priorities for follow-up on indeterminate cases and check progress from previous period.

References

1. CDC. Guidelines for national human immunodeficiency virus case surveillance, including monitoring for human immunodeficiency virus infection and acquired immunodeficiency syndrome. *MMWR* 1999; 48(RR-13): 1–31.

2. Council of State and Territorial Epidemiologists. CSTE: Position Statement ID-6. Pediatric HIV Infection: Addition to the National Public Health Surveillance System (NPHSS); 1995.

3. Council of State and Territorial Epidemiologists. CSTE: Position Statement ID-4. Surveillance for perinatal HIV exposure; 2002.

4. American Academy of Pediatrics. Surveillance of pediatric HIV infection. *Pediatrics* 1998; 101:315–319.

5. Lindegren ML, Fleming P, Steinberg S, et al. Implementation of United States Public Health Service (USPHS) recommendations to prevent perinatal HIV transmission: pediatric HIV case surveillance, US. In: Program and Abstracts of the Interscience Conference on Antimicrobial Agents and Chemotherapy, Toronto, Canada; September 1997 (Abstract I-118).

6. CDC. Update: perinatally acquired HIV/AIDS—United States, 1997. *MMWR* 1997; 46(46):1086–1092.

7. CDC. U.S. Public Health Service recommendations for human immunodeficiency virus counseling and voluntary testing for pregnant women. *MMWR* 1995; 44(RR-7): 1–15.

8. CDC. Revised recommendations for HIV screening of pregnant women. *MMWR* 2001; 50(RR-19):59–86.

9. CDC. 1994 Revised classification system for human immunodeficiency virus infection in children less than 13 years of age. *MMWR* 1994; 43(RR-12):1–10.

10. Perinatal HIV Guidelines Working Group. Public Health Service Task Force recommendations for use of antiretroviral drugs in pregnant HIV-1-infected women for maternal health and interventions to reduce perinatal HIV-1 transmission in the United States. Rockville (MD): U.S. Public Health Service; November 17, 2005.

11. Dunn DT, Brandt CD, Krivine A, et al. The sensitivity of HIV-1 DNA polymerase chain reaction in the neonatal period and the relative contributions of intrauterine and intrapartum transmission. *AIDS* 1995; 9:F7–F11.

12. CDC. Guidelines for the use of antiretroviral agents in pediatric HIV infection. *MMWR* 1998; 47(RR-4):1–31.

13. Nesheim S, Lee F, Kalish ML, et al. Diagnosis of perinatal human immunodeficiency virus infection by polymerase chain reaction and p24 antigen detection after immune complex dissociation in an urban community hospital. *J Infect Dis* 1997; 175:1333–1336.

14. Steketee R, Abrams EJ, Thea DM, et al. Early detection of perinatal human immunodeficiency virus (HIV) type 1 infection using HIV RNA amplification and detection. *J Infect Dis* 1997; 175:707–711.

15. McIntosh K, Pitt J, Brambilla D, et al. Blood culture in the first 6 months of life for the diagnosis of vertically transmitted human immunodeficiency virus infection. *J Infect Dis* 1994; 170:996–1000.

16. Bremer JW, Lew JF, Cooper E, et al. Diagnosis of infection with human immunodeficiency virus type 1 by a DNA polymerase chain reaction assay among infants enrolled in the Women and Infants Transmission Study. *J Pediatr* 1996; 129:198–207.

17. Delamare C, Burgard M, Mayaux M, et al. HIV-1 RNA detection in plasma for the diagnosis of infection in neonates. *J Acquir Immune Defic Syndr Hum Retrovirol* 1997; 15:121–125.

18. Rich KD, Janda W, Kalish L, et al. Immune complex-dissociated p24 antigen in congenital or perinatal HIV infection: role in the diagnosis and assessment of risk of infection in infants. *J Acquir Immune Defic Syndr Hum Retrovirol* 1997; 15:198–203.

19. McIntosh K, Comeau A, Wara D, et al. The utility of IgA antibody to HIV-1 in early diagnosis of vertically transmitted infection. *Arch Pediatr Adolesc Med* 1996; 150:598–602.

20. Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science* 1996; 271:1582–1586.

21. Simonds RJ, Brown TH, Thea DM, et al. Sensitivity and specificity of a qualitative RNA detection assay to diagnose HIV infection in young infants. *AIDS* 1998; 12:1545–1549.

22. Young NL, Shaffer N, Chaowanachan T, et al. Early diagnosis and viral dynamics in HIV-1 infected infants in Thailand using RNA and DNA PCR assays sensitive to non-b subtypes. In: Program and Abstracts of 6th Conference on Retroviruses and Opportunistic Infections, Chicago, Illinois; February 1999 (Abstract 181).

23. CDC. Revision of the CDC surveillance case definition for acquired immunodeficiency syndrome. *MMWR* 1987; 36 (suppl 1): 1–15.

24. CDC. 1993 Revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *MMWR* 1992; 41(RR-17).

25. CDC. Classification system for human immunodeficiency virus (HIV) infection in children under 13 years of age. *MMWR* 1987; 36:225–230. 235–236.

Technical Guidance for HIV/AIDS Surveillance Programs

HIV Incidence Surveillance

HIV Incidence and Case Surveillance Branch
Atlanta, Georgia

Notes

Contents — HIV Incidence Surveillance

Policies and Procedures for HIV Incidence Surveillance 12-5
Background 12-5
Introduction 12-5
Structural Requirements 12-6
Policies and Procedures 12-6
The BED Assay for STARHS 12-7
Testing and ARV Use History Data 12-8
Staffing Needs 12-8
Process Standards 12-10
Obtaining Testing and ARV Use History Data 12-11
Obtaining Remnant HIV-Positive Specimens for Testing Using STARHS 12-12
Determining the Disposition of Specimens and Communicating with the Public Health and CDC STARHS Laboratories 12-13
Entering Data into Surveillance Databases 12-14
Creating the HIS Data Set and Transferring Data to CDC 12-15
Ensuring Security and Confidentiality 12-16
Analyzing and Disseminating Data 12-16
Training Staff 12-17
Outcome Standards 12-17
References 12-18

Epidemiologic Flow Chart for HIV Incidence Surveillance 12-19

Operational Public Health Laboratory Flow Chart for HIV Incidence Surveillance 12-20

National Center for HIV, STD, and TB Prevention's Non-research Determination for HIV Incidence Surveillance 12-21

Standard HIV Incidence Surveillance Data Elements 12-23

Guidance for the Remnant HIV-Positive Specimen Transportation Activities for HIV Incidence Surveillance 12-26
Purpose 12-26
Introduction 12-26
Laboratory Types 12-26
Laboratory Types 12-27
Specimen Information 12-27
Type of Specimens Shipped to CDC STARHS Laboratory 12-27
Specimen Volume 12-27
Sample Storage 12-27
Specimen Numbering 12-28
Specimen Retention 12-28
Packaging and Shipping Procedures 12-28
Shipping Guidance 12-28

Frequency of Shipments 12-28
Shipping Couriers 12-28
Coordinating Shipments 12-29
Additional Information for Commercial Laboratories Only 12-29
Tracking Shipments 12-29
Additional Information Related to APHL-Contracted Commercial Laboratories Only 12-29
Sample Rejection Criteria 12-30
Confidentiality and HIPAA Regulations 12-30
Funding for Specimen Handling 12-30
Specimen Transport Options 12-31
Option A: Specimen Originated at Private Laboratory and Is Sent Directly to the CDC STARHS Laboratory 12-31
Option B: Specimen Originated at or Sent Via State / Local Public Health Laboratory 12-34
Responsibilities 12-38
Private Laboratories 12-38
Selection of a Model Type 12-38
Additional Laboratory Report Information 12-38
Public Health Laboratories 12-39
Sample Storage and Retention 12-39
Aliquoting and Sample Shipment 12-39
State/Local HIV Incidence Surveillance Coordinator 12-39
Sample Disposition 12-39
Data Management 12-39
CDC STARHS Laboratory 12-39
Sample Rejection Criteria 12-39
Sample Storage and Retention 12-39
Results Reporting 12-40
References 12-40

Guidance for Processing, Storage, and Shipping of Specimens to the CDC STARHS Laboratory 12-44
Purpose 12-44
Introduction 12-44
CDC STARHS Laboratory 12-44
Setting and Personnel for Specimen Processing 12-44
Materials 12-45
Specimen Collection and Processing 12-45
Shipping 12-46
Packing Procedures for Shipping to the CDC STARHS Laboratory 12-46

HICSB Incidence Surveillance STARHS Specimen Submission Form 12-49

Training and Certification for Shipping Infectious Substances 12-50

Technical Guidance for HIV/AIDS Surveillance Programs — Policies and Procedures for HIV Incidence Surveillance

Background

Introduction

Although historically AIDS surveillance data have been of great value, current data do not represent the entire population affected by the HIV epidemic. Unlike AIDS data, HIV data provide a window into the epidemic at an earlier stage of disease, thereby allowing public health officials to monitor the epidemic more effectively and completely, allocate resources, and plan and implement programs, particularly prevention programs. In the past, however, biomedical technology did not discriminate between recent and chronic HIV infection; as a result the incidence of HIV infection in the United States could not be measured directly. The serologic testing algorithm for recent HIV seroconversion (STARHS)[1] is performed on remnant serum specimens from confirmed HIV antibody positive tests and consists of a series of two tests, a standard, sensitive, HIV antibody test currently followed by a test to determine the normalized optical density (OD_N) level of the concentration of HIV-specific antibodies to total antibodies. STARHS distinguishes between recent and long-standing HIV-1 infection on a population level and should allow the estimation of local and national HIV incidence.

HIV incidence surveillance (HIS) is the aspect of the national HIV/AIDS surveillance system that uses STARHS results, as well as data on the history of testing and use of medications with antiretroviral (ARV) properties for each case reported to HIV/AIDS surveillance programs, to generate an HIV incidence estimate. HIS will give a more representative picture of the HIV epidemic, its trends, and its impact on the public's health. The Centers for Disease Control and Prevention's (CDC's) human subjects protection process has determined that the implementation of HIS programs using STARHS, like other public health surveillance activities, is not research (see National Center for HIV, STD, and TB Prevention's Non-research Determination for HIV Incidence Surveillance).

The primary functions of HIS are to:

- Incorporate STARHS into routine HIV surveillance activities by testing HIV-seropositive specimens obtained from persons with a new HIV diagnosis using an assay approved for this purpose
- Collect HIV testing and ARV use history information for persons with newly reported HIV infections as part of routine HIV surveillance
- Apply a statistical model(s) to estimate HIV incidence locally and nationally using a combination of STARHS results and information from surveillance case reports, and testing and ARV use history information for all persons with a new diagnosis of HIV infection
- Use incidence data to assist with local HIV prevention program planning and evaluation using incidence data.

The tasks to achieve the functions of HIS include:

- Elicit collection and reporting of testing and ARV use history information by all providers reporting HIV cases to the health department
- Ensure that an aliquot of each confirmed seropositive specimen from a newly diagnosed case is shipped to the CDC STARHS laboratory
- Determine the disposition of remnant HIV-positive specimens and coordinate communication between the Incidence Surveillance Coordinator (ISC) and the public health laboratory or the CDC STARHS laboratory regarding specimens shipped to and stored at the these laboratories
- Enter STARHS and testing and ARV use history information into the designated HIS database
- Electronically transfer data to CDC
- Ensure that HIS data handling procedures comply with all security and confidentiality guidelines as described in *Technical Guidance for HIV/AIDS Surveillance Programs, Vol. III: Security and Confidentiality Guidelines*
- Analyze and disseminate data locally
- Train staff in the policies and procedures of the HIV/AIDS surveillance system as well as those of the HIS system

The prerequisites (structural requirements), best practices (process standards), and outcome standards for HIS are described next.

Structural Requirements

All persons with a new diagnosis of HIV infection who were tested confidentially should be reported to the HIV surveillance system in accordance with *Technical Guidance for HIV/AIDS Surveillance Programs, Vol. I: Policies and Procedures*. In areas incorporating HIS into their HIV/AIDS surveillance systems, information regarding HIV testing history and ARV use, as well as the STARHS result of a remnant of the diagnostic HIV-positive specimen, should be included for all cases. HIV/AIDS surveillance case report data, in combination with these data elements, will be used to calculate population-based HIV incidence estimates.

Policies and Procedures

HIS is a fully integrated component of HIV/AIDS surveillance; therefore, documentation of HIS activities should be incorporated into locally tailored policies and procedures manuals developed for HIV/AIDS surveillance (*Technical Guidance for HIV/AIDS Surveillance Programs, Vol. I: Policies and Procedures*) to establish standardization, maintain continuity of meaning, document changes over time, and develop training programs. All manuals describing policies and procedures of the local surveillance program should address the needs of HIV/AIDS surveillance as well as the specific

policies related to HIS. In addition to the information listed in *Technical Guidance for HIV/AIDS Surveillance Programs, Vol. I: Policies and Procedures*, HIS specific policies and procedures should include information related to

- Training of testing providers in collection of additional data used for HIV incidence estimation
- Laboratory contacts
- Determining which specimens should be tested using STARHS (and which should be discarded)
- Specimen aliquoting
- Specimen shipping guidance

The BED Assay for STARHS

The assay STARHS currently uses is the BED HIV-1 Capture enzyme immunoassay (EIA) manufactured by Calypte Biomedical Corporation. The principle of the BED HIV-1 Capture EIA is based on the observation that the ratio of HIV-specific IgG to total IgG increases with time after HIV infection[2]. The BED HIV-1 Capture EIA is applied to the diagnostic HIV-positive specimen, and the assay is sensitive to the length of time since the infection (i.e., antibody level present). The time from when a specimen would first be reactive on the standard EIA to the time when the serum or plasma, if tested with the BED HIV-1 Capture EIA, reaches an optical density (OD) level predetermined to distinguish recent from nonrecent infections is defined as the STARHS mean window period. Although the mean STARHS window period may vary slightly by HIV subtype, the mean window period for calculating population-based incidence estimates in the United States is 156 days when the BED HIV-1 Capture EIA is used.

The BED HIV-1 Capture EIA for STARHS is performed only on HIV antibody positive sera[2] and is not approved as a diagnostic test. Because of the variability in antibody development in individuals, the predictive value of an individual's STARHS result is low; the results are reliable only as part of the population-based HIV incidence estimate. The Food and Drug Administration (FDA) has ruled that the BED HIV-1 Capture EIA be labeled "For Surveillance use only. Not for diagnostic or clinical use." Under FDA regulations, results of STARHS performed for purposes of HIS cannot be returned to individuals or their health care providers or used for clinical management. As with earlier assays used for STARHS, data show that the BED HIV-1 Capture EIA can produce a substantial number of false-positive and false-negative classifications on the individual level[3]. At the population level, the number of false positives is approximately equal to the number of false negatives thus effectively "canceling" each other out. However, the number of misclassifications can be large, and each of the misclassified individual results would receive an incorrect interpretation. STARHS results may also be misclassified because of the use of ARV therapies or late stage of the disease. Evaluation of the BED HIV-1 Capture EIA has determined that the specimens of persons with low HIV-1-specific antibodies resulting from ARV therapy or disease progression (i.e., AIDS) could lead to the incorrect conclusion that these persons were recently infected. When a person has AIDS, this is thought to be due to a loss of immune response as immune deficiency

progresses. When a person is taking ARV therapy, this result is thought to be due to suppression of the HIV viral load, which in turn reduces antigenic stimulation and the quantity of circulating HIV-1-specific antibodies. The effect of ARVs taken for post-exposure prophylaxis or for concurrent hepatitis B infection, for example, is not known. Theoretically, it could take longer to develop a full immune response to HIV infection. CDC data reliably support using STARHS for estimating incidence at the population level only.

Testing and ARV Use History Data

Information on testing behavior is needed, such as recency of testing and testing frequency. Additionally, history of ARV use (for example, pre- or post-exposure prophylaxis or treatment for hepatitis) and immunological status (CD4 cell counts and viral loads) must be included for all cases reported to the surveillance system.

Information needed for HIV incidence estimation is available as part of a standard case report, and nearly all testing and ARV use history information is gathered as part of a comprehensive HIV counseling session. However, not all of the required HIS data elements have been collected uniformly, and many have not previously been recorded. Therefore, a standard set of HIV testing and ARV use history data elements needed for the HIV incidence estimate has been developed (see Standard HIV Incidence Surveillance Data Elements), and providers of HIV testing should be trained in the appropriate reporting of those data.

Staffing Needs

Implementation of the HIS system requires personnel with specific skills and dedicated time to integrate HIS into the existing core HIV/AIDS surveillance system effectively. Generally, HIS staff should have

- An understanding of HIS and the characteristics of the HIV/AIDS surveillance in their area
- Good communication skills
- Strong leadership skills
- Enthusiasm about disease reporting for public health purposes
- Dedication to the successful implementation of HIS
- Ability to work closely with CDC, other states, local sites, private providers, and laboratories

The recommended staffing plan including roles and responsibilities follows. In terms of personnel time, CDC recommends that one full-time equivalent (FTE) be dedicated to the ISC position. Successful implementation and integration of HIS requires a full-time ISC dedicated to implementing and maintaining the system. Other personnel assigned to HIS may vary depending on the implementation phase, prevalence of HIV/AIDS, and available resources.

ISC

- Provide overall management of the HIS system
  - Serve as lead on the area specific implementation of the HIS guidance
- Oversee data collection processes
  - Determine the disposition of HIV-positive specimens reported to the surveillance system
  - Receive STARHS results
  - Oversee collection of HIV testing and treatment history information from public HIV testing sites and private providers
- Collaborate with other HIS staff
  - HIV incidence epidemiologist
    - Development of training materials and courses
    - Data collection procedures
  - HIV incidence data manager
    - Data collection methods
    - Data entry and quality assessment
    - Data editing and file correction
    - Data transport procedures
    - Preparation of monthly reports
    - Security and confidentiality procedures
  - HIV incidence laboratory liaison
    - Specimen transfer
    - Specimen tracking
- Manage any employee or other service contracts related to HIS
- Serve as the primary point of contact for CDC on HIS
- Participate in CDC site visits, trainings, and workshops

Epidemiologist/Trainer

- Serve as lead on training HIV testing providers and laboratories on HIS, including development/modification of surveillance area-specific training materials
- Coordinate HIS and epidemiology activities with the ISC
- Participate in the development or modification of testing and ARV use history data elements
- Participate in data dissemination activities
  - Collaborate with stakeholders to determine data needs and frequency of reporting
  - Identify results and surveillance issues for review and dissemination
  - Develop a data dissemination plan in collaboration with the ISC and CDC
- Participate in CDC site visits, trainings, and workshops as appropriate

Laboratory Liaison

- Act as the liaison between the public health, private, and community laboratories and the ISC
- Oversee preparation and shipping of public health laboratory specimens to the CDC STARHS laboratory
- Monitor quality control procedures outlined for preparing specimens for testing using STARHS
- Monitor security and confidentiality of specimens and STARHS results
- Track specimens identified for testing using STARHS (all laboratories)
- Participate in CDC site visits, trainings, and workshops as appropriate

Data Manager

- Assist ISC with daily management of HIS data
- Conduct data collection from surveillance sites
  - Serve as subject matter expert on HIV incidence data elements and data management programs
  - Receive data transfer from other health department entities and the CDC STARHS lab, and incorporate those data into the HIS database and data sets for transfer to CDC
  - Conduct data quality assessments
- Conduct data management
  - Modify CDC's generic data management programs for use at the area level
  - Develop and implement edit checks and conduct data cleaning
  - Collaborate with the ISC, epidemiologist, and other area surveillance and prevention staff, as needed, on data cleaning, data entry, and data set preparation
  - Prepare data sets for local analysis
  - Collaborate with CDC on data set preparation for national incidence estimates
  - Prepare HIV incidence data reports for local use in collaboration with the ISC, epidemiologist, and CDC
- Maintain security and confidentiality of HIV incidence data
- Participate in CDC site visits, trainings, and workshops as appropriate

Process Standards

HIS involves the following processes:

- Obtaining testing and ARV use history data from providers
- Identifying laboratories that perform HIV-related tests and obtaining remnant specimens for testing using STARHS

- Determining specimen disposition as it relates to testing using STARHS
- Establishing a schedule for contact between the ISC and the public health laboratory and the CDC STARHS laboratory to communicate regarding shipping, testing, and discarding remnant specimens that are housed at the laboratory
- Entering data into the HIS database
- Electronically transferring data to CDC
- Ensuring that data handling procedures comply with Security and Confidentiality Guidelines (see *Technical Guidance for HIV/AIDS Surveillance Programs, Vol. III: Security and Confidentiality Guidelines* and Model State Public Health Privacy Act)
- Analyzing and disseminating data locally
- Training staff in HIV/AIDS surveillance and HIS methods

Obtaining Testing and ARV Use History Data

The primary purpose of gathering HIV testing and ARV use history is to calculate a statistical weight (see Statistical Method for Generating Population-Based HIV Incidence Estimates) that will allow inference to the general population. The weight reflects the probability that an individual will be tested during the STARHS window period and is related to the data elements listed in Standard HIV Incidence Surveillance Data Elements.

Testing and ARV use history data are reported by providers and surveillance staff using the area standard reporting procedures or other procedures meeting the routine security and confidentiality guidelines for HIV/AIDS surveillance. These data

- Should be included for all adult/adolescent (≥13 years at diagnosis) HIV/AIDS cases newly reported to the HIV/AIDS surveillance system by all providers of HIV testing
- May be collected through client/patient interview and/or chart abstraction
- If based on patient self-report, should be collected when an individual presents for an HIV test or returns for the results
  - Takes advantage of the individual's ability to recall information that is more proximal to the event
  - Should be recorded and reported to the surveillance system on the basis of the patient's self-report within 3 months of the HIV diagnosis
  - Longer intervals may increase the risk of recall bias, yet this consideration should not prevent efforts to obtain the information even after 3 months if necessary

CDC has assisted in developing materials for use in training providers to collect HIV testing and ARV use history data. These materials are available at all sites, or upon request from CDC.

Obtaining Remnant HIV-Positive Specimens for Testing Using STARHS

To be most useful, testing using STARHS should be performed on the HIV-positive diagnostic serum or plasma specimen. For the purposes of this guidance, the HIV-positive diagnostic specimen is the HIV-positive specimen from the diagnostic test that resulted or should have resulted in the case being reported to the HIV/AIDS surveillance system. Remnant specimens for all confirmed HIV-positive diagnostic specimens should be tested using STARHS.

- Laboratories performing routine diagnostic confirmatory HIV testing by Western blot, indirect fluorescent antibody (IFA) tests, immunological status tests such as CD4, or viral load counts should report to the state/local health department surveillance program per existing requirements.
- In each surveillance area, all laboratories should be identified from a review of local HIV surveillance data and laboratory licensing records and must be approached to request that remnants of all diagnostic specimens be made available for testing using STARHS
  - Surveillance areas should maintain a directory of laboratory contacts at all reporting laboratories to facilitate communication in the event that reporting or shipping of specimens is disrupted or that changes in policy or procedures need to be communicated
  - Originating laboratories are those to which a specimen is first sent for testing
  - Reference laboratories are those to which a specimen is sent for confirmatory testing when the originating laboratory does not do confirmatory testing
- All remnant specimens from HIV-diagnostic Western blot or IFA tests must be shipped to the CDC STARHS laboratory in New York for testing using STARHS.
  - A minimum of 0.5 mL HIV-positive serum or plasma specimen is necessary for testing using STARHS
  - Private or community laboratories performing HIV diagnostic testing should choose one of two options for shipping the remnant HIV-positive serum or plasma specimen to the CDC STARHS laboratory (see Guidance for the Remnant HIV-Positive Specimen Transportation Activities for HIV Incidence Surveillance)
    - Ship the specimen directly to the CDC STARHS laboratory
    - Ship the specimen to the state public health laboratory affiliated with the health department that receives the new HIV case report for processing
      - State public health laboratories can then batch and ship all specimens identified for testing using STARHS to the CDC STARHS laboratory
  - State public health laboratories conducting HIV diagnostic testing should ship their own HIV-positive specimens identified for testing using STARHS directly to the CDC STARHS laboratory

Specimen availability for testing using STARHS depends on the testing needs for the specimen. Uses of the remnant of specimens should follow the CDC HIS recommended hierarchy (described below) for specimen aliquoting to ensure adequate specimen volume

for multiple diagnostic tests. When multiple tests must be performed on a collected serum or plasma specimen, the aliquots must be made available with the following hierarchy in mind:

1. HIV diagnostic testing
2. Testing with STARHS
3. HIV drug resistance genotyping (known as Variant, Atypical, and Resistant HIV Surveillance [VARHS])

Aliquots made following this hierarchy will ensure that adequate specimen volume is available according to the priority for data determined by CDC.

Determining the Disposition of Specimens and Communicating with the Public Health and CDC STARHS Laboratories

A specimen should be held at the state public health laboratory or the CDC STARHS laboratory (i.e., specimens shipped directly from private or commercial laboratories) until the area ISC, using routine HIV/AIDS surveillance reporting procedures (i.e., HIV/AIDS Reporting System [HARS/eHARS]), determines whether the specimen represents the person's first reported positive HIV test result in the HIS area. A specimen should be tested using STARHS if

1. the specimen represents the diagnostic specimen (the HIV-positive specimen that led, or should have led, a case to be reported to HARS/eHARS) **or**
2. the diagnostic specimen is unavailable and the specimen was drawn within 3 months of the diagnostic specimen **and**
3. the specimen was drawn for an HIV-related test (viral load, polymerase chain reaction [PCR] test, CD4 level).

A specimen should not be tested using STARHS if

1. the specimen is not the diagnostic specimen that led, or should have led, the individual to be reported to HARS/eHARS **and**
2. the individual had a previous specimen that was tested using STARHS **or**
3. the individual did not have a previous specimen tested using STARHS but the specimen was drawn more than 3 months after the diagnostic specimen.

Because a remnant sample of every Western blot positive blood specimen will be shipped by originating or reference laboratories to either the state or local public health laboratory or to the CDC STARHS laboratory, the HIS program must inform the appropriate laboratory of the disposition of the specimen.

- Specimens for cases not previously reported to HARS/eHARS (or those drawn within 3 months of a diagnostic specimen that is unavailable) will constitute the test list. A test list should
  - Be compiled for the state or local public health laboratory and for the CDC STARHS laboratory

    - Include those specimens located at the individual laboratory that should be tested using STARHS
    - Be cumulative
- Specimens that are neither diagnostic specimens nor drawn within 3 months of a diagnostic specimen that is unavailable will constitute the toss list. A toss list should
    - Be compiled for the state or local public health laboratory and for the CDC STARHS laboratory
    - Include those specimens held at the laboratory that should not be tested using STARHS
    - Be cumulative

Specimens should be handled, packaged, and shipped according to the CDC STARHS laboratory shipping protocol (see Guidance for the Remnant HIV-Positive Specimen Transportation Activities for HIV Incidence Surveillance). Specimens shipped as diagnostic specimens and packed in dry ice must be handled according to the procedures for packing and shipping specimens with dry ice (see Guidance for Processing, Storage, and Shipping of Specimens to the CDC STARHS Laboratory).

The ISC should regularly, at an interval to be determined locally (e.g., monthly), inform area laboratory designees (e.g., at the public health laboratory) and the CDC STARHS laboratory of all stored specimens, or remnant specimens, with a positive HIV diagnostic test to be tested using STARHS (the test list) and those to be discarded (the toss list).

- For those specimens on the test list
    - An aliquot of blood to be used for STARHS is drawn from the specimen
        - At the local public health laboratory before shipping
        - At the CDC STARHS laboratory if the specimen was shipped directly to the CDC STARHS laboratory
    - The aliquot is relabeled with a unique STARHS identification number (SID)
    - The SID is paired with the corresponding specimen number and is sent to the ISC with no other identifying data
    - After STARHS, the results are returned to the ISC with results identified by SID only
- Specimens on the toss list should be discarded according to routine laboratory protocols for HIV-positive serum or plasma specimens.
- The ISC should inform the CDC STARHS lab to discard all specimens that have been tested with STARHS. If the ISC would like to have these specimens returned at the surveillance area's expense, he/she should make arrangements with the CDC STARHS lab.

Entering Data into Surveillance Databases

Data elements needed for the calculation of statistical weights used to make population-based HIV incidence estimates fall into one of four categories.

- Demographic data
  - Age
  - Sex
  - Race/ethnicity
  - HIV risk factors
- HIV testing and ARV use history data
- Clinical data
  - CD4 count
  - Viral load
- Laboratory data
  - Specimen collection date
  - SID
  - STARHS results

All data elements needed for HIS are included in eHARS, but could not be added to HARS. Only demographic data and clinical data are included in the HARS database. As a result, an HIS Access database was developed to store the additional data related to HIS. Until a surveillance area begins using eHARS and eHARS accommodates this information, the jurisdiction should enter data into both HARS/eHARS and the HIS Access database as appropriate. Testing and ARV use history information and laboratory data related to the diagnostic HIV test (including specimen collection date and SID) are variables that should be entered into the separate HIS Access database along with the unique state number (HARS/eHARS "stateno") assigned to each case. A list of data elements necessary for HIS, but not included in HARS can be found in Standard HIV Incidence Surveillance Data Elements.

Creating the HIS Data Set and Transferring Data to CDC

HIS data entry and management take place at state or local health departments by using one of three data management systems:

- HIS Access database (until conversion to eHARS)
- eHARS
- Software that is compatible with CDC software

Information is merged into the data management systems from other sources:

- Excel spreadsheet containing STARHS results identified by SID only from the CDC STARHS laboratory
- Program Evaluation Monitoring System (PEMS)
- Other databases

For states that have not transitioned to eHARS, data are merged by using unique identifiers reported with each case for data transfer to CDC:

- the HARS/eHARS stateno links HARS/eHARS records to corresponding HIS Access database records

When eHARS has the ability to import HIS data, merging data sets before transferring data to CDC will no longer be necessary.

Before the 15th of each month, the complete data set from the preceding month (see *Access Database 3.1a Users Manual* and *Access Database Version 3.1 Data Dictionary*) should be transmitted to CDC over the Secure Data Network (SDN). Data transmitted to CDC must include no personal identifiers and must be encrypted and password protected according to *Technical Guidance for HIV/AIDS Surveillance Programs, Vol. III: Security and Confidentiality Guidelines*.

Ensuring Security and Confidentiality

HIV testing is a medical procedure. Therefore, policies and procedures are in place to protect the confidentiality of tested persons and their medical records. STARHS should be performed only on specimens that have tested positive for HIV. HIS data are considered part of routine HIV surveillance data and should be held to the standards of security and confidentiality for HIV/AIDS surveillance outlined in *Technical Guidance for HIV/AIDS Surveillance Programs, Vol. III: Security and Confidentiality Guidelines*. Policies and procedures, based on these guidelines and local laws, are already in place at state and local health departments and are used to secure hard copy and electronic information to protect the confidentiality of persons reported as having HIV infection. These measures must be extended to protect the STARHS information held locally. Access by HIS staff to information in HARS, HIV testing and ARV use history, and STARHS data is governed by the same security and confidentiality requirements.

Analyzing and Disseminating Data

Data from persons who choose to have a confidential HIV test and who test positive will be used to estimate the incidence of HIV nationally and in participating areas, including the incidence of undiagnosed HIV infection. HIV incidence estimates can be used to assess current HIV prevention programs locally, regionally, and nationally. HIS data will be stratified by selected factors such as demographic or behavioral factors, thus creating subpopulation data at the national and local levels. If the sampling procedure has sufficient statistical power, this stratification will allow comparisons between areas and among groups with different risk factors. The methods used to generate the population-based incidence estimate are described in Statistical Method for Generating Population-Based HIV Incidence Estimates, which introduces the methods, statistical formulas, and different groups for which incidence estimates will be made.

CDC will have the primary responsibility for analyzing interpreting, and disseminating these data; surveillance areas will contribute as appropriate. Results from the aggregate CDC database will be analyzed regularly and feedback will be provided to participating areas. Aggregate results will also be published in CDC's HIV/AIDS Surveillance Reports. Area-specific analyses should be conducted at the discretion of participating areas. As

appropriate, results will be presented at conferences and published in peer-reviewed journals. The number of representative authors from participating areas and CDC will be determined for each presentation or paper.

In addition to the variables needed to estimate HIV incidence, the following data elements will be used to evaluate the performance of the BED HIV-1 Capture EIA for determining of estimates of HIV incidence:

- Whether AIDS has been diagnosed and if so, the date of diagnosis
- At the time of HIV diagnosis,
  - whether HIV ARV agents have been used for post-exposure prophylaxis or for any other medical condition (e.g., lamivudine for treatment of hepatitis B), and if so, the name(s) of the agent, and dates and duration of use
  - where available, CD4 count, viral load, and HIV-1 subtype along with the type of test used to determine the subtype

As a result, all HIV-positive diagnostic specimens should be tested using STARHS irrespective of the time to AIDS diagnosis for an individual, or evidence of previous ARV use.

Training Staff

Because HIS is a fully integrated component within the HIV/AIDS surveillance system, all HIS staff should receive training in the local policies and procedures for core surveillance including

- active and passive surveillance methods
- laboratory reporting mechanisms
- data management processes

In accordance with *Technical Guidance for HIV/AIDS Surveillance Programs, Vol. III: Security and Confidentiality Guidelines*, HIS staff must also receive training in security and confidentiality procedures, and should sign a confidentiality statement upon being hired and annually thereafter.

Outcome Standards

Outcome standards described in the “Introduction to Policies and Procedures” and “Data Quality” sections of *Technical Guidance for HIV/AIDS Surveillance Programs, Vol. I: Policies and Procedures* can be applied to HIS. These sections address issues of completeness of case ascertainment, timeliness of reporting, evaluation of standard data edits, and missing/unknown information. Meeting core surveillance standards for case ascertainment and timeliness is essential for HIS to be successful given the time-sensitive nature of HIS data elements, including testing and ARV use history data and STARHS. The quality of the HIV incidence estimate depends on the quality of data included in the HIS system. All outcome standards for HIS relate only to cases that reside within the surveillance area at the time of diagnosis.

- The minimum standard for passing standard data edits related to HIS data is 97% with a target of 100%
- At least 85% of newly reported HIV/AIDS cases for a diagnosis year should have testing history and ARV use data within 12 months of the date of the initial HIV/AIDS case report, measured at 12 months after the close of the diagnosis year
- At least 85% of cases newly reported to the surveillance system and diagnosed using a serum/plasma specimen or having a follow-up HIV-related test conducted on a serum/plasma specimen within 3 months of the diagnosis should have a specimen transported to the CDC STARHS laboratory within 12 months of diagnosis assessed for the most recent diagnosis year at 12 months after that diagnosis year

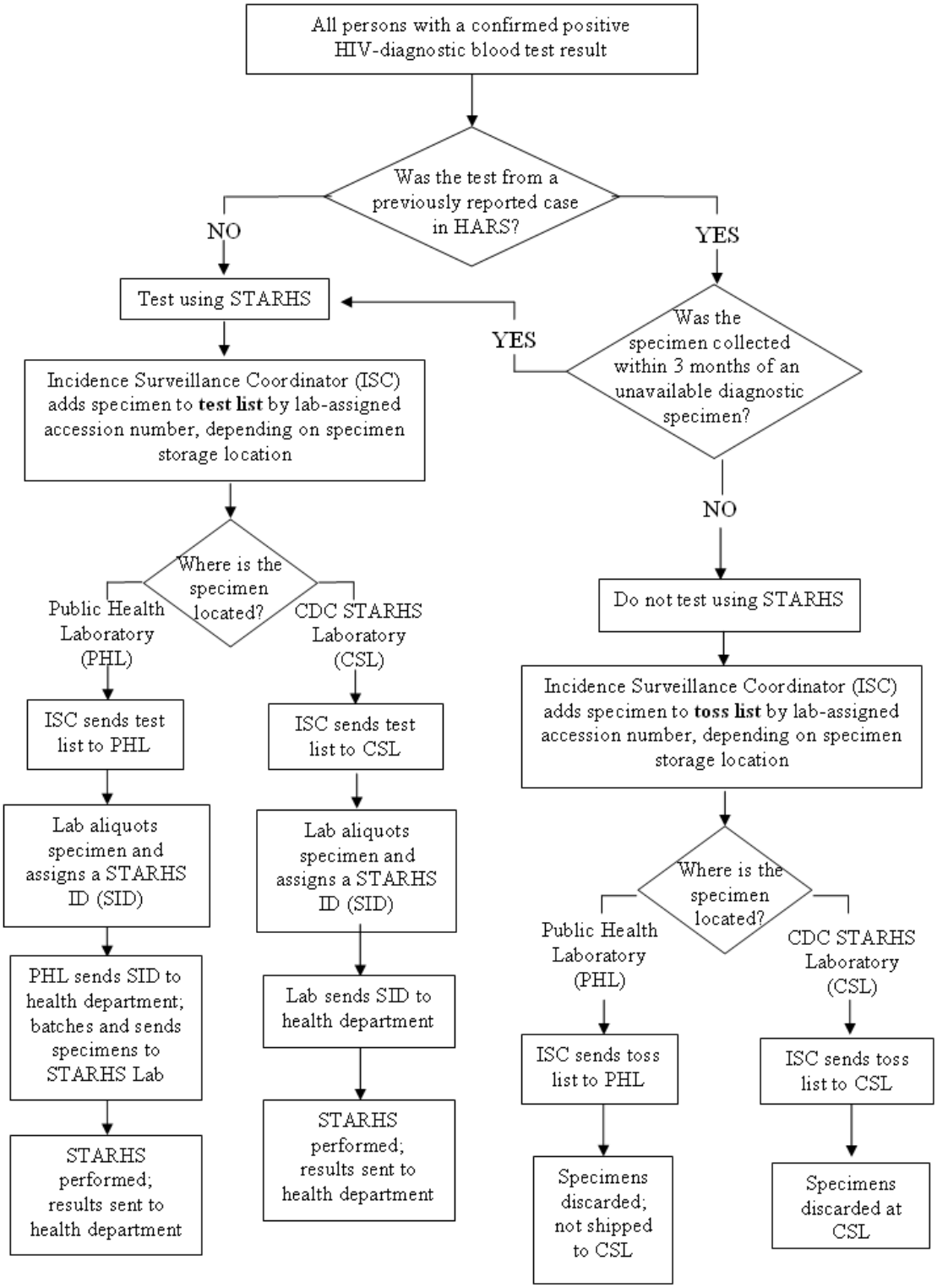
References

1. Janssen RS, Satten GA, Stramer SL, et al. New testing strategy to detect early HIV-1 infection for use in incidence estimates and for clinical and prevention purposes. *JAMA* 1998; 280:42–48.

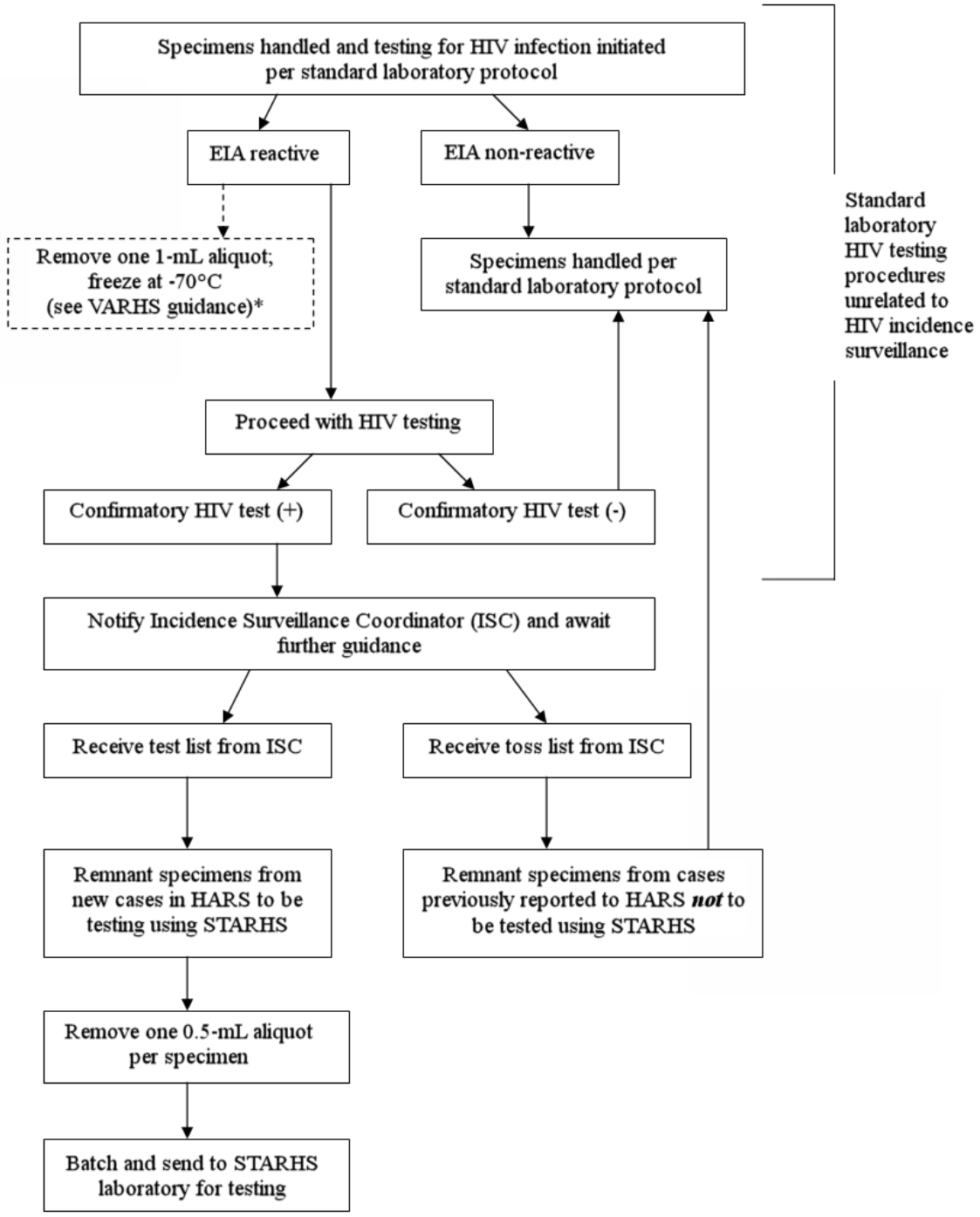
2. Parekh B, Kennedy MS, Dobbs T, et al. Quantitative detection of increasing HIV type 1 antibodies after seroconversion: a simple assay for detecting recent HIV infection and estimating incidence. *AIDS Research & Human Retroviruses* 2002;18(4):295–307.

3. Linley L, Reed C. Applicability of Population-Based STARHS HIV Incidence Measure in Determining Recency of Individual Infection among Patients Attending STD Clinics. In Program and abstract of the 11th Annual Conference on Retroviruses and Opportunistic Infections; February 2004; San Francisco. Abstract 854.

Epidemiologic Flow Chart for HIV Incidence Surveillance



Operational Public Health Laboratory Flow Chart for HIV Incidence Surveillance



* Applies only to laboratories participating in Variant, Atypical, and Resistant HIV Surveillance (VARHS)

National Center for HIV, STD, and TB Prevention's Non-research Determination for HIV Incidence Surveillance

NCHSTP Research/Non-research Determination

(Request to Classify Project as Not Involving Human Subjects or Research)

This form should be used to submit to NCHSTP ADS materials for projects involving CDC investigations that are not subject to human subjects regulations. Projects are eligible for this classification either as "non-research" projects (primary intent is not to generate generalizable knowledge) or as research projects that do not involve identifiable human subjects. Such projects do not require submission to the CDC Human Subjects Office for IRB review. Do **NOT** use this form for IRB "EXEMPT" research.

Project Title___ Standard HIV Incidence Surveillance Procedures Guidance _____

_____

Project Locations/Sites: _____33 funded incidence sites, eventually all HIV/AIDS Surveillance areas___

Project Officer(s)____Lisa M. Lee, PhD __ Division:__ DHAP-SE _____ Telephone: __404.639.2052

Proposed Project Dates: Start: 03/__01__/_2005_ End: N/A, routine surveillance _____/_____/_____

Categories of data collection that do not constitute human subjects research include are listed below. Please check appropriate category:

_X__I. **Activity is not research** . Primary intent is public health practice or a disease control activity.

- ___**A**. Epidemic/endemic **disease control** activity; collected data directly relate to disease control needs.
- _X__**B**. Routine **disease surveillance** activity; data used for disease control program or policy purposes.
- ___**C**. **Program evaluation** activity; data are used primarily for that purpose.
- ___**D**. **Post-marketing surveillance** of efficacy and/or adverse effects of a new regimen, drug or device.
- ___E. **Activity is purely administrative** (e.g., purchase orders or contracts for services or equipment) and not related to research [this category I-E may be determined by Divisional ADS]

-OR-

___**II. Activity is research but does NOT involve identifiable human subjects.**

- ___**A.** Activity is research involving collection/analysis of data about health facilities or other organizations or units which are not individual persons....**or**...
- ___**B.** Activity is research involving data and/or specimens from deceased persons...**or**...
- ___**C.** Activity is research using unlinked anonymous data or specimens: **All** (1-4) of the following are required:
  - ___ 1. No contact with human subjects is involved for the proposed activity...**and**...
  - ___ 2. Data or specimens are/were collected for another purpose...**and**...
  - ___ 3. No extra data/specimens are/were collected for **this** purpose...**and**...
  - ___ 4. Identifying information either was not obtained **or** has been removed so that data cannot be linked or re-linked with identifiable human subjects. (Note: under certain conditions, research *may* qualify as non-human subjects when identifiers are removed by local staff; contact NCHSTP ADS office for details.)

Attach project description (standard format at end of this form) in enough detail to clarify its non-human subject research nature. Submit through division ADS/Director to: NCHSTP ADS, Attn: Janella Dodson (MS E-07)

☐ Check here if this request is an **amendment** of an existing non-research determination.

Approval initials: _____ **Branch/Section Chief** 3/2/05 Date _____ **ADS or Div. Director** 3/2/05 Date

res-nonres.det.form.0704

Page 2

Project Title_____ Standard HIV Incidence Surveillance Procedures Guidance _____

NCHSTP ADS Review **Date rec'd in NCHSTP ADS Office:**_____

✓ **Concur, project does not constitute human subjects research**

or

___**Project constitutes human subjects research**, submission for Human Subjects review required

this project is national HIV/AIDS surveillance & not research.

Comments/Rationale:

Additional Comments:

1. This form cannot be used to document "IRB Exempt Research," which must instead be submitted to the CDC IRB. (Please contact the NCHSTP ADS Office for details).
2. Although CDC Human Subjects (IRB) review is not required in this instance, investigators/project officers are expected to adhere to ethical principles and standards by respecting and protecting to the maximum extent possible the privacy, confidentiality and autonomy of participants. All applicable State and Federal privacy laws must be followed.
3. Although this project does not constitute human subjects research, informed consent may be appropriate. Information disclosed in the consent process should address the eight standard consent elements.
4. Other:

Signed: _____ 3/9/05

Terence Chorba, MD, MPH, MPA, MA
Acting Associate Director for Science, NCHSTP
National Center for HIV, STD, and TB Prevention

Date

res-nonres.det.form.0704

Standard HIV Incidence Surveillance Data Elements

	Required Surveillance Data Elements
	Required Laboratory/Specimen Related Variables

Variable Description	Valid Value
Antiretroviral use	0 = no 1 = yes 7 = refused 9 = don't know
Antiretroviral (ARV) medications taken	22 = Agenerase (amprenavir) 30 = Aptivus (tipranavir, TPV) 32 = Atripla (efavirenz/emtricitabine/tenofovir DF) 24 = Combivir (lamivudine/ zidovudine) 06 = Crixivan (indinavir, IDV) 11 = Emtriva (emtricitabine, FTC) 03 = Epivir (lamivudine, 3TC) 28 = Epzicom (abacavir/lamivudine) 25 = Fortovase (saquinavir, SQV) 10 = Fuzeon (enfuvirtide, T20) 19 = Hepsera (adefovir) 02 = Hivid (zalcitabine, ddC) 23 = Hydroxyurea 18 = Invirase (saquinavir, SQV) 16 = Kaletra (lopinavir/ ritonavir) 31 = Lexiva (fosamprenavir, 908) 07 = Norvir (ritonavir, RTV) 88 = Other 33 = Prezista (darunavir, DRV) 09 = Rescriptor (delavirdine, DLV) 26 = Retrovir (zidovudine, ZDV, AZT) 15 = Reyataz (atazanavir, ATV) 08 = Saquinavir (Fortavase, Invirase) 21 = Sustiva (efavirenz, EFV) 13 = Trizivir (abacavir/lamivudine/zidovudine) 27 = Truvada (tenofovir DF/emtricitabine) 99 = Unspecified 01 = Videx (didanosine, ddl) 14 = Videx EC (didanosine, ddl) 17 = Viracept (nelfinavir, NFV) 05 = Viramune (nevirapine, NVP) 12 = Viread (tenofovir DF, TDF) 04 = Zerit (stavudine, d4T) 20 = Ziagen (abacavir, ABC)

Variable Description	Valid Value
Date Highly Active Antiretroviral Treatment (HAART) use began	yyyymmdd
Date HAART use ended	yyyymmdd
Date information is extracted either from the client by interview or the medical chart by abstraction	yyyymmdd
Date of first positive HIV test	yyyymmdd
Date of last negative HIV test	yyyymmdd
Ever tested negative	0 = no 1 = yes 7 = refused 9 = don't know
Number of HIV tests in the 2 years before first positive test	1–99 R = refused D = don't know
Clinical Laboratory Improvement Amendments (CLIA) code for source lab where specimen originated	text
Date of STARHS test	yyyymmdd
Date specimen was obtained	yyyymmdd
Specimen approved for STARHS	0 = no 1 = yes 2 = pending
Laboratory ID	33D0654341 = NYST 33D0654341 = CDCSTARHS 33D0654341 = NY 33D0654341 = CDCSTAR 21D0649758 = MARY01 50D0661430 = WASH
Optical density	text
Reason STARHS not performed	1 = quantity not sufficient (QNS) 2 = specimen never received at public lab 3 = broken in transit 4 = other
Results received	0 = no 1 = yes
Specimen identification number from source lab	text
STARHS identification number (SID)	text
STARHS regional laboratory specimen identification number (same as STARHS laboratory imported variable SPECIMEN ID)	text

Variable Description	Valid Value
STARHS test result	01 = long term 02 = recent 91 = QNS 92 = not received by STARHS lab 93 = broken 94 = other
State laboratory CLIA code	text
State laboratory specimen ID number	text
Test assay	BED = BED BVLS = BVLS (Vironostika LS) OTLS = OTLS (Vironostika LS) OTV = OTV (Vironostika LS) AVID = AVID
Type of test performed on specimen (Logical Observation Identifiers Names and Codes [LOINC])	5220-9 = EIA/ELISA 21009-6 = Western blot 5472-6 = CD4 25835-0 = Viral Load (NASBA) 5017-9 = Viral Load (bDAN) 25836-8 = Viral Load (RT-PCR)
Type of specimen obtained	1 = blood finger stick 2 = blood venipuncture 3 = blood spot 4 = oral mucosal transudate 5 = urine 8 = other 9 = unknown

Guidance for the Remnant HIV-Positive Specimen Transportation Activities for HIV Incidence Surveillance

Purpose

This guidance provides an overview of the incidence surveillance specimen transport activities and describes two possible specimen transport models that originating laboratories may use to ship remnant diagnostic serum specimens to the CDC STARHS laboratory for testing for recent HIV-1 infection using the serologic testing algorithm for recent HIV seroconversion (STARHS). Originating laboratories may choose to select either model but must clearly communicate their choice to and coordinate with the state/local HIV Incidence Surveillance Coordinator (ISC), who will be responsible for managing the results.

Introduction

In December 2004, an expert consultation was convened by the Centers for Disease Control and Prevention (CDC) and the Association of Public Health Laboratories (APHL). The purpose of the 5th HIV Incidence Consultation on Laboratory and Specimen Transport was to discuss the best approaches for transporting remnant HIV-positive sera from private and public laboratories to the CDC STARHS laboratory for testing using STARHS. The consultation participants included HIV incidence surveillance staff from CDC and state/local areas, and personnel from commercial, private, university, and public health laboratories, APHL, and the American Clinical Laboratory Association. The goal for the meeting was to gather input from stakeholders for developing an infrastructure for shipping specimens from private (including university and/or medical center), commercial, and public health laboratories to the CDC STARHS laboratory.

Participants concluded that two models were acceptable for shipping specimens from testing laboratories to the CDC STARHS laboratory. This guidance describes both specimen transportation models. The models differ in 1) the extent of the testing laboratories' involvement in aliquoting/labeling samples for STARHS; 2) the physical storage location of the samples until the ISC determines specimen disposition (i.e., whether to be tested using STARHS or to be discarded); and 3) the frequency of shipments to the CDC STARHS laboratory.

Each testing laboratory may choose either model, but this choice should be clearly communicated to the ISC.

Laboratory Types

For the purposes of this guidance, there are three laboratory types. Although each testing laboratory may independently decide which specimen transport model will work best for that facility, CDC has provided suggestions based on the type of laboratory and that facility's relationship with the state/local public health laboratory.

Laboratory Types

1. **Private Laboratories**:
   a. Larger **commercial laboratories** that process samples from many states and/or jurisdictions (examples in this category are Quest Diagnostics Inc, Laboratory Corporation of America [LabCorp], ARUP Laboratories, Specialty Laboratories, and Mayo Clinic)
   b. Smaller **private**/**university**/**hospital or medical center laboratories** that provide service primarily at the state or local level, but may also process samples for more than one state and/or jurisdiction
2. **Public health laboratories** (PHLs)

Specimen Information

Type of Specimens Shipped to CDC STARHS Laboratory

HIV-positive serum from diagnostic samples confirmed by Western blot (WB) or immunofluorescence assay (IFA) will ultimately be shipped to the CDC STARHS laboratory, depending on the specimen transport model chosen by the originating laboratory. Detailed information about which samples will be shipped is included in the model descriptions of this guidance (see Specimen Transport Options).

Specimen Volume

The optimal quantity of serum required for STARHS testing is 0.5mL per aliquot. However, if less than 0.5mL of the remnant sample is available for testing using STARHS, the sample should still be sent to the CDC STARHS laboratory. The CDC STARHS laboratory is the only laboratory that should determine whether a sample is rejected because of insufficient quantity.

Sample Storage

Short-term (less than 1 week) storage of samples in the refrigerator (temperatures ranging from 2° to 8°C) is acceptable, but for long-term storage (more than 1 week), samples must be frozen at -20°C or colder. This includes any period that the samples are kept at the originating/testing laboratory or the “pass-through” public health laboratory before shipment to the CDC STARHS laboratory or the interim period while STARHS disposition is being determined. Effort should be made to avoid repeated freezing and thawing of samples, as this may give erroneous results.

- If not already in practice, a daily temperature log should be kept to ensure the freezer is operating properly
- The freezer should be housed in a location with proper ventilation to avoid overheating and freezer failure
- The freezer should contain adequate space to store specimens

Specimen Numbering

The specimen number on the samples shipped to the CDC STARHS laboratory will either be the original laboratory-assigned specimen accession number or the STARHS identification number, depending on the transport model selected by the originating laboratory. Detailed information about specimen numbering is included in the model descriptions of this guidance (see Specimen Transport Options).

Specimen Retention

The ISC must coordinate with the laboratory storing HIV-positive remnant sera (the CDC STARHS laboratory and/or their state/local PHL) to identify samples that should be tested using STARHS. However, not all stored samples will be tested using STARHS, and those that will not be tested will have to be identified for disposal. The ISC should regularly notify the storage laboratory about which samples should be tested using STARHS and which should be disposed of by submitting a list of laboratory-assigned specimen accession numbers with “test” or “toss” for each specimen according to the decision reached. The state/local ISC and the storing laboratory should communicate regularly (every 1–3 months) to discuss any specimens for which no disposition has been communicated to determine whether the sample can be disposed of or whether further investigation is needed. Samples should not be destroyed or disposed of until the disposition is definitively determined.

Packaging and Shipping Procedures

Shipping Guidance

Shipping procedures are described in detail in Guidance for Processing, Storage, and Shipping of Specimens to the CDC STARHS Laboratory. Specimens may be shipped from originating laboratories to the state PHL as a pass-through facility or to the CDC STARHS laboratory as diagnostic specimens. However, because of the requirement for dry ice, all laboratories shipping HIV-positive samples must be certified to ship dangerous goods.

Frequency of Shipments

The frequency of specimen shipments to the CDC STARHS laboratory or the pass-through facility should be on a regular schedule, every 1–3 months, and will be determined by the shipping laboratory, considering factors such as specimen retention policies and freezer/storage space, and in consultation with the ISC and the receiving laboratory.

Shipping Couriers

Specimens must be shipped on dry ice by same-day or overnight delivery service to ensure that specimens do not thaw in transit. The shipping laboratory may decide which courier service to use for specimen transport.

Shipments from private laboratories to the PHL may be shipped by Federal Express (or a similar commercial courier) or an established local courier service. Funding permitting, program areas may elect to set up a billing account with Federal Express (or a similar commercial courier) to pay for shipping costs incurred by the private laboratory to either the state/local PHL or the CDC STARHS laboratory (see Funding for Specimen Handling).

Coordinating Shipments

For larger private laboratories that process samples from multiple states and jurisdictions, program areas should collaborate to coordinate specimen shipment mechanisms. A summary of the participating commercial laboratories, primary laboratory contact, primary ISC contact, and shipping arrangements, is available on the HIV Incidence and Case Surveillance Branch (HICSB) password accessible website at:

https://team.cdc.gov/team/cdc/dispatch.cgi/hicsb_Incidence/folderFrame/100027/0/def/61fd

Additional Information for Commercial Laboratories Only

APHL set up a Federal Express billing account for the large commercial laboratories participating through an APHL contract (ARUP Laboratories, LabCorps, and Mayo Clinic) to defer costs of shipping samples to the CDC STARHS laboratory. The APHL Federal Express billing account is available through the end of 2007 only for those commercial laboratories under contract to APHL.

Tracking Shipments

The shipping laboratory should notify the receiving laboratory (state PHL or CDC STARHS laboratory) by fax or email when specimens are shipped, *including the name of the courier and the tracking number of the shipment*. The receiving laboratory will be responsible for tracking the shipments and will notify the originating laboratory if the specimens are not received.

Note: Private laboratories shipping specimens for multiple jurisdictions should provide shipment tracking information to the designated primary ISC responsible for shipping oversight, following procedures established by agreement between the laboratories and participating agencies.

Additional Information Related to APHL-Contracted Commercial Laboratories Only

As part of the contract between APHL and its participating commercial laboratories, APHL will track shipments from these commercial laboratories to the CDC STARHS laboratory. The contracted commercial laboratories must provide APHL with a list of sample numbers sent to the CDC STARHS laboratory. The CDC STARHS laboratory must notify APHL of any shipments sent from an APHL-contracted laboratory that are received by the CDC STARHS laboratory. This notification is needed for billing purposes at APHL.

Sample Rejection Criteria

Sample rejection due to thawing, breakage, insufficient quantity, or lost-in-transit status will be determined and recorded by the CDC STARHS laboratory. The CDC STARHS laboratory will include sample rejection information with the STARHS results report that is transmitted to the ISC.

Confidentiality and HIPAA Regulations

STARHS must ensure that confidentiality is protected and maintained to meet standards for HIV surveillance. The Privacy Rule of the Health Insurance Portability and Accountability Act (HIPAA) regulations permit protected health information to be shared for the purposes of public health surveillance activities[1]. This protection allows the originating laboratories to send specimens labeled with their laboratory-assigned accession number to either the state PHL or the CDC STARHS laboratory, where samples to be tested using STARHS will be reassigned a unique STARHS identification number before testing. This process will minimize relabeling errors and simplify the shipment procedures for private laboratories. However, the state/local public health department must have the laboratory accession number to link the test result to the patient information in the surveillance record. Therefore, the laboratory accession number must be included on the HIV laboratory report sent to the state/local public health department by the originating laboratory.

Funding for Specimen Handling

As a rule, surveillance is not a remunerated activity. However, through a Cooperative Agreement with APHL (# U60/CCU303019-17), funds were made available to provide a predetermined fee to offset personnel, administrative, and handling costs incurred by participating high-volume, multijurisdictional commercial laboratories (ARUP Laboratories, Mayo Clinic, and LabCorps) for an initial start-up period. The APHL Cooperative Agreement was effective through June 2006 for covering specimen handling fees for these participating laboratories, and was extended through the end of 2007 to cover only shipping expenses for these laboratories. Reimbursement for these laboratories beyond the initial start-up period and for all other private laboratories is not covered by the APHL Cooperative Agreement; funding for specimen handling costs may be made available through the state's Cooperative Agreement with CDC for HIV/AIDS Incidence Surveillance (Program Announcement 04017), but handling fee reimbursement is not recommended.

Specimen Transport Options

Option A: Specimen Originated at Private Laboratory and Is Sent Directly to the CDC STARHS Laboratory

Transportation Overview

In this transportation model, the originating private laboratory performing the confirmatory testing will send ***all*** confirmed HIV-positive diagnostic specimens directly to the CDC STARHS laboratory (bypassing the state/local PHL) by overnight shipping in accordance with the procedures described in Guidance for Processing, Storage, and Shipping of Specimens to the CDC STARHS Laboratory. The CDC STARHS laboratory will store specimens until specimen disposition is determined by the state/local ISC, at which point, the CDC STARHS laboratory will pull samples to be tested using STARHS, aliquot, relabel them with a STARHS identification number, and perform STARHS. All samples that should not be tested using STARHS (i.e., samples that are not the diagnostic specimen) will be discarded.

In this model, the originating laboratory would continue to submit laboratory report information in the current manner to the appropriate jurisdiction, but must also include the laboratory-assigned specimen accession number and the collection date on the report.

Figure 1 graphically depicts the flow of specimens and reports when samples originate at a private laboratory and are then shipped directly to the CDC STARHS laboratory.

Procedures for Specimens Sent Directly from a Private Laboratory to the CDC STARHS Laboratory

- The ISC notifies the CDC HIV Incidence Surveillance (HIS) Coordinator about each private laboratory that plans to send specimens directly to the CDC STARHS laboratory. The ISC should send the following information to the CDC HIS Coordinator for each laboratory:
  - Name of laboratory and laboratory point of contact;
  - Full contact information, including mailing address, phone number, fax, and email; and
  - Estimated number of positive samples expected per year.
- The CDC HIS Coordinator will provide the CDC STARHS laboratory with this information from private laboratories. This information is important for the CDC STARHS laboratory to plan for storage capacity.

- The state/local ISC should provide the private laboratory with a copy of the HICSB Incidence Surveillance STARHS Specimen Submission Form. The form should be prefilled with the laboratory contact information, the name and address of the person who will receive the STARHS results (ISC), and the appropriate check box marked for Incidence Surveillance:

  ☑ **INCIDENCE SURVEILLANCE (HICSB)**

- The submitting private laboratory must include a copy of the HICSB Incidence Surveillance STARHS Specimen Submission Form with a list of all laboratory-assigned accession numbers included in the shipment. If more than one specimen box is included in the shipping container, then each box should contain its own STARHS Specimen Submission Form. If possible, an encrypted electronic version of the list of specimen accession numbers should also be included in the shipment; this version will help the CDC STARHS laboratory log the samples with minimal chance of entry errors. ***Note: the manifest should not contain any patient identifiers other than specimen accession numbers.*** At the time of shipment, the submitting laboratory should also mail a copy of the shipping manifest to the ISC, notifying him/her of the shipment. This notification is critical for the ISC to be able to track specimens. As previously noted, identifying information for specimens should not be faxed or emailed, even if encrypted.
- The CDC STARHS laboratory will not provide the private laboratories with any shipping materials, labels or cryovials, but will return the shipping container if a prepaid return air bill is included in the shipment. Surveillance sites may provide the private laboratories with prepaid shipping labels or shipping account numbers (i.e., Federal Express) to cover shipping expenses.
- Specimens shipped from private laboratories directly to the CDC STARHS laboratory will be stored frozen, indicated only by their original laboratory-assigned accession number. On a monthly basis, the ISC will send to the CDC STARHS laboratory a written list of specimens (identified by the original laboratory-assigned accession number, and, if known, the name of the originating lab) that are to be tested using STARHS (test list) or to be disposed of (toss list). The CDC STARHS laboratory will not assign a STARHS identification number unless the ISC notifies the laboratory that the sample is to be tested using STARHS. The CDC STARHS laboratory will continue to hold specimens that are not on one of these two lists.
- The CDC STARHS laboratory will test the specimens and send the STARHS results back to the designated ISC listed on the HICSB Incidence Surveillance STARHS Specimen Submission Form. Samples will be tested and reported by the newly assigned STARHS identification number.
- Periodically, the CDC STARHS laboratory will review the stored specimens to reconcile the status of any samples that have been stored for a lengthy period. This review will reveal any specimens that the ISC never ordered to be tested or discarded. The length of time specimens must be held will vary widely by surveillance site depending on such factors as reporting delays, etc.

Roles of Parties Involved

Role of Private Laboratories

The private laboratories are responsible for forwarding two items for HIV incidence testing: (1) a laboratory report to the public health surveillance department per local requirements, including the collection date, the laboratory-assigned specimen accession number, and identification information about the testing facility; and (2) remnant HIV-positive serum from WB- or IFA-confirmed diagnostic samples labeled with the laboratory-assigned specimen accession number.

The private laboratories may elect to aliquot 0.5 mL of the remnant sera for shipment to the CDC STARHS laboratory so that any additional portion of the remnant sera may be stored at their facility; or they may send the entirety of their remnant sera, without any further manipulation, to the CDC STARHS laboratory.

Before sending shipments to the CDC STARHS laboratory, private laboratories should carefully review Guidance for Processing, Storage, and Shipping of Specimens to the CDC STARHS Laboratory to ensure proper shipping and handling of specimens.

Role of CDC STARHS Laboratory

The CDC STARHS laboratory must store all remnant HIV-positive serum samples received until specimen disposition has been determined by the appropriate jurisdiction's ISC. The ISC will provide the CDC STARHS laboratory with a list of all samples to be tested using STARHS (test list) and a list of all samples to be discarded (toss list), listed by specimen accession number. The samples on the toss list should be discarded according to established laboratory methods.

The CDC STARHS laboratory will pull all samples on the test list and aliquot them into the designated cryogenic vial for testing. The CDC STARHS laboratory will simultaneously relabel the samples to be tested using STARHS with a STARHS identification number. The CDC STARHS laboratory must provide the appropriate ISC with a link between the STARHS identification number and the original specimen accession number. After the ISC has been provided with the linkage information, the CDC STARHS laboratory will destroy the laboratory copy of the specimen accession information.

The CDC STARHS laboratory will test all samples on the test list by the STARHS identification number and send results to the ISC from the appropriate jurisdiction. The STARHS results are for surveillance purposes only therefore results will not be reported back to the originating laboratory, provider, or client.

Role of the State/Local HIV Incidence Surveillance Coordinator

The ISC from the jurisdiction where the specimen originated will determine the disposition of the specimen and coordinate with the CDC STARHS laboratory to ensure that the specimen is either tested or discarded as appropriate. The ISC will also maintain the link between the original specimen accession number and the STARHS number, and will manage the STARHS results.

Specimen Numbering

Specimens will be stored at the CDC STARHS laboratory by the original laboratory-assigned specimen accession number. Once a specimen appears on the test list the sample will be assigned a unique STARHS identification number and will be tested using STARHS. For all subsequent procedures, only the STARHS identification number will be used.

Theoretical Laboratory Types for This Transportation Model

The laboratories that would best use this model are high-volume, multijurisdictional commercial laboratories. However, other private laboratories may also choose this specimen transport model.

Note: The testing laboratory may choose either of the two transport models. With the exception of public health laboratories, the examples listed in this section are merely suggestions, not requirements, for the types of laboratories that may choose this model.

Option B: Specimen Originated at or Sent Via State / Local Public Health Laboratory

Transportation Overview

In this transportation model, confirmatory testing will have been performed at either the PHL or a private laboratory. For samples originating at a private laboratory, that laboratory will send ***all*** confirmed HIV-positive diagnostic specimens to the state/local PHL. The state/local PHL will store all specimens received from the private laboratories until sample disposition is determined by the ISC. All specimens on the test list (those to be tested using STARHS) will be pulled, aliquoted into the designated cryogenic vials provided by the CDC STARHS laboratory, relabeled with a STARHS identification number, and shipped to the CDC STARHS laboratory by overnight shipping in accordance with the procedures described in Guidance for Processing, Storage, and Shipping of Specimens to the CDC STARHS Laboratory. All specimens that are on the toss list (those that are not to be tested using STARHS) will be pulled and discarded according to existing laboratory procedures.

In this model, the originating laboratory would continue to submit laboratory report information in the current manner, but must also include the laboratory-assigned specimen accession number, other relevant specimen identifiers, and testing laboratory identification on the report.

Many laboratories send enzyme immunoassay (EIA) positive specimens to a reference laboratory for confirmatory WB or IFA, which usually results in different laboratory accession numbers. In this case, care must be taken to ensure that the appropriate specimen accession numbers are associated with the correct surveillance report.

Figure 2a graphically depicts the flow of specimens and reports when samples originate at the PHL and are then shipped to the CDC STARHS laboratory. **Figure 2b** describes the flow of specimens and reports when samples originate at a private laboratory and are sent to the PHL for storage before shipment to the CDC STARHS laboratory.

Procedures for Specimens Sent from a Private Laboratory through a State Public Health Laboratory to the CDC STARHS Laboratory

- The ISC works with each private laboratory to set up procedures for shipping specimens to the state PHL.
- The ISC should provide the private laboratory with a copy of the HICSB Incidence Surveillance STARHS Specimen Submission Form. The form can be prefilled with the state PHL contact information, or a new form may be developed in agreement with the PHL. Some private laboratories already have existing mechanisms for transporting specimens to the PHL. These procedures may also be used.
- The submitting private laboratory must include a list of all laboratory-assigned specimen accession numbers included in the shipment to the PHL. At the time of shipment, the private laboratory will mail a copy of the list to the ISC, notifying him/her of the shipment. The private laboratory should also notify the PHL of the shipment by calling or emailing the PHL with the shipment tracking number, if applicable, and the number of samples sent. This information is critical for both parties to be able to track specimens.
- The PHL will store the specimens from the private laboratory, holding them until specimen disposition is determined by the ISC. Specimens should be stored frozen and according to the original laboratory-assigned specimen accession number. On a regular basis, the ISC will notify the PHL which specimens they are storing should be pulled for STARHS testing (test list) and those that can be discarded (toss list).
- The PHL will discard all specimens on the toss list and will prepare all specimens on the test list for testing using STARHS.
    - All specimens to be tested using STARHS will be pulled, thawed, and aliquoted into the designated cryovials provided by the CDC STARHS laboratory.
    - Using labels provided to the PHL by the CDC STARHS laboratory, the PHL will relabel the samples to be tested using STARHS with a unique STARHS identification number.

- The PHL will send the ISC the linkage information between the original laboratory-assigned specimen accession number and the new unique STARHS identification number.
- The PHL will ship all relabeled specimens to the CDC STARHS laboratory according to the procedures described in Guidance for Processing, Storage, and Shipping of Specimens to the CDC STARHS Laboratory.
- The PHL should provide the CDC STARHS laboratory with a completed HICSB Incidence Surveillance STARHS Specimen Submission Form, listing all samples in the shipment by the newly assigned STARHS identification number. The PHL should also include an encrypted electronic version of the specimen list in the shipment; this version will help minimize data entry errors at the CDC STARHS laboratory. The list of STARHS identification numbers on the HICSB Incidence Surveillance STARHS Specimen Submission Form serves as verification from the ISC that all samples in the shipment are to be tested using STARHS.
- A copy of the completed HICSB Incidence Surveillance STARHS Specimen Submission Form should also be mailed to the ISC as notification of the shipment.
- The PHL should also notify the CDC STARHS laboratory of the shipment by calling or emailing the laboratory to provide the shipment tracking number and number of samples sent.

- The CDC STARHS laboratory will test all samples received from the PHL with a preassigned and labeled STARHS identification number and send STARHS results back to the designated ISC listed on the HICSB Incidence Surveillance STARHS Specimen Submission Form.
- Periodically the PHL should review their stored specimens to reconcile the status of any samples that have been stored for a lengthy period. This review will reveal any specimens that the ISC did not order to be tested or discarded. The length of time specimens must be held will vary widely by surveillance site depending on such factors as reporting delays, etc.

Roles of Parties Involved

Role of Private Laboratories

The private laboratories are responsible for forwarding two items for HIV incidence testing: (1) a laboratory report to the public health surveillance department per local requirements with specimen identifiers, the laboratory-assigned specimen accession number, and identification information about the testing facility; and (2) remnant HIV-positive serum from WB- or IFA-confirmed diagnostic samples labeled with the laboratory-assigned specimen accession number and testing laboratory identification information.

The private laboratories may elect to aliquot 0.5 mL of the remnant sera to send to the PHL so that any additional portion of the remnant sera may be stored at their facility, or they may send the entirety of their remnant sera, without any further manipulation, to the PHL.

Before sending shipments to the CDC STARHS laboratory, private laboratories should carefully review Guidance for Processing, Storage, and Shipping of Specimens to the CDC STARHS Laboratory to ensure proper shipping and handling of specimens.

Role of Public Health Laboratories

The PHL must store all remnant HIV-positive serum samples (tested in their own facility or shipped from a private laboratory) until specimen disposition has been determined by the ISC. The ISC will provide the PHL with a list of all samples to be tested using STARHS (test list) and a list of all samples to be discarded (toss list) listed by specimen accession number.

The PHL will pull all samples that will not be tested using STARHS and discard them according to existing laboratory procedures.

The PHL will pull all samples to be tested using STARHS and aliquot them into the designated cryogenic vial provided by the CDC STARHS laboratory. Using labels provided to the PHL by the CDC STARHS laboratory, the PHL will simultaneously relabel the samples with a STARHS identification number. The PHL must also provide the ISC with a link between the STARHS identification number and the original specimen accession number. The PHL will ship all samples to be tested using STARHS, labeled only with the STARHS identification number, to the CDC STARHS laboratory according to the procedures described in Guidance for Processing, Storage, and Shipping of Specimens to the CDC STARHS Laboratory.

Role of State/Local HIV Incidence Surveillance Coordinator

The state/local ISC from the jurisdiction where samples originated will determine the disposition of all samples stored at the PHL and coordinate with the PHL to ensure that specimens are either tested using STARHS or discarded as appropriate. The ISC will also maintain the link between the original specimen accession number and the STARHS number, and will manage the STARHS results.

Role of CDC STARHS Laboratory

Using the STARHS identification number, the CDC STARHS laboratory will test all samples received from a PHL. Once testing is complete, the CDC STARHS laboratory will return results to the appropriate jurisdiction's ISC.

Specimen Numbering

Specimens will be stored at the PHL labeled with the original laboratory-assigned specimen accession number. Once specimen disposition is determined, each sample to be tested using STARHS will be assigned a unique STARHS identification number by the PHL before shipment to the CDC STARHS laboratory. For all subsequent procedures, only the STARHS identification number is used.

Theoretical Laboratory Types for this Transportation Model

The laboratory types that would best use this model are single-jurisdiction laboratories such as hospital, medical center, university, or small independent reference laboratories, or local branches of large commercial laboratories. In many cases, these laboratories already have working relationships and established procedures for submitting samples to their state and/or local PHL and would prefer not to change their existing practices. All state/local PHLs performing confirmatory testing for HIV also fall into this category, except they would simply hold samples until the ISC determines specimen disposition.

Note: The testing laboratory may choose either of the two transport models. With the exception of public health laboratories, the examples listed in this section are merely suggestions, not requirements, for the types of laboratories that may choose this model.

Responsibilities

Private Laboratories

Selection of a Model Type

Each private laboratory performing confirmatory testing of HIV diagnostic specimens must select one of the two specimen transport model types and inform the ISC which model was chosen. The private laboratory must send remnant sera from confirmed HIV-seropositive samples to either the state/local PHL or to the CDC STARHS laboratory.

Additional Laboratory Report Information

The private laboratory must include the laboratory-assigned specimen accession number on the laboratory report form sent to the HIV surveillance department, per state/local disease reporting requirements.

Many laboratories send EIA-positive specimens to a reference laboratory for confirmatory WB or IFA. In this case, care must be taken to ensure that the appropriate specimen accession numbers are associated with the correct surveillance report.

Public Health Laboratories

Sample Storage and Retention

The PHL will often serve a dual function as a testing laboratory or a pass-through facility for private laboratories. The PHL will store all WB- or IFA-confirmed positive samples and/or all samples received from private laboratories until sample disposition is determined by the ISC.

Aliquoting and Sample Shipment

Once sample disposition has been determined by the ISC, the PHL will be responsible for pulling the identified samples, aliquoting samples into the appropriate tubes, and relabeling the samples with a STARHS identification number for testing. The PHL will send the ISC the linkage information between the laboratory-assigned specimen accession number and the STARHS identification number. The PHL will ship all samples to be tested using STARHS (test list) to the CDC STARHS laboratory and discard all samples on the toss list according to existing laboratory procedures.

State/Local HIV Incidence Surveillance Coordinator

Sample Disposition

The ISC will determine sample disposition for all HIV-seropositive diagnostic samples tested in the jurisdiction. The ISC will coordinate with the PHL and/or the CDC STARHS laboratory to ensure the proper samples are tested.

Data Management

To ensure that STARHS results can be matched to surveillance data, the ISC will retain the linkage information between the laboratory-assigned specimen accession number and the STARHS identification number. The ISC will send cumulative incidence data to CDC monthly on or before the 15th of each month. If the ISC is in a local jurisdiction, then results should be sent to the state ISC for matching purposes before submitting data to CDC.

CDC STARHS Laboratory

Sample Rejection Criteria

Sample rejection due to thawing, breakage, insufficient quantity, or lost-in-transit status will be determined and recorded by the CDC STARHS laboratory. The CDC STARHS laboratory will include sample rejection information with the STARHS results report that is transmitted to the ISC.

Sample Storage and Retention

All samples that are shipped directly from a private laboratory and not a state PHL must be stored (see Sample Storage) at the CDC STARHS laboratory until sample disposition is determined by the ISC. Storage time may vary from state-to-state depending on the state's surveillance practices. Once sample disposition has been

determined by the ISC, the CDC STARHS laboratory will be responsible for pulling all samples on the ISC's test and toss lists. The samples on the toss list will be discarded. The samples on the test list will be aliquoted into the appropriate tubes and relabeled with a STARHS identification number for testing. The CDC STARHS laboratory will apply a label to the sample tube and then to a line listing of specimen accession numbers received from the ISC or PHL for those samples to be tested using STARHS. The CDC STARHS laboratory will send the ISC the linkage information and then destroy the linkage information held at the CDC STARHS laboratory. All subsequent testing and results will refer only to the STARHS identification number and will no longer include the original specimen accession number.

Results Reporting

The CDC STARHS laboratory will report STARHS results by the STARHS identification number only to the ISC with jurisdiction over the sample as designated by the HICSB Incidence Surveillance STARHS Specimen Submission Form. The results reporting mechanism will adhere to the methods agreed upon between CDC and the CDC STARHS laboratory.

References

1. CDC. HIPAA Privacy Rule and public health: guidance from CDC and the U.S. Department of Health and Human Services. *MMWR* 2003;52:1–12. Available at http://www.cdc.gov/mmwr/preview/mmwrhtml/m2e411a1.htm.

Figure 1. Specimen originates at national commercial laboratory or private laboratory and is sent directly to the CDC STARHS laboratory

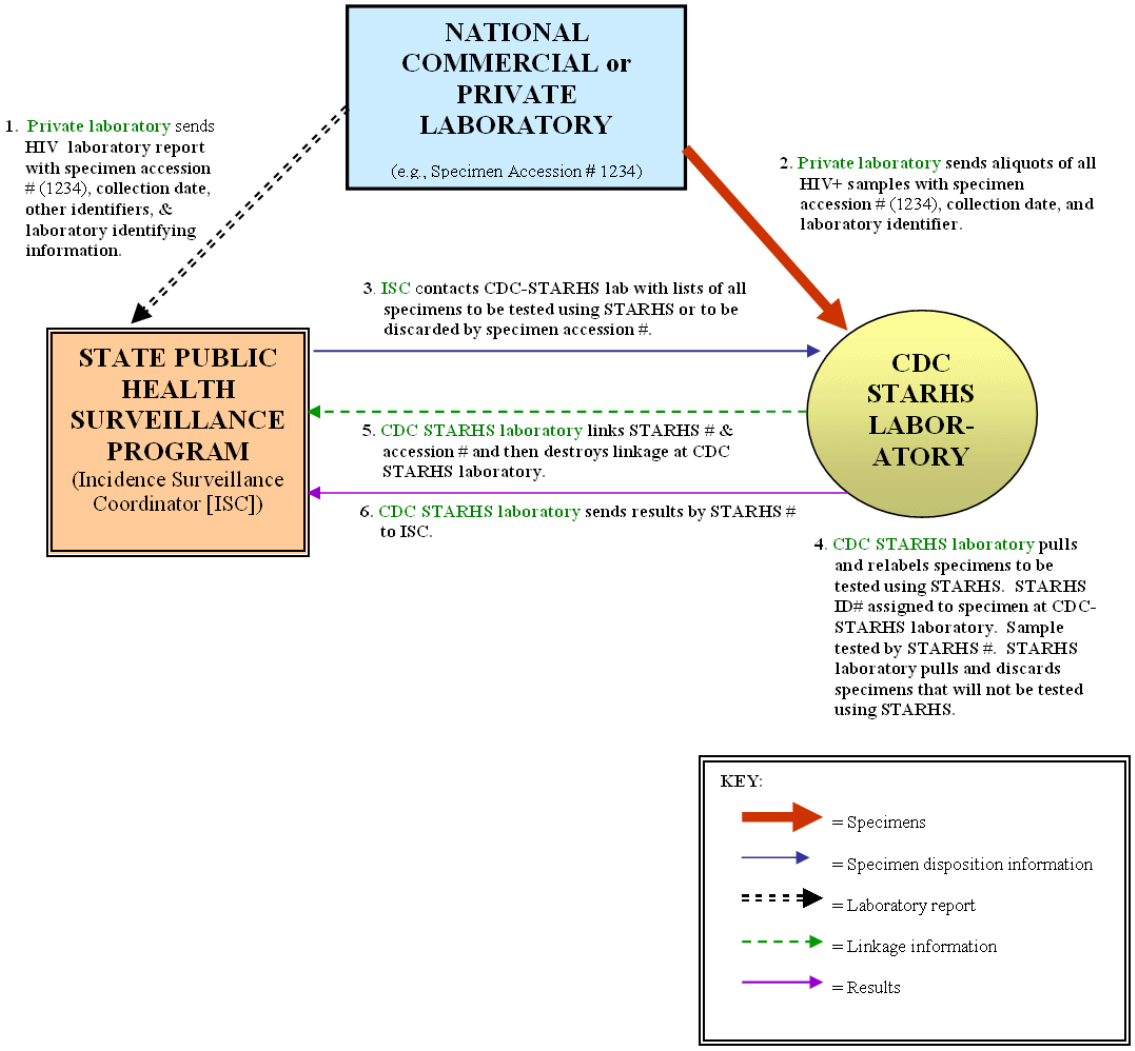


Figure 2a. Specimen originates at a public health laboratory (PHL performed the confirmatory HIV testing)

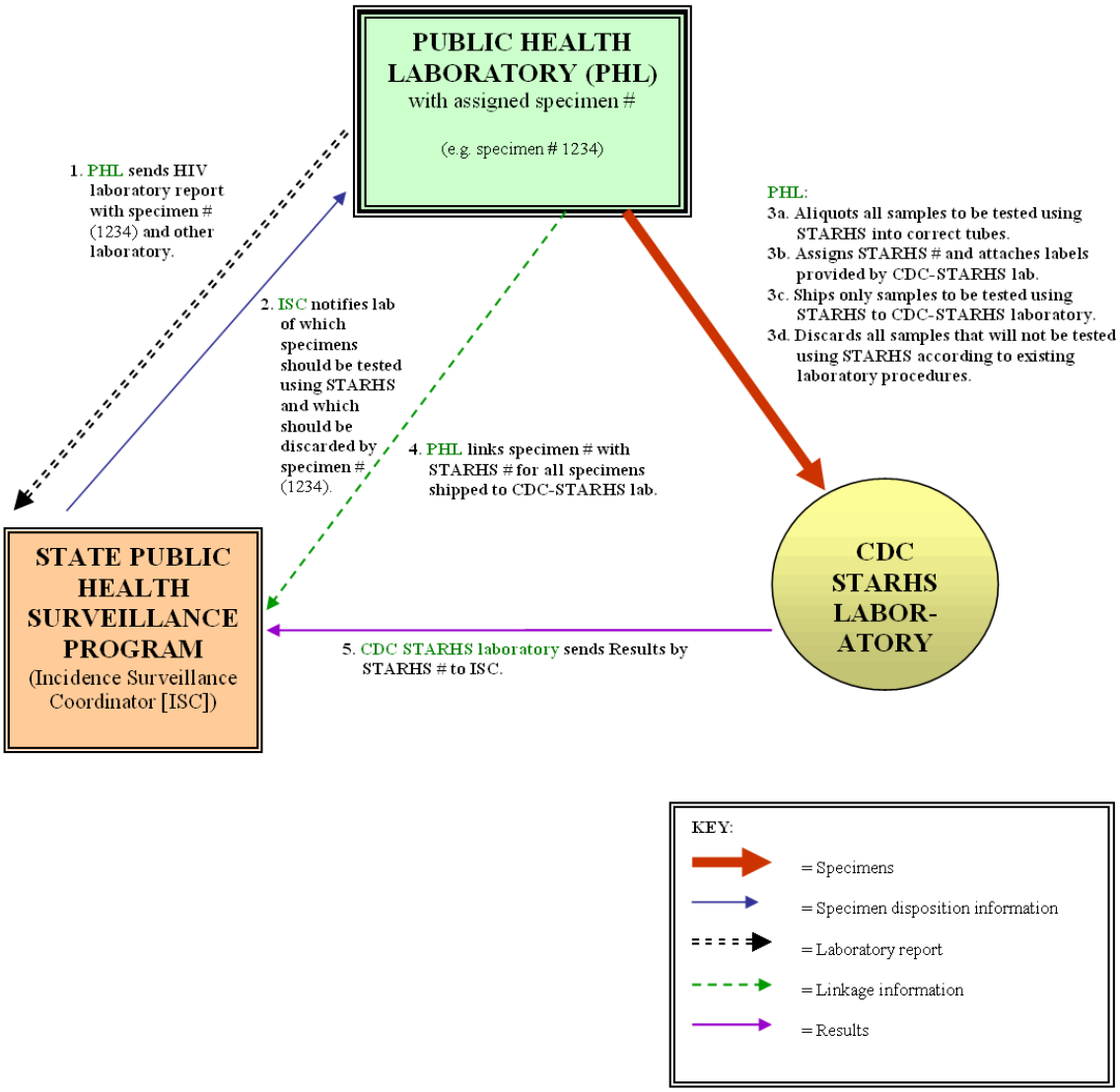
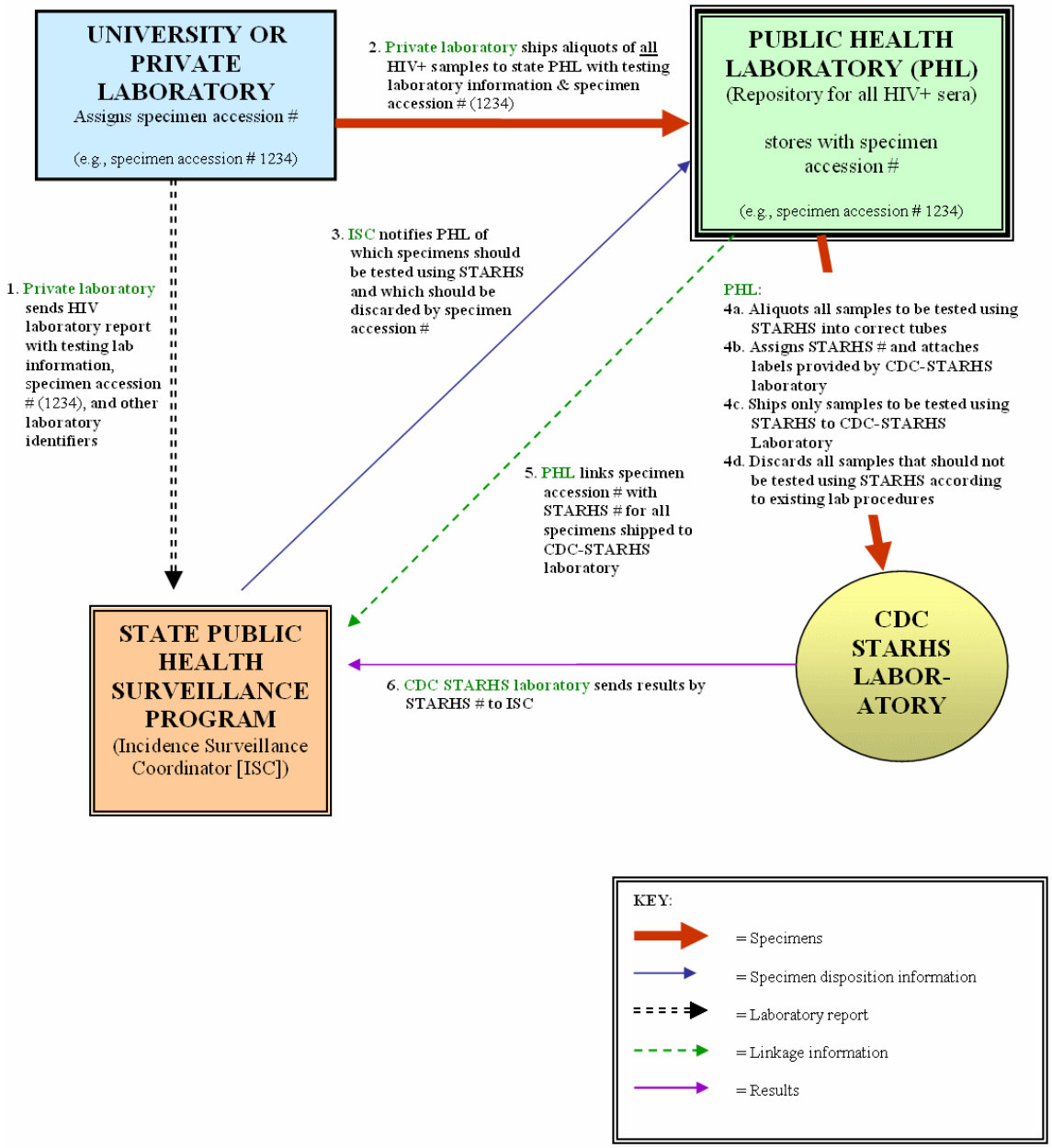


Figure 2b. Specimen originates at a private laboratory (for example, a university hospital laboratory, regional or local independent commercial laboratory) and sample is sent to state public health laboratory (serves as a pass-through facility)



Guidance for Processing, Storage, and Shipping of Specimens to the CDC STARHS Laboratory

Purpose

This standard operating procedure describes methods for the handling, storing, and shipping serum specimens that will be tested for recent HIV-1 infection using STARHS. Results from these tests will help estimate HIV incidence.

Introduction

Remnant serum from HIV-positive diagnostic specimens is to be collected and frozen by using vials and labels specified or supplied by the CDC STARHS laboratory. Ideally, 0.5 mL should be collected for each aliquot. Frozen serum will be shipped to the CDC STARHS laboratory for testing.

CDC STARHS Laboratory

The CDC STARHS laboratory is the Wadsworth Center Retroviral Immunology Diagnostic HIV Testing Laboratory which is part of the New York State Department of Health. Frozen aliquots will be shipped to:

NYSDOH Wadsworth Center
Axelrod Institute
Diagnostic HIV Testing Lab: STARHS
120 New Scotland Avenue
Albany, New York 12208
Attn: N'ko Lea Ali-Napo

Setting and Personnel for Specimen Processing

- Centrifugation, aliquoting, and shipping should be performed at or under the auspices of a laboratory that is certified under the Clinical Laboratory Improvement Amendments (CLIA) for handling HIV+ specimens.
- All personnel handling specimens should receive blood borne pathogens training. See the **Occupational Safety and Health Administration (OSHA) Bloodborne Pathogens Standard**:

  http://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=STANDARDS&p_id=10051
- Personnel handling or processing specimens should have appropriate laboratory training in the relevant laboratory techniques for handling HIV+ specimens and for performing the specific tasks required.

- The setting in which centrifugation, aliquoting, and shipping occurs should meet Biosafety Level 2 specifications required by the U.S. Department of Health and Human Services for handling of specimens containing HIV:
  - *Biosafety in Microbiological and Biomedical Laboratories* (BMBL), 4th ed. Washington: 1999. p. 20–27, 171–175. Available from URLs: http://www.cdc.gov/od/ohs/biosfty/bmbl4/bmbl4toc.htm and http://www.cdc.gov/od/ohs/pdffiles/4th%20BMBL.pdf

Materials

- Cryogenic vials—Supplied by CDC STARHS laboratory
- Specimen labels—Supplied by CDC STARHS laboratory: label will identify sample (barcode, number, etc.) by STARHS identification number
- Cardboard storage boxes for cryogenic vials—Can be supplied by CDC STARHS laboratory if requested
- Freezer—STARHS samples can be refrigerated at 2–8°C, but for long-term storage and shipping, samples should be frozen at -20°C
  - If not already in practice, a daily temperature log should be kept to ensure the freezer is operating properly
  - The freezer should be housed in a location with proper ventilation to avoid overheating and freezer failure
  - Staff must be certain there is adequate space in freezer to store specimens
- A supply of dry ice in pellet form
- Insulated shipping containers certified to ship frozen diagnostic specimens (HIV+ serum and dry ice)
- Shipping courier air bills
- Materials for shipper packing—See Packing Procedures for Shipping to the CDC STARHS Laboratory in this document
- HICSB Incidence Surveillance STARHS Specimen Submission Form

Specimen Collection and Processing

All processing of specimens should be done by personnel qualified to handle HIV+ specimens under the auspices of a laboratory equipped for the handling of HIV+ specimens [*Biosafety in Microbiological and Biomedical Laboratories* (BMBL), 4th ed. Washington: 1999. p. 20–27, 171–175. Available from URLs: http://www.cdc.gov/od/ohs/biosfty/bmbl4/bmbl4toc.htm and http://www.cdc.gov/od/ohs/pdffiles/4th%20BMBL.pdf].

- Aliquot the serum (0.5 mL per cryogenic vial). Use labels to identify the specimen and record this information in the proper setting (specimen log for eventual transfer to HIV Incidence Surveillance database).
- Store aliquots in refrigerator or freezer until specimen disposition has been determined and scheduled shipping date has arrived.

Shipping

- Specimens for STARHS should be sent to the CDC STARHS laboratory at the address on the HICSB Incidence Surveillance STARHS Specimen Submission Form and in the CDC STARHS Laboratory section of this document. All specimens will be shipped as diagnostic specimens according to the International Air Transport Association (IATA) Packing Instructions 650. Dry ice will be included with each shipment per IATA Packing Instructions 904.
  - Because samples will be shipped with dry ice, shipping personnel must be trained and certified to ship dangerous goods. See Training and Certification for Shipping Infectious Substances for a list of companies that provide training.
  - Establish contact with Lea N'ko Ali-Napo (nla01@health.state.ny.us) at the CDC STARHS laboratory.
  - Ensure that adequate STP320 or equivalent shipping containers are available. The CDC STARHS laboratory will return them to the submitting laboratory if a return air bill is included in the shipment. The shippers are expensive and need to be re-used.
  - Ensure that you have an adequate supply of shipping courier air bills, which can be obtained free of charge from most couriers.

Packing Procedures for Shipping to the CDC STARHS Laboratory

All of the following steps should be read and understood **before** starting the preparation of the actual shipment:

- Bring the STP320 shipper or equivalent that is to be used for the shipment and materials needed for packing the specimens into the area in which the shipment will be prepared.
- If the shipper is new and being used for the first time, check to be sure that it includes the following items:
  - Two (2) sheets of bubble wrap
  - Two (2) STP 710 or equivalent certified secondary containers
  - Two (2) 250-mL absorbent strips
  - Class 9 label and dry ice quantity label
  - Other hazard and handling labels
  - One (1) instruction sheet

- For a diagram of the above contents, refer to the Saf-T-Pak catalog.
- Use only what is needed of the above contents for each individual shipment. Save leftover supplies for future shipments.
- If the shipper is being re-used, the proper labels will already be in place on the outer cardboard container. Ensure that adequate supplies of the other materials listed above are on hand.
- Put on personal protective equipment.
- Remove cryogenic vials from freezer and accurately record the specimen accession or STARHS identification numbers. The specimen numbers can either be written directly onto the STARHS Specimen Submission Form or on a separate list that will be attached to the form. Return them to the freezer. Repeat the process until all specimen numbers have been recorded for each vial that is going to be shipped.
- **These specimens should remain frozen at all times and therefore should not be removed from a freezing temperature environment for more than a few minutes**.
- Prepare 3 copies of the HICSB Incidence Surveillance STARHS Specimen Submission Form listing or attaching the specimen number for each vial to be shipped.
  - Copy 1 (original) should be sent with the specimens in the shipment.
  - Copy 2 should be mailed to ISC as notification of shipment.
  - Copy 3 should be retained by the submitting laboratory for its records.
- If possible, on a floppy disk or CD, also include an encrypted electronic version of the list of specimen numbers in the shipment. This will help the CDC STARHS laboratory staff minimize the amount of data entry they have to do when logging in the samples, thereby minimizing errors.
- Prepare the shipping courier air bill that the CDC STARHS laboratory will use to return the shipper back to the submitting laboratory for re-use. The air bill MUST be completely filled in with the return address, the CDC STARHS laboratory address, and the proper billing number.
- If dry ice is in another location which requires leaving the area in which the shipment is prepared, use a separate container to bring the dry ice that is needed for shipping back into the shipping area at this time.
- Bring the specimens to the area in which the shipment is prepared. Work quickly, keeping in mind that **these specimens should remain frozen at all times and therefore should not be removed from a freezing temperature environment for more than a few minutes**.
- Re-check the screw-cap lids on the specimen vials and tighten if necessary.

- Place the specimens into the secondary leak-proof container and make sure samples are surrounded by bubble wrap and absorbent strips. The vials should not move around or rattle inside the vessel.
- Place the secondary vessel into the inner box and place the inner box into the polystyrene cooler.
- Pack dry ice pellets in the shipper and around the inner box. The STP320 shipper will hold ~8 kg of dry ice (~10 lb) and, if packed completely, will keep the contents frozen for greater than 80 hours.
- DO **NOT** PUT DRY ICE INSIDE THE INNER BOX.
- Place the lid on the polystyrene cooler.
- Place one copy of the completed HICSB Incidence Surveillance STARHS Specimen Submission Form on top of the shipping box return form with the completed *return* FedEx air bill stapled to it. Fold in half and place on top of the polystyrene lid.
- Fold over the top flaps and seal the shipping container with clear shipping tape.
- The outer box must have a mark in the form of a square set at an angle of 45° (diamond shaped). The mark must be at least 2 inches by 2 inches and include the UN 3373 designation. The proper shipping name—"Diagnostic specimens"—must be marked on the outer package adjacent to the diamond-shaped mark. Labels can be purchased to place on the outer box that fulfill this requirement.
- Apply the Class 9 Hazard Label over the lower diamond-shaped outline on the box.
- Apply the net quantity dry ice label to the outlined area adjacent to the Class 9 Hazard Label. Write the approximate amount (in kg) of dry ice you used to pack the container.
- Prepare the shipping courier paper work addressed to the CDC STARHS laboratory. Select the overnight shipping option.
- Call or email the CDC STARHS laboratory to notify them of the shipment. Provide the CDC STARHS laboratory with the shipment tracking information and the total number of samples in the shipment.

Note: ***Do not fax or email laboratory-assigned specimen accession numbers or STARHS identification numbers.***

HICSB Incidence Surveillance
STARHS Specimen Submission Form

Please complete this form and send it with each shipment. Specimens should be sent to

NYSDOH Wadsworth Center
Axelrod Institute
Diagnostic HIV Testing Lab: STARHS
120 New Scotland Avenue
Albany, NY 12208
Attn: N'ko Lea Ali-Napo

SHIPPING FACILITY INFORMATION:

Name: _____

Address: _____

_____

Phone Number: _____

Fax: _____

Email: _____

Contact Person: _____

RESULTS SENT TO:

Name: _____

Address: _____

_____

Phone Number: _____

Fax: _____

Email: _____

Mark the box for the appropriate surveillance activity for these specimens:

❑ **INCIDENCE SURVEILLANCE (HICSB)**

Note: List of which specimens are appropriate to test for HIV incidence surveillance may be sent separately.

❑ **BEHAVIORAL SURVEILLANCE (BCSB)**

❑ **EVALUATION OF DRIED FLUID SPOT SURVEILLANCE (DFS)**

RANGE OF SPECIMEN NUMBERS SENT (OR ATTACH LIST):

_____

_____

_____

_____

Please identify any specimens on your list that are collected under a research protocol and that should not be tested using the BED HIV-1 Capture EIA.

01/08/2007

Training and Certification for Shipping Infectious Substances

FedEx 800-GO-FEDEX

- 3-day, IATA-based training covers all hazardous materials. Cost is $650.

Saf-T-Pak 800-814-7484

- Specifically for infectious and diagnostic substances and dry ice. Three options: one-day seminar, on-site programs, or interactive CD (staff can be trained in 3 to 5 hours using interactive CD). Certificate is good for 2 years OR until regulations change. Cost is ~$250.

Viking Packing Specialist (Oklahoma; David Weilert, President) 800-788-8525

- Seminars conducted monthly in Tulsa. Cost is $300 per person. Covers all nine classes of hazardous materials, covers shipping under IATA, and certificate is good for 2 years. Will do group classes in local area ($3,000 plus travel costs).

These are some companies that provide training for dangerous goods shipping. The Centers for Disease Control and Prevention does not endorse any particular company.

Appendix A

Statistical Method for Generating Population-Based HIV Incidence Estimates

This document is in development and will be added soon.

Technical Guidance for HIV/AIDS Surveillance Programs

Variant, Atypical, and Resistant HIV Surveillance (VARHS)

HIV Incidence and Case Surveillance Branch
Atlanta, Georgia

Notes

Contents — Variant, Atypical, and Resistant HIV Surveillance (VARHS)

Policies and Procedures for Variant, Atypical, and Resistant HIV Surveillance (VARHS) ... 13-5
- Background ... 13-5
- Objectives ... 13-6
- Methods ... 13-7
- Public Health Benefit ... 13-7
- Structural Requirements ... 13-8
  - Policies and Procedures ... 13-8
  - Laboratory Procedures ... 13-10
  - Staffing Needs ... 13-11
- Process Standards ... 13-13
  - Determining Eligibility ... 13-14
    - Inclusion criteria ... 13-14
    - Exclusion criteria ... 13-14
  - Identifying Laboratories and Obtaining Remnant Specimens or Sequence Data ... 13-15
  - Specimen Handling and Transport Procedures ... 13-16
    - Tube type for HIV diagnostic specimens ... 13-16
    - Specimen handling ... 13-16
    - Serum and plasma processing ... 13-16
    - Cryovials and aliquots ... 13-17
    - Diagnostic and clinical testing and freezing of serum or plasma aliquots ... 13-17
    - Specimen tracking data elements to be recorded in VARHS specimen log or database ... 13-18
    - Shipment to the VARHS genotyping laboratory ... 13-18
  - Required Data Elements ... 13-19
    - Data used to determine VARHS eligibility ... 13-19
    - Demographic and clinical data ... 13-20
    - Specimen tracking data ... 13-20
    - Laboratory data ... 13-20
  - Returning Genotyping Results ... 13-21
  - Ensuring Security and Confidentiality ... 13-22
  - Data Management ... 13-23
    - CDC software ... 13-23
    - Identification of VARHS specimens ... 13-23
    - Data encryption and transfer ... 13-24
  - Data Analysis and Data Dissemination ... 13-25
- Outcome Standards ... 13-25
- References ... 13-26

Appendix A 13-29
Appendix B 13-30
Appendix C 13-31
Appendix D 13-32
Appendix E 13-33
Appendix E.1 Training and Certification for Shipping Infectious Substances... 13-42
Appendix E.2 CDC VARHS Shipping Manifest 13-43
Appendix E.3 Stanford Laboratory's Return of Shipping Container 13-44
Appendix F 13-45

Technical Guidance for HIV/AIDS Surveillance Programs — Policies and Procedures for Variant, Atypical, and Resistant HIV Surveillance (VARHS)

Background

The Centers for Disease Control and Prevention (CDC) is responsible for maintaining a national surveillance system that provides data on the HIV/AIDS epidemic that can be used for national, state, and local public health HIV/AIDS prevention program planning and evaluation. Clinical and laboratory testing information is collected to characterize the epidemic and guide public health action. HIV-1 genetic sequence data based on the *pol* region (protease and reverse transcriptase genes) have been incorporated into the HIV/AIDS surveillance system to evaluate the distribution of HIV-1 mutations associated with HIV drug resistance and subtypes among persons with a new diagnosis of HIV infection.

In the late 1990s, several new nucleoside reverse transcriptase inhibitors (NRTIs), nonnucleoside reverse transcriptase inhibitors (NNRTIs), and protease inhibitors (PIs) were approved for treating HIV infection in the United States. These newer drugs, along with the NRTIs already available, provide clinicians with a variety of choices for initiating and changing antiretroviral treatment for patients infected with HIV-1. An international panel with expertise in antiretroviral research and HIV patient care, convened by the International AIDS Society–USA and an interagency work group of the Department of Health and Human Services (DHHS) and the Public Health Service, provide recommendations for prophylaxis or therapy that include all of the antiretroviral drugs approved by the Food and Drug Administration (FDA) and in use in the United States (1).

The therapeutic purposes of antiretroviral drugs include prophylaxis after known occupational exposure (postexposure prophylaxis), vertical transmission prophylaxis, treatment of primary infection (4 to 7 weeks after infection), initial treatment from early (little or no immunologic damage) to late (substantial immunologic damage) infection and changes in treatment regimens depending on virologic and immunologic response (1–5). Clinical trials are being conducted to evaluate the use of antiretroviral drugs for preexposure prophylaxis. Studies have demonstrated that the results (genotype and phenotype) of HIV drug-resistance testing can be used to predict clinical outcome and to guide drug treatment choices (5). Also, studies have shown that the results of *pol* region sequencing that indicate the presence of mutations associated with drug resistance predict phenotypic sensitivity to antiretroviral drugs and clinical response (1, 5). Studies have also demonstrated that the HIV-1 protease and reverse transcriptase genes of the *pol* region can be used to characterize HIV-1 subtypes and that a high prevalence of HIV-1 subtypes other than subtype B in a geographic area has implications for the appropriate selection of HIV diagnostic and clinical tests for populations and individuals (6).

Given the public health concerns about primary HIV drug resistance, the clinical impact on persons with a new diagnosis of HIV infection and the uncertain clinical significance of resistant HIV strains, CDC funded a project called Antiretroviral Drug Sentinel Surveillance to Examine Trends in Prevalence of Drug-Resistant Strains of HIV in 2002.

This demonstration project showed that it was feasible to use remnant sera for genotyping; in turn, this led to the inception of variant, atypical, and resistant HIV-1 surveillance (VARHS).

On June 25, 2004, VARHS received a non-research determination from the CDC Human Subjects Office (Appendix A). The document that serves as the basis for the federal regulations for protecting human research participants (Title 45 of the Code of Federal Regulations part 46, 45 CFR 46) is available at http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.htm). Subpart A, subsection 46.101c states that "the [federal] Department or Agency heads retain final judgment as to whether a particular activity is covered by this policy." The policy refers to the requirements for human subjects protection review for research protocols under the 45 CFR 46 regulations. *Disease surveillance* is one of the 4 major public health practice activities that usually involve data collection but are not, according to 45 CFR 46, research and thus do not need review by an institutional review board (IRB). Although CDC determined that VARHS is a disease surveillance activity by CDC (IRB review not required), this determination does not supersede state or local laws or regulations that may require IRB notification or review for public health surveillance activities.

On July 1, 2004, VARHS was incorporated nationally into routine HIV surveillance, and 22 surveillance areas were funded to participate in VARHS (Appendix B).

Objectives

The primary objectives of VARHS are to

1. Incorporate surveillance of transmitted strains of variant, atypical, and resistant HIV-1 into routine HIV surveillance activities by
    - amplifying and sequencing the *pol* region (protease and reverse transcriptase genes) from persons with a new diagnosis of HIV infection who have been reported to the HIV/AIDS Reporting System (HARS)
    - estimating the prevalence of mutations associated with HIV drug resistance, resistant patterns, and trends in dissemination of drug-resistant strains
    - determining the prevalence of genetic subtypes to gain insight into the genetic diversity of strains and distribution of B and non-B variants among populations
2. Provide data on HIV-1 drug resistance, subtypes, and factors associated with resistance to assist HIV treatment, prevention, and program planning and evaluation

Methods

VARHS is used to evaluate the prevalence of HIV-1 drug resistance and the distribution of HIV-1 subtypes among persons with a new diagnosis who have been reported to HARS by state, county, or city departments of health. Specimens from these persons should be included by using either of 2 methods.

First, if sufficient volume is available, aliquots of remnant diagnostic specimens should be set aside for HIV-1 drug-resistance testing from each blood specimen drawn for HIV testing from eligible persons in participating areas. For persons meeting VARHS criteria, HIV genetic sequencing (i.e., genotyping) will be performed on the protease and reverse transcriptase genes to detect the presence of mutations associated with HIV-1 drug resistance. HIV-1 subtypes will also be identified on the basis of these genes.

HIV-1 drug-resistance testing for VARHS is performed with standard tests, which are widely used clinically. Like drug-resistance testing in other infectious disease surveillance systems (e.g., tuberculosis, urinary tract infections, and sexually transmitted diseases), testing diagnostic specimens for HIV-1 drug resistance and HIV-1 subtype surveillance is not experimental and does not require informed consent (7–9). Areas may choose to include information in their HIV testing forms to inform persons that HIV drug-resistance testing is part of HIV diagnostic testing in participating areas. For persons with a new diagnosis who have been reported and whose specimens are successfully amplified and genotyped through the VARHS system, the health department will provide a hard copy of the results to the health care provider(s) designated by the persons tested.

Second, if private-sector laboratories performed genotyping on specimens from persons eligible for inclusion in VARHS, surveillance areas should request electronic sequence data from the laboratories.

Genotyping results and information from the HIV surveillance case report will be used to estimate the prevalence of HIV-1 drug resistance and the distribution of HIV-1 subtypes among persons with a new diagnosis who have been reported as having HIV infection. Prevalence will also be estimated for relevant demographic groups and HIV transmission categories. HIV sequence information may also be used to track the spread and clustering of atypical HIV strains that are of interest nationally.

Public Health Benefit

Although previous surveys have been based on convenience samples of specimens, VARHS is the first surveillance system in the United States designed to evaluate all persons who have been given a diagnosis of HIV infection and reported to HARS/eHARS from participating areas. VARHS will determine the distribution of drug-resistant strains and subtypes among persons with a new diagnosis who have been reported as having HIV-1, thus supporting efforts to characterize and track the HIV epidemic nationally. VARHS will also provide information on trends in the transmission of drug-resistant strains and the factors associated with resistance. Analyses will support evaluation of first-line

antiretroviral drugs and provide information that will be useful in developing new antiretroviral drugs. Strategies for using specific prophylactic drugs will be enhanced by information from the surveillance of HIV drug resistance.

VARHS will support evaluation of the usefulness of testing baseline clinical HIV-1 drug resistance in specific geographic areas. Current DHHS guidelines for the treatment of HIV-1 infection recommend baseline resistance testing for all HIV patients before the initiation of antiretroviral therapy. Even when therapy is deferred, the guidelines suggest that baseline testing should be considered: "In the absence of therapy, resistant viruses may decline over time to less than the detection limit of standard resistance tests but may still increase the risk of treatment failure when therapy is eventually initiated" (1). The DHHS guidelines cite a 2005 cost-effectiveness analysis that concluded, "A strategy of genotype-resistance testing at initial diagnosis of HIV infection increased per-person quality-adjusted life expectancy by 1.0 months, with an incremental cost-effectiveness ratio of $23,900 per quality-adjusted life-year gained, compared with no genotype testing. The cost-effectiveness ratio for resistance testing remained less than $50,000 per quality-adjusted life-year gained, unless the prevalence of resistance was ≤1%, a level lower than those reported in most regions of the United States and Europe…. Genotype-resistance testing of chronically HIV-infected, antiretroviral-naive patients is likely to improve clinical outcomes and is cost-effective…. Resistance testing at the time of diagnosis should be the standard of care" (10).

VARHS will also provide public health and clinical personnel with information on the distribution of HIV-1 subtypes in the United States over time. If the prevalence of HIV-1 non-B subtypes increases, this finding could have implications for vaccine studies. Also, should specific drug-resistant mutations be associated with some non-B subtypes, there could be implications for treatment guidelines.

In the surveillance of other organisms, such as *Mycobacterium tuberculosis*, molecular surveillance has allowed identification of atypical strains of special interest (9, 11). Although no such HIV strains have been identified, routine surveillance based on sequencing and phylogenetic analyses of the HIV *pol* region will allow such strains to be detected. Routine surveillance will also provide information on multidrug-resistant strains in various geographic areas and specific population subgroups. HIV-1 genetic sequencing surveillance data collected over time will also contribute to analyses of the evolution of HIV-1 in the United States.

Structural Requirements

Policies and Procedures

All persons with a new diagnosis of HIV infection after confidential testing should be reported to HARS in accordance with *Technical Guidance for HIV/AIDS Surveillance Programs, Vol. I: Policies and Procedures* (12). VARHS is an extension of the national population-based HIV/AIDS case surveillance system, and state and local health

department partners and CDC will use the case surveillance infrastructure to collect the information necessary to estimate the prevalence of transmitted HIV-1 drug resistance and HIV-1 subtypes.

VARHS will be implemented in areas that also conduct HIV incidence surveillance. CDC recommends that HIV incidence surveillance procedures be in place before VARHS implementation is planned. In some circumstances, however, the simultaneous implementation of both surveillance systems may be beneficial. Before implementing VARHS, surveillance areas must

- Consult their CDC HIV incidence surveillance epidemiologist and the CDC VARHS project officer or coordinator about the appropriate time to plan and implement VARHS
- Develop VARHS guidance outlining local policies and procedures, which must be reviewed and approved by CDC
- Host a site visit by CDC
- Document CDC approval before collecting specimens for VARHS

Because VARHS is an integrated component of HIV/AIDS surveillance, documentation of VARHS activities should be incorporated into locally tailored policy and procedures manuals developed for HIV/AIDS surveillance (12) to establish standardization, maintain continuity of meaning, document changes over time, and develop training programs. In addition to the information in *Technical Guidance for HIV/AIDS Surveillance Programs, Vol. I: Policies and Procedures* (12), VARHS-specific policies and procedures should include information related to

- Responsibilities of staff
- Training in VARHS procedures for health care providers, counselors, and other staff at participating testing sites and clinics
- Identification of participating laboratories
- Eligibility determination
- Specimen handling and transport procedures
- Collection of VARHS-specific specimen tracking and laboratory data elements
- Entering and importing data into the Incidence and Viral Resistance (IVR) database
- Data management practices
- Return of genotyping results
- Monthly data transmission to CDC
- Security and confidentiality
- Local analysis and dissemination plan
- Expansion plan (including a timeline for jurisdiction-wide implementation of VARHS)

At the end of some sections in this document, a box (similar to the one headed Sample VARHS Implementation Task) is used to present task-specific guidance that the surveillance area will need to address in order to implement VARHS. The local VARHS guidance should include a description of local plans and policies and procedures related to each task, as the completion of these tasks is key in the successful implementation of VARHS.

Sample VARHS Implementation Task

VARHS areas should complete each implementation task that appears at the end of selected sections in this document.

Laboratory Procedures

Specimen handling and transport procedures are stringent for VARHS, because proper specimen handling and transport contribute to the rate of successful amplification for genetic sequencing. Specimen handling and transport procedures should be decided by the public health laboratory and local sites. If necessary, procedures can be modified when new aspects of the surveillance system are instituted. At a minimum, surveillance areas planning to implement VARHS need to take the following laboratory procedures into consideration.

- If volume seems sufficient for the purposes of all basic laboratory tests and other local priorities, such as archiving and HIV incidence surveillance, a 1 mL (optimal) aliquot from all enzyme immunoassay (EIA)–reactive specimens should be frozen at –70ºC as soon as possible after the first reactive EIA.
- The following will maximize chances for successful amplification for genetic sequencing
  - Plasma and serum should be separated as quickly as possible, and neither should be refrigerated until after separation.
  - HIV diagnostic testing or other relevant clinical testing should begin immediately after separation of plasma or serum.
  - Sera or plasma should be kept on the bench at room temperature for as short a time as possible before being returned to the refrigerator or placed on ice.
  - Once frozen, specimens for genotyping **should not** be thawed until they reach the genotyping laboratory. A freeze-thaw cycle will reduce the chance of successful amplification for genotyping.
  - Specimens should be shipped on dry ice to the genotyping laboratory.
- A system must be in place to collect all required data elements for specimen tracking and to transmit them to CDC (see Appendix F).

Staffing Needs

Implementation of VARHS requires personnel with specific skills and dedicated time to effectively integrate VARHS into the core HIV/AIDS surveillance system. In general, personnel who work primarily on core HIV/AIDS surveillance or HIV incidence surveillance will develop work plans to integrate VARHS into HIV/AIDS surveillance.

VARHS staff should have

- An understanding of VARHS and the characteristics of the local HIV/AIDS surveillance system, including HIV incidence surveillance
- An understanding of how HIV diagnostic laboratories function and how they are related to the public health system
- Good communication skills
- Strong leadership skills
- An ability to work closely with CDC, other states, local HIV diagnostic and clinical sites, private providers, and laboratories

CDC recommends a full-time VARHS coordinator position dedicated to implementing and maintaining the system. The number of other personnel assigned to VARHS may vary, depending on the implementation phase, prevalence of HIV/AIDS, and available resources.

VARHS coordinator

- Provide overall management of VARHS
  - Develop local VARHS guidance (to be reviewed and approved by CDC)
  - Serve as the lead for area-specific implementation of the VARHS technical guidance
  - Serve as the primary CDC contact for VARHS
  - Manage employee or other service contracts related to VARHS
- Develop training materials and courses, including presentations for local clinical and diagnostic sites and laboratories to introduce VARHS
- Train providers, counselors, and other staff at participating clinical and diagnostic sites
- Oversee data collection processes
  - Directly manage or oversee the process for determining the eligibility of specimens for VARHS
  - Work with the genotyping laboratory to develop procedures for shipment of specimens and receipt of genotyping results
  - Develop and manage the system for returning genotyping results to designated providers
- Collaborate with other VARHS staff
- Participate in CDC site visits, trainings, and workshops

Epidemiologist

- Plan and implement the integration of VARHS activities into HIV core surveillance and HIV incidence surveillance activities
- Develop a VARHS analysis plan
- Develop systems to ensure data quality, analyze local VARHS data, and produce reports
- Participate in data dissemination activities
  - Collaborate with stakeholders to determine data needs and frequency of reporting
  - Identify results and surveillance issues for review and dissemination
  - Develop a data dissemination plan in collaboration with the VARHS coordinator and CDC
- Participate in CDC site visits, trainings, and workshops

Laboratory liaison

- Serve as the liaison between the public health department and laboratories from which specimens are being shipped
- Oversee transfer of specimen tracking data from laboratories to the IVR database at the health department
- In collaboration with liaisons based at participating laboratories, develop local procedures for processing and shipping specimens to the genotyping laboratory
- In collaboration with liaisons based at participating laboratories, develop quality control procedures for preparing specimens
- Develop and oversee procedures to maintain security and confidentiality of specimen identifiers and results
- Participate in CDC site visits, trainings, and workshops

Laboratory liaison based at a participating HIV diagnostic laboratory

- Work with the VARHS coordinator and the health department laboratory liaison to develop the laboratory-specific plan for processing, storing, determining eligibility, tracking, and shipping specimens to the genotyping laboratory
- Oversee preparation and shipping of VARHS specimens to the genotyping laboratory
- Monitor quality control procedures for preparing VARHS specimens
- Record specimen tracking data in the log or database

Data manager

- Assist the VARHS coordinator with daily management of VARHS data
- Serve as subject matter expert on VARHS data elements and data management programs

- Receive data from other health department entities and the genotyping laboratory, and incorporate those data into the IVR database and data sets for transfer to CDC
- Assess data quality
- Obtain a Secure Data Network (SDN) certificate and renew the certificate annually to allow for data transmission to CDC
- Identify at least 1 backup person to hold an SDN certificate and maintain the list of persons with certification, notifying CDC of changes
- Oversee VARHS data management
  - Assess data collection methods
  - Modify CDC's generic data management programs for use at the local level
  - Maintain the IVR database
  - Run SAS statistical software programs to produce monthly data sets
  - Develop and implement edit checks and edit data
  - Collaborate with the VARHS coordinator, epidemiologist, and other area surveillance and prevention staff on data editing, data entry, and data set preparation
  - Prepare monthly data set for transmission to CDC
  - Transfer data through SDN to CDC
  - Prepare data sets for local analysis
  - Collaborate with CDC on data set preparation for national resistance estimates
  - Prepare HIV resistance data reports for local use in collaboration with the VARHS coordinator, epidemiologist, and CDC
- Maintain security and confidentiality of data
- Participate in CDC site visits, trainings, and workshops

VARHS Implementation Task

Areas should identify the staff needed to implement VARHS and define staff roles, responsibilities, and the time each staff member should spend on VARHS.

Process Standards

VARHS involves the following processes:

- Determining eligibility
- Identifying laboratories and obtaining remnant specimens or sequence data
- Developing specimen handling and transport procedures
- Collecting required data elements

- Returning genotyping results
- Ensuring security and confidentiality
- Managing data
- Analyzing and disseminating data

Determining Eligibility

Specimens confirmed as HIV-1–positive by Western blot or other methods acceptable to the department of health are potentially eligible for VARHS. All specimens yielding indeterminate or negative Western blot results in the absence of other proof of HIV infection should be handled according to standard laboratory procedures for handling, storage, and discarding specimens.

The VARHS coordinator will identify the confirmed positive frozen aliquots that are eligible for VARHS by matching laboratory and HARS data (or data from an equivalent HIV/AIDS reporting system). The VARHS coordinator also will identify cases with electronic sequence data that are eligible for VARHS. Eligibility should be established according to the following inclusion and exclusion criteria at the case and the specimen level.

Inclusion criteria

- The specimen is the diagnostic specimen (the HIV-1–positive specimen that led to the case report to HARS).
- If the diagnostic specimen is unavailable, a later specimen is eligible provided that
  - the later specimen was drawn <3 months after the date of the diagnostic specimen **and**
  - the later specimen was drawn for an HIV-related test (viral load, polymerase chain reaction, CD4 count).
- Electronic sequence data are eligible if the specimen used for genotyping was drawn <3 months after the date of the diagnostic specimen.

Exclusion criteria

- The specimen is not from the diagnostic specimen that led to the report to HARS.
- The specimen is from an HIV-related test (viral load, polymerase chain reaction, CD4 count) or a genotyping test and was drawn >3 months after the date of the diagnostic specimen.
- The person had received antiretroviral drugs before the VARHS specimen was collected.
- There is no method to link the person to HARS (i.e., the person was tested anonymously).
- An earlier specimen had been sequenced and documented in VARHS.
- The person’s diagnosis was HIV-2 infection.
- The specimen was collected before the surveillance area implemented VARHS.

How often HARS is checked for eligibility is left to the discretion of the surveillance area; however, at a minimum, HARS should be checked to determine VARHS eligibility

- before sending specimens for genotyping **and**
- before each monthly data transfer to CDC.

All specimens that are not eligible for VARHS should be handled according to local standard laboratory procedures.

Identifying Laboratories and Obtaining Remnant Specimens or Sequence Data

In each surveillance area, laboratories that could participate in VARHS will be identified from a review of local HIV surveillance data and laboratory licensing records. In practice, VARHS will usually begin in public health laboratories, which collaborate closely with the health department.

The preferred default practice for VARHS is to have eligible specimens shipped by the central public health laboratory to the CDC-contracted genotyping laboratory for VARHS testing. Deviations should be discussed with CDC. Shipping costs associated with sending specimens from state public health laboratories to the genotyping laboratory should be paid from the funds provided through CDC's cooperative agreement with each participating area. Surveillance areas are responsible for including estimated shipping costs in their request for funds and for budgeting accordingly.

To minimize the possibility that a person with a new diagnosis might begin taking antiretroviral drugs before a specimen is collected, HIV genetic sequencing for VARHS must be performed on an HIV-1–positive specimen collected at the time of HIV diagnosis or at the time of a specimen collected at a follow-up visit no more than 3 months after diagnosis. If volume and logistics permit, a remnant specimen for VARHS will be obtained from all eligible, confirmed HIV-positive diagnostic specimens. Specimen types suitable for VARHS are serum, plasma, and whole blood.

The use of dried fluid spots for VARHS is being evaluated as an alternative type of specimen collection but has not been approved for use in VARHS.

National commercial reference laboratories and private laboratories where HIV diagnostic or clinical testing is performed can also participate in VARHS, if resources are available, to allow remnant specimens to be processed within the time frames described in the following sections or to allow time for the receipt of genotyping results (i.e., sequence data). Reimbursements to local commercial and private laboratories will be determined, on the basis of local policies, by each surveillance area.

When sequence data are available from commercial or private laboratories for cases that are eligible for VARHS, the following should be done:

- Secure the fasta files (standard text-based format for representing nucleic acid sequences) containing the nucleotide sequence.
- Match the sequence data with case information in the IVR database.
- Indicate in the appropriate field in the IVR database that the specimen was handled through a commercial or private laboratory.

VARHS Implementation Task

Identify laboratories that are required to report HIV-positive test results, the total number of new diagnoses made annually in each laboratory, and the number of HIV-positive specimens suitable for VARHS (plasma, serum, or whole blood).

Specimen Handling and Transport Procedures

Specimen handling and transport procedures should be decided by the public health laboratory and local sites. If necessary, procedures can be modified when new aspects of the surveillance system are instituted. The following sections describe tube types, transport times, and processing times that are optimal for VARHS. Before implementing VARHS, surveillance areas should take these factors into consideration when developing local procedures.

In general, 1 or more central public health laboratories will process and ship VARHS specimens to the genotyping laboratory. (See Appendix E for detailed shipping procedures for serum and plasma specimens.)

Tube type for HIV diagnostic specimens

The decision on the type of tube to use for HIV diagnostic and follow-up blood draws should be made on the basis of local diagnostic and surveillance needs and available resources. If changes in practice are being considered, note that the HIV diagnostic or the follow-up blood draw should be 8 to 10 mL to ensure sufficient volume for additional HIV surveillance uses, including VARHS.

Specimen handling

After the blood draw and before serum or plasma separation, the tube should be stored at room temperature. The blood draw and specimen transport should be timed so that specimens arrive at the HIV testing laboratory and are processed (separated, aliquoted, and frozen) within 96 hours or as soon as possible after the blood draw. The likelihood of amplification may be reduced when specimen processing takes longer.

Serum and plasma processing

For HIV testing and VARHS, it is recommended that specimens be centrifuged to remove red blood cells and prevent hemolysis. Successful HIV amplification for sequencing is less likely with hemolyzed specimens. After separation of the serum or plasma, specimens should be maintained in a refrigerator at 4ºC or kept on ice until the specimens are EIA tested and eligible aliquots are frozen for potential use in VARHS.

Cryovials and aliquots

Cryovials for sera or plasma aliquoted for VARHS should be 2 mL, polypropylene, with screw caps and external threads. If labels are not used, the tube should have a writing area.

After the VARHS identification number (VARHS ID) is generated locally, each specimen should be labeled with the appropriate VARHS ID, by either writing the ID number on the tube with a permanent marker or attaching a label to the tube. Labels that adhere to frozen tubes can be supplied at no cost by CDC upon request.

Ideally, an aliquot of 1 mL per specimen should be prepared for HIV genetic sequencing and sent to the genotyping laboratory. When volume is limited, priority is given to (1) the laboratory's standard operating procedures for HIV diagnostic and clinical testing and serum archiving, (2) HIV incidence surveillance activities, and (3) VARHS. If less than 1 mL is available for genetic sequencing when all other needs have been met, the specimen may still be sent for sequencing after consultation with the genotyping laboratory.

Diagnostic and clinical testing and freezing of serum or plasma aliquots

The following will maximize chances for successful amplification for genetic sequencing.

- HIV diagnostic testing or other relevant clinical testing should begin immediately after separation of plasma or serum.
- Sera or plasma should be kept on the bench at room temperature for as short a time as possible before being returned to the refrigerator or placed on ice.
- If volume appears sufficient for the purposes of all basic laboratory tests and other local priorities, such as archiving and HIV incidence surveillance, a 1 mL aliquot from all EIA-reactive specimens should be frozen at –70ºC as soon as possible after the first reactive EIA.

  Note: Optimally, the specimen should be frozen at –70ºC within 96 hours after the blood draw. After 96 hours, the likelihood of amplification may be reduced, but the specimens should still be sent for genetic sequencing provided that amplification results are followed up carefully.

- Once frozen, specimens for genotyping should not be thawed until they reach the genotyping laboratory. A freeze-thaw cycle will reduce the chance of successful amplification for genotyping.

After the first reactive EIA, the aliquot is still considered an HIV diagnostic specimen; it will not be defined as a VARHS specimen until it has been confirmed as HIV-positive and until it is clear that specimen needs for all higher priority laboratory tests, including diagnostic HIV testing and HIV incidence testing, have been fulfilled. Aliquots frozen as potential VARHS specimens can be thawed at any time if they are needed for higher priority tests. The freeze-thaw cycle will not affect the HIV diagnostic test results or STARHS (serologic testing algorithm for recent HIV seroconversion) results for HIV incidence testing.

Specimen tracking data elements to be recorded in VARHS specimen log or database

VARHS data are considered part of routine HIV surveillance data and should be held to the standards of security and confidentiality for HIV/AIDS surveillance outlined in *Technical Guidance for HIV/AIDS Surveillance Programs, Vol. III: Security and Confidentiality Guidelines* (13). The diagnostic testing laboratory should maintain a VARHS specimen tracking log or database to provide information for evaluation if problems with amplification, specimen mixup, or specimen contamination occur. (See Table 2, Appendix F, for a list of the minimum required data elements for specimen tracking.) The specimen tracking data elements listed in Table 3 are optional unless the amplification rates for a particular surveillance area or local site are low (<90%). If the rates are low, all of the specimen tracking data elements listed in Table 3, in addition to those in Table 2, will be **required**. A Microsoft Access database for storage of the required specimen tracking and laboratory data is supplied by CDC.

Specimen tracking and laboratory data may be recorded first on person laboratory slips or specimen labels. Batch slips may be used for the initial recording of information applicable to batches of specimens, such as site of the blood draw, lists of patient identification numbers for specimens transported in a batch, date/time of receipt in the diagnostic laboratory, range of laboratory accession numbers for a batch, date/time of centrifugation, date/time of separation, date/time of first positive EIA result, date/time of aliquoting, and date/time aliquots were frozen. All or some of this information may also be captured from in-house laboratory databases. Information will be transferred to the IVR database for each specimen that is confirmed as HIV-positive and **not** identified as **ineligible**. Laboratory or health department personnel may work with CDC to develop methods of information transfer to minimize duplicate recording of information.

Shipment to the VARHS genotyping laboratory

Shipments should be made to the genotyping laboratory at least monthly. The frequency of shipments per month should be determined locally: the goal is to return resistance test results to designated health care providers as soon as possible. (For handling, packaging, and shipping of specimens, see Appendix E.)

Surveillance staff will compile a list of VARHS–eligible specimens to be packaged and shipped by the state or local public health laboratory to the genotyping laboratory. VARHS ID numbers should be recorded on the specimen manifest included in the shipment. Two additional copies of the specimen manifest should be made: 1 should be sent to sent to CDC by the US Postal Service or through SDN; the other should be filed with the VARHS coordinator (see Appendix E).

VARHS Implementation Task

VARHS areas should identify all laboratories that will transport specimens to a central public health laboratory, which will ship specimens to the genotyping laboratory. The local version of this guidance should include descriptions of the procedures by which the VARHS coordinator will receive monthly information about specimen tracking, receive shipping manifests whenever shipments are made, and communicate at least quarterly with laboratories making the shipments.

VARHS areas should include detailed laboratory procedures in the local version of this guidance (e.g., information on obtaining and processing specimens, determining eligibility, recording specimen tracking data for local use, storing specimens, and shipping specimens). All procedures should be developed by the VARHS coordinator in consultation with participating laboratories.

Required Data Elements

VARHS data are considered part of routine HIV surveillance data and should be held to the security and confidentiality standards for HIV/AIDS surveillance outlined in *Technical Guidance for HIV/AIDS Surveillance Programs, Vol. III: Security and Confidentiality Guidelines* (13).

Data used to determine VARHS eligibility

HARS or equivalent HIV/AIDS reporting systems will serve as the primary means of determining eligibility by using the date of the first positive HIV test result. Information on the prior use of antiretroviral drugs may be obtained from the incidence surveillance Testing and Treatment History Data or other sources, provided that information is collected on whether antiretroviral drugs have **ever** been used (i.e., no time limit should be placed on the history of antiretroviral drug use).

The elements of specimen tracking and laboratory data are linked with the HIV surveillance database to determine whether the person was reported to HARS more than 3 months before the relevant blood draw. Given the commitment to make reports available to providers in “real time,” the shipment of specimens for VARHS should not be delayed more than 30 days for the purposes of such evaluation (if possible, specimens should be shipped within 6 weeks after the HIV diagnostic test or other blood draw).

The dates of HIV diagnosis and the prescription of antiretroviral drugs may not be available in HIV/AIDS surveillance reports until after VARHS specimens have been shipped, tested, and entered in the IVR database. If additional data makes a case ineligible, the reason for ineligibility should be recorded in the database.

Demographic and clinical data

VARHS does not require the collection of additional demographic or clinical data. Instead, relevant data are supplied by HARS or the local equivalent. (See Table 1 of Appendix F for the minimum set of data elements required for determining HIV-1 drug resistance and HIV-1 subtype prevalence in the population and in population subgroups.) Demographic data used to estimate the prevalence of HIV-1 drug resistance may include age, sex, race/ethnicity, country of origin, and transmission categories associated with HIV infection. These elements will contribute to the calculation of prevalence estimates for national subgroups; local prevalence may be estimated in areas where the number of cases is sufficient.

Specimen tracking data

The required data elements for specimen tracking (i.e., must be collected for every specimen) are listed in Table 2 of Appendix F, and the optional data elements for specimen tracking are listed in Table 3 of Appendix F.

Note: CDC will analyze amplification rates stratified by surveillance area. If the amplification rates in a surveillance area are low (<90%), all the data elements for specimen tracking listed in Tables 2 and 3 will be **required**. These additional data must be collected until the amplification problem has been resolved or until CDC concludes that the problem is not associated with the procedures for handling or storing specimens.

Laboratory data

Data transmitted to CDC must not include personal identifiers and must be encrypted and password protected as specified in *Technical Guidance for HIV/AIDS Surveillance Programs, Vol. III: Security and Confidentiality Guidelines* (13).

For specimens sent to the CDC contract and state laboratories, the genotyping laboratory will do the following:

- Transmit information on amplification to the health department and to CDC
- Transmit the complete sequence of the protease gene and at least the first 240 codons of the reverse transcriptase gene, both as nucleotides and amino acids
- Transmit a separate list of all mutations (including those associated with drug resistance and those not associated with drug resistance) of any strain that differs from the reference strain used at the laboratory
- Provide hard copy of laboratory reports (to the appropriate health department only)

VARHS Implementation Task

VARHS areas should review Appendix F and determine the completeness and quality of the data elements used for analysis in VARHS. Because all of the demographic and clinical data elements used are key elements for general HIV surveillance, areas may wish to explore methods for increasing the reporting of these elements (e.g., some areas have added specific elements, such as country of birth, to their counseling and testing laboratory request forms).

Returning Genotyping Results

HIV genetic sequencing is usually performed for clinical purposes for persons receiving treatment for HIV. Physicians who treat HIV-infected persons are familiar with the formats in which results are reported. The genotyping laboratory sends the hard copy of a report (similar to the reports of drug-resistance testing, with which physicians are familiar) to the health department no more than 30 days after the genotyping laboratory receives the specimen. Because the interpretation of these results requires familiarity and training, the health department should not return genotyping results directly to the patient, but to the patient's designated health care provider.

A person with HIV-1–positive results who returns for posttest counseling or follow-up should

- Receive a brief explanation of ancillary HIV diagnostic testing, including genotyping
- Receive instruction on how to designate a health care provider to receive, at that time or later, the genotyping result if the patient wishes to do so
- Receive a brief explanation on when the health department will send the results to the health care provider

The health department is responsible for

- Determining the method(s) by which the patient can designate a health care provider(s) who will receive the genotyping result
- Developing a method for the health care provider to obtain the genotyping results
- Sending the hard copy of reports to the health care provider of the patient's choosing
- Acting as a repository for the hard copy of reports for a locally specified time

If the blood draw that produced a VARHS specimen was performed at a clinical site to which a patient is returning for medical consultation or care, the health department may arrange for the genotyping report to be returned directly to the patient's provider at that clinical site.

VARHS Implementation Task

Before VARHS is implemented, staff should determine the method(s) by which the patient can designate a health care provider(s) who will receive the genotyping results. Methods should be put in place to allow the patient to designate a health care provider(s) at the time of the initial diagnosis or later.

Ensuring Security and Confidentiality

HIV testing is a medical procedure. Therefore, policies and procedures are in place to protect the confidentiality of tested persons and their medical records. VARHS will be performed only on specimens that have tested positive for HIV or specimens used for the follow-up of a confirmed HIV-positive result.

VARHS data are considered part of routine HIV surveillance data and should be held to the standards of security and confidentiality for HIV/AIDS surveillance outlined in *Technical Guidance for HIV/AIDS Surveillance Programs, Vol. III: Security and Confidentiality Guidelines* (13). Current policies and procedures, based on these guidelines and local laws, are used by state and local health departments to secure hard copy and electronic information to protect the confidentiality of persons reported as having HIV infection. These measures must be extended to protect VARHS data held locally. Access by VARHS staff to information in HARS, HIV incidence surveillance data, and VARHS data is governed by the same security and confidentiality requirements.

All staff responsible for the transmission and receipt of specimen shipping manifests, VARHS sequence data, or reports of results must be trained in the security and confidentiality procedures for HIV/AIDS surveillance.

Information that could identify a person (e.g., name, Social Security number, full date of birth, address, ZIP code) will not be included in data sets transmitted from local surveillance areas to CDC. Data transmitted to CDC must be encrypted and password protected according to *Technical Guidance for HIV/AIDS Surveillance Programs, Vol. III: Security and Confidentiality Guidelines* (13).

VARHS Implementation Task

VARHS areas should review their security and confidentiality requirements and if necessary, draft specific procedures for handling VARHS data or add VARHS components to the local protocol. VARHS areas should also review their protocol and outline how VARHS data will be stored and secured. Personnel with access to the data should receive annual training in security and confidentiality procedures and should sign a confidentiality statement that outlines the procedures and the consequences of violating them.

Data Management

All data will be considered part of routine HIV surveillance data. Data entry and management will take place at state or local health departments by the use of HARS or eHARS and software developed by CDC or software that is compatible with CDC software.

CDC software

The IVR database supplied by CDC is a Microsoft Access database, an adjunct to HARS and eHARS, with tables for specimen tracking and HIV drug-resistance data. An additional table is provided for local data entry and for use by local HIV surveillance staff for the purposes of determining eligibility or for returning results.

CDC also provides a SAS program that facilitates the merging of specimen tracking and drug-resistance data (from the IVR database) with selected demographic and clinical data (from HARS or eHARS). The SAS program produces a data quality report along with a cumulative SAS data set to be submitted to CDC.

Identification of VARHS specimens

A VARHS ID will be assigned to each specimen sent for genotyping and used for importing HIV genetic sequencing information to the IVR database. The 14 digits of the VARHS ID are assigned as follows:

- Digits 1–4: These 4 digits represent the Federal Information Processing Standard (FIPS) code of the surveillance area.
- Digits 5–8: These 4 digits represent the site where the blood draw was performed.
- Digits 9–10: These 2 digits identify the last 2 digits of the year the blood was drawn.
- Digits 11–14: These 4 digits are a sequence assigned by the health department.

The IVR database includes functions to allow areas to enter their FIPS code only once and to create a menu of blood drawing sites and their codes. The VARHS ID will be created automatically in the IVR database when the 4 components listed above are entered for each specimen.

All data from the health department, local laboratories, genotyping laboratory, and HARS or an equivalent reporting system will be associated with either a VARHS ID number or a HARS State No. and then will be merged on the basis of these IDs. In addition to the VARHS ID, a specimen accession number may be entered in the database for local use.

Data encryption and transfer

Data transmitted to CDC must not include personal identifiers and must be encrypted and sent through SDN, as specified in *Technical Guidance for HIV/AIDS Surveillance Programs, Vol. III: Security and Confidentiality Guidelines* (13). Either SEAL or Pretty Good Privacy (PGP) encryption software must be used to encrypt all data being sent through SDN (13). SEAL software, for encrypting data sent to CDC, is provided by CDC at no cost. VARHS surveillance areas must purchase PGP software with at least 128-bit encryption to encrypt communications with genotyping laboratories. The PGP Personal Desktop edition is suitable; a perpetual license is recommended. All persons using PGP to encrypt and decrypt data must exchange PGP encryption keys. Any encryption software other than SEAL or PGP must be approved by CDC.

Electronic data sent from the genotyping laboratory to VARHS areas must be encrypted with PGP (or equivalent software). Please note that to receive genetic sequencing data from the CDC contract laboratory, VARHS areas must purchase PGP encryption software. The genotyping laboratory and the VARHS area must exchange PGP encryption keys.

CDC will use PGP software to encrypt quality assurance data and will send those data to VARHS areas through SDN (13).

Before the 15th day of each month, the complete VARHS data set from the preceding month should be transmitted to CDC.

VARHS Implementation Task

The local guidance should include the procedure for entering specimen tracking, genotyping, and local data into the VARHS database, exporting these data from the VARHS database, and using SAS to merge these data with selected demographic and clinical variables from HARS. Before transmitting the monthly VARHS data set to CDC, VARHS staff must have encryption software and a CDC-approved SDN certificate and must understand CDC's security and confidentiality guidelines (13). If PGP is used, all persons who encrypt and decrypt data must exchange encryption keys.

Data Analysis and Data Dissemination

To standardize the national resistance data set, CDC uses a regularly updated program to analyze resistance data transmitted from the genotyping laboratories. This program translates the nucleotide sequence and incorporates information on individual mutations of interest, the level of resistance to each antiretroviral drug in common use, and HIV-1 subtype.

The prevalence of resistance to at least 1 antiretroviral drug and the distribution of HIV-1 subtypes among persons with a new diagnosis of HIV infection should be reported annually at the national and local levels. The prevalence of resistance to individual antiretroviral drugs and categories of commonly used antiretroviral drugs (currently NRTIs, NNRTIs, and PIs) may also be reported. These data will be stratified by demographic factors or transmission category for subpopulation analyses at the national and local levels. If adequate case counts are available, stratification will allow comparisons of differing geographic areas and of differing transmission categories.

Aggregate data from the participating surveillance areas will be analyzed by CDC, and results will be disseminated. Each surveillance area will be responsible for conducting analyses at the local level.

As appropriate, results will be presented at conferences and published in peer-reviewed journals. The authors representing the surveillance areas and those representing CDC will be determined for each presentation or paper.

VARHS Implementation Task

VARHS areas should describe, in the local version of the VARHS guidance, the proposed plan for analyzing and disseminating VARHS data.

Outcome Standards

Outcome standards described in the Introduction to Policies and Procedures and Data Quality sections of *Technical Guidance for HIV/AIDS Surveillance Programs, Vol. I: Policies and Procedures* (12) should be applied to VARHS. Given the time-sensitive nature of VARHS data elements, meeting the core surveillance standards for case ascertainment and timeliness is essential to the success of VARHS. The quality of the HIV resistance estimate depends on the quality of data in the VARHS system. All outcome standards for VARHS relate only to cases in persons who resided in the surveillance area at the time of diagnosis.

- The minimum standard for collecting the required specimen tracking and laboratory data elements is 85% (goal, 100%).
- Each year, for at least 50% of the persons in a jurisdiction who have a new diagnosis and have been reported to HARS, a specimen should be sent for genotyping and documentation in the IVR database (50% of cases diagnosed in a year should have genotyping results 12 months after the diagnosis year).

References

1. DHHS Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. October 10, 2006. Available at http://www.aidsinfo.nih.gov/ContentFiles/AdultandAdolescentGL.pdf. Accessed April 12, 2007.

2. Centers for Disease Control and Prevention. US Public Health Service Task Force recommendations for use of antiretroviral drugs in pregnant HIV-1–infected women for maternal health and interventions to reduce perinatal HIV-1 transmission in the United States. *MMWR* 2002;51(RR-18):1–38.

3. Centers for Disease Control and Prevention. Updated U.S. Public Health Service guidelines for the management of occupational exposures to HBV, HCV, and HIV and recommendations for postexposure prophylaxis. *MMWR* 2001;50(RR-11):1–42.

4. Centers for Disease Control and Prevention. Antiretroviral postexposure prophylaxis after sexual, injection-drug-use, or other nonoccupational exposure to HIV in the United States. *MMWR* 2005;54(RR-2):1–19.

5. Hirsch MS, Brun-Vezinet F, Clotet B, et al. Antiretroviral drug resistance testing in adults infected with human immunodeficiency virus type 1: 2003 recommendations of an International AIDS Society–USA panel. *Clin Infect Dis* 2003;37:113–128.

6. Bennett DE. HIV genetic diversity surveillance. *J Infect Dis* 2005;192:4–9.

7. Interagency Task Force on Antimicrobial Resistance. A public health action plan to combat antimicrobial resistance, part 1: domestic issues. Available at http://www.cdc.gov/drugresistance/actionplan/. Accessed July 10, 2007.

8. Centers for Disease Control and Prevention. Reported Tuberculosis in the United States, 2005. Atlanta, Ga: US Department of Health and Human Services; 2006. Available at http://www.cdc.gov/tb/surv/surv2005/. Accessed July 16, 2007.

9. Lewis B, Granich R, Courval J, et al. Epidemiology of multi-drug-resistant *Mycobacterium tuberculosis* in the United States, 1996–2000. Annual meeting of the American Public Health Association; November 2003; San Francisco, California. Abstract 105.

10. Sax PE, Islam R, Walensky RP, et al. Should resistance testing be performed for treatment-naïve HIV-infected patients? a cost-effectiveness analysis. *Clin Infect Dis* 2005;41:1316–1323.

11. Munsiff SS, Nivin B, Sacajiu G, Mathema B, Bifani P, Kreiswirth BN. Persistence of a highly resistant strain of tuberculosis in New York City during 1990–1999. *J Infect Dis* 2003;188;356–363.

12. Centers for Disease Control and Prevention and Council of State and Territorial Epidemiologists. Technical guidance for HIV/AIDS surveillance programs, vol 1: policies and procedures. 2005. Available at https://team.cdc.gov/team/cdc/dispatch.cgi/w.nchstp_dhap_hicsb/tart/hicsb_TechGuid/.frame [password-protected Web site].

13. Centers for Disease Control and Prevention and Council of State and Territorial Epidemiologists. Technical guidance for HIV/AIDS surveillance programs, vol 3: security and confidentiality guidelines. 2006. Available at http://www.cdc.gov/hiv/topics/surveillance/resources/guidelines/guidance/. Accessed July 10, 2007.

Appendixes

Appendix A National Center for HIV, STD, and TB Prevention's Non-research Determination for HIV Variant, Atypical, and Resistant HIV-1 Surveillance

Appendix B Participating Surveillance Areas

Appendix C Epidemiologic Flowchart for Variant, Atypical, and Resistant HIV-1 Surveillance

Appendix D Laboratory Flowchart for Variant, Atypical, and Resistant HIV-1 Surveillance

Appendix E Technical Guidance for the Processing, Storage, and Shipment of Variant, Atypical, and Resistant HIV-1 Surveillance (VARHS) Specimens to Genotyping Laboratories for Antiretroviral Drug-Resistance Testing

Appendix E.1 Training and Certification for Shipping Infectious Substances

Appendix E.2 CDC VARHS Shipping Manifest

Appendix E.3 Stanford Laboratory's Return of Shipping Container

Appendix F VARHS Data Elements

Appendix A

National Center for HIV, STD, and TB Prevention's Non-research Determination for HIV Variant, Atypical, and Resistant HIV-1 Surveillance

NCHSTP
Research/Non-research Determination
(Request to Classify Project as Not Involving Human Subjects or Research)

This form should be used to submit to NCHSTP ADS materials for projects involving CDC investigations that are not subject to human subjects regulations. Projects are eligible for this classification either as "non-research" projects (primary intent is not to generate generalizable knowledge) or as research projects that do not involve identifiable human subjects. Such projects do not require submission to the CDC Human Subjects Office for IRB review. Do **NOT** use this form for IRB "EXEMPT" research.

Project Title: Surveillance of variant, atypical, and resistant strains of HIV
(under PA 1194, 4017, 4118, and future program announcements supporting genotyping for this purpose)

Project Locations/Sites Current sites: Piloted for feasibility under Protocol 3575 (as research, because of pilot status, but consent was waived by all involved IRBs) in Colorado, Illinois, Maryland, and Seattle/King County.
We now propose to reclassify this activity from pilot status to routine HIV surveillance. Departments of Health in Chicago, Colorado, Florida, Illinois, Indiana, Louisiana, Massachusetts, Michigan, New York City, New York State, North Carolina, Pennsylvania, Puerto Rico, Seattle/King County, South Carolina, and Virginia have received funding under PA 4017 for this HIV surveillance activity.
Departments of Health in all other states performing HIV surveillance are eligible under a new program announcement (4118).

Project Officer(s) Diane Bennett Division: DHAP-SE Telephone: 404-639-5349

Proposed Project Dates: Proposed Start (as non-research routine surveillance): 7/1/2004 Ending: ongoing HIV surveillance activity

Categories of data collection that do not constitute human subjects research include are listed below. Please check appropriate category:

___I. **Activity is not research**. Primary intent is a public health practice disease control activity.
- ___A. Epidemic/endemic **disease control** activity; collected data directly relate to disease control needs.
- XB. Routine **disease surveillance** activity; data used for disease control program or policy purposes.
- ___C. **Program evaluation** activity; data are used primarily for that purpose.
- ___D. **Post-marketing surveillance** of efficacy and/or adverse effects a new regimen, drug or device.

NCHSTP ADS Review **Date rec'd in NCHSTP ADS Office:**_____

✓ **Concur, project does not constitute human subjects research**

or

___**Project constitutes human subjects research**, submission for Human Subjects review required

Additional Comments:
1. This form cannot be used to document "IRB Exempt Research," which must instead be submitted to the CDC IRB.
2. Although CDC Human Subjects (IRB) review is not required in this instance, investigators/project officers are expected to adhere to ethical principles and standards by respecting and protecting to the maximum extent possible the privacy, confidentiality and autonomy of participants. All applicable State and Federal privacy laws must be followed.
3. Although this project does not constitute human subjects research, informed consent may be appropriate. Information disclosed in the consent process should address the eight standard consent elements.
4. Other:

Signed:_____ 6-25-04
Andrew Vernon, MD, MHS
Associate Director for Science
National Center for HIV, STD, and TB Prevention
Date

Appendix B

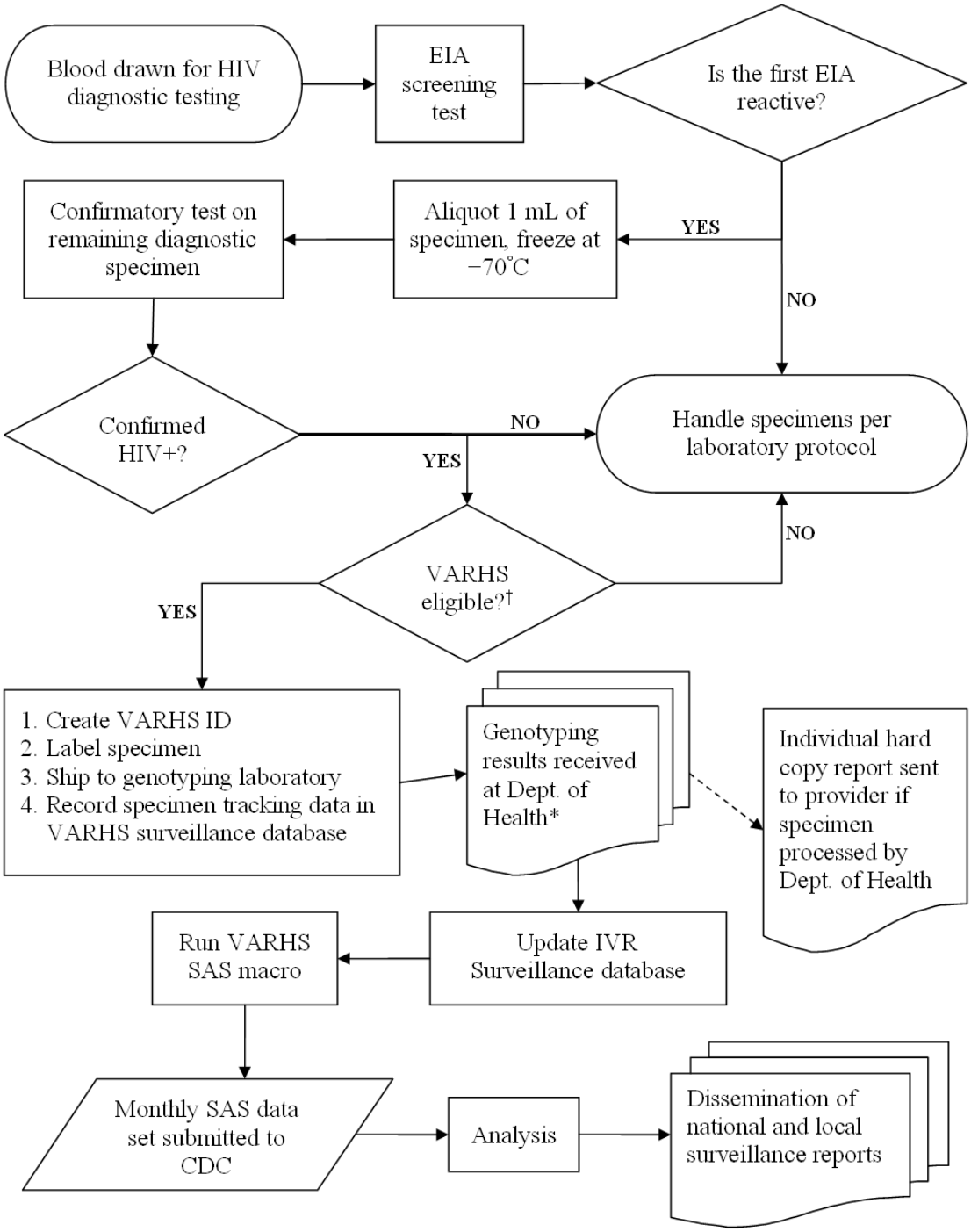
Participating Surveillance Areas

The following surveillance areas were funded to participate in VARHS (including sites that participated in the antiretroviral drug resistance testing [ARVDRT] evaluation project) as of 2004.

1. Chicago Department of Public Health
2. Colorado Department of Public Health and Environment
3. District of Columbia Department of Health
4. Florida Department of Health
5. Illinois Department of Public Health
6. Indiana State Department of Health
7. Louisiana Office of Public Health
8. Maryland Department of Health and Mental Hygiene
9. Massachusetts Department of Public Health
10. Michigan Department of Community Health
11. Mississippi State Department of Health
12. New Jersey Department of Health and Senior Services
13. New York City Department of Health and Mental Hygiene
14. New York State Department of Health
15. North Carolina Department of Health
16. Pennsylvania Department of Health
17. Puerto Rico Department of Health
18. Seattle and King County Public Health
19. South Carolina Department of Health
20. Texas Department of State Health Services
21. Virginia Department of Health
22. Washington State Department of Health

Appendix C

Epidemiologic Flowchart for Variant, Atypical, and Resistant HIV-1 Surveillance

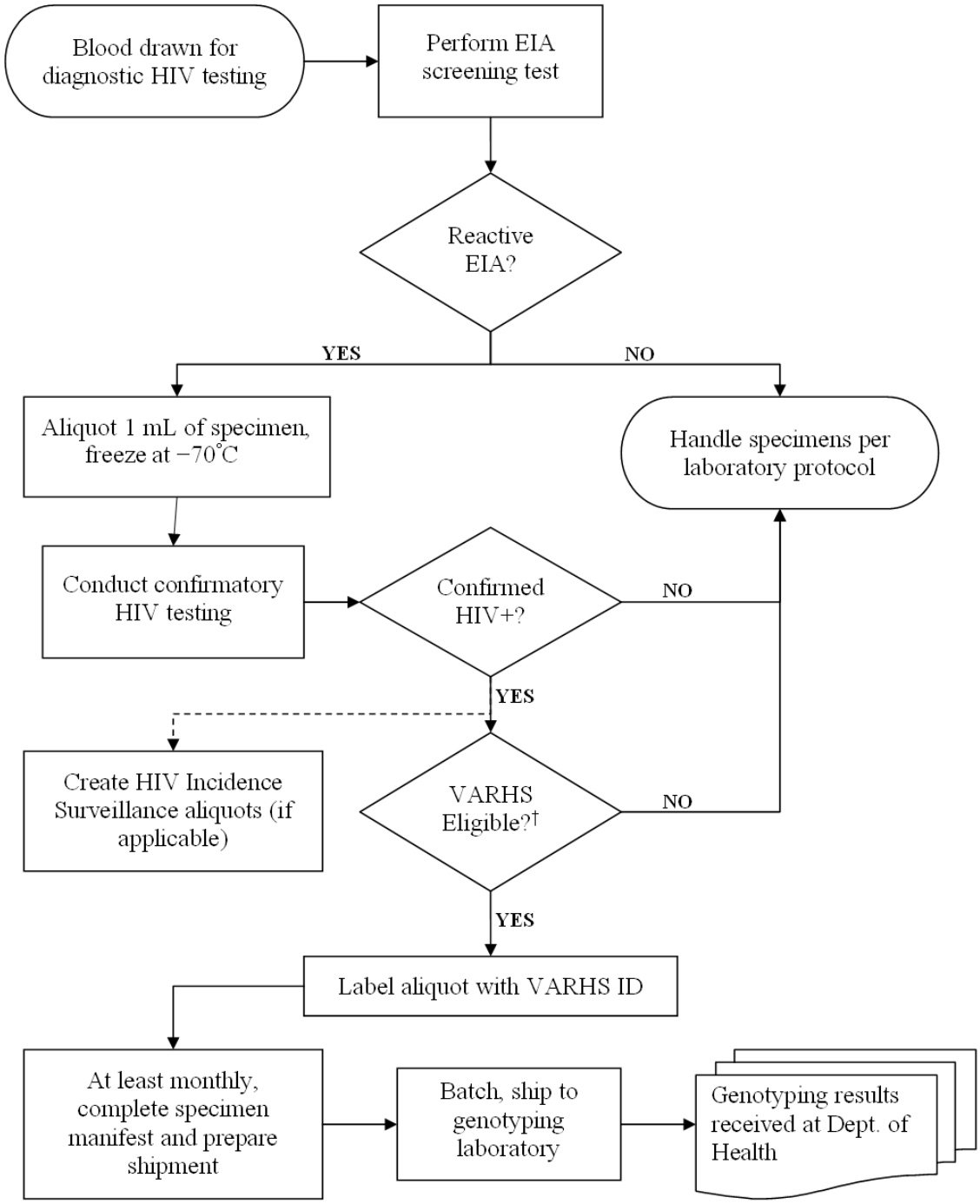


† See technical guidance section on determining eligibility.

* If sequences are received from commercial laboratories, Dept. of Health should determine eligibility.

Appendix D

Laboratory Flowchart for Variant, Atypical, and Resistant HIV-1 Surveillance



† See technical guidance section on determining eligibility.

Appendix E

Technical Guidance for the Processing, Storage, and Shipment of Variant, Atypical, and Resistant HIV-1 Surveillance (VARHS) Specimens to Genotyping Laboratories for Antiretroviral Drug-Resistance Testing

Purpose

This technical guidance describes the methods for the processing, storage, and shipment of VARHS serum and plasma specimens that will be tested for HIV antiretroviral drug resistance and HIV-1 subtyping.

Introduction

Serum or plasma from HIV diagnostic specimens should be collected and frozen at –70ºC. For the purpose of resistance testing, serum should ideally be separated within 48 hours after the blood draw and frozen within 96 hours after the blood draw. Frozen serum or plasma will be shipped to a CDC-designated testing laboratory for genotypic analysis.

Note: Surveillance areas may elect to store a backup aliquot for use in the event that something happens to the original aliquot sent to the laboratory or if a specimen needs to be retested for any reason.

Setting and Personnel Required for Specimen Processing

- Centrifugation, aliquoting, and shipping should be performed at or under the auspices of a laboratory that has been certified by CLIA (Clinical Laboratory Improvement Amendments of 1998) to handle HIV-positive specimens.

  Note: All personnel handling specimens should receive training in handling bloodborne pathogens. (See OSHA's [Occupational Safety and Health Administration's] standards regarding occupational exposure to bloodborne pathogens at http://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=STANDARDS&p_id=10051.)

- Personnel handling or processing specimens should receive training in the relevant laboratory techniques for handling HIV-positive specimens and for performing the specific tasks required.
- The setting in which centrifugation, aliquoting, and shipping are done should meet Biosafety Level 2 specifications, required by the US Department of Health and Human Services, for the handling of HIV specimens (http://www.cdc.gov/od/ohs/biosfty/bmbl4/bmbl4s2.htm).

Materials

The following materials are required for the collection and shipment of VARHS specimens:

- Cryogenic vials: 1.5 to 2 mL, with screw cap, external threads, O-ring, and made of polypropylene.

- Freezer labels that will remain on the tubes upon freezing. Many cryogenic vials bear a label area that is suitable for writing with a permanent marker. Writing on the printed area, instead of affixing a label, is acceptable.
- Cardboard storage boxes for cryogenic vials: 81 spaces per box.
- Low temperature freezer: –70ºC.
  - Keep a daily temperature log to ensure that the freezer is operating properly.
  - House the freezer in a location with proper ventilation to avoid overheating and freezer failure.
  - Be sure that the –70ºC freezer has enough space for the storage of VARHS specimens.
- A supply of dry ice in pellet form.
- Saf-T-Pak (http://www.saftpak.com) STP 320 insulated shipping containers certified for frozen diagnostic specimens (i.e., certified for HIV-positive specimens and dry ice).
- Courier's airbills.
- Materials for packing the shipping container (see 2.4).

Specimen Collection and Processing

1. All processing of specimens should be done by personnel qualified to handle HIV-positive specimens under the auspices of a laboratory equipped to handle HIV-positive specimens (http://www.cdc.gov/od/ohs/biosfty/bmbl4/bmbl4s2.htm).

2. If the amount of serum seems to be sufficient for all diagnostic needs and HIV incidence testing, transfer an aliquot from the collection tube to the corresponding vial for HIV drug-resistance testing as soon as possible after separation (1 mL per cryogenic vial is optimal).

3. Ship the aliquot to the Stanford University laboratory or to the locally contracted genotyping laboratory. Surveillance areas may elect to store a backup aliquot for use in the event that something happens to the original aliquot sent to the laboratory or if a specimen needs to be retested for any reason.

4. Place the aliquots in cardboard boxes for cryogenic vials in a –70ºC freezer. Optimally, freeze the aliquot within 96 hours after the blood draw.

5. The aliquots frozen as potential specimens for HIV drug-resistance testing remain HIV diagnostic specimens until all HIV diagnostic tests and any other basic laboratory tests have been completed. If additional serum is needed for diagnostic testing or HIV incidence testing, the aliquot for potential HIV drug-resistance testing should be thawed and used for those purposes. If a backup specimen was stored, the backup specimen should be used first. If more volume is required, the specimen designated for the genotyping laboratory may also be used.

6. After a positive result from an HIV confirmatory test or at the time determined by standard laboratory procedures, label each cryogenic vial with the appropriate VARHS specimen ID. The VARHS ID is a 14-digit number that is assigned to each specimen sent for genotyping.

The 14-digit VARHS specimen ID consists of the following:

- The 4 digits that denote the **surveillance area** (digits 1–4), which correspond to the Federal Information Processing Standard (FIPS) code for the surveillance area (state or city)
- The 4-digit **site number** (digits 5–8), which designates where the blood draw was performed
- The last 2 digits of the **year of the blood draw** (digits 9–10)
- The 4-digit **sequence number** (digits 11–14), which is assigned by the local health department

7. Every specimen that is eligible for VARHS should be entered into a specimen handling and processing log or a database. Each surveillance area is expected to develop its own log or database or to use the Microsoft Access database provided by CDC. In addition, each surveillance area is responsible for ensuring that all information for each specimen is logged before the specimen is shipped.

Shipping

1. Preliminary procedures for shipping to the Stanford University genotyping laboratory (CDC contract laboratory)

1.1. Specimens for genotypic resistance testing should be sent to the Stanford University laboratory. All specimens should be shipped as diagnostic specimens, according to International Air Transport Association (IATA) Packing Instruction 650. Dry ice should be included with each shipment, as specified in IATA Packing Instruction 904.

1.2. Because shipping specimens involves using dry ice, shipping personnel must be trained and certified to ship dangerous goods. (See Appendix E.1 for a list of companies that provide training.)

1.3. Establish contact with the point person at the Stanford University laboratory.

- Stanford Virology Main Lab: 650-498-5575
- Stanford HIV Genotyping Laboratory: 650-725-7165
- Christina Trevino: 650-725-7165
- Mary Arroyo: 650-725-4146

1.4. A "test" shipment must be sent before the first shipment of VARHS specimens to ensure that all procedures are in place. The test shipment should exactly duplicate a real shipment (i.e., ship frozen liquid in cryogenic vials on dry ice) and should be sent only once. (See the procedures outlined in 1.5–1.8.)

The purpose of the test shipment is to familiarize the sender with the processes for notifying the laboratory about the shipment and for packing and shipping the diagnostic specimens on dry ice. Shipping frozen water on dry ice without the infectious substance labels will accomplish the purpose of the test shipment.

1.5. Arrange for the preparation of the test shipment and the initial shipment to be overseen by laboratory staff experienced in the shipment of comparable specimens.

1.6. Notify the Stanford University laboratory that the test shipment is on the way by calling the main laboratory at 650-725-7165 or by e-mailing (do not e-mail or fax the shipping manifest).

In the call or e-mail, include the number of specimens being shipped and if applicable, the FedEx tracking number. Per CDC's *Technical Guidance for HIV/AIDS Surveillance Programs, Vol. III: Security and Confidentiality Guidelines* (13), do not e-mail or fax the shipping manifest (the manifest is included in the shipment).

1.7. Be sure to have an adequate number of STP 320 shipping containers or the equivalent are on hand. The Stanford laboratory will return these containers to the shipping laboratory (they are expensive and should be reused).

1.8. Be sure to have an adequate supply of the courier's airbills on hand.

2. Packing procedures for shipping to the Stanford University laboratory

2.1. Specimens must be shipped on the same day that they are packed. Plan to begin the packing process early so that the container(s) will be ready for the courier's last pick-up of the day. ***Pack and ship only Monday through Wednesday. Never pack and ship the week of Thanksgiving, Christmas, New Year's, or July Fourth***.

2.2. Personnel in the shipping laboratory should read and walk through all of these steps before starting to prepare the actual shipment in order to be familiar with the requirements.

2.3. Wear gloves and a lab coat.

2.4. Bring the STP 320 shipping container and other packing materials into the area where the shipment is being prepared. Check to be sure that the container includes the following items:

- 2 sheets of bubble wrap
- 2 STP 710 2-part secondary pressure vessels or equivalent certified containers
- Two 250 mL absorbent strips
- Class 9 label and label for the quantity of dry ice
- Other hazard and handling labels
- 1 instruction sheet

2.5. For a diagram of the contents of the shipping container, see the Saf-T-Pak catalog (http://www.saftpak.com). Of the listed contents, use only what is needed for each shipment. Save leftover supplies for future shipments.

2.6. If the STP 320 shipping container is being reused, the labels will already be in place on the outer cardboard container.

2.7. Be sure that adequate supplies of the other materials listed in 2.4 are on hand.

2.8. Prepare 3 copies of the shipping manifest (Appendix E.2). On the manifest, write the VARHS ID number on each vial to be shipped, along with the date the specimen was drawn and the date the specimen was frozen (freeze date).

Note: ***The freeze date entered here should be the last freeze date for the aliquot being shipped.***

Indicate (circle) whether the specimens are serum or plasma, and record the volume being shipped if it is less than 1 mL. Bring the copies of the manifest into the area where the shipment is being prepared.

- Copy 1 of the shipping manifest should be included in the shipment to the Stanford laboratory.
- Copy 2 of the shipping manifest should be sent to CDC (see 3.2). Send through SDN (preferred) or the US Postal Service.
- Copy 3 of the shipping manifest should be kept by the surveillance area's VARHS coordinator or the laboratory sending the specimens to the Stanford laboratory.

2.9. Prepare the courier's airbill (Appendix E.3); Stanford will use the airbill to return the shipping container for reuse. On the airbill, fill in the shipping laboratory's complete return address, Stanford University's address, and the billing number. Staple the airbill to the shipping box return form. Bring the copy of the airbill into the area in which the shipment is being prepared.

2.10. If dry ice is in another location, which requires leaving the area where the shipment is being prepared, put the dry ice needed for this shipment in a separate container and bring it to the shipping area.

2.11. Go to the freezer and remove the entire 2-inch freezer box containing the specimens to be sent.

2.12. Bring the specimens to the area where the shipment is being prepared. ***Specimens are to remain frozen at all times and therefore should not be removed from a –70°C environment for more than a few minutes.***

2.13. Recheck the screw-cap lids on the specimen vials and tighten if necessary.

2.14. Place the freezer box containing the specimens in the secondary pressure vessel (leak-proof) and make sure the specimens are surrounded by bubble wrap and absorbent strips. The vials should not move around or rattle inside the vessel.

2.15. Place the secondary pressure vessel in the inner box and place the inner box in the polystyrene cooler.

2.16. **Do not** put dry ice inside the inner box.

2.17. Pack pelleted dry ice in the shipping container and around the inner box. The STP 320 shipping container will hold ~8 kg of dry ice (~10 lbs); if the cooler is packed completely, this quantity of dry ice will keep the contents frozen for more than 80 hours.

2.18. Place the lid on the polystyrene cooler.

2.19. Fold 1 copy of the VARHS shipping manifest in half and place it on top of the shipping box return form with the airbill that Stanford laboratory will use to return the shipping container to the shipping laboratory (for reuse), and place on top of the polystyrene lid.

2.20. Fold over the top flaps of the outer box and seal with clear shipping tape.

2.21. Place the category B label (“UN 3373”) over the 2½" × 2½" diamond-shaped outline on the outer box.

2.22. Write “Diagnostic Specimens” on the outer box, adjacent to the category B label.

2.23. Place the Class 9 hazard label over the 4" × 4" diamond-shaped outline.

2.24. Place the label for the net quantity of dry ice over the rectangular outline adjacent to the Class 9 hazard label. On the label, write the approximate amount (in kg) of dry ice used to pack the container.

Note: The labels (category B designation, Class 9 hazard, and quantity of dry ice) can be purchased.

2.25. Prepare the courier’s paperwork as directed by the training and certification course and select the overnight shipping option.

2.26. If the preceding steps are not completed before the courier’s last pickup of the day, unpack the specimens and put them back in the –70ºC freezer and begin the process again on the next appropriate day.

3. Procedures for shipping to the Stanford University laboratory

3.1. Ship only Monday through Wednesday. Never ship the week of Thanksgiving, Christmas, New Year's, July Fourth, or major local holidays.

3.2. Send the second copy of the shipping manifest to the CDC Laboratory Liaison, Richard Kline, through the SDN (preferred) or the US Postal Service. On the manifest, indicate the date that the specimens were shipped to the Stanford laboratory. ***Do not e-mail or fax the shipping manifest to CDC or to the Stanford laboratory.***

Mailing Address:

Attn: Richard Kline
Centers for Disease Control and Prevention
1600 Clifton Rd, NE, MS E-47
Atlanta, GA 30333
Phone: 404-639-4958

3.3. Keep the third copy of the shipping manifest.

3.4. Notify the Stanford laboratory that a shipment is on the way by calling the HIV genotyping laboratory at 650-725-7165 or by e-mailing (do not e-mail or fax the shipping manifest).

In the call or e-mail, include the number of specimens being shipped, the date the shipment will be sent, and if applicable, the FedEx tracking number. ***Do not e-mail or fax the shipping manifest (the manifest is included in the shipment).***

3.5. Track the shipment by using the FedEx tracking number (10–12 digits) on the airbill. Tracking can be done through the Web site (http://www.fedex.com) or by calling 1-800-GO-FEDEX.

If a problem is identified, please notify

Stanford HIV Genotyping Laboratory: 650-725-7165
and
Richard Kline (CDC): 404-639-4958

3.6. The Stanford laboratory will contact the shipping laboratory if expected shipments are not received.

4. Preliminary procedures for shipping to a locally contracted genotyping laboratory

Follow the same procedures, substituting the locally contracted laboratory's contact information for the Stanford laboratory's contact information (see 1.3 and 1.6).

5. Packing procedures for shipping to a locally contracted genotyping laboratory

Follow the procedures listed above, substituting local procedures as applicable and substituting the locally contracted laboratory for the Stanford laboratory.

6. Procedures for shipping from the processing laboratory to a locally contracted genotyping laboratory

Follow the procedures listed above, substituting local procedures as applicable and substituting the locally contracted laboratory for the Stanford laboratory.

7. Procedures for local transport from the processing laboratory to a locally contracted genotyping laboratory

7.1. Follow the procedures listed above or substitute local procedures as discussed with Richard Kline at CDC to ensure that specimens remain frozen during transport and until ready for genotyping at the genotyping laboratory. Always include a copy of the shipping manifest in the shipment.

7.2. Whenever specimens are transported, send a copy of the specimen manifest to Richard Kline at CDC through the SDN (preferred) or the US Postal Service. On the manifest, indicate the date that the specimens were shipped to the genotyping laboratory. ***Do not e-mail or fax the shipping manifest to CDC or the genotyping laboratory***.

Mailing Address:

Attn: Richard Kline
Centers for Disease Control and Prevention
1600 Clifton Rd, NE, MS E-47
Atlanta, GA 30333
Phone: 404-639-4958

7.3. Transport specimens only Monday through Thursday, unless local guidelines specify otherwise. If local HIV surveillance or laboratory staff are not personally handling the transport, do not transport specimens during the weeks of Thanksgiving, Christmas, New Year's, July Fourth, or major local holidays.

7.4. Track the shipment through FedEx's or another courier's tracking system.

If there are problems with the transfer, please notify

Richard Kline (CDC): 404-639-4958
and
the appropriate personnel at the genotyping laboratory

Develop an acknowledgment system so that the genotyping laboratory informs the shipping laboratory when the shipment has arrived.

8. Local transfer from the processing area to the genotyping area within one laboratory (i.e., the HIV diagnostic testing and genotyping for HIV drug resistance are performed within the same laboratory)

If HIV testing and genotyping are performed in the same laboratory, diagnostic testing is complete, and aliquots are identified as VARHS specimens, do the following:

8.1. Send a copy of the shipping manifest (Appendix E.2) to Richard Kline at CDC through the SDN (preferred) or the US Postal Service. On the manifest, indicate the date that the specimens were shipped to the genotyping laboratory. ***Do not e-mail or fax the shipping manifest to CDC or the genotyping laboratory.***

Mailing Address:

Attn: Richard Kline
Centers for Disease Control and Prevention
1600 Clifton Rd, NE, MS E-47
Atlanta, GA 30333
Phone: 404-639-4958

8.2. Ideally, the initial freezing of specimens should be done in a freezer convenient to the genotyping laboratory so that it will not be necessary to transfer frozen specimens.

8.3. If frozen specimens are to be transferred from a freezer in one area to a freezer in another area, a biohazard bag, cooler, and dry ice or cool packs should be ready before the specimens are removed from the freezer. Place the specimens (in their storage box) in the biohazard bag. Place the bag in the cooler with dry ice or cool packs and transport as quickly as possible to the other freezer.

8.4. If there are problems with the transfer, please notify

Richard Kline (CDC): 404-639-4958
and
the appropriate personnel at the genotyping laboratory

Appendix E.1 Training and Certification for Shipping Infectious Substances

The following are some of the companies that provide training in shipping dangerous goods. CDC does not endorse any particular company.

- **FedEx**
  - 1-800-GO-FEDEX
  - 3-day IATA-based training
  - Covers all hazardous materials
  - Cost: ~$550

- **Saf-T-Pak**
  - 1-800-814-7484
  - Specifically for infectious and diagnostic substances and dry ice
  - 3 options: 1-day seminar, on-site program, or interactive CD (can be completed in 3–5 hours)
  - Certificate: valid for 2 years or until regulations change
  - Cost: ~$250

- **Viking Packing Specialist** (Oklahoma)
  - 1-800-788-8525; Contact: David Weilert
  - Monthly seminars in Tulsa
  - Covers all 9 classes of hazardous materials
  - Covers shipping according to IATA instructions
  - Certificate: good for 2 years
  - Cost: ~$300 per person
  - Group classes in local area: ~$3,000 plus travel costs

Appendix E.2 CDC VARHS Shipping Manifest

Stanford Hospital and Clinics
300 Pasteur Drive
Stanford, CA 94035

CDC VARHS Shipping Manifest

Clinical Virology Laboratory @ Hillview
(650) 498-5575
Directors: R. Sibley, MD / B. Patterson, MD

Date specimens sent to Stanford: _____

Link project number (VARHS ID number)	**Draw Date**	**Freeze Date**	**Serum or Plasma** (CIRCLE)	**Volume** (if < 1 mL)	**Stanford Laboratory number** (STANFORD USE)
			Serum or Plasma		
			Serum or Plasma		
			Serum or Plasma		
			Serum or Plasma		
			Serum or Plasma		
			Serum or Plasma		
			Serum or Plasma		
			Serum or Plasma		
			Serum or Plasma		
			Serum or Plasma		
			Serum or Plasma		
			Serum or Plasma		
			Serum or Plasma		
			Serum or Plasma		
			Serum or Plasma		
			Serum or Plasma		
			Serum or Plasma		
			Serum or Plasma		
			Serum or Plasma		
			Serum or Plasma		

Shipping Instructions: 1) Call Stanford Virology Laboratory (650) 725-7165 when package is sent (keep a record of tracking number)
2) Complete this manifest form and include a copy with the specimens
3) Package and ship on Monday, Tuesday, or Wednesday only
4) Ship to: Attn: Virology Laboratory Specimen Processing
Stanford Clinical Lab at Hillview
3375 Hillview Ave
RM 2401
Palo Alto, CA 94304-1204

Mail hard copy of the report to: Attn: _____

Location _____

Street Address _____

City_____

State _____ ZIP code_____

Contact person name and phone number: _____; ( ) __ __ __ - __ __ __ __

Stanford Virology Test Code: AVRT

For office use only: **Billing Done**: _____

Original: **Box #**_____

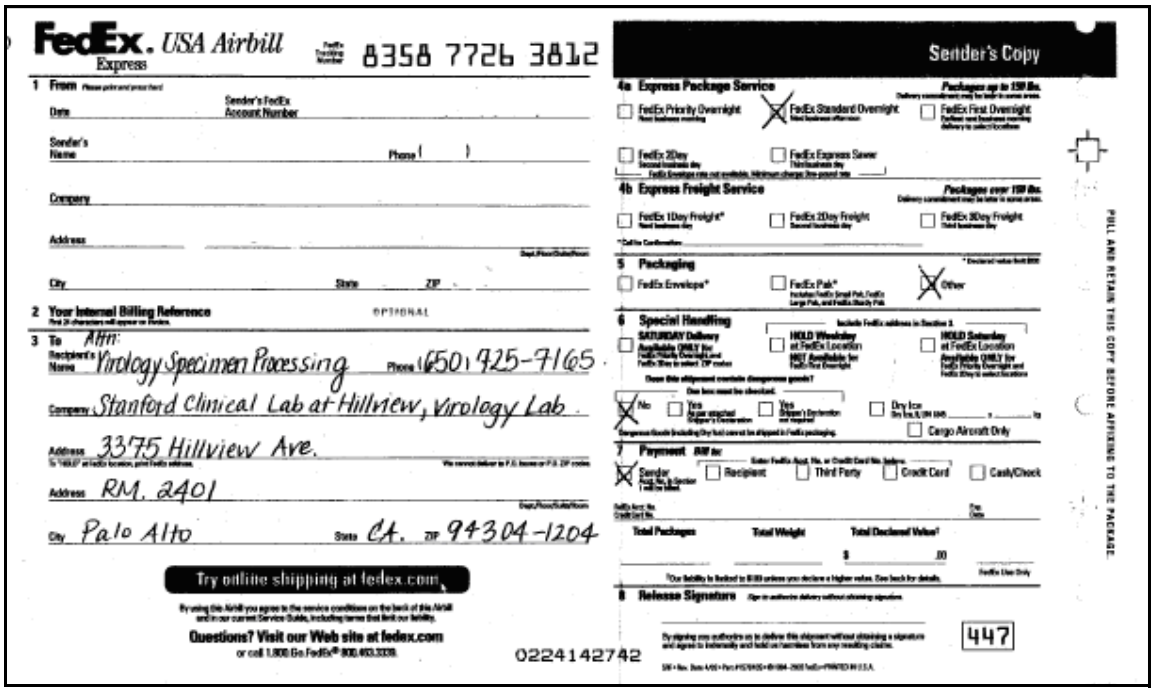
Extract:: **Box #**_____

Product: **Box #**_____

Appendix E.3 Stanford Laboratory's Return of Shipping Container

Sender: Attach return FedEx airbill. Complete all sections except the date, which will be completed at the Stanford laboratory. See example below.

Stanford Virology Receiving Desk: Give the empty shipping container, packing materials, and the return FedEx airbill to Hina, in Virology, for return to sender.


FedEx. USA Airbill Express
8358 7726 3812
Sender's Copy
1 From
Date
Sender's FedEx Account Number
Sender's Name
Phone ( )
Company
Address
City
State
ZIP
2 Your Internal Billing Reference
OPTIONAL
3 To
Recipient's Name Attn: Virology Specimen Processing
Phone (650) 725-7165
Company Stanford Clinical Lab at Hillview, Virology Lab
Address 3375 Hillview Ave.
Address RM. 2401
City Palo Alto
State CA. ZIP 94304-1204
Try online shipping at fedex.com
Questions? Visit our Web site at fedex.com
0224142742
4a Express Package Service
Packages up to 150 lbs.
FedEx Priority Overnight
FedEx Standard Overnight
FedEx First Overnight
FedEx 2Day
FedEx Express Saver
4b Express Freight Service
Packages over 150 lbs.
FedEx 1Day Freight*
FedEx 2Day Freight
FedEx 3Day Freight
5 Packaging
FedEx Envelope*
FedEx Pak*
Other
6 Special Handling
SATURDAY Delivery
HOLD Weekday at FedEx Location
HOLD Saturday at FedEx Location
No
Yes
Dry Ice
Cargo Aircraft Only
7 Payment Bill to:
Sender
Recipient
Third Party
Credit Card
Cash/Check
Total Packages
Total Weight
Total Declared Value
$ .00
8 Release Signature
447
PULL AND RETAIN THIS COPY BEFORE APPLYING TO THE PACKAGE.

Appendix F

VARHS Data Elements

Table 1

The following data elements are used for estimations of the national prevalence and incidence of transmitted HIV-1 drug resistance and distribution of HIV-1 subtypes and for making estimates for major subpopulations.

Note: Most of the data elements are collected through routine HIV case surveillance.

Data Element	HIV-1 Drug Resistance Prevalence	HIV-1 Subtype Distributions
Demographic Data		
Age	X	X
Sex	X	X
Race/ethnicity	X	X
Transmission category for HIV infection	X	X
Country of origin	X	X
Current state of residence	X	X
State of residence at HIV diagnosis	X	X
Laboratory Data		
HIV pol gene nucleotide sequence	X	X
Mutation-specific assays	X	X
Previous HIV Testing Data		
Date of first HIV test	X	
Date of first positive HIV test	X	
Date of last negative HIV test	X	
Clinical Data		
Date of AIDS diagnosis	X	
CD4 counts	X	
Dates of CD4 counts	X	
Viral load	X	
Dates of viral loads	X	
Antiretroviral agents used	X	
Start/end dates of antiretroviral agent use	X	
Opportunistic infection(s) diagnosed	X	
Date(s) of opportunistic infection diagnosis	X	

Table 2. Required Data Elements for Specimen Tracking

The following required laboratory data elements are used for local tracking of specimens, evaluating problems with amplification of HIV for genotyping, supporting plans to optimize specimen handling processes for surveillance purposes, and evaluating problems with contamination.

Name	Label
SPECIMEN ID	VARHS ID number
DRAWDTTM	Date/time of blood draw
SPECSOURCE	Was specimen from a VARHS site or a commercial/private laboratory?
ACCESNUM	Laboratory accession number (if applicable)
TUBE TYPE	Blood collection method
BLDCOMP	Type of specimen sent from collection site
BLDLAB	Type of specimen sent for resistance test
VOLGENO	Volume collected for genotype aliquot
GFRZ1DTM	Date/time of first freeze of aliquot
NOT SENT	Was specimen sent for genotyping?
AMPLIFY	Was this specimen amplified?

Table 3. Optional Data Elements for Specimen Tracking

The following laboratory data elements are optional **unless** low amplification rates (<90%) are seen in a particular surveillance area. In this case, all of the specimen tracking data elements listed in Table 3, in addition to Table 2, will be required. The additional elements must be collected until the amplification problem has resolved or until CDC can conclude that the problem is not associated with specimen handling/storage procedures.

Name	Label
COURDTTM	Date/time of courier pick-up from collection site
GRESDATE	Date/time genotyping results received at DOH
DRECDTTM	Date/time of receipt in diagnostic laboratory
CENTDTTM	Date/time of centrifugation
SEPDTTM	Date/time of separation
ALIQDTTM	Date/time of aliquoting
VOLUME	Volume of serum/plasma after separation
COND	Specimen condition
SPECIMEN FROZEN	Was this specimen frozen prior to aliquoting?
GFRZ2+YN	Was the genotyping aliquot thawed or refrozen?
GTHW1DTM	Date/time of first thaw of genotyping aliquot
GFRZ2DTM	Date/time of second freeze of genotyping aliquot
GSHPDATE	Date of shipment to genotyping laboratory
BFRZ1DTM	Date/time of first freeze of local back-up aliquot
BTHW1DTM	Date/time of first thaw of local back-up aliquot
BFRZ2DTM	Date/time of second freeze of local back-up aliquot
VOLBACK	Volume of local back-up aliquot

Technical Guidance for HIV/AIDS Surveillance Programs

Data Analysis and Dissemination

HIV Incidence and Case Surveillance Branch
Atlanta, Georgia

Contents — Data Analysis and Dissemination

Structural Requirements 14-3
Minimum performance standards 14-3
Sufficient capacity for analyzing HIV/AIDS surveillance data 14-3
Outcome Standards 14-4
Process Standards 14-4
Statistical report 14-5
Recommended analyses for annual or more frequent reports 14-7
Suggested additional analyses 14-9
Linking HIV/AIDS surveillance data with other data sources 14-10

Appendixes
Appendix A Terms, Definitions, and Calculations Relevant to HIV/AIDS Surveillance Data 14-12
Appendix B Techniques for Analyzing and Displaying HIV/AIDS Surveillance Data 14-17
Appendix C Integrated HIV/AIDS Epidemiologic Profile 14-22
Appendix D Tools and Resources for Analyses 14-23

References 14-24

Technical Guidance for HIV/AIDS Surveillance Programs — Data Analysis and Dissemination

Decisions regarding public health are dependent on high-quality data. Accurate surveillance data are central to the effective monitoring of trends in HIV infection and AIDS, the identification of behavioral risk factors in populations, and the successful development and evaluation of HIV/AIDS intervention, prevention, and care programs. It is also important that the presentation of surveillance data facilitate the use of this information for public health action. Therefore, it is essential that HIV/AIDS surveillance data meet specific criteria for quality before being analyzed and disseminated.

For more information about data quality, see "Data Quality" (page 10-3) in *Technical Guidance for HIV/AIDS Surveillance Programs, Volume I: Policies and Procedures* (https://team.cdc.gov/team/cdc/dispatch.cgi/_admin).

1. Structural Requirements

1.1 Minimum performance standards

Before analysis, HIV/AIDS surveillance data should meet minimum performance standards. A framework for evaluation and evaluation standards for HIV/AIDS surveillance data have been described by Hall and Mokotoff [1], and the standards are also listed in "Data Quality" (page 10-3) of *Technical Guidance for HIV/AIDS Surveillance Programs, Volume I: Policies and Procedures* (https://team.cdc.gov/team/cdc/dispatch.cgi/_admin). Additionally, any report or presentation of the data should include a discussion of the quality and the limitations of the data.

1.2 Sufficient capacity for analyzing HIV/AIDS surveillance data

HIV/AIDS surveillance programs need the capability to properly analyze and interpret data. Programs need either their own staff who are proficient in conducting statistical analyses or access to other people with the necessary skills (through an agreement with another group in the health department, an academic institution, contractor, or peer-to-peer consultation). All personnel involved with the analysis of HIV/AIDS surveillance data must comply with all appropriate security and confidentiality guidelines (http://www.cdc.gov/hiv/topics/surveillance/resources/guidelines/guidance/index.htm).

CDC provides licenses for SAS statistical software to its surveillance grantees and provides SAS code to assist with analysis. Surveillance programs need the most recent version of SAS to ensure functionality with analysis programs available from CDC.

To obtain relevant SAS programs to better analyze your data, contact Coordinating Center for Infectious Diseases (CCID) Informatics Customer Support by telephone at 877-659-7725 or by e-mail at ccidinformatics@cdc.gov. If your program does not have sufficient staff to support data analysis (e.g., your program is very small), please contact CDC to request technical support.

2. Outcome Standards

At a minimum, an HIV/AIDS surveillance report should be published annually to present descriptive HIV/AIDS surveillance data (1) to other units of the health department and community planning groups that use HIV/AIDS surveillance data to direct or prioritize services for HIV prevention and patient care; (2) to those who report the data; and (3) to the public. Routine reports that are responsive to the needs of various data users decrease the number of individual requests for data. The frequency of publication of standard statistical reports should be determined by each area but should be done at least annually.

3. Process Standards

Surveillance programs should have clear policies and procedures for analyzing and disseminating data. These policies and procedures should include

- Guidance to prevent the release of information that could identify an individual.
- Rules on the minimum cell sizes for reporting. For example, cell sizes of 5 or more may be released. However, decisions whether to provide breakdowns with fewer than the established threshold in a given category should be made in consultation with staff supervisors. Decisions about reporting cell size should be based on the size of the denominator (e.g., geographic area, time period).

  For additional guidance on cell size, see “Important considerations for presenting data in tables and figures” in Appendix B: Techniques for Analyzing and Displaying HIV/AIDS Surveillance Data.

- Rates calculated from numerators of less than 20 are considered unreliable and thus should be denoted as such in a footnote. In addition, it is recommended that numerators of less than 5 be suppressed in accordance with local procedures.
- Guidance that addresses data requests for which a meaningful interpretation of the data is not possible: data requests that are epidemiologically inappropriate because of small cell sizes, incomplete or uncertain data, or data unsuitable for the type of inquiry).
- A policy for review of products (reports, guidance documents, presentations).
- A list of job positions involved in analysis of data and the position of the person who approves the release of data.
- A mechanism that ensures that all staff comply with security and confidentiality procedures and policies.
- A policy on authorship.
- A policy on access to data by researchers outside the health department.

3.1 Statistical report

The statistical report should present descriptive statistics on the epidemiology of the local HIV/AIDS epidemic and should complement, but not repeat, the information in the epidemiologic profiles produced in the jurisdiction. This report should be published regularly (at least annually).

Analyses of the variables listed in 3.1.1 should be included in each statistical report. How to display them should be discussed with data users. For example, the HIV care part of the health department may want age groupings that match how it guides programs or is funded.

3.1.1 Race/ethnicity

In 1997, the Office of Management and Budget (OMB) announced the Revisions to the Standards for Classification of Federal Data on Race and Ethnicity (http://www.whitehouse.gov/omb/fedreg/1997standards.html). For comparability, it is recommended that local surveillance programs use these standards. Implementation by January 1, 2003, was mandated for federal agencies. What that means is that the old classifications may be used for data collected before that date; the new classifications must be used for data collected after that date. There are, however, situations in which surveillance programs may have begun using the new classifications before 2003.

Areas should present race/ethnicity data in a way that is pertinent and useful to their jurisdictions. However, it must be possible to collapse all race/ethnicity groupings into the OMB's revised classifications.

The new OMB race categories are as follows:

- American Indian or Alaska Native
- Asian
- Black
- Native Hawaiian or Other Pacific Islander
- White

Two ethnicity categories (Hispanic and not Hispanic) should be collected regardless of race. For more information, consult the U.S. Census Bureau Guidance on the Presentation and Comparison of Race and Hispanic Origin Data (http://www.census.gov/population/www/socdemo/compraceho.html).

If large numbers of persons of other races or ethnic groups reside in a jurisdiction, rates of HIV/AIDS among these groups may differ from the rates among other racial/ethnic groups. This may include locally relevant populations, such as foreign-born persons, which may have an impact on a regional epidemic. To improve the quality of information available to local prevention efforts, programs should consider collecting and displaying this information.

To obtain relevant SAS programs to better analyze your race/ethnicity data, contact CCID Informatics Customer Support by telephone at 877-659-7725 or by e-mail at ccidinformatics@cdc.gov.

3.1.2 Age

Data by age are often displayed in groups; data are also commonly displayed for the categories “adults and adolescents” and “children.” The age category for adults and adolescents comprises persons aged 13 years and older; the age category for children comprises persons younger than 13 years. Age at the time of diagnosis is most often used for prevention purposes; current age is most often used for care purposes.

3.1.3 Sex

Sex is most commonly expressed as assigned sex at birth. *Male* and *female* are the standard designations for sex; however, areas may choose to analyze and display data for subcategories based on local needs; for instance, areas may wish to collect and analyze data for transgender populations.

3.1.4 HIV transmission category

Transmission category is the term for the hierarchical classification that summarizes a person’s possible HIV risk factors; the summary classification results from selecting, from the presumed hierarchical order of probability, the risk factor on the case report form that is most likely to have been responsible for transmission.

Data on risk factors are typically displayed in the following transmission categories:

- Male-to-male sexual contact
- Injection drug use (IDU)
- Male-to-male sexual contact and IDU
- High-risk heterosexual contact (defined as heterosexual contact with a person known to have, or to be at high risk for, HIV infection)
- Perinatal exposure (from HIV-infected mother to infant)
- Other (This category typically includes transmission through blood transfusion or hemophilia; since the advent of stringent blood screening tests in 1985, the number of such transmissions has become negligible.)
- No reported risk factor (NRR): cases in which risk factor information is absent from the initial case report
- No identified risk factor (NIR): cases for which epidemiologic follow-up has been conducted and sources of data have been reviewed, but no risk factor information has been identified (Any case without a reported risk factor 12 months after report date is considered NIR.)

  Areas should consider whether it is relevant or helpful to data users to keep the NRR and NIR categories separate.

For other ways to display information about HIV transmission, review the position statements of the Council of State and Territorial Epidemiologists at (http://www.cste.org/dnn/AnnualConference/PositionStatements/tabid/191/Default.aspx).

Risk factor redistribution: Because recently reported cases of HIV or AIDS are more likely to be reported without sufficient risk factor information and risk factor information may never be reported for some cases, some transmission categories will be underestimated unless adjustments are made. Historically, CDC has adjusted risk factor information in analyses of the national HIV/AIDS surveillance data set based on risk factor redistributions of cases that were diagnosed 3–10 years earlier and initially classified as NRR, but that were later reclassified because a risk factor was found through chart review or follow-up investigation. In the future, CDC will use *multiple imputation* to redistribute HIV risk factors for cases that are reported without risk factor information [2]. As part of this activity, CDC will provide analysis programs and technical assistance for implementing multiple imputation locally. The most important determinants of whether it is reasonable for a state or locality to redistribute risk factors are the overall number of cases reported, the proportion reported without risk factor, and the initial risk factor distribution.

3.1.5 Geographic area

Surveillance programs should, at a minimum, analyze and display data for their entire jurisdiction (e.g., state) and, in most areas, by designated geographic subdivision, such as health district, county, or census unit. Areas should conform to security and confidentiality guidelines when displaying small cell sizes.

For further assistance, see *Technical Guidance for HIV/AIDS Surveillance Programs, Volume III: Security and Confidentiality Guidelines* (https://team.cdc.gov/team/cdc/dispatch.cgi/_admin).

3.2 Recommended analyses for annual or more frequent reports

Spectrum of Disease: Analyses and reports should describe various aspects of the spectrum of HIV disease.

- **HIV/AIDS**: The term *HIV/AIDS* is used to refer to 3 categories of HIV diagnoses collectively:
  1. a diagnosis of HIV infection (not AIDS)
  2. a diagnosis of HIV infection and a later diagnosis of AIDS
  3. concurrent diagnoses of HIV infection and AIDS (i.e., diagnoses of HIV infection and AIDS during the same calendar month)
- **AIDS**: An updated HIV case definition (including AIDS) was published in 2008 [3]. AIDS represents the end-stage of HIV disease and is less representative of the current state of the HIV epidemic than are HIV diagnoses. However, AIDS data still provide valuable information about late testing, inadequate care, and potential differences in access to HIV testing and care services.

  For more information on this topic, see http://www.cdc.gov/mmwr/pdf/rr/rr5710.pdf.

- **HIV incidence** (if available): HIV incidence is defined as all cases newly infected with HIV during a specified period of time, whether or not they have been diagnosed and reported to the surveillance system. CDC is developing technical assistance to help reporting areas produce local HIV incidence estimates.

  For more information on this topic, see Appendix A and see also "HIV Incidence Surveillance" (page 12-5) in *Technical Guidance for HIV/AIDS Surveillance Programs, Volume I: Policies and Procedures* (https://team.cdc.gov/team/cdc/dispatch.cgi/_admin).

- Number, percentage, and rate of new diagnoses of HIV/AIDS and AIDS in most recent calendar year and the number and percentage of prevalent HIV infections (not AIDS), AIDS and HIV/AIDS (this is the sum of persons living with HIV infection [not AIDS] and persons living with AIDS), overall and by
  - Race/ethnicity
  - Sex
  - Age group
  - Transmission category
  - Transmission category for each sex and race/ethnicity (may not be applicable for all areas or for all racial or ethnic groups, depending on morbidity)
  - Additional cross-tabulations (as appropriate)
- Trends
  - Analyzing data over time is useful for monitoring changes in various measures of HIV/AIDS surveillance. Trend analyses are often used to evaluate progress in prevention, testing, and treatment programs. HIV/AIDS surveillance data are most commonly compared by calendar year (although local programs may find trend analyses by other periods [e.g., by quarter] to be of value).
  - Trend analyses can be performed for case counts, rates, percentages. (One method for analyzing trends is to calculate the estimated annual percent change; see Appendix A.) However, special considerations must be taken into account when interpreting trend analyses. Analysts need to be aware of significant changes that may have occurred during the analysis period and that might affect the data of interest. For example, analysis of the 10-year trend in the number of annual HIV diagnoses could be skewed if a major new testing program was initiated during year 5. If that information is not taken into account, an increase in the number of new infections after year 5 could be misinterpreted. In this instance, analysis of the rate of diagnoses per 100 persons tested could help explain differences before and after the testing initiative.

    For information about trends, prevalence and prevalence rates, incidence and incidence rates, and how to analyze them, see Appendixes A and B.

3.3 Suggested additional analyses

- **Unmet need or not in care**: Data on persons with a diagnosis but who are not accessing available services or not currently receiving medical care can be used to identify barriers to care and services. Areas can establish different parameters for defining unmet need or not in care. For example, unmet need for HIV primary medical care might be defined as no evidence within a 12-month period of (1) CD4 cell count, (2) viral load testing, or (3) prescription for prophylactic or antiretroviral medications (if that information has been collected). For the CD4 cell counts and viral load levels, the area will need to obtain the laboratory results of these tests, including reports of undetectable viral load. Analyses of adequate care and unmet need should be stratified by basic demographic variables (e.g., race, age, sex).

  For more information, go to ftp://ftp.hrsa.gov/hab/unmetneedpracticalguide.pdf

- **Late testers**: Knowing the proportion of persons who receive a diagnosis of HIV infection and then receive a diagnosis of AIDS within the next 12 months will be helpful in focusing routine testing in many settings (e.g., clinical settings, counseling and testing programs at health departments, community-based organizations). Stratifying the analyses by race/ethnicity, sex, transmission category, age group, or country of origin may also be useful.
- **Diagnoses, by country of origin**: Several states have reported increases in the number of diagnoses among foreign-born persons. HIV-infected foreign-born persons may differ from HIV-infected U.S.-born persons in age, transmission category, race/ethnicity, and sex. The subpopulations in states may differ as well, so some states may want to include analyses of subpopulations. Changes in the rate of immigration of HIV-positive persons can influence trends among diagnoses in some populations. Auxiliary data sources (such as the U.S. Census Bureau) should be reviewed to allow complete and accurate interpretation of the data.
- **Topics of local interest**: In addition to collecting standard HIV/AIDS surveillance data, local areas may collect information on additional topics of interest. For example, a potential topic of local interest is transgender persons with a diagnosis of HIV infection.

  For examples of HIV/AIDS surveillance analyses and the display of data, see Appendix B. For a list of the Web addresses for reports of state or local HIV and AIDS surveillance, see the last page of CDC's HIV/AIDS Surveillance Report (http://www.cdc.gov/hiv/topics/surveillance/resources/reports/).

  For information on epidemiologic profiles, see Appendix C. (*Integrated Guidelines for Developing Epidemiologic Profiles* is also available at http://www.cdc.gov/hiv/epi_guidelines.htm.)

3.4 Linking HIV/AIDS surveillance data with other data sources

Many persons with HIV/AIDS have other medical conditions that can affect their service needs. Examining comorbid conditions may provide further information on service needs, service utilization, and quality of life. Linking may also provide missing information. For instance, race/ethnicity or risk factors may be found in the record for another disease that is not in the surveillance record. For further information on this topic, see "Record Linkage" (page 9-3) in *Technical Guidance for HIV/AIDS Surveillance Programs, Volume I: Policies and Procedures* (https://team.cdc.gov/team/cdc/dispatch.cgi/_admin).

Appendixes

Appendix A Terms, Definitions, and Calculations Relevant to HIV/AIDS Surveillance Data

Appendix B Techniques for Analyzing and Displaying HIV/AIDS Surveillance Data

Appendix C Integrated HIV/AIDS Epidemiologic Profile

Appendix D Tools and Resources for Analyses

Appendix A

Terms, Definitions, and Calculations Relevant to HIV/AIDS Surveillance Data

Case: "In epidemiology, a person in the population or study group identified as having the particular disease, health disorder, or condition under investigation" [4], such as HIV infection (i.e., an HIV case) or AIDS (i.e., an AIDS case).

Confidence interval (CI): A range of values for a measure that is believed to contain the true value at a specified level of statistical certainty (most commonly 95%).

Cumulative cases: The total number of cases of a disease reported or diagnosed during a specified time regardless of current vital status. Cumulative cases therefore include cases in persons who have already died.

> *Example*: Assume that 9,000 AIDS cases had been diagnosed in a state from the beginning of the epidemic through the year 2001. Of the 9,000 persons with AIDS, 4,000 had died. The cumulative number of AIDS cases diagnosed in that state through 2001 would be 9,000.

Cumulative incidence: The proportion of a group of people who experience the onset of a health-related event during a specified period of time. Cumulative incidence is calculated by dividing the number of new cases during a specified period by the population at risk for the health-related event during that period.

$$\text{cumulative incidence} = \frac{\text{number of new cases during a given period of time}}{\text{population at risk}}$$

It is important to pay attention to the time period to which the cumulative incidence relates. Cumulative incidence is commonly calculated for annual time periods, but it can also be calculated for shorter or longer periods. Use caution when comparing incidence calculations for differing periods.

Incidence: The number of new infections in a defined population during a specific period, often 1 year, which can be used to measure disease frequency. It is important to understand the difference between HIV incidence and newly reported HIV diagnoses. Incident HIV cases refer to all cases in persons infected with HIV during a specified period. New HIV diagnoses, however, not only include cases in persons whose diagnosis was made after recent infection but can include cases in persons who were infected earlier but whose diagnosis was recent.

CDC is currently funding selected surveillance programs to conduct HIV incidence surveillance. Data from these areas will be used to provide national estimates of new HIV infections (not just the number of cases newly diagnosed and reported to surveillance).

For more detailed information on incidence surveillance practices in the United States, see "HIV Incidence Surveillance" (page 12-5) in *Technical Guidance for HIV/AIDS Surveillance Programs, Volume I: Policies and Procedures* (https://team.cdc.gov/team/cdc/dispatch.cgi/_admin).

Mean: The sum of values for a variable, a group, or other category divided by the total number of values (e.g., in a data set). The mean is what many people refer to as an average.

Example: Assume that people in a given service area in 2001 are the following ages at diagnosis of HIV infection: 18, 18, 19, 20, 22, 23, 26, 31, 41. The total of the 9 ages = 218 years.

$$\frac{218 \text{ years}}{9} = \text{mean age } 24.2 \text{ years}$$

Median: The middle value in a data set: approximately half the values will be higher and half will be lower. The median is useful when a data set contains a few unusually high or unusually low values, which can affect the mean. It is also useful when data are skewed, meaning that most of the values are at one extreme or the other.

Example: Assume the following ages at diagnosis of HIV in the year 2001 data for a given service area: 18, 18, 19, 20, 22, 23, 26, 31, 99. Although the mean age is 30.7, the median age is 22. In this instance, the median age better reflects the central value of age for the population.

No identified risk factor (NIR): This is an NRR (no reported risk) case for which an HIV risk factor cannot be identified or confirmed 1) although all available data sources have been reviewed or contacted, or 2) epidemiologic follow-up was either not initiated or not completed, but 12 months have elapsed since the date of the initial case report. All cases without a reported risk factor 12 months after the report date are considered as NIR.

No reported risk factor (NRR): A case is initially classified as NRR if it is reported without any risk factor information or with unconfirmed COPHI risk factor information. However, regardless of follow-up, if a risk factor has not been identified within 12 months after being classified as NRR, the case is considered NIR.

Percentage: A proportion of the whole, in which the whole is 100.

Example: Assume that 15 of the 60 cases of AIDS in a given year in a state occurred in women.

$$\frac{15}{60} = .25 \times 100 = 25\%$$

Prevalence: The proportion of cases in a population at risk, measured at a given point in time (often referred to as point prevalence).

Prevalence can also be measured over a period of time (e.g., a year). This second type of prevalence—a combination of point prevalence and incidence—is called period prevalence.

Prevalence data provide an indication of the extent of a condition and may have implications for the services needed in a community. Both measures of prevalence are proportions—as such, they are dimensionless and should not be described as rates [5].

$$\text{prevalence} = \frac{\text{\# of existing cases*}}{\text{population at risk*}}$$

* during specified period

Source: (http://mchp-appserv.cpe.umanitoba.ca/viewConcept.php?conceptID=1036)

Prevalence does not indicate how long a person has had a disease. It can, however, be used to estimate the probability that an individual in a population will have a disease at a point in time. For HIV/AIDS surveillance, prevalence refers to persons living with HIV infection or AIDS regardless of the time of infection or the date of diagnosis.

To obtain relevant SAS programs to better analyze your prevalence data, contact CCID Informatics Customer Support by telephone at 877-659-7725 or by e-mail at ccidinformatics@cdc.gov

Proportion: A portion of a population or a data set, usually expressed as a fraction or a percentage of the population or the data set.

Example: Assume that 12 of 20 HIV prevention programs in a given service area are school-based programs.

Calculation of the proportion as a fraction:

$$\frac{12}{20} = .6 = 6/10 = 3/5$$

Calculation of the proportion as a percentage:

$$\frac{12}{20} = .6 \times 100\% = 60\%$$

Qualitative data: Information from sources such as narrative behavior studies, focus groups, open-ended interviews, direct observations, ethnographic studies, and documents. Findings from these sources are usually described in terms of common themes and patterns of response rather than numerically or statistically. Qualitative data are useful as supplements to surveillance data to help explain behavioral risk factors and associated factors in specific locales or populations that may not be well represented in routine surveillance data.

Quantitative data: Numeric information (e.g., numbers, rates, and percentages).

Range: The smallest and the largest values in a series.

Example: Assume the following ages at diagnosis of HIV in the year 2001 in a given service area: 18, 18, 19, 20, 22, 23, 26, 31, and 41. The range in values is 18–41.

Rate: A measure of the frequency of an event compared with the number of persons at risk for the event. When rates are being calculated, it is usual for the denominator to be the general population rather than the population potentially exposed to HIV infection by various high-risk behaviors. The size of the general population is known from data from the U.S Census Bureau, whereas the size of a population at high risk is usually not known.

$$\frac{\text{number of HIV diagnoses*}}{\text{population*}} \times 100{,}000 = \text{population rate of HIV diagnoses}$$

*during a given period

For ease of comparison, the multiplier (100,000) is used to convert the resulting fraction to number of cases per 100,000 population. Although arbitrary, the choice of 100,000 is standard practice.

Example: Assume that 16 cases of HIV infection were diagnosed during 2000 in a service area and that 400,000 persons lived in the area in 2000.

Calculation of the rate:

$$\frac{16}{400{,}000} \times 100{,}000 = 4 \text{ per } 100{,}000$$

Sample: Often, it may not be feasible or practical to collect data on every member of a population. However, we may be able to use information from a sample of the population to make inferences about the wider population. In a **random sample**, an individual has an equal chance of being selected from the population. If a **stratified sampling** scheme is used, the rate at which individuals from several subsets are sampled can be varied so as to produce a greater representation of some classes than of others [4]. A **convenience sample**, on the other hand, is a sample chosen on the basis of accessibility, expediency, cost, efficiency or other reason not directly concerned with sampling parameters. When a convenience sample is used, inferences to the general population cannot be made.

Seroprevalence: The number of persons in a defined population who test positive for HIV infection (based on HIV testing of blood specimens). Seroprevalence is often presented as a percentage of the total specimens tested.

Stratification: The separation of a population into subgroups according to predetermined criteria, such as age group, sex, socioeconomic status. Stratification is used to control confounding effects and to detect modifying effects.

Trend: An assessment of changes over time. A simple linear trend could be described by calculating how much the quantity being measured increased (or decreased) from the beginning value (at the beginning of the period) to the ending value (at the end of the period). The trend could be further described by calculating a time-rate of change in the quantity measured. This is the difference between the beginning and ending values, divided by the number of time units (e.g., years) for which the trend is measured. This calculation yields the amount of increase (or decrease) per time unit. Another key factor is the statistical significance of the trend, which evaluates whether the trend is likely to be real or due to chance alone. Trends should not be calculated or presented for small

numbers because small numbers are subject to wide fluctuations. An effective test for trends is the Cochran-Armitage test for trend. (For information on how to analyze data using this test, go to http://umanitoba.ca/medicine/units/mchp/).

It is recommended that data on cases of HIV infection (regardless of AIDS diagnosis) be reported from areas that have had HIV reporting long enough (usually 4 years) to allow for stabilization of data collection. The most recent 5-year period is often used when displaying data over time.

The estimated annual percent change (EAPC) and 95% confidence intervals in the number of annual diagnoses or diagnosis rates can be determined by fitting a linear regression line to the natural logarithm of the number of diagnoses (rate), using calendar year of diagnosis as an independent variable. The significance of a trend is assessed by determining whether the 95% confidence interval for the EAPC includes zero.

For more on how to calculate EAPC, go to http://seer.cancer.gov/seerstat/WebHelp/seerstat.htm#Trend_Algorithms.htm.

Appendix B

Techniques for Analyzing and Displaying HIV/AIDS Surveillance Data

Surveillance data can summarize and display complex issues clearly and effectively and can emphasize specific points. Proper analysis and interpretation help make sense of the data and communicate findings to various audiences (e.g., community planning group, care providers). However, poorly designed analyses or confusing data can mislead users or distract them from the true findings of the data.

Analyzing Data

There are several ways to analyze HIV/AIDS surveillance data. It is important to understand the purpose and limitations of each of the analysis techniques. The following are examples of common data analyses.

Case count/number of cases/measures of frequency

The most basic measure of disease frequency is a simple count of affected individuals—also known as case count. Case counts can be displayed in aggregate (for the total population) or for groups (e.g., by sex, race/ethnicity). As an example, surveillance program X had 30 cases of AIDS diagnosed in 2000: 25 among men and 5 among women.

Count data alone have limited usefulness in describing details of the local epidemic because the size of the source population is not considered in calculating simple count data. Consequently, rates are commonly used.

Rates

Rates are calculated by dividing the number of events (numerator) by the size of the population (denominator) and including a measure of time. Rates calculated from numerators of less than 20 should be denoted in a footnote as unreliable. In addition, numerators of less than 5 should be suppressed in accordance with local procedures.

When comparing rates between populations, it is typical to standardize the denominator in order to make direct comparisons. The standardization is usually expressed as a factor of 100; that is, the number of events per population of 100 (10^2), 1,000 (10^3), 10,000 (10^4) or 100,000 (10^5). This standardization will depend on the magnitude of the local surveillance data—for national data, the population size is most often standardized to 100,000.

For example, if 25 cases of AIDS were diagnosed among 80,000 men in the year 2000, we would say that the diagnosis rate for men in 2000 is 25 per 80,000. Similarly, if 5 cases of AIDS were diagnosed among 75,000 women in the year 2000, we would say that the diagnosis rate for women in 2000 is 5 per 75,000. The standardized rate for men is calculated as 25/80,000 × 100,000 = 31.2 cases per 100,000 population; and the

standardized rate for women is 5/75,000 × 100,000 = 6.7 cases per 100,000 population. Comparing the standardized rates in this example, it would be correct to say that the rate for men is about 4.5 times greater than the rate for women.

Percentages

Percentages are calculated as the part of the total represented by various data elements. Added, the percentages of the elements equal 100%.

In our example (above), the total number of diagnosed AIDS cases in 2000 is 30: 25 cases in men and 5 in women. The percentages of cases, by sex, would be calculated as follows:

25/30 = 0.833 × 100 = 83.3% of the total cases were in men

5/30 = 0.167 × 100 = 16.7% of the total cases were in women

Estimating risk factor data

Recently reported cases of HIV infection are more likely than other cases to be reported without an HIV risk factor; thus, some transmission categories will be underestimated unless an adjustment is made. Multiple imputation is a statistical method for assigning values to data that are missing from large data sets [2]. Multiple imputation comprises 3 steps: (1) Impute missing variables by using a statistical model that utilizes random variation. (2) Perform statistical analyses on each of the data sets. (3) Combine all of the results from the data sets to calculate the standard error and an overall multiple imputation estimate. This method maintains the original variability of the missing data by creating imputed values, which are based on variables correlated with the variable to be imputed and with missingness. Uncertainty is accounted for by generating iterations of the missing data and observing the variability between the imputed data sets [6].

Multiple imputation is based on 3 assumptions: the data must be missing at random, the model that is used to generate the imputed values must be “correct” (include all anticipated predictor values), and the model that is used in the analysis and imputation must match (that is, use mostly the same variables). Multiple imputation offers several advantages that offset the uncertainty inherent in approximating missing data: it produces less biased subgroup estimates and offers a semiautomated approach.

Presenting and Displaying Data

Summarizing your data and presenting them in tables or figures is critical to getting the intended message across to the audience because raw data are difficult to understand, visualize, aggregate, or use in detecting trends. Tables and figures can be used to summarize and display complex data clearly and effectively and can emphasize specific points.

These tools can help identify and present distributions, trends, and relationships among the data. They help make sense of the data and communicate findings to the audience. However, poorly designed or executed tables and figures can mislead users or distract them from the message.

Tables may be the only presentation format needed when the data are few and relationships are straightforward (tables are the best choice when the display of exact values is important). Figures (e.g., line and bar graphs, pie charts) make more sense for trends and for comparing populations, especially when you wish to show populations stratified into subsets, such as males and females or age groups. The key points of tables and figures should always be explained in the accompanying narrative.

Important considerations for presenting data in tables and figures

- The table or figure should be an integral part of the text but should also be able to stand alone (i.e., the reader should understand the table or figure without reference to the text). Ideally, a table or figure should convey one main point.
- The table or figure should explain the who, what, when, and where of your data. For example, a figure (perhaps a line or bar graph) is useful for showing gender or racial differences, geographic differences, or trends.
- Consider the number of tables and figures you display. You should have enough to clearly summarize and display your data, but not so many that they are confusing and difficult to understand in terms of the text (tables and figures should be comprehensible to all users, regardless of their technical background).
- Write short, informative titles. Include year(s), geographic area, and groups or populations.
- For figures, write clear and consistent labels, and label all elements to avoid misunderstanding. For tables, write clear and consistent column headings and row entries (use consistent terms).
- Avoid clutter. Include only what you need to communicate the point. Eliminate unnecessary words that can detract from the message (e.g., footnotes to tables and notes to figures need not be expressed in complete sentences).
- Maintain scale and balance by keeping the width and height of the table or figure in proportion (i.e., for a figure, the length of the vertical [y] axis should be approximately two-thirds the length of the horizontal [x] axis; in general, tables are longer than they are wide).
- Name the source of your data.
- Discuss the key points of the table or figure in your text.
- Consider how your material will be produced. Often, handouts are photocopied rather than professionally printed. If a color document is photocopied in black and white, the data elements (e.g., bars in a chart or slices of a pie chart) will probably be difficult to distinguish. In bar charts and other figures (such as maps), shades of gray are preferable to cross-hatching or other patterns, but the shades must differ by at least 30%, or the gray elements will not be clearly distinguished in the original or in the copies (even if the document is professionally printed).

- Consider the preferences of your audience. If you have an opportunity, find out how they would like to see the data presented. That will help you determine the types of presentation that are easiest for them to understand and use.
- Consider the best way to present your data:
  - Ensure that your presentation does not inadvertently stigmatize the demographic groups to which the data refer. Consultation with the affected communities can help avoid this problem.
  - In situations in which the presentation of data on larger groups would overwhelm the presentation of data on smaller groups, you can present the data on the smaller groups separately. In the explanation below the figure, point out the differences between the larger and smaller groups.
  - When the numbers for a group are small, observe restrictions on cell size to protect confidentiality. Technical assistance may be required to interpret the data when analyzing small numbers of cases because small, absolute changes in the number of cases can produce large changes in rates. For example, rates calculated from numerators of less than 20 should be denoted in a footnote as unreliable. In addition, case counts of less than 5 should be suppressed.

    For more information on presenting data with small numbers, see "Areas with Low Morbidity and Minimal Data" in chapter 5 of *Integrated Guidelines for Developing Epidemiologic Profiles: HIV Prevention and Ryan White CARE Act Community Planning* (http://www.cdc.gov/hiv/epi_guidelines.htm).

  - Aggregation: The aggregation of values may be appropriate for data elements with large numbers of possible values, such as age, dates, and geographic areas. Also, it is customary to collapse a data set in order to create tables without small numbers as denominators or numerators in cells [7]. The following table shows some examples.

		Granularity: Aggregation		
Field	**Type**	**Small**	**Medium**	**Large**
Age	Continuous	Year of birth	5-year age group	10-year age group
Date of occurrence	Continuous	Day	Month	Year
Diagnosis	Nominal	Complete ICD code	3-digit ICD	"Selected cause" tabulation
Geography	Ordinal (spatial)	Zip code, census tract	Sub-county area	County

Source: Washington State Department of Health. *Guidelines for Working with Small Numbers*. Olympia, WA: Washington State Department of Health; 2006. http://www.doh.wa.gov/Data/Guidelines/SmallNumbers.htm#Guidelines.

- Cell suppression: When it is not possible, or not desirable, to create a table without small numbers as denominators or numerators in cells, then cell suppression is used, together with complementary suppression. "Primary" cell suppression is used to withhold the data (numerator) in the cell that fails to meet the threshold, followed by suppression of cells as appropriate in order to avoid inadvertent disclosure through back-calculation [7]. Additionally, cell suppression is useful in preventing data users from calculating rates that might be unstable because of small numerators [8].

Age (yr)	Black	White	Other	Total
0–34	S	30	S	60
35–64	S	60	S	150
≥65	70	90	80	240
Total	**120**	**180**	**150**	**450**

Source: adapted from Washington State Department of Health. *Guidelines for Working with Small Numbers*. Olympia, WA: Washington State Department of Health. http://www.doh.wa.gov/Data/Guidelines/SmallNumbers.htm#Guidelines.

S = data suppressed.

- Other methods: When neither of these methods (aggregation of data values to increase granularity, and cell suppression) is satisfactory, two alternatives to reduce the risk of a confidentiality breach remain. The first, and better choice, is to combine multiple years of data (which is a form of aggregation). The effect will be to increase the effective population size, because the (usually unstated) denominator is actually person-time in rate calculations, and the numerators are likely to rise correspondingly as well. The second alternative is to omit certain fields from analysis. The omission of fields is demonstrated in the following hypothetical example of asthma data. It was not possible to achieve adequately large cell denominators in annual county-level data showing both age-specific and gender-specific counts and rates. An advisory group opted to omit the gender-specific data and to display only tables of age-specific data, on the grounds that no intervention programs targeted groups differently on the basis of gender but that most intervention programs targeted age groups differently.

  For additional information on statistical methods for dealing with small cell sizes, go to http://www.fcsm.gov/working-papers/SPWP22_rev.pdf.

As you develop your product and determine which kind of display to use, ask yourself these questions:

- Can the audience understand what I want to convey by looking at this type of display, or would another type be better?
- Given the needs of the audience, is this presentation of the data logical?

Appendix C

Integrated HIV/AIDS Epidemiologic Profile

The Integrated Epidemiologic Profile (IEP) describes the HIV/AIDS epidemic in various populations in a defined geographic area. A prevention-oriented epidemiologic profile is a requirement of local Community Prevention Planning. More recently, epidemiologic profiles that integrate the needs of care providers have become standard. These two purposes may require differing displays of data. For example, prevention planners usually want to know the age of infected persons at the time of HIV diagnosis, while providers of HIV-related care and services usually find current age more useful. An area may need to consider displaying data several ways in an IEP.

The IEP identifies characteristics of the general population, HIV-infected populations, and non-infected (and untested) persons whose behaviors place them at risk for HIV infection. The IEP consists of information describing the effect of HIV/AIDS on an area in terms of sociodemographic, geographic, behavioral, and clinical characteristics. Data sources for preparing the IEP include the U.S. Census Bureau; HIV/AIDS surveillance data; data from special studies (such as the Behavioral Risk Factor Surveillance System [BRFSS], Youth Risk Behavior Survey [YRBS], the National HIV Behavioral Surveillance System [NHBS], and the Medical Monitoring Project [MMP]); and services data. The IEP serves as a source of quantitative data from which HIV prevention and care needs are identified and priorities set for a given jurisdiction.

An IEP is designed to

- Provide a thorough description of the HIV/AIDS epidemic among the various populations (and subpopulations) in a service area
- Describe the current status of HIV/AIDS cases in the service area and provide some understanding of how the epidemic may look in the future
- Identify characteristics of the general population and of populations who are living with, or at high risk for, HIV/AIDS in defined geographic areas and who need primary and secondary prevention or care services
- Provide information required to conduct needs assessments and gap analyses

Before creating an IEP, surveillance programs should consult their prevention and care planning councils so that their needs can be considered. For example, most statistical reports from a specific area describe only the HIV-infected persons living in the jurisdiction at the time of diagnosis. However, prevention and care planners need information on all infected persons currently living in the jurisdiction. HIV surveillance staff may decide to include data on all HIV-infected persons living in their jurisdiction.

For information on preparing an epidemiologic profile, see *Integrated Guidelines for Developing Epidemiologic Profiles: HIV Prevention and Ryan White CARE Act Community Planning* (http://www.cdc.gov/hiv/epi_guidelines.htm).

Appendix D

Tools and Resources for Analyses

CCID Informatics Customer Support (formerly known as the DHAP Help Desk) is available for assistance with SAS programs and other technical assistance.

A listserv has been established: Epidemiological Capacity Building and Technical Assistance for Integrated HIV/AIDS Epi-Profiles (HIV-IEP-CBATA@LISTSERV.CDC.GOV). The listserv is primarily for jurisdictions separately funded for this activity. However, all surveillance programs are encouraged to submit queries and comments related to the analysis and dissemination of HIV/AIDS surveillance data. To obtain access to the listserv, contact CCID Informatics Customer Support by telephone at 877-659-7725 or by e-mail at ccidinformatics@cdc.gov.

References

1. Hall HI, Mokotoff, ED. Setting standards and an evaluation framework for human immunodeficiency virus/acquired immunodeficiency syndrome surveillance. *JPHMP* 2007;13(5):519–523.
2. McDavid Harrison K, Kajese T, Hall HI, Song R. Risk factor redistribution of the national HIV/AIDS surveillance data: an alternative approach. *Public Health Rep* 2008;123:618–627.
3. Centers for Disease Control and Prevention. Revised Surveillance Case Definitions for HIV Infection Among Adults, Adolescents, and Children Aged <18 Months and for HIV Infection and AIDS Among Children Aged 18 Months to <13 Years—United States, 2008 (*MMWR* 2008; 57(RR10):1–15).
4. Last JM, ed. *A Dictionary of Epidemiology.* 4th ed. New York: Oxford University Press; 2001.
5. Friis RH, Sellers TA. *Epidemiology for public health practice*. Gaithersburg, MD: Aspen Publishers; 1999.
6. Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. New York: John Wiley & Sons; 1987.
7. Washington State Department of Health. *Guidelines for Working with Small Numbers*. Olympia, WA: Washington State Department of Health; 2006. http://www.doh.wa.gov/Data/Guidelines/SmallNumbers.htm#Guidelines.
8. Delcher PC, Edwards KT, Stover JA, Newman LM, Groseclose SL, Rajnik DM. Data suppression strategies used during surveillance data release by sexually transmitted disease prevention programs. *JPHMP* 2008,14(2): E1–E8.