State of California—Health and Human Services Agency
California Department of Public Health

KAREN L. SMITH, MD, MPH
Director & State Health Officer

EDMUND G. BROWN JR.
Governor

August 31, 2015

TO: Participants in the July 2015 Proficiency Test in Forensic Alcohol Analysis

SUBJECT: Assigned Values and Acceptable Ranges for the July 2015 Proficiency Test in Forensic Alcohol Analysis

Attached is a summary of the descriptive statistics for the July 2015 proficiency test in forensic alcohol analysis. Included here are the target formulation values, the true values as determined by the Department's analyses, the peer-group or consensus values and the standard deviations, and a graphical summary of the distribution of participant results.

Historically, the Department has determined the acceptable limits of performance based on reported results that are within the range representing ±5% of the 99% confidence interval of the peer group mean, where the range has been truncated to two significant figures (Table 1). This range is described as the "Tier #2 interval." The Department also calculates a "Tier #1 interval," which represents the range of reported results that are within ±5% of the 95% confidence interval of the peer group mean where the range is based on the results reported to three significant figures. Tier #1 is expected to include those laboratories demonstrating a high degree of accuracy. The second, wider tier would include those laboratories not as close to the central tendency as the first tier, but still accurate and therefore adequately competent. Again, historically, the Department has used the wider second tier to evaluate the laboratories' results.

The IUPAC International Harmonized Protocol for the Proficiency Testing of Analytical Chemistry Laboratories (Harmonized Protocol) recommends the use of z-scores for evaluating proficiency test data. However, the Harmonized Protocol notes that that the interpretation of the z-scores is based on the normal distribution of reported results, in which case the z-scores can be expected to follow the standard normal distribution. As indicated in Table 2, the results for Pools 06235 & 06295 in this proficiency test were not found to be normally distributed: departures from normality are indicated on the Normality plots, Figure 4. Accordingly, the use of z-scores may not be completely appropriate, but they still may be useful to identify outlier and/or warning level results. The expression for calculating a z-score is included in Table 2. Generally a score between -2 and +2 ($|z| \leq 2$) is considered satisfactory or acceptable. A score outside the range -3 to +3, inclusive ($|z| \geq 3$) is considered unsatisfactory or unacceptable and the laboratory must take corrective actions. Z-scores between -3 and -2 or +2 and +3 ($2 < |z| < 3$) are considered questionable and these two ranges should be used as warning limits. Scores within the warning limit ranges in two or more consecutive test events could be considered unacceptable.

The proficiency test results expressed as *z*-scores for the participants whose results were used to determine the peer group mean and statistics in the July 2015 test are summarized in Figure 5. Participants are identified by codes. An attachment to this letter provides codes for participants from your laboratory. The figure is provided for educational purposes only and was not used to formally evaluate a laboratory's performance.

Another approach for evaluating proficiency test data, which is non-parametric and does not require the data to be converted to a standard normal form, divides the test data at regular intervals or quantiles. The quartile is a type of quantile: the first quartile (Q_1) is defined as the middle value between the lowest value and the median of the data set. The second quartile (Q_2) is the median of the data set. The third quartile (Q_3) is the middle value between the median and the highest value of the data set. The interquartile range (IQR), a measure of the dispersion of the data, is the difference between the upper and lower quartiles ($IQR = Q_3 - Q_1$). Boundaries (called fences) are set at $Q_1 - 1.5$ IQR (lower fence) and $Q_3 + 1.5$ IQR (upper fence) to identify potential outliers in the tails of the distribution. In Figure 3, the data from the two pools are presented as box and whisker or Tukey plots with the quartiles and fences shown. The median of the data is shown by a black line and the mean of the data is shown by a red line inside the box. These figures can be used by the participants to evaluate their data.

This report is also available from the Forensic Alchol Program website. Go to http://www.cdph.ca.gov/programs/DFDRS/Pages/FDLB-ForensicAlcoholProgram.aspx and click desired document under the section, Forensic Alcohol Proficiency Test Results.

Sincerely,

Clay Larson, Chief
Abused Substances Analysis Section
Food and Drug Laboratory Branch

Statistical Data for July 2015 Proficiency Test in Forensic Alcohol Analysis

Table 1 CDPH Tier #1 and Tier #2 Acceptable Ranges

Pool	Peer Group Mean	Tier #1	Tier #2
#1	0.109	0.102 – 0.116	0.10 – 0.12
#2	0.238	0.224 – 0.252	0.22 – 0.25

Table 2 Summary of Test Pool Data

Parameter	POOL 1 (06235)		POOL 2 (06295)	
Pre-distribution Data[1]	Target Value True Value Standard Deviation	0.11% 0.108 0.0004	Target Value True Value Standard Deviation	0.24% 0.236 0.0022
Descriptive statistics	Mean Adjusted Mean[2] Standard Deviation Median Standard Error[3] Minimum Maximum Count	0.109 0.109 0.0023 0.109 0.0002 0.100 0.115 103	Mean Adjusted Mean[2] Standard Deviation Median Standard Error[3] Minimum Maximum Count	0.237 0.238 0.0049 0.238 0.0005 0.218 0.259 109
Descriptive statistics (non-parametric, box plot), Figure 3	Q1 (25%) Q3 (75%) IQR Lower Fence Upper Fence	0.108 0.110 0.003 0.124 0.136	Q1 (25%) Q3 (75%) IQR Lower Fence Upper Fence	0.236 0.240 0.004 0.230 0.246
Histogram	Figure 1		Figure 2	
Normal distribution?[4]	NO		NO	
Robust mean, X*[5]	0.109		0.238	
Robust standard deviation[5], σ_{rob}	0.0016		0.0032	
Fitness-for-purpose standard deviation, σ_p[6]	0.0032		0.0059	
Consensus value (X_a)[7]	0.1090		0.2379	
Uncertainty of the consensus value, X_a , S.E.[8]	0.0002		0.0003	
$X_a \pm$ S.E.	0.1090 ± 0.0002		0.2379 ± 0.0003	
z-score calculation formula	$z = \frac{X - X_a}{\sigma_p}$		$z = \frac{X - X_a}{\sigma_p}$	

[1] Based on CDPH's Headspace Gas Chromatographic Method
[2] Mean determined from participant data after the removal of outlier(s)
[3] Standard Error of the Mean
[4] Shapiro-Wilk test used at 0.05 significance level (See also Normal Probability plot, Figure 4)
[5] Calculated per *Statistical methods for use in proficiency testing by interlaboratory comparisons ISO 13528 (Algorithm A)*
[6] σ_p "describes the standard uncertainty that is most appropriate for the application area of the results of the analysis" see *The International Harmonized Protocolfor the proficiency testion of analytical laboratories, IUPAC technical Report (3.1.1)*
[7] Determined as Mode ($\mu_{1/2}$) of Gaussian Kernel distribution, see also *M. Thompson. "Bump-hunting for the proficiency tester: Searching for multimodality", Analyst 127,1359–1364 (2002)*
[8] Determined as Standard Error of Mode using bootstrap simulation technique with bandwidth of $0.75*\sigma_p$

Figure 1

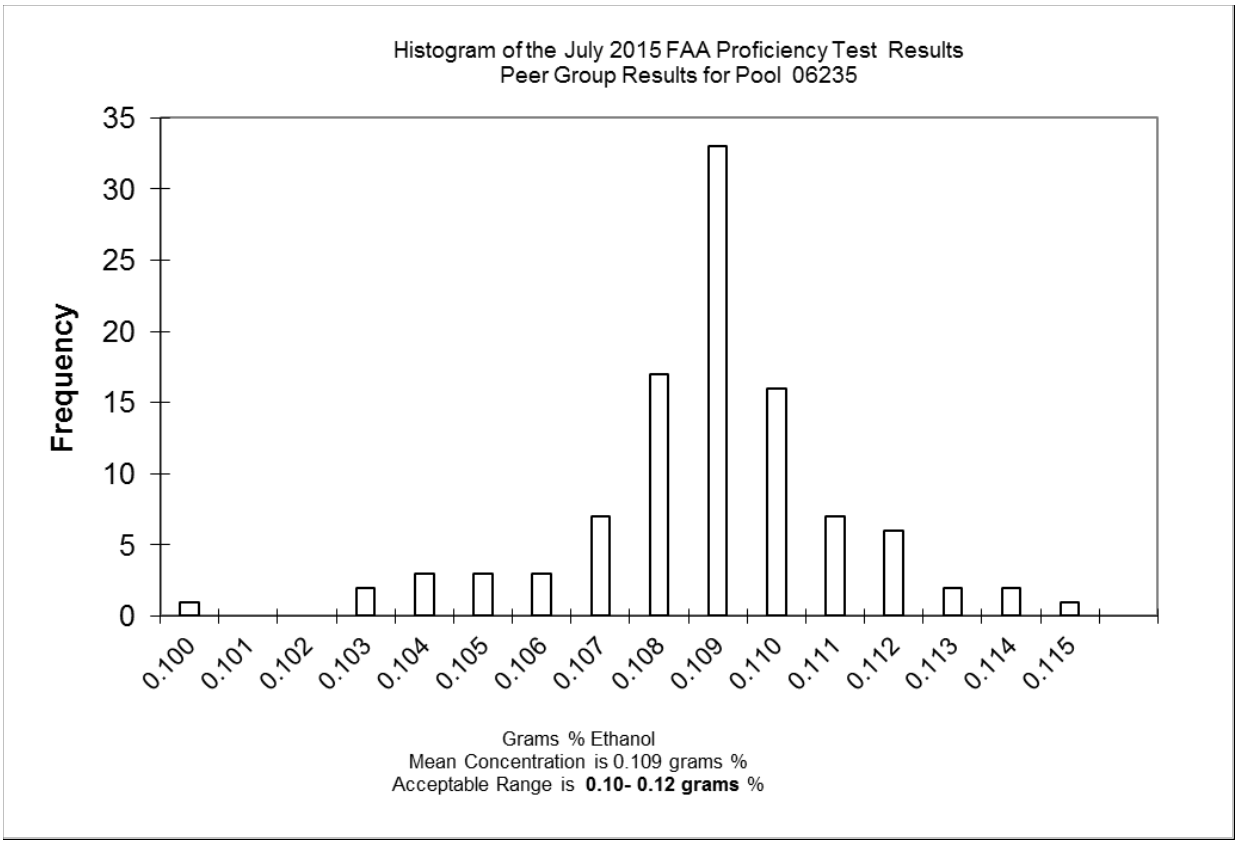


Figure 2

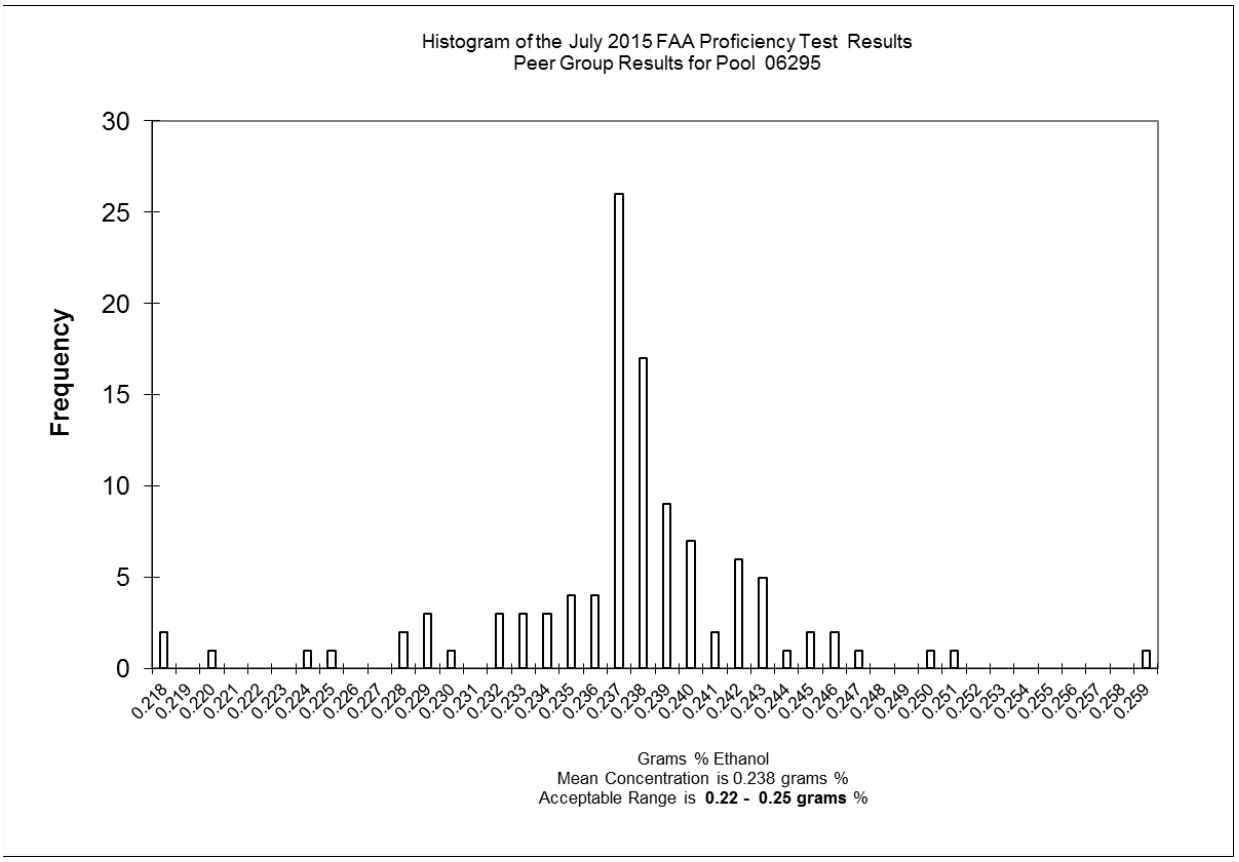


Figure 3 SigmaPlot analysis of pools 06235 & 06295

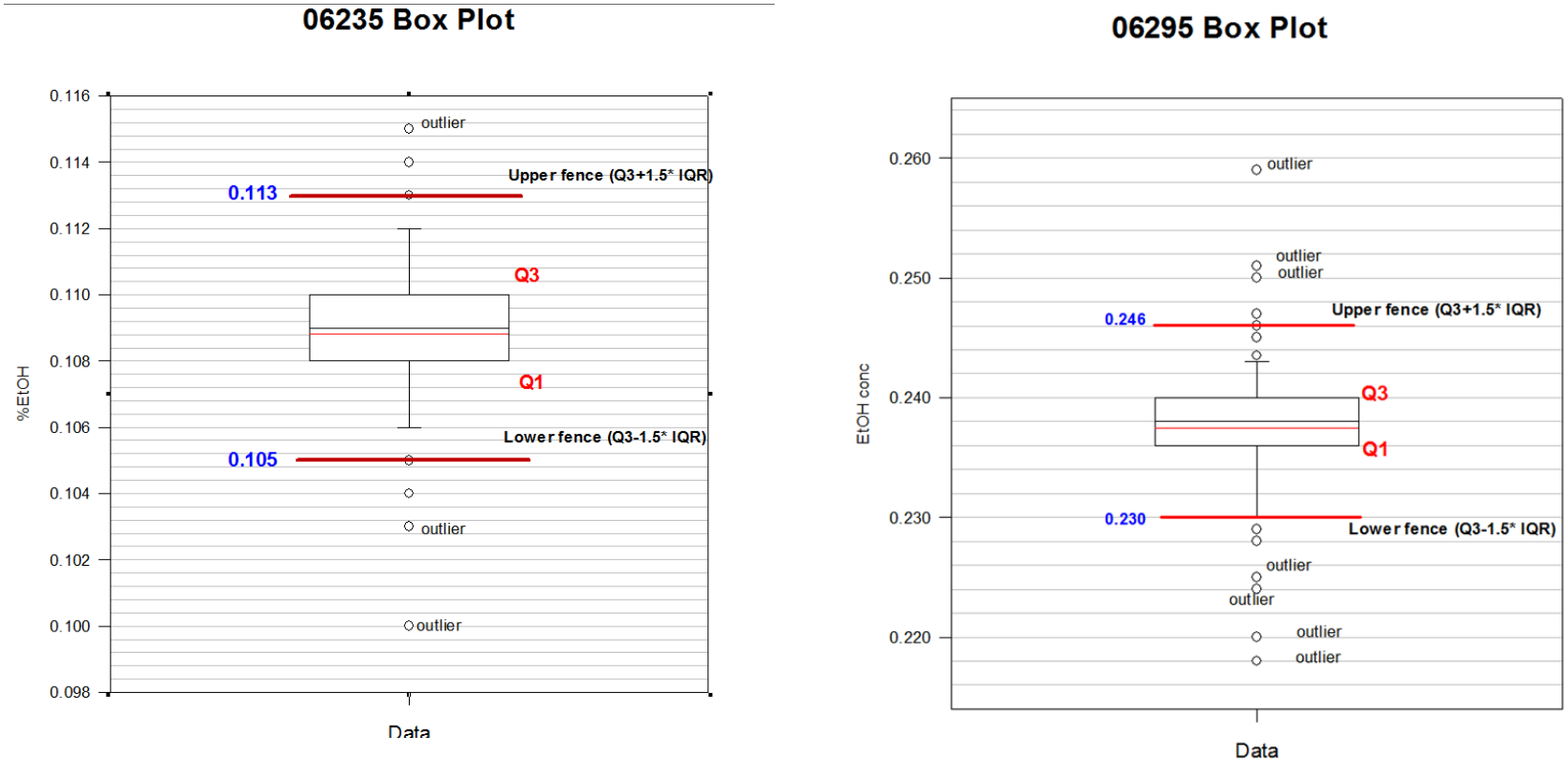
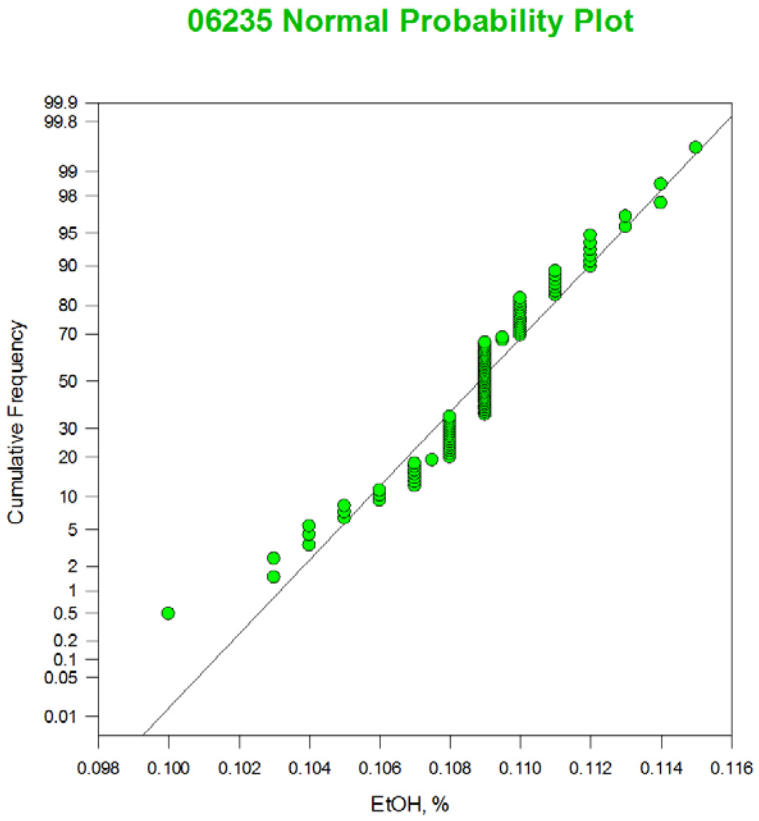


Figure 4 Normal Probability Plots and Histograms of pools' data, binning size is ½ of pools standard deviation.



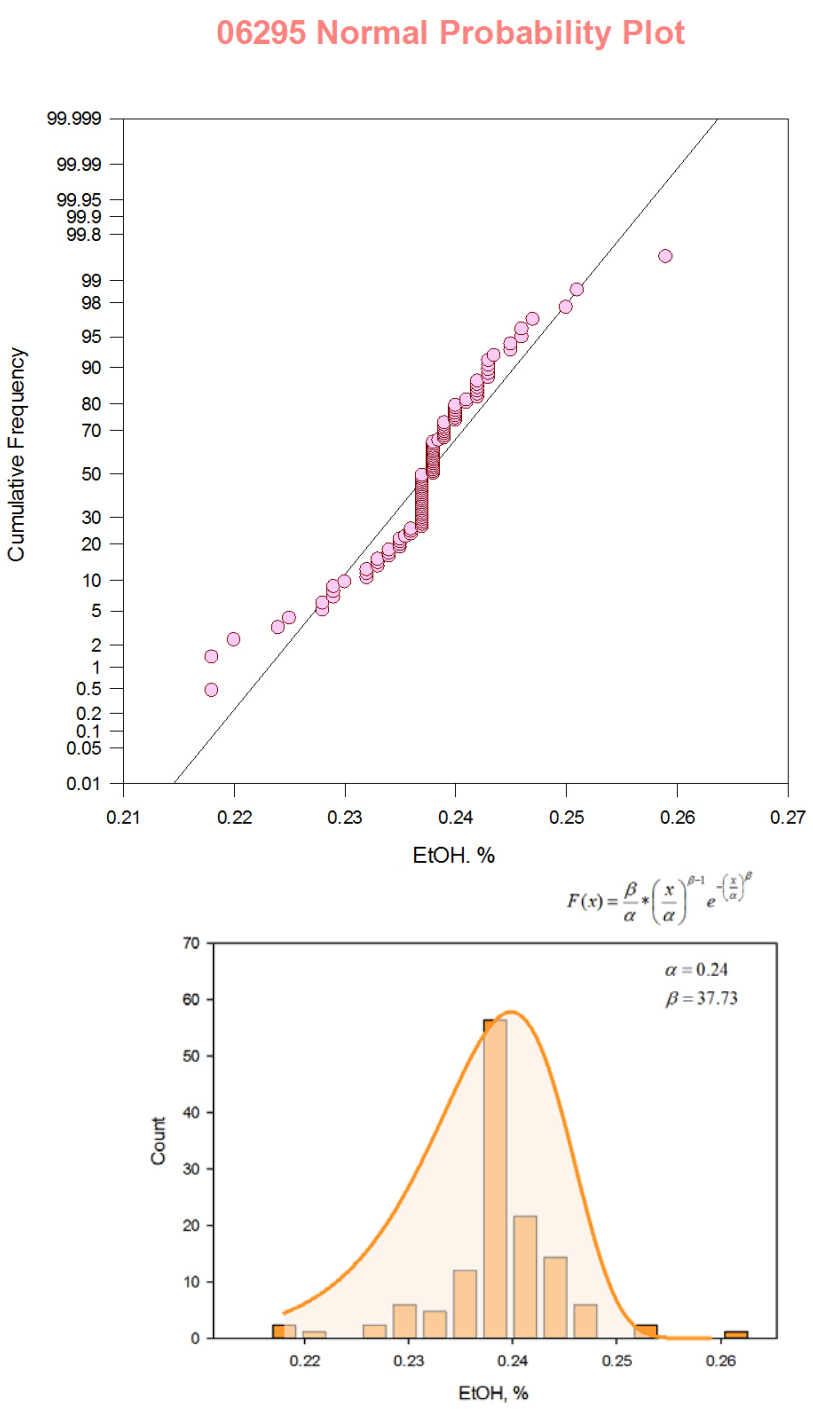
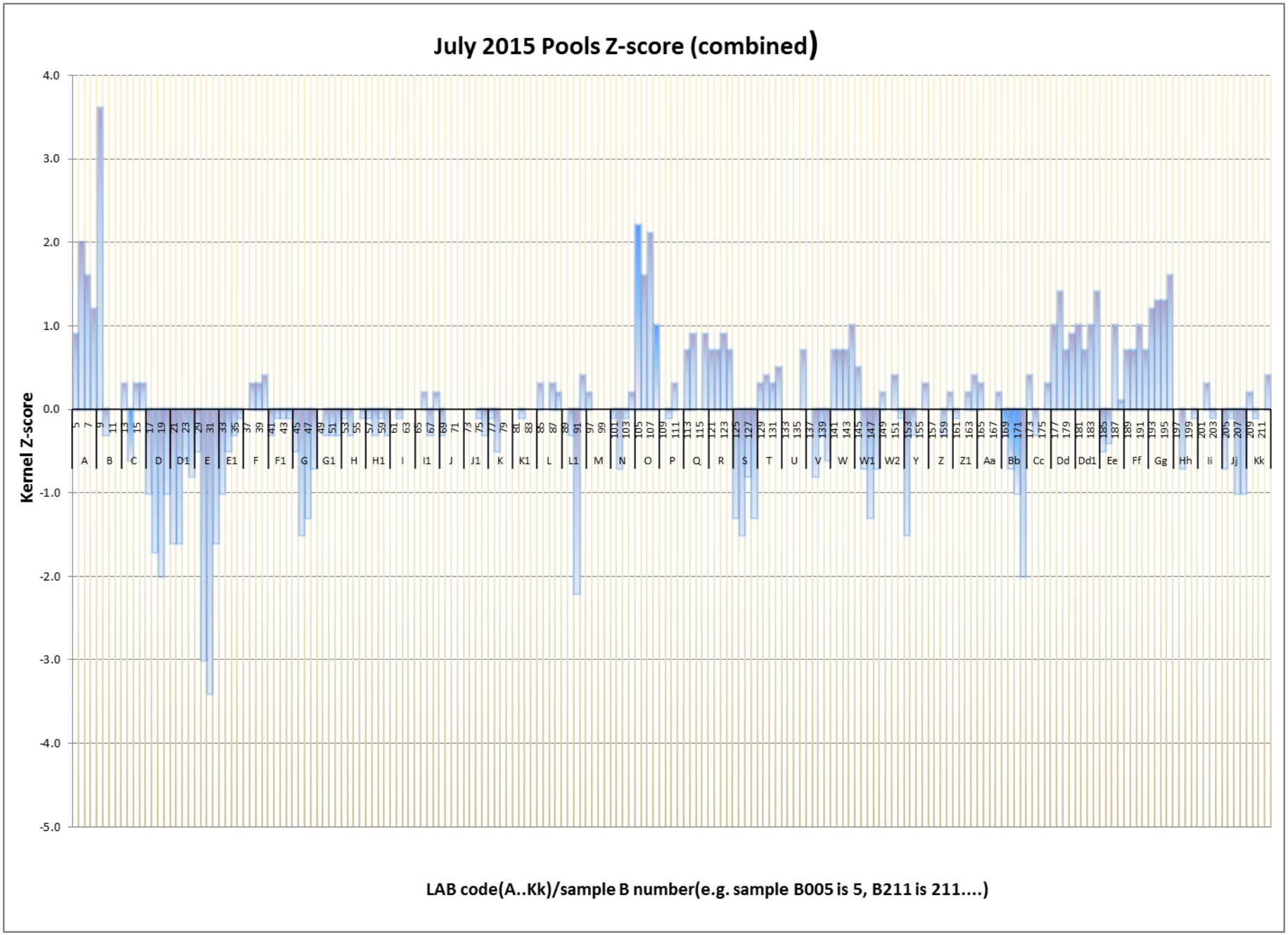


Figure 5


July 2015 Pools Z-score (combined)
Kernel Z-score
4.0
3.0
2.0
1.0
0.0
-1.0
-2.0
-3.0
-4.0
-5.0
A B C D D1 E E1 F F1 G G1 H H1 I I1 J J1 K K1 L L1 M N O P Q R S T U V W W1 W2 Y Z Z1 Aa Bb Cc Dd Dd1 Ee Ff Gg Hh Ii Jj Kk
LAB code(A..Kk)/sample B number(e.g. sample B005 is 5, B211 is 211....)