



Ron Chapman, MD, MPH  
Director & State Health Officer

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



EDMUND G. BROWN JR.  
Governor

January 22, 2015

TO: Forensic Alcohol Analysis Laboratories

SUBJECT: Assigned Values and Acceptable Ranges for December 2014 Proficiency Test in Forensic Alcohol Analysis

Attached is a summary of the descriptive statistics for the December 2014 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  $\pm 5\%$  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  $\pm 5\%$  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 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 this proficiency test were not normally distributed. 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.

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 number between the lowest number 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.

Sincerely,

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

## Statistical Data for November 2014 Proficiency Test in Forensic Alcohol Analysis

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

<u>Pool</u>	<u>Peer Group Mean</u>	<u>Tier #1</u>	<u>Tier #2</u>
#1	0.150	0.140 – 0.160	0.13 – 0.16
#2	0.261	0.245 – 0.277	0.24 – 0.28

**Table 2**      **Summary of Test Pool Data**

Parameter	POOL 1 (11174)		POOL 2 (11244)	
Pre-distribution Data	Target Value	0.15%	Target Value	0.26%
	True Value <sup>1</sup>	0.147	True Value <sup>2</sup>	0.257
	Standard Deviation <sup>2</sup>	0.0009	Standard Deviation <sup>2</sup>	0.0016
Descriptive statistics	Mean	0.151	Mean	0.261
	Adjusted Mean <sup>2</sup>	0.150	Adjusted Mean <sup>2</sup>	0.261
	Standard Error <sup>3</sup>	0.0005	Standard Error <sup>3</sup>	0.0009
	Median	0.151	Median	0.260
	Standard Deviation	0.0033	Standard Deviation	0.0063
	Minimum	0.147	Minimum	0.248
	Maximum	0.164	Maximum	0.285
	Count	49	Count	47
Descriptive statistics (box plot)	Q1 (25%)	0.149	Q1 (25%)	0.257
	Q3 (75%)	0.153	Q3 (75%)	0.265
	IQR	0.004	IQR	0.008
	Lower Fence	0.143	Lower Fence	0.245
	Upper Fence	0.159	Upper Fence	0.277
Histogram	Figure 1		Figure 2	
Normal distribution? <sup>4</sup>	NO		NO	
Box Plot <sup>5</sup>	Figure 3		Figure 3	
Robust mean, X*	0.151		0.261	
Robust standard deviation, $\sigma_{rob}$	0.0019		0.0048	
Fitness-for-purpose standard deviation, $\sigma_p$	0.0038		0.0063	
Consensus value ( $X_a$ ) determined as Mode ( $\mu_{1/2}$ ) of Gaussian Kernel distribution	0.1507		0.2602	
Uncertainty of the consensus value, $X_a$ , S.E. <sup>6</sup>	0.0004		0.0011	
$X_a \pm$ S.E.	$0.1507 \pm 0.0004$		$0.2602 \pm 0.0011$	
z-score	$z = \frac{X - X_a}{\sigma_p}$		$z = \frac{X - X_a}{\sigma_p}$	

<sup>1</sup> Based on CDPH's Headspace Gas Chromatographic Method

<sup>2</sup> Mean determined from participant data after the removal of outlier(s)

<sup>3</sup> Standard Error of the Mean

<sup>4</sup> Shapiro-Wilk test used.

<sup>5</sup> If the pool is found to be not normally distributed (Shapiro-Wilk test) the box plot data are provided for this pool

<sup>6</sup> Determined as Standard Error of Mode using bootstrap simulation technique with bandwidth of  $0.75 * \sigma_p$

Figure 1

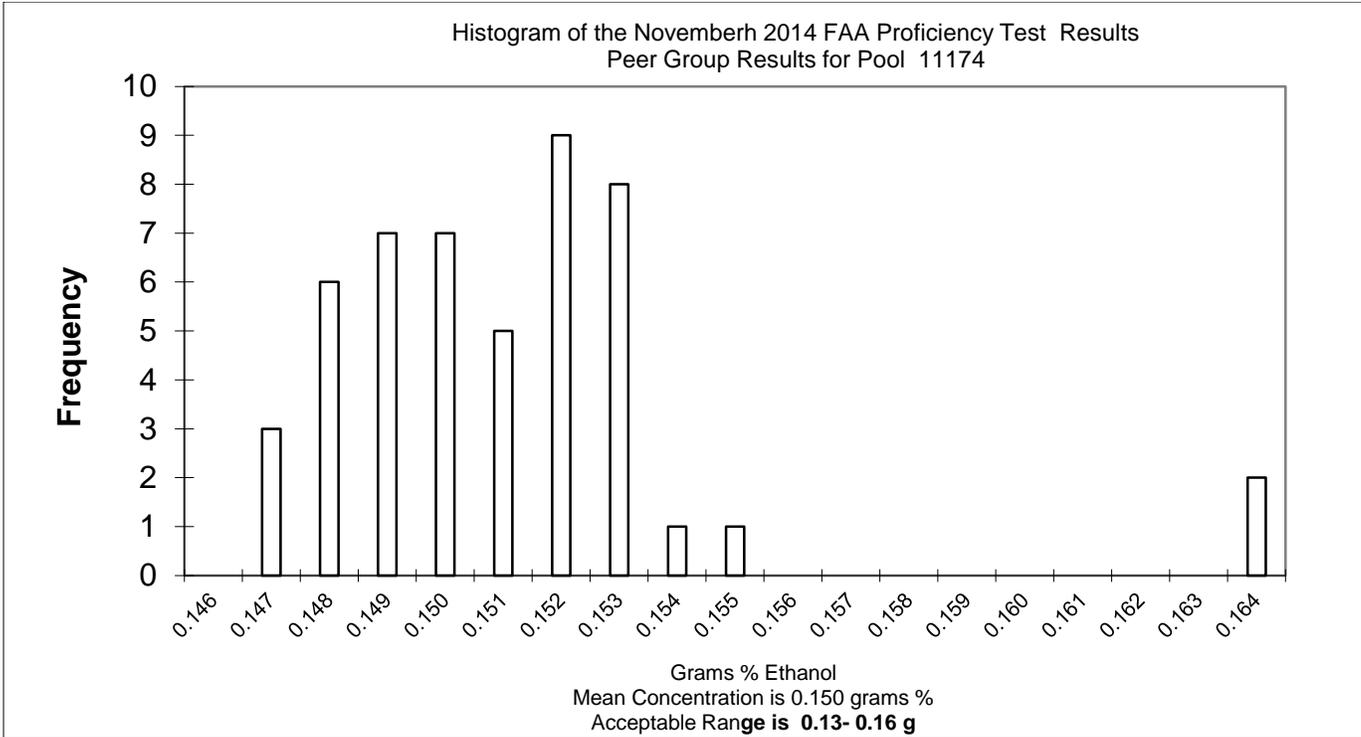


Figure 2

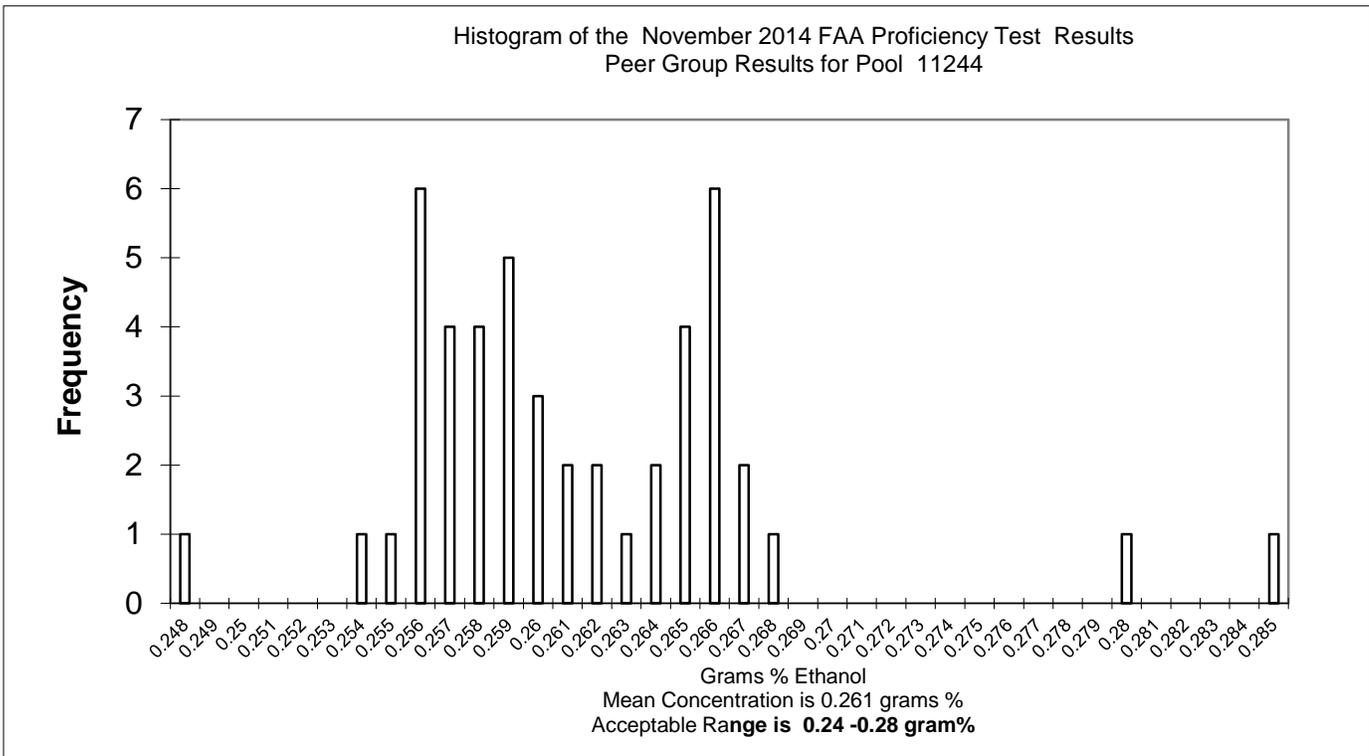
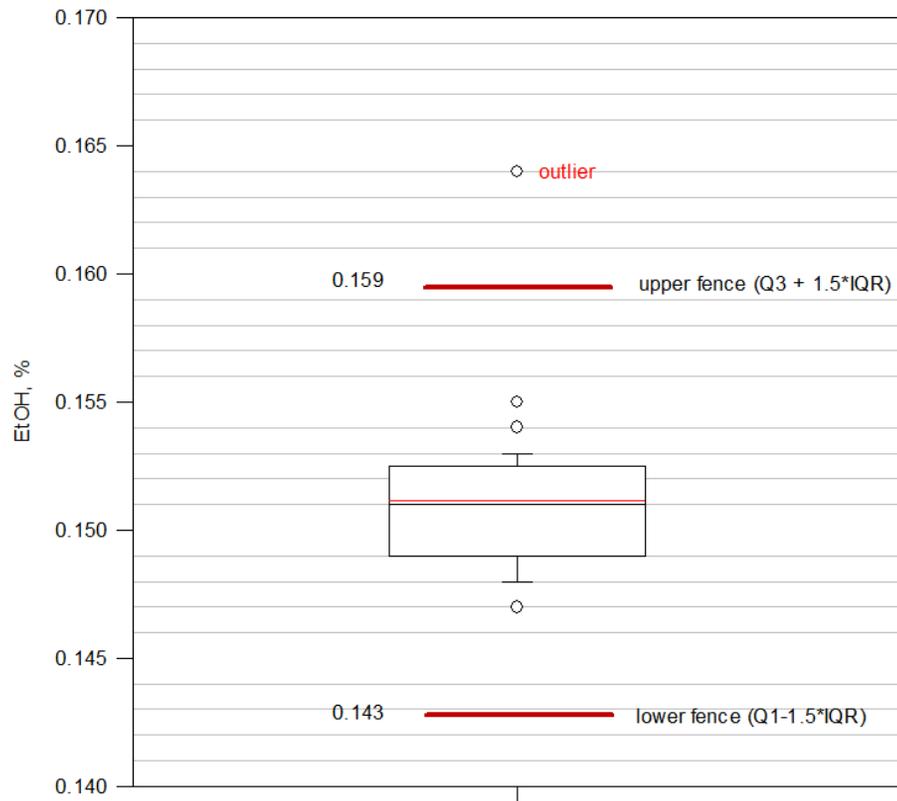


Figure 3 SigmaPlot analysis of pools 11174 & 11244

11174 Box Plot



11244 Box Plot

