

**SUMMARY OF THE
TNI CHEMISTRY EXPERT COMMITTEE MEETING**

FEBRUARY 3, 2015

The Committee met during the Forum on Laboratory Accreditation in Crystal City, VA, on Tuesday, February 3, 2015, at 8:00 am EST. Chair Richard Burrows led the meeting.

1 – Roll call

Richard Burrows, Test America (Lab)	Present
Francoise Chauvin, NYC DEP (Lab)	Absent
Brooke Connor (Other)	Present
Gale Warren, NYSDOH (Accreditation Body)	Absent
Colin Wright, Florida DEP (Lab)	Present
JD Gentry, ESC (Lab)	Present (by telephone)
Nancy Grams, Advanced Earth Technologists, Inc. (Other)	Present
Anand Mudambi, USEPA (Other)	Present
John Phillips, Ford Motor Co. (Other)	Present
Scott Siders, IL DEP (AB)	Absent
Gary Ward, OR DPH (AB)	Present
Ken Jackson, Program Administrator	Present

2 – Introductions

Richard welcomed the audience and the Committee members introduced themselves. The agenda for the day's meeting was presented, showing the morning session being devoted to Detection and Quantitation, and comments on the Calibration Interim Standard being presented and discussed during the afternoon session. Richard said, if time permitted later in the afternoon, John would present some data to support the proposed criterion for LOQ.

3 – Detection and Quantitation Working Draft Standard

Richard used the first 30 minutes of the session to present slides describing: why an LOD and LOQ are required; current definitions and their comparison with Currie's L_c , L_d , and L_q ; proposed new requirements; key components of the LOQ and LOD procedures; specific procedural details; and "big picture" items from the comments received on the Working Draft Standard (WDS).

The rest of the morning was devoted to the following "big picture" items from the WDS.

Should we make the EPA MDL mandatory for determining the TNI LOD (in the absence of method or regulatory constraints to the contrary)?

This would be the revised MDL procedure that the committee had developed and presented to EPA. It was about to be published in the Federal register. Richard said the committee would know by the time the standard was completed if EPA was to adopt the new MDL procedure. Mitzi Miller commented that the standard should be silent on this, because there are too many definitions of LOD already being required. Carl Kircher did not want to restrict a laboratory that chose to use a more rigorous approach to LOD, though he admitted most laboratories he assessed were using the EPA MDL. Nancy agreed with Carl, saying laboratories must be allowed to use other methods if they have a regulatory requirement to do so. Barbara Escobar felt there were too many choices for LOD. John Gumpper was against requiring any one procedure, because ISO 17025 says a laboratory must meet its customer requirements. However, Mike Miller emphasized the revised MDL is a huge contribution to get better science into the standard, so it should be required. An audience member pointed out it would be a precedent for a standard to mandate something that is not part of that standard. Gary said having a recommendation in the standard is a waste of time. He said laboratories have clients requiring them to report quantitative data down to the MDL and this needs fixing. Scott Hoatson suggested the laboratory SOP should explain the need to choose a specific procedure. John said most laboratories already do the EPA MDL, so he was in favor of requiring it.

Should we change the name of the LOD to the MDL?

Richard said this would remove the conflict with Department of Defense language and would allay a great deal of confusion. He added this would only make sense if the EPA MDL procedure is required. Mike Miller said it should be changed, because that would tie in with the EPA Office of water requirement. John, Anand, and Nancy agreed “LOD” should be got rid of to clear up the mis-use and confusion of that term. Nilda Cox also agreed, but emphasized MDL must then be defined. Richard summarized the general consensus on this question was “yes”.

Should 3x LOD be the minimum criterion for LOQ? What are the benefits of that and what are the alternatives? Pros and cons of alternatives.

Richard explained the following is expected from an LOQ: known precision; known accuracy; the ability to detect and report; freedom from false positives; and freedom from false negatives. He added that the current TNI LOQ fails to meet any of these criteria. Richard showed the alternative requirements the committee had considered for LOQ, and the pros and cons were considered briefly. John’s data were presented to show how the committee had arrived at the 3xLOD criterion.

Should LOD be required if not reporting below LOQ? Pros and Cons. If not, what alternative criterion would be best for LOQ? (I think a RSD criterion would be best).

Marlene Moore asked why anything would be reported below the LOQ, because of the uncertainty of LOQ and uncertainty is not reported. Richard added this would at least be a step towards uncertainty reporting.

What is the best language to convey when LOD/LOQ needs to be repeated in the event of a change in methodology or instrument sensitivity?

John said there are many cases of minor changes that do not affect these parameters, so perhaps it should just be specified when a laboratory does not have to do it. However, Richard argued the opposite may be the case. Nancy said the laboratory should have to do an appropriate check if it has changed something. Anand suggested listing the major ones that are cut and dried, and Gary agreed. Anand said a laboratory should take responsibility for what is a significant change and should be able to show this to the auditors. John Gumper offered suggested language for specific changes that require repeating the LOD/LOQ. Dan Hautman suggested the laboratory should be required to run a CCV periodically, but another commenter thought that would be too restrictive.

When is LOD/LOQ required?

This question did not generate discussion.

Do we need more clarification for gravimetric, titrimetric, colorimetric methods?

Steve Arms said he would send Richard a link to a list on the State of Florida website. He said it works for them.

At this point Richard summarized the discussion. He then presented and discussed a flow chart of the proposed LOD/LOQ procedure.

The meeting was adjourned for lunch at 12:00 pm

4 – Calibration Interim Standard

The afternoon session was opened at 1:00 pm, and the following comments on the Calibration Interim standard were presented and discussed.

1.7.1.1 (d) (ii) *“I believe that for tests with a large number of analytes the requirement to always remove an entire calibration level (all analytes) from the interior of a calibration curve is excessive. To me it makes more sense for you to remove all potentially effected analytes. To better clarify, an example that we have at our lab would be our standard 8260 list which contains 105 analytes. A calibration curve contains 11 points and each point is spiked with five standards. That is 55 individual spikes per curve. It does happen where only one of the five standards was spiked incorrectly or a spiking error occurred in one of the calibration points. In this case currently we would document that all compounds from standard "X" were removed due to spiking error. There is absolutely nothing wrong with all the other compounds in the calibration level, and when an incident like this occurs it is clearly identifiable. In my opinion the above stated practice with the proper documentation should not be excluded from the standard. We should not be required to remove data points that are clearly acceptable due to the failure of one spike mix or any other clearly identifiable cause for specific compound/compound group failure. An error of this nature is bound to happen from time to time due to the high number of analytes and individual spikes required.”*

Richard said this could happen, but rarely, and allowing it would be challenging. Paul Junio suggested having the laboratory justify why it does it. Richard countered by saying first the laboratory could get into an argument with its assessor, and second those who voted in favor would object to changing it.

Following the discussion, the comment was ruled Non-Persuasive because no Committee member moved for it to be persuasive.

1.7.1.1 (d). *“As an assessor, we must have written procedures when removal or replacement of points occur. This should not have been removed - having this allows assessors to clearly write this as a finding.”*

The committee had accidentally left out the part about the laboratory needing a written procedure. It was moved by John and seconded by Gary that the comment was persuasive and the language would be put back in. All Committee members were in favor.

1.7.1.1 (d). *“This section has been revised to remove the requirement for a laboratory to have a written procedure to address removal/replacement of calibration standards. Even though the standard is more clear regarding removal/replacement, a written procedure is still necessary so that the laboratory can define its actual procedures and address how and where the required documentation will take place. The procedure is also essential to staff training and consistent application of the use of removal/replacement.”*

This was the same comment as the preceding one, and was therefore Persuasive.

1.7.1.1 (e). *“It is our opinion that it is not necessary to increase the minimum number of calibration standards. Since the degrees of freedom portion of the table is deleted there is no need for Note (b). Delete Note (b).”*

Six more similar comments had been received, and all were considered together. Scott Hoatson suggested there could be language to allow the method to be followed if it specified fewer standards. However, it was noted the new drinking water methods will require that many standards. This change had been made to the standard as a result of several persuasive comments at the Voting Draft standard stage. In the absence of a motion to rule persuasive, the comment was Non-persuasive.

1.7.1.1 (j). *“It is our opinion that calculating a %RE or %RSE as a means of assessing the acceptability of calibration curves would not be cost effective and as an alternative, we suggest using similar language as used in UCMR3 Methods. As an example Method 524.3, rev 1.0, June 2009. This procedure is easily implemented offers EPA acceptance, and consistency with acceptance criteria across all laboratories.”*

This was grouped with two other identical comments. Richard said their suggested procedure is the same as RE. John Gumper wondered if the commenters wanted harmonization in the way it was expressed. In the absence of a motion to rule persuasive, the comment was Non-persuasive.

1.7.1.1 (j). *“If there is no specification for RSE or RSD in the method, then the RSE shall be specified in the laboratory SOP.”*

A majority of the inorganic methods do not have criteria for RSD or RSE. Without any guidance being provided by either the published method or regulatory guidelines for either RSD or RSE, how can an assessor consistently audit laboratories to the calibration section of the standard. Each lab could have

very different acceptance criteria for either parameter (RSE or RSD) and as long as the laboratory meets its defined requirement, an assessor could not cite for inadequacy of one lab's practice against another. Therefore, the consistency of the evaluation process is compromised due to incomplete guidance from the TNI Standard.

"If no criterion for the lowest calibration level is specified in the method, the criterion and the procedure for deriving the criterion shall be specified in the laboratory SOP."

As stated in above comment how can an assessor consistently evaluate laboratories if each laboratory chooses their own criteria instead of using criteria with guidance from regulatory agencies or the TNI Standard."

Richard said it would be too difficult to have specific criteria, because different methods have different criteria. He added the same argument could be made for LCS which has no specific criteria. In the absence of a motion to rule persuasive, the comment was Non-persuasive.

1.7.1.1 (j). *"Criteria for %RE and %RSE are not included in any of the published methods historically used for environmental analysis. I understand that it has been included in the updated version of EPA 8000; however not all states have moved to the updated versions of the SW-846 methods. This could cause a lapse in assessor understanding and application of the concepts given.*

Drift and %Difference are normally listed for daily calibration verification use, but appear to be much the same as the equations given in the proposed standard for %RE. Why confuse the nomenclature by introducing different terminology for these concepts when they are clear in the published methods. In addition, since no criteria is given for acceptance of these concepts in the draft standard, using the acceptance criteria in the methods along with the terminology in the methods seems more appropriate for the evaluation of the mid-level standard. The inclusion of the low level standard is a variation from method protocols and no supporting data is available that demonstrates that assessing the low level standard to the same criteria as a mid-level standard would be practical in real world applications.

With no acceptance criteria given in the standard, you are allowing each laboratory to determine a criteria for evaluation would provide inconsistency of data between laboratories and would force assessors to develop independent acceptance criteria for each regulatory program that would then remove consistency in the evaluation process that was the initial and ultimate goal of the TNI program."

Richard said "Drift" and "% Difference" could be used and asked if anyone thought it would be better than RE. Anand said "Drift" is only for CCV. The standard could say "%RE (or %Difference)", but that could get confusing. John and Anand suggested modifying the definition to say that RE and drift are mathematically the same, but it was agreed not to change the standard. On the last point Richard said there is clearly a demonstrable problem with existing calibration criteria.

1.7.1.1 (j). *"Remove the requirement to calculate %RE or %RSE and allow this to be driven by the referenced methods (i.e. 8260C, 8270D)."*

In the absence of a motion to rule persuasive, the comment was Non-persuasive.

1.7.1.1 (j). *“Most metals and general chemistry methods do not require an additional RE or RSE for the mid point and low standard in the curve. This will cause a lot of either manual calculation because most data stations doing regression, do not have this built in as an addition. While the calculations are easy, it will require more manual or spread sheet calculations for many organizations. Can the committee clarify what this will buy with respect to quality. If one really wants to evaluate linearity then the low and high should be used not a mid and low point.”*

Richard said no problems have been seen with the high point, and the committee wanted to make it easier because the mid-point is already being used by most laboratories. It was also noted this would be a new requirement that could not be considered at the Interim Standard stage. In the absence of a motion to rule persuasive, the comment was Non-persuasive.

1.7.1.1 (j). *“This is a new requirement with no criteria. While laboratories are accustomed to meeting similar requirements for Organic analytical methods, this criteria is new to most inorganic/metals methods. A requirement with no criteria provides no benefit to the standard.*

Historical practice for calibration curve evaluation relates linearity to a correlation coefficient of >0.995 or a coefficient of determination of >0.99 . While this alone does not represent true linearity, adding RE/RSE with no historical basis adds little value. It is possible to have good linearity according to r or r^2 and have a bad RE/RSE. But with no historical basis or criteria, who determines what a bad RE/RSE really is? In addition one can have a “bad” RE/RSE at one portion of the curve and a good RE/RSE at a different portion of the curve. Who determines what is acceptable? Is it okay to have a value at the low end of the curve at 40% and a value of 10% at the mid-point of the curve? This would indicate tilting of the curve, but with no acceptance criteria, what determines acceptable and unacceptable?

Why pick low and mid points? What is the basis? Seems as though low and high points would be more representative. Curves typically tilt at the mid-point and therefore tend to be more accurate at the mid-point due to the nature of linear “curve-fit”.

This point had already been made and ruled non-persuasive. In the absence of further discussion the comment was ruled Non-Persuasive.

1.7.1.1 (m). *“The intent of the statement that “...all initial calibrations shall be verified with a standard obtained from a second independently prepared lot or from a second manufacturer” needs clarification. To avoid the possible interpretation by an auditor that someone other than the analyst who prepared the calibration standards must prepare the ICV, I suggest that the clause should be revised to read “...all initial calibrations shall be verified with a standard obtained from a second manufacturer or a separate lot prepared independently by the manufacturer.”*

Carl Kircher commented he assessed a laboratory that had two sources from the same manufacturer with different lot numbers but the same date. An audience member questioned what is meant by a “lot”. It was moved Anand and seconded by Brooke to rule the comment Persuasive and to adopt the language suggested by the commenter. All were in favor.

1.7.1.1 (m). *“ISO17025 already requires initial calibration standards that are traceable to a national standard. There is added cost to purchasing these standards. Adding it as an ICV requirement in the TNI Standard is not necessary since the calibration standards are already traceable and the ICV is checked against those. It would be a redundant check with increased cost to the laboratories.”*

It was suggested to remove “traceability to a national standard” and say initial calibration standards are traceable. Jeff Lowry said the initial calibration just needs to be to a CRM. He offered to supply proposed language for the committee to consider.

1.7.2 (c) and (d). *“We often receive client requirements in QAPPs that stipulate the concentration range for the CCV that differ from the $\leq 1/2$ the high cal std requirement stated in c). There are also some methods and client requirements for a low and high level LCS. Would d) iii. allow using these LCS standards for the CCV when the processes are the same for the two standards? Is it acceptable to use project specific criteria for the CCV levels if it differs from the range stipulated in the TNI standard?”*

Richard said this has been in the standard for quite a long time. It already says an LCS can be used for a CCV. In the absence of a motion to rule persuasive, the comment was Non-persuasive.

1.7.2 (d). *“This whole section is confusing to me. Any clause that contains the word "except" tends to invite further thought or scrutiny. If the Committee would please consider and use the following revisions, this section would read much more clearly (at least to me) on what the laboratory is required to do (changes and additions underlined):*

d) Instrument continuing calibration verification shall be performed at the beginning and end of each analytical batch, normally with the same standard source as used for the initial calibration, and at the frequency defined in the method except:

i. if an internal standard calibration procedure is used, calibration verification needs only to be performed at the beginning of each analytical batch, and at the frequency defined in the method;

ii. a calibration verification with a second source standard that passes the continuing calibration verification criteria may be used in place of a continuing calibration verification made with the same initial calibration standard source.

iii. a laboratory control sample (LCS) may be used in place of a continuing calibration verification (but not as a replacement for a failing CCV) for methods where the calibration goes through the same process as the LCS and the LCS results pass the continuing calibration verification acceptance criteria.”

Richard said the committee would look at the proposed language and then decide.

1.7.2 (f). *“1. It is not TNI's right or obligation to determine what data is "valid" for individual laboratory, industry, or regulatory use. This section implies that if a continuing calibration verification fails during an analysis the data is not reported unless it meets the special conditions in 1.7.2.f.iii.a and 1.7.2.f.iii.b. This section should be modified to allow for reporting the data with qualification unless the*

special conditions in these 2 sections are met. Let the end-user determine if the data is "valid" for their purposes.

2. Most instruments are automated in today's society and these sections do not give options for the course of action to take once the analysts determines what occurred during an automated analytical run where the samples and other QC have already been run before the analyst was aware of the failing continuing calibration verification. 1.7.2.f.i refers to running a second calibration verification sample immediately, but that is not realistic for automated analyses. And especially for a "missed autosampler injection", at least one sample would be in the process of being analyzed prior to the analyst discovering that an injection was missed."

On the first point, Richard said the language was put in at the request of LASEC, so the committee would be hesitant to change it. However, they would ask LASEC if they have a problem with the change. On the second point, Richard said if you let the run go on automatically, you will have missed the CCV. In the absence of a motion to rule persuasive, the comment was Non-persuasive.

1.7.2 (f) (ii). *"The allowance to run two passing CCVs after a failed CCV and then continuing with analysis has been removed. This should be added back in."*

Richard said there were a lot of negative votes on that allowance. In the absence of a motion to rule persuasive, the comment was Non-persuasive.

1.7.2 (f) (iii). *"The current language reads: "if samples are analyzed using a system on which the calibration has not been verified, the results shall be qualified. Data associated with an unacceptable calibration verification may be reported under the following special conditions, unless "prohibited by the client, a regulatory program or regulation". The language needs to be more clear that in such circumstances, if data are to be reported they must also be qualified.*

Proposed revised language:

"if samples are analyzed using a system on which the calibration has not been verified, the results shall be qualified. Data associated with an unacceptable calibration verification may be reported with qualifiers under the following special conditions, unless prohibited by the client, a regulatory program or regulation. ""

Richard said the intent was that there was no need to qualify. The committee would look at the language and attempt to make it clearer. They would then ask the NELAP Accreditation Council for an opinion.

This completed the discussion of comments. An additional hour of the afternoon session was devoted to John's presentation of additional data to support the LOQ criteria.

5 – Adjournment

The meeting was adjourned at 4:30 pm EST.