JCTLM
Accurate results for patient care

Database Newsletter

https://www.jctlmdb.org/

Special Report | May 2025

# Successes and challenges delivering higher order references for laboratory medicine

**Graham Jones[1], Vincent Delatour[2], Tony Badrick[3]**
**[1]SydPath, St Vincent's Hospital, Sydney, NSW, Australia**
**[2]Laboratoire national de métrologie et d'essais, Paris, France**
**[3]RCPA Quality Assurance Programs, St Leonards, Sydney, Australia**

Laboratory medicine commonly involves measuring the concentrations of substances in biological fluids or tissues. The substances include elements, proteins, metabolites, fats, nucleic acids, cells, drugs and toxins, and the fluids include plasma, cells, CSF and urine. Laboratory testing is part of medical management, involving a series of steps in the process. The process starts with the doctor raising a clinical question, typically either to assist with a diagnosis or to monitor clinical progress, then selection of a test known to help answer the question. A sample is then collected, transported to the laboratory, prepared and analyzed. The result is then reported with the required supporting information and returned to the doctor for interpretation, discussion with the patient, and planning the next step in management. Each step in these simple sentences covers a wide range of basic and applied science. Each must be performed correctly for the optimal contribution of pathology testing to the patient's clinical outcome.

For numerical results, all interpretation is made by comparison. There are four commonly used comparators. These are population reference intervals, clinical decision points, previous results from the same patient, and the experience of the treating doctor. Each of these comparators has been derived from other measurements of the same measurand. For reference intervals, it is the method used to determine the intervals; for decision points, it is the methods used in the clinical studies where the evidence for the decision point was derived; and for clinical experience, it is the method(s) used where the doctors have had clinical experience.

In an ideal world, all results for a measurand will be "the same, every time, every place". This means every laboratory where a previous result for the patient may have been generated and where the doctors may have worked. It also means every method used in the relevant clinical research may be from almost any country in the world and may have been very recent or many years ago.

To achieve this ideal, results from all methods in both routine and research laboratories must be metrologically traceable to the same higher-order references, with low enough measurement uncertainty, and associated with the same measurand. Put in plain language – all results from all laboratories must be unbiased relative to each other.

**The current reality**

There are a range of laboratory tests where good metrological traceability has been achieved. The most notable examples are HbA1c and serum creatinine, where ongoing international activities have led to changes in traceability and uncertainty by manufacturers, delivering (largely) globally standardized results. (1,2) There are other tests where the major manufacturers deliver results with sufficiently low bias that common reference intervals may be shared or indeed that patients may be safely monitored when moving from one laboratory to another across an entire country. An Australian study showed acceptable "total country CVA" for fifteen common biochemistry measurands and acceptable between method bias for nineteen such measurands. (3)

The reality, however, is often far from the ideals expressed above. Different laboratories and different methods often produce variable results for the same test. If this is not recognized, wrong clinical decisions can be made. This may be due to wrongly applying research data, making errors in monitoring a patient or using clinical experience based on different values for results.

When standardization has not been achieved, or even when it is achieved, but older data remains in use, results from unstandardized assays need to be

recognized as such and considered separately. Users need to be aware of method/assay differences and use this information to minimize the risk of error. This requires agreed, accessible terminology so that data can be validly compared when it is valid, but inappropriate comparisons between biased results are not made. If between-method differences are stable and known, mathematical correction may assist with data comparison.

A range of responses to these problems can be made at different levels in the process. Clearly aligning results with correct metrological traceability to internationally agreed higher-order references is the preferred option. Given the global nature of information and the reach of the manufacturers, a global solution is needed.

Even when results are comparable, work must be done to fully take advantage of this status. It needs to be known that results are equivalent and supported by unified units of measure, reference intervals and clinical interpretation, for example, against clinical decision points. For doctors and patients to get the benefits of traceable assays, these reporting issues must also be harmonized.

**The role of the JCTLM**

The JCTLM, now in its 23rd year, as well as supporting traceability through education and promotion, provides lists of Reference Materials (RM), Reference Measurement Procedures (RMP) and Reference Measurement Services that are validated fit for use as higher-order references for laboratory medicine by application of relevant ISO standards. (4)

At present, the JCTLM database lists 290 RMs for approximately 166 measurands, 146 for pure materials and 49 for matrix-matched materials, including some with a range of concentrations supplied. (5) 238 RMPs are listed for approximately 107 analytes in a range of matrices, making 119 measurands. There are 293 listed RMSs covering 42 analytes and 56 measurands. Of these measurands, over 22 have five or more laboratories offering RMS testing.

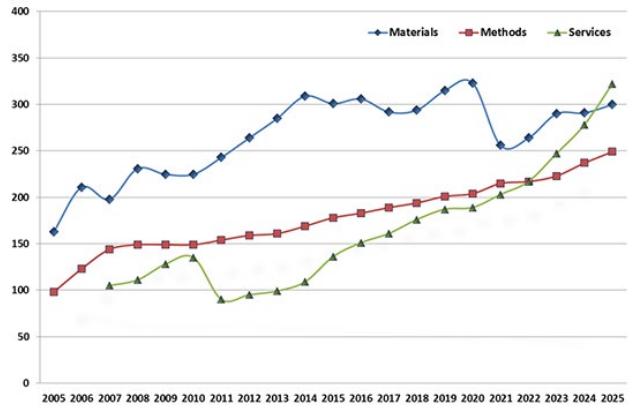


Figure 1. Changes in reference materials (RM), reference measurement procedures (RMP), and reference measurement services that have been listed on the JCTLM website over time.

**The challenges**

Although RMs, RMPs and RMSs exist for a large fraction of the most frequently measured analytes, several thousand clinical analytes exist, and higher-order reference materials and RMPs are not available for many important measurands. Below are considerations regarding possible causes and ongoing initiatives to overcome the observed challenges.

The lack of RMPs and RMSs is mainly due to i) the time and resources needed to develop and validate high-accuracy methods with fit-for-purpose measurement uncertainties, limits of quantification and selectivity, ii) the cost to operate such labour-intensive methods with laborious sample preparation and high hand-on time and iii) the lack of primary reference materials (RMs) with well-characterized purity that are needed to calibrate RMPs and establish the traceability of the results to the SI units. (6)

A major contributor to the lack of matrix-matched RMs is the difficulty and cost of properly assessing their commutability. (7) This property describes an RM's ability to properly mimic the behaviour of actual patient samples, an absolute requirement for producing valid traceability chains. The paramount importance of commutability has only been recognized recently. The causes of materials' non-commutability remain largely unknown. (7) An improved understanding of key common causes of limiting materials' commutability would be extremely valuable to materials producers. One response to this is the COMET project, which aims to identify manufacturing processes enabling consistent production of calibration and quality control materials of high commutability levels. To achieve this objective, the commutability of various materials consisting of different formats (for example, frozen, lyophilized, spiked or not with exogenous compounds such as preservatives and/or analyte of interest) will be evaluated and compared to identify key common causes affecting commutability. Although commutability assessment frameworks have recently been established by the IFCC working group on Commutability (8), evaluating commutability remains cumbersome (7). As demonstrating commutability constitutes a major component (as much as 70 %) of resources required for CRM production, CRM producers may delay or cancel the production of replacement batches of an existing CRM running out of stock. Additionally, allocating substantial resources to the commutability assessment of existing CRMs' replacement batches and budget and resource constraints may lead to delays or cancellations of projects to develop new CRMs for other measures. The IFCC Working Group on Commutability in Metrological Traceability recently published recommendations describing a procedure facilitating

the commutability assessment of replacement batches of existing secondary CRMs, in which commutability was successfully demonstrated in a full commutability assessment. (9) These recommendations encourage CRM producers to produce replacement batches and nominate those for listing in the JCTLM database. Improved availability of replacement batches is expected to help ensure the ongoing availability of critical CRMs necessary to sustain a viable calibration infrastructure over the long term to benefit all stakeholders, including medical laboratories, patients, healthcare providers, CRM producers and assay manufacturers.

Commutability testing is costly and cumbersome, so the COMET project will explore additional strategies to make commutability assessment more straightforward and cost-effective. These include i) simultaneous commutability assessment of a large number of CRMs and EQA materials for a panel of measurands, ii) use of commutability panels consisting of frozen pools in which commutability was qualified in a first study, iii) use of automated and/or multiplexed RMPs for high-throughput analysis of study materials, iv) automated statistical approaches to analyse commutability data much faster and, v) mutualizing the resources and capabilities of a coordinated network of reference laboratories sharing the work to assign reference method target values to all study materials jointly.

**Measuring Success – EQA/PT**

Measurement traceability is "the property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties."(10) Traceability helps assure the accuracy and comparability of a patient's lab results. "Establishing metrological traceability satisfies the basic requirements of evidence-based laboratory medicine. Thus, it improves patient care, disease control, and prevention, saving money by allowing the pooling of clinical trial data rather than repeating studies. However, because of variations in manufacturing processes and reference measurement procedures, there will always be a need to identify when these can cause a traceable assay to deviate from specifications. The best tool to detect any such deviation is routine External Quality Assurance (EQA), also referred to as Proficiency Testing (PT).

EQA schemes using commutable sample materials provide information from different laboratories about how different IVD MDs compare. If an EQA uses higher-order reference method value assignment, information on the success of the metrological traceability of end-user IVD MD results is provided by EQA data. EQA monitors the diagnostic equivalence of the end-user measurement results and can provide information to assess the effectiveness of metrological traceability. (11)

EQA organizations often use stabilized EQA samples due to their advantages despite the critical and frequent lack of commutability (14), but providing commutable samples is challenging. (12,13)

While EQA primarily aims to assess the diagnostic equivalence of measurement results, EQA organizers should also strive to develop and maintain EQA schemes for assessing metrological traceability since traceability is a cornerstone in achieving equivalence of measurement results and minimizing patient risks. (11) The European Organization for External Quality Assurance Providers in Laboratory Medicine (EQALM) and the "International Consortium for Harmonization of Clinical Laboratory Results" (ICHCLR) are collaborating to combine the results from EQA providers that use commutable samples to assess the harmonization of measurement procedure results from different IVD MDs used in medical laboratories (https://www.eqalm.org/cooperations/halma).

**Conclusions**

Metrologcally traceable results are a vital component of optimal health care delivery. Despite some successes, there remains much work to be done. This requires active participation from RM producers, RMP developers, RMS laboratories, manufacturers, routine laboratories, clinical and laboratory professional organizations, accreditation bodies, clinical researchers, IT professionals as well as overarching bodies such as the JCTLM.

**REFERENCES**


1. Penttilä I., Penttilä K., Holm P., et al. Methods, units and quality requirements for the analysis of haemoglobin A 1c in diabetes mellitus. *World J Methodol.* 2016 **6**(2):133-142.
2. Piéroni L., Delanaye P., Boutten A., et al. A multicentric evaluation of IDMS-traceable creatinine enzymatic assays. *Clin. Chim. Acta* 2011 **412**(23–24):2070-2075.
3. Farrell C.J.L., Jones G.R.D., Sikaris K.A., et al. Sharing reference intervals and monitoring patients across laboratories - findings from a likely commutable external quality assurance program. *Clin. Chem. Lab. Med.* 2024 **62**(10):2037-2047.
4. Jones G.R.D., Jackson C. The Joint Committee for Traceability in Laboratory Medicine (JCTLM) - its history and operation. *Clin. Chim. Acta* 2016 **453**: 86-94.

5. JCTLM. Joint Committee of Traceability in Laboratory Medicine. Database of higher-order reference materials, measurement methods/procedures and services.
6. Josephs R.D., Martos G., Li M., et al. Establishment of measurement traceability for peptide and protein quantification through rigorous purity assessment - A review. *Metrologia* 2019 **56**(4)
7. Miller W.G., Greenberg N., Budd J., et al. The evolving role of commutability in metrological traceability. *Clin. Chim. Acta* 2021 **514**:84-89.
8. Miller W.G., Schimmel H., Rej R., et al. IFCC working group recommendations for assessing commutability part 1: General experimental design. *Clin. Chem*. 2018 **64**(3):447-454.
9. Deprez L., Johansen J. V., Keller T., et al. Recommendations for assessing commutability of a replacement batch of a secondary calibrator certified reference material. *Clin. Chim. Acta* 2025 **567**.
10. BIPM. JCGM 200:2008 International vocabulary of metrology - Basic and general concepts and associated terms (VIM). BIPM [Internet]. 2008;(third edition):90. Available from: http://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2008.pdf
11. Theodorsson E., Meijer P., Badrick T., External quality assurance in the era of standardization. *Clin. Chim. Acta* 2024 **557**.
12. Van Der Hagen E.A.E., Weykamp C., Sandberg S., et al., Feasibility for aggregation of commutable external quality assessment results to evaluate metrological traceability and agreement among results. *Clin. Chem. Lab. Med*. 2020 **59**(1):117-125.
13. Stavelin A., Petersen P.H., Sølvik U., et al., External quality assessment of point-of-care methods: Model for combined assessment of method bias and single-participant performance by the use of native patient samples and noncommutable control materials. *Clin Chem*. 2013 **59**(2):363-371.