Historical developments

Authored by Elvar Theodorsson (elvar.theodorsson@liu.se, +46736209471) and reviewed by the JCTLM Traceability Education & Promotion Working group and by other beneficiaries of the JCTLM

BIPM in Paris in 2022

This is a “living document”, version 2022-03-23. Suggestions for improvements are gratefully received.

JCTLM Secretariat

Bureau International des Poids et Mesures
Pavillon de Breteuil
92312 Sèvres
Cedex France
Tel: +33 1 45 07 70 70
Fax: +33 1 45 34 20 21
Email: jctlm@bipm.org

Historical developments underpinning traceability in Laboratory Medicine

The theory and practice of measurements developed slowly during the last three millennia, but a rapid development phase started with the French revolution in 1789 with emphasis on standardization of measurement units for length and weight. Physicists have since then built solid theoretical and practical foundations for metrology, most concretely formulated, and implemented in the International System of Units (SI) (1-4) launched in 1960.

The humanistic sciences developed a theoretically separate "science of measurement" (examinations) since the late 1930s (5-7). It is, however, increasingly apparent that the practice of clinical medicine needs both *measurements*, for example measurements of the concentrations of biomarkers and *examinations* when diagnosing for example infections and cancer from images and in obtaining understanding the effects of diagnostic- and therapeutic procedures on the quality of life of patients and groups of patients (8-15).

From the outset, metrology focused on the measurement of very concrete and visible characteristics, e.g., the quantities length and mass. Quantities such as voltage and strength of light, which joined later, are not visible with the human eye alone and their quantification (making the quantity measurable or visible to the naked eye) is therefore dependent on the use of measuring systems. When chemistry entered the metrological community in the 1960s (16) leading to the introduction of the SI unit of mole in chemistry in 1971, practically all measurements demanded the use of increasingly complex measuring systems making invisible to the naked human eye the relation between what can be seen and directly measured and demanding the use of increasingly theoretical measurement methods and systems.

The *General Conference on Weights and Measures* (*CGPM*) established in 1875 through the Metre Convention, is the highest authority of the International Bureau of Weights and Measures (www.bipm.org). Initially it dealt only with the kilogram and the metre, but its scope was in 1921 extended to accommodate all physical measurements and all aspects of the metric system.

The *International Committee for Weights and Measures* (*CIPM*) works under the auspices of the CGPM to maintain and promote the International System of Units (SI) which was established in 1960.

The *Consultative Committee for Amount of Substance* (*CCQM*) working under the auspices of the CIPM is responsible for the metrology in Chemistry and Biology. It was

established in 1993 to establish and maintain world-wide traceability for measurements in the highest metrological levels of chemistry and biology.

The *Co-operation on International Traceability in Analytical Chemistry* (*CITAC*) was also started in 1993 to establish fitness for purpose in chemical measurement, to improve comparability of measurements made in different laboratories in different countries, and to achieve the first and the second by establishing traceability to recognized reference materials/methods and where appropriate to SI (17).

1993 – the year of metrology and traceability in Laboratory Medicine

All this means that the year 1993 represents an important year in the metrology of Analytical Chemistry and in Laboratory Medicine. It is also the year when the metrology of Analytical Chemistry and in Laboratory Medicine became officially recognized (18-23). The term "*traceability*" was also formally defined in 1993 in the International Vocabulary of General and Basic Terms in Metrology (*VIM*). The same year, the Cooperation on International Traceability in Analytical Chemistry (*CITAC*) was established to encourage the realization of traceability in analytical chemistry. With the advent in 1998 of the European Union Directive on in vitro diagnostic systems (IVD) (24), a regulatory framework establishing traceability of measurements performed with in vitro diagnostic devices became mandatory in the EU and has since influenced similar regulatory authorities in Laboratory Medicine globally.

Notably, already two years later – in 1995 - at the 20th General Conference on Weights and Measures, resolution 7 stated that "formidable difficulties exist in establishing international traceability for measurements in Chemistry" (25, 26). The full realization of differences in establishing traceability between physics and chemistry had become fully and painfully apparent. The CGPM therefore called for broad international collaborations to establish and maintain world-wide traceability in chemistry. Several organizations around the globe including national metrology institutes headed the call, amongst them the *International Organization for Standardization* (*ISO*), founded in 1947, the *Clinical and Laboratory Standards Institute* (*CLSI,* originally formed as the National Committee for Clinical Laboratory Standards (NCCLS)) in the U.S.A., founded in 1967 and *Eurachem* founded in 1989 (27, 28).

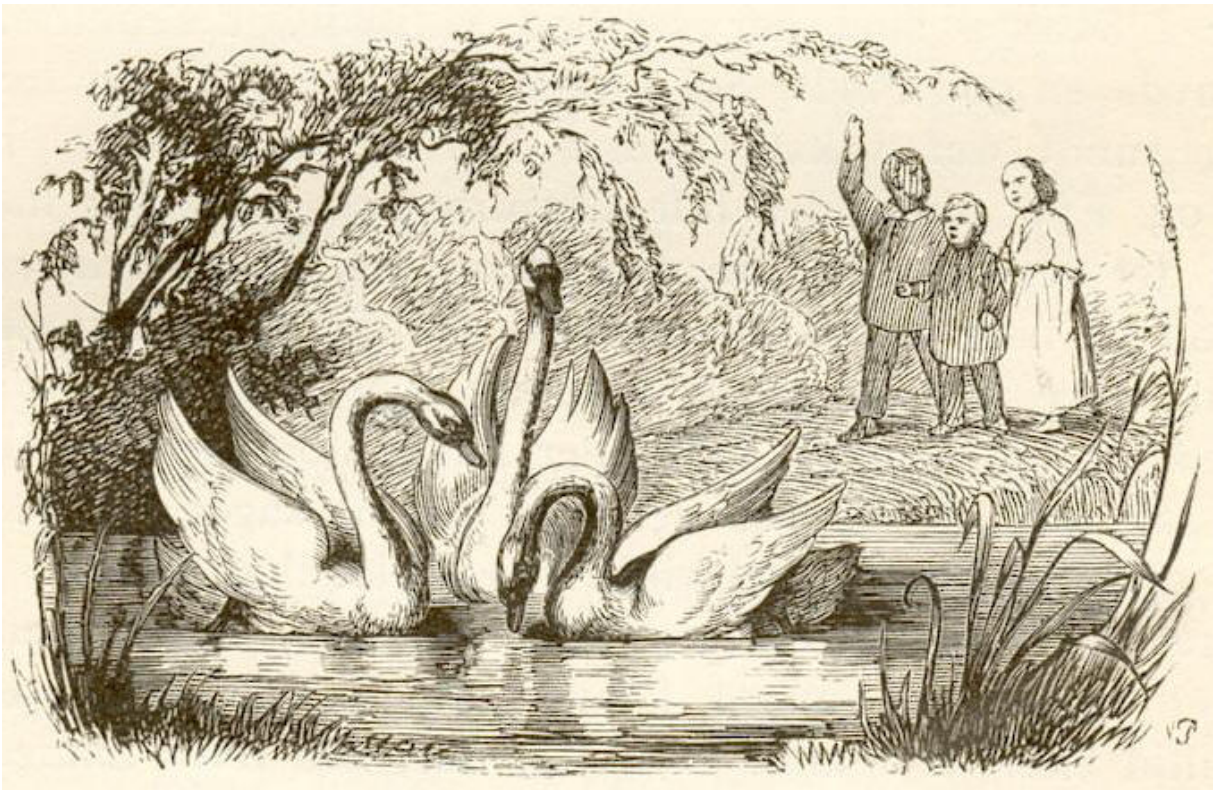

Figure X: The "Ugly Duckling" (Danish: Den grimme ælling) is a literary fairy tale by Danish poet and author Hans Christian Andersen (1805–1875). The lesson being that a strange and conceivably ugly bird on the pond may in time turn into a Swan.

Analytical Chemistry and Laboratory Medicine are amongst the "ugly ducklings" of international metrology due to the presence of matrix effects and since traceability to SI is only feasibly for a minority of their measurands. Alternative traceability hierarchies leading to standardization were accepted as ISO standards in Laboratory Medicine in 2020 (29, 30). It may be a "small step for man" but may prove to become "giant leap" for Laboratory Medicine since the current primary focus is now on broad practical *equivalence* of measurement results rather than on the few measurands optimally traceable to SI. The formal meaning of "equivalence" in this context is "agreement of measured values among different measurement systems intended to measure the same measurand, where the differences in measured values on the same human samples do not affect clinical interpretation" (30). According to an old Italian saying cited by Voltaire in 1770 "Perfect is the enemy of good". It is consequently of primary importance to develop and apply methods for improving equivalence of measurement results in Laboratory Medicine for example through improved traceability hierarchies even though these methods yet do not reach the rigor of traceability to the SI.

However, with time, as medical and scientific knowledge and measurement technologies improve, we will hopefully be able to identify which epitopes of

macromolecules, sequences of genes are the best objective markers of disease and most appropriate for follow-up of treatment effects. This opens the optimal doors to proper SI traceability. The rewards are better clinical performance including equivalence, traceability, and less measurement- and diagnostic uncertainty. Amongst the best examples of such a development is the standardization of the measurements of glycated hemoglobin (HbA1c) (31-38). The work took about 20 years, but the benefits for the patients are unquestionable.

The unfortunate schism that initially played out in the United Kingdom in the 1930s (39-41) regarding *the nature of measurements* between the natural sciences and the humanities still lingers on and hampers developments in metrology. The "Ferguson committee" (Ferguson, A. Myers, R. J. Bartlett, H. Banister, F. C. Bartlett, W. M. Brown, Campbell, N.R. Drever, J. Guild, R. A. Houstoun, J. C. Irwin, Kaye, G.W.C. Philpott, S.J.F. Richardson, L.F. Shaxby, J.H. Smith, T. Thouless, R.H. Tucker, W.S.) was appointed in 1932 to "Consider and Report upon the Possibility of Quantitative Estimates of Sensory Events". The members were unable to reach consensus and made in the end their separate sub-reports. Stevens theory on "scales of measurement" (42, 43) catalyzed the development of a separate science of measurement in e.g. psychology, sociology and pedagogics (44-51).

It is probable that the pure white swans of physics at the helm of the metrology family will soon find themselves as an important principle- and pace-setting minority group as measurement-oriented scientists in the humanistic sciences are welcomed to the flock as "ugly ducklings" together with Analytical Chemistry and Laboratory Medicine. There is a rather small step in introducing sciences relying on analogous mental constructs including Psychology, Sociology, Pedagogics etc. to the Metrology flock (8, 9, 13-15) when Analytical Chemistry and Laboratory Medicine have been there since 1993.

The economic volume of the global in vitro diagnostic (IVD) market in 2019 was almost 70 billion US dollars and has increased substantially since then due to the Covid-19 pandemic. Laboratory Medicine which constitutes the knowledge base of IVD activities, consists of several disciplines with different historical roots and an arsenal of methodologies. Pathology originally studied organs from the dead and subsequently samples from living tissues under the microscope by examinations. Currently, pathology also uses measurements based on immunochemical and molecular biology techniques, which are very much a part of the typical methodological arsenal of all specialties of laboratory medicine. Similarly, microbiology, which originated in cultivating bacteria and viruses, also uses examinations to identify microbiota species and measurements based on molecular biology and mass spectrometry to determine species and treatment options.

Challenges in traceability in Analytical Chemistry and Laboratory medicine

Most concepts and practical solutions applied in metrology have their roots in physics and physical metrology. The metrology of Analytical Chemistry and Laboratory medicine has strived to adapt to these general principles and contributed substantially to the coherent development of international metrology, e.g., the creation of the VIM (52). However, it is essential to understand that traceability in the different specialties of Laboratory Medicine has a long way to go before reaching the degree of worldwide traceability already achieved in physics regarding length, mass, temperature, etc.

The concept of "analyte" is commonly used in Laboratory Medicine for the ideal idea of the molecule intended to be measured. As detailed below, most measurements in Laboratory Medicine measure *surrogate markers* ("measurands" = the quantity intended to be measured) for the intended "analytes."

Optimally, quantitative measurements in Laboratory Medicine (analogous to quantitative physical measurements) should be made traceable to the SI unit of amount of substance when the "analyte" can be uniquely identified, e.g. by a chemical structure, sequence of nucleic acids, etc. This means that measurements in Laboratory Medicine are optimally expressed as "amount of substance," the relevant basic quantity in the international measuring system (SI). Unfortunately, this is possible only for a minority of the measurands in regular use in clinical practice and Laboratory Medicine since the "analytes" corresponding to the measurands need to be available in a pure form and the same stable condition in humans. Macromolecules crucial for the proper function of the human organism are usually present *in vivo* as several molecular forms pose the risk of resulting in different quantity values for different measurands.

It is possible to establish traceability of measurands to "*international conventional reference materials,"* for example WHO reference materials and to agreed reference measurement methods. Still, calibration hierarchies to SI are impossible for such measurands since the "analyte" cannot be uniquely identified. Traceability for all non-SI traceable measurands in Laboratory Medicine must be handled separately, commonly through *harmonization protocols* since the imagined "analyte" may only partially comprise molecules represented in quantity measured in the measurand.

For example, measurement in laboratory medicine of the concentration of molecules is rarely direct. Instead, it relies on chemical, immunochemical, and molecular biology reactions combined with the measurement of physical quantities, which – together with the chemical reactions - are sufficiently characteristic of the molecules intended to be measured and to be fit for the intended use. Expressed in other words – measurements

in Laboratory Medicine are commonly performed using *surrogate markers* (measurands) for the intended "analytes."

Several in vitro diagnostic measurement measuring systems and methods claim to measure the same "analytes" but base their measurements on different chemical principles, resulting in different observed values for the same human sample or reference material. The most likely reasons are *differences in measurement selectivity characteristics,* including tertiary molecular structures, microheterogeneity, or chemical configurations of the intended "analyte." Measures must therefore be taken at all levels of the calibration hierarchy to prevent problems caused by differences or changes in the measured quantity among the different measurement methods at the various levels in the calibration hierarchy. It is crucial to recognize and minimize the differences between the quantity being measured and the quantity intended to be measured (measurand). Cases with variable microheterogeneity of the analyte (e.g., isoforms and posttranslational modifications) within the calibrators or human samples are especially important.

Metrological traceability problems often occur when the principle of the measurement system is based on the *detection of a surrogate for the analyte of interest,* e.g., a peptide epitope in a large protein rather than the entire protein molecule or a fragment of the protein molecule. Similarly, the measurement system's calibrator may contain an analyte that is a surrogate for the analyte found in human samples. Two or more diagnostic measurement systems with immunochemical methods claim to measure the amount of substance concentration of a single protein hormone (e.g., prostate-specific antigen [PSA]). If different measurement systems using immunochemical measuring methods recognize and react to varying extents with various epitopes of PSA, values for other although related quantities are generated by each measuring system and method, possibly leading to a lack of equivalence in the final measured quantity values in specific human samples.

Non-equivalence of values among different in vitro diagnostic measuring systems may be observed among very selective (but different) measurement principles (e.g., a mass-spectrometric measurement procedure vs. an immunoassay procedure for a protein hormone in patient plasma). Each diagnostic measurement system is targeted toward the detection of different isoforms or fragments of the same protein. Still, different values are commonly determined because different quantities are being measured with each diagnostic measurement system, e.g. due to binding of the selective antibodies to different epitopes of the intended molecules that also may show molecular heterogeneity due to post-translational processing.

Amongst the unique challenges facing the specialties of Laboratory Medicine in producing traceable measurement results are the following:

- There are “matrix factors” (substances and factors in the sample except for the analyte of interest).
- Inability to produce the substance in a pure form that can be weighed.
- Molecular heterogeneity, for example, transferrin, LH, FSH, TSH.
- Selectivity for different epitopes of the molecule of interest.
- Lack of knowledge of which epitopes of molecules are medically most relevant, for example, most substantial biological activity or best diagnostic properties.
- Changes in posttranslational modification of molecules, for example LH and FSH, during the ovarian cycle.

Another challenge in laboratory medicine is that the concentration is measured in a biological organism that changes over time in response to internal and external factors – biological variation. The concentration measured in the sample from a person is thus a snapshot in time and must be interpreted in that context.

Similarities and differences between chemical reference materials and physical standards

The general principles and nomenclature of metrology are evidently applied both for physics and the Analytical Chemistry and Laboratory Medicine (52). This means that a reference can be a “measurement standard” or a “reference material” traceability is a concept used in all areas of metrology, and SI units are used whenever possible in all areas in countries that have comprehensively implemented the SI system. However, reference materials in Laboratory Medicine, not even certified reference materials, realize the SI unit to the extent that physical measurement standards do because of the presence of “influence quantities” in the calibrators and the patient samples. Physical quantities such as length and mass can commonly be measured without significant influence from unwanted influences. However, the “analytes” measured in, for example human plasma samples, can never be measured without the risk of effect of surrounding molecules which commonly are present in order of magnitude higher concentration than the “analyte” of interest.

References

1. Quinn T. The Metre Convention, the BIPM and the SI. P Int Sch Phys. 2013;185:1-14.
2. Kaarls R. Evolving Needs for Metrology in Trade, Industry and Society and the Role of the BIPM. 2003.
3. De Bievre P. The multilateral recognition arrangement between National Measurement Institutes (NMI) is signed at BIPM. Accredit Qual Assur. 2000;5(2):43-.
4. Page CH, Vigoureux Pe. International Bureau of Weights and Measures 1875–1975. U. S. Department of Commerce: National Bureau of Standards; 1975.
5. Ferguson A. Quantitative estimates of sensory events. Inerim report of the committee appointed to consider and report upon the possibility of quantitative estimates of sensory events. British Asvancement of Sciencesociation for the Ad; 1938.
6. Michell J, editor History and philosophy of measurement: a realist view. 10th IMEKO TC7 International Symposium; 2004; Saint-Petersburg, Russia, June 30-July 2.
7. Michell J. The logic of measurement: A realist overview. Measurement. 2005;38(4):285-94.
8. Maul A, Mari L, Wilson M. Intersubjectivity of measurement across the sciences. Measurement. 2019;131:764-70.
9. Mari L, Maul A, Wilson M. Can there be one meaning of “measurement” across the sciences? Journal of Physics: Conference Series. 2019;1379.
10. Mari L, Maul A, Wilson M. On the existence of general properties as a problem of measurement science IOP Conf Series: Journal of Physics: Conf Series1065 2018.
11. Mari L, Maul A, Irribarra DT, Wilson M. Quantities, Quantification, and the Necessary and Sufficient Conditions for measurement. Measurement. 2017;100:115-21.
12. Mari L, Maul A, Irribarra DT, Wilson M. A meta-structural understanding of measurement. Journal of Physics. 2016;772:Conference Series 772 (2016) 012009.
13. Mari L, Wilson M. An introduction to the Rasch measurement approach for metrologists. Measurement. 2014;51:315-27.
14. Mari L, Maul A, Irribarra DT, Wilson M. Quantification is Neither Necessary Nor Sufficient for Measurement. 2013 Joint Imeko (International Measurement Confederation) Tc1-Tc7-Tc13 Symposium: Measurement across Physical and Behavioural Sciences. 2013;459.
15. Pendrill L. Quality assured measurement : unification across social and physical sciences. Switzerland: Springer; 2019.
16. De Bievre P. Looking back at two decades of "Metrology in Chemistry". Accredit Qual Assur. 2011;16(12):591-6.

17. Walker R. CITAC – international co-operation to achieve analytical comparability. Accred Qual Assur. 1997;2(1):40-3.

18. Adams F. Traceability and analytical chemistry. Accredit Qual Assur. 1998;3(8):308-16.

19. White GH. Metrological traceability in clinical biochemistry. Annals of clinical biochemistry. 2011;48(Pt 5):393-409.

20. Siekmann L. Establishing measurement traceability in clinical chemistry. Accredit Qual Assur. 2004;9(1-2):5-17.

21. Valcarcel M, Rios A. Traceability in Analytical-Chemistry. Analyst. 1995;120(8):2291-7.

22. Kessler A. Mass spectrometry - a key technique for traceability in clinical chemistry. Trac-Trends in Analytical Chemistry. 2016;84:74-9.

23. De Bievre P, Dybkaer R, Fajgelj A, Hibbert DB. Metrological traceability of measurement results in chemistry: Concepts and implementation (IUPAC Technical Report). Pure Appl Chem. 2011;83(10):1873-935.

24. EU. Directive 98/79/EC of the European Parliament and of the Council of 27 October 1998 on in vitro diagnostic medical devices. Eur-Lex, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31998L0079:EN:NOT; 1998.

25. Quinn TJ. News from the BIPM. Metrologia. 1996;33(1):81-4.

26. CGPM. The 20th Conférence Générale des Poids et Mesures. Paris: BIMP; 1995.

27. Williams A. EURACHEM - the first 10 years - Part 1. Accredit Qual Assur. 2000;5(3):121-3.

28. Williams A. EURACHEM - the first 10 years - Part 2. Accredit Qual Assur. 2000;5(4):153-5.

29. ISO. ISO 21151:2020 In vitro diagnostic medical devices — Requirements for international harmonisation protocols establishing metrological traceability of values assigned to calibrators and human samples. Technical Committee : ISO/TC 212 Clinical laboratory testing and in vitro diagnostic test systems. Geneva, Switzerland: International Organization for Standardization; 2020.

30. ISO. ISO 17511:2020 In vitro diagnostic medical devices — Requirements for establishing metrological traceability of values assigned to calibrators, trueness control materials and human samples. Technical Committee : ISO/TC 212 Clinical laboratory testing and in vitro diagnostic test systems Geneva, Switzerland: International Organization for Standardization; 2020.

31. Jeppsson JO, Kobold U, Barr J, Finke A, Hoelzel W, Hoshino T, et al. Approved IFCC reference method for the measurement of HbA(1c) in human blood. Clinical Chemistry and Laboratory Medicine. 2002;40(1):78-89.

32. Finke A, Kobold U, Hoelzel W, Weykamp C, Miedema K, Jeppsson JO. Preparation of a candidate primary reference material for the international standardisation of HbA1c determinations. Clinical chemistry and laboratory medicine : CCLM / FESCC. 1998;36(5):299-308.

33. Eckerbom S, Bergquist Y, Jeppsson J-O. Improved method for analysis of glycated haemoglobin by ion exchange chromatography. Ann Clin Biochem. 1994;31:355-60.

34. Kilpatrick ES. Toward a Global Overview of HbA1c Test Performance. Clinical chemistry. 2018;64(8):1131-2.

35. Molinaro RJ. Targeting HbA1c: standardization and clinical laboratory measurement. MLO: medical laboratory observer. 2008;40(1):10-4, 6-9.

36. Miedema K. Standardization of HbA1c and Optimal Range of Monitoring. Scand J Clin Lab Invest Suppl. 2005;240:61-72.

37. Young RJ. The clinical need for HbA1c standardization. Ann Clin Biochem. 2000;37(47-48).

38. Marshall SM, Barth JH. Standardization of HbA1c measurements: a consensus statement. Ann Clin Biochem. 2000;37(45-46).

39. Quantitative Estimates of Sensory Events. Nature. 1939;144:973.

40. Ferguson A. Quantitative estimates of sensory events. Nature. 1932;130(3279):334-6.

41. Ferguson A. Quantitative estimates of sensory events. Inerim report of the committee appointed to consider and report upon the possibility of quantitative estimates of sensory events. British Association for the Advancement of Science Ad 1938; 1938.

42. Stevens SS. Mathematics, measurement, and psychophysics. In: Stevens SS, editor. Handbook of Experimental Psychology. New York: John Wiley; 1951. p. 1-49.

43. Stevens SS. On the theory of scales of measurement. Science. 1946;103:677-80.

44. Salkind NJ. Tests & measurement for people who (think they) hate tests & measurement. 2nd ed. Thousand Oaks, Calif.: SAGE Publications; 2013. xxiii, 399 p. p.

45. Salkind NJ, Rasmussen K. Encyclopedia of measurement and statistics. Thousand Oaks, Calif.: SAGE Publications; 2007.

46. Salkind NJ, Rasmussen K. Encyclopedia of educational psychology. Thousand Oaks, Calif.: Sage Publications; 2008.

47. Michell J. An Introduction to the Logic of Psychological Measurement. London and New York: Routledge; 1990.

48. Michell J. Measurement in Psychology. A critical History of a Methodological Concept. Cambridge: Cambridge University Press; 1999.

49. Krantz DH. Foundations of measurement. New York ; London : : Academic Press; 1971.

50. Suppes P, Luce RD. Foundations of measurement: Academic Press; 1971.

51. Kurtz KH. Foundations of psychological research : statistics, methodology, and measurement. Boston, Mass.: Allyn & Bacon; 1965.

52. JCGM. International vocabulary of metrology — Basic and general concepts and associated terms (VIM 3): Bureau International des Poids et Mesures; 2012 [3

edition:[Available from: https://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2012.pdf.