Essential concepts and terms

For clarity, this document intends to adhere to the terminology described in the International Vocabulary of Metrology (VIM3) (1) and in the VIN (2). Necessary terms are summarized in the glossary which is provided as an appendix and the most essential concepts and terms for the present presentation are mentioned only briefly here.

Measurements in Laboratory Medicine for example of the concentration of medically relevant molecules are rarely direct. Instead, measurements rely on chemical, immunochemical, and molecular biology *reactions* combined with the measurement of physical *quantities* which – together with the chemical reactions - are sufficiently characteristic for the molecules intended to be measured 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".

The concept "analyte" should be avoided in the metrology of Laboratory Medicine since it does not refer to a quantity. If used the term "analyte" refers to the component which quantity (e.g. concentration) is to be measured. Since this is rarely possible, the quantity which is intended to be measured in practice is called a "measurand". The *measurand* refers to a quantity that can be measured in practice (3-10).

The term *measurement procedure* refers to a *written* specification for how a measurement is performed, including a technical description of reagents, calibrators, equipment, instrument, and other details necessary to create and operate a measurement that implements those specifications. A *measuring system* is the entire *physical* in-vitro diagnostic system manufactured according to the specifications of the measurement procedure and used to perform measurements of measurands in patient samples to produce quantity values that are used for diagnosis, monitoring of treatment effects and for screening for risk factors and for diseases. A measuring system comprises the physical instrument and includes calibrators, reagents and any necessary auxiliary equipment.

Matrix effect is influence of a property of the sample, independent of the presence of the analyte, on the measurement and thereby on the measured quantity value. Matrix effects are present both in natural patient samples and in reference materials and are crucial for the commutability of reference materials. A crucial difference between natural patient samples on one hand and calibrators and control materials on the other is that natural patient samples are commutable by definition. *Commutability* is a nominal property (a material is either commutable or not commutable) of reference materials, demonstrated by the equivalence of the mathematical relationships among

the results of different measuring systems for a reference material and for representative samples of the patient samples intended to be measured. The conclusions reached regarding commutability of a certain measuring system is that these conclusions are assumed to be applicable to all other measuring systems that are implementations of the same measurement procedure.

Equivalence is primarily a functional/clinical concept "Agreement of measured values among different in vitro diagnostic measurement devices intended to measure the same measurand, where the differences in measured values on the same human samples do not affect clinical interpretation" (11), but limits also enter into the concept as follows "NOTE 1: A conclusion of equivalence of measured values for the same human samples among two or more measuring systems is based on the differences in measured values being within a pre-defined margin or limit (11).

Harmonization is defined as follows "Achievement of equivalent measured quantity values (within clinically meaningful limits) for human samples examined for a stated measurand among two or more measuring systems by applying an *international consensus protocol* in their calibration hierarchies when fit-for-purpose higher order reference materials or reference measuring systems are not available

Note 1: Harmonization is one of the calibration hierarchy models described in ISO-17511:2020 (11) to achieve metrologically traceable quantity values for human samples.

Note 2: Harmonization is a special case of non-SI traceable standardization where the calibration of two or more measuring systems is traceable to an international harmonization protocol that defines the highest level of metrological traceability for the stated measurand, but with no traceability to SI.

Note 3: Harmonized is the condition in which harmonization (equivalence among quantity values) is achieved among two or more measuring systems." (11).

Since the publication of ISO-17511:2020 (11) and ISO-21151:2020 (12) harmonization should not anymore be regarded as an alternative to standardization, but rather as one of the tools for reaching standardization used in calibration hierarchies #3 to #6 (11).

A qualitative concept measurement *trueness* is the "closeness of agreement between the average of an infinite number of replicate measured quantity values and a reference quantity value" (Figures 2 and 3). It is quantitatively expressed as *bias*. Another qualitative concept measurement *accuracy* describes the "closeness of agreement between a measured quantity value and a true quantity value of a measurand. It includes both systematic and random error components.

A more accurate result has a smaller measurement error. It is on the average more true when the bias is small and more precise when the random error is small. *Precision* is expressed quantitatively as its opposite – imprecision using the unit of standard deviation.

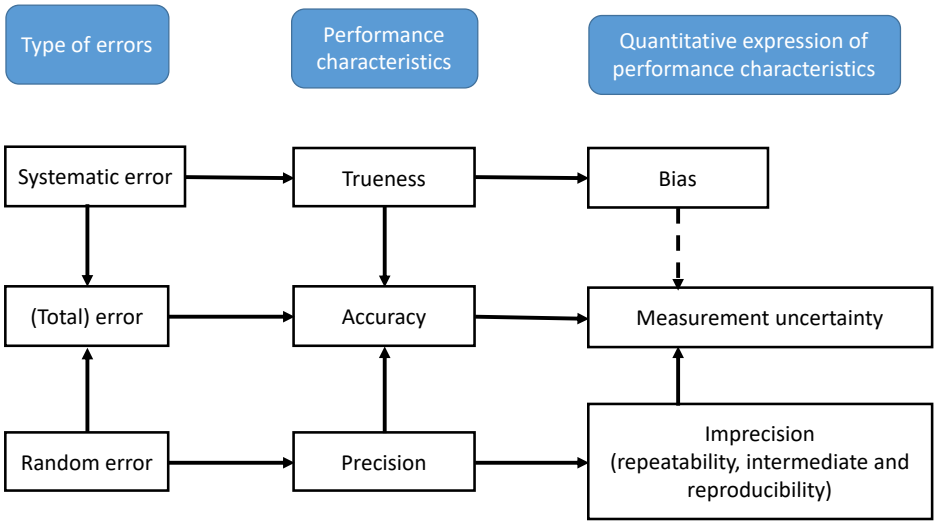

Figure 2: Concept diagram, adapted from Menditto et al. (13), explaining the relations between concepts describing, random and systematic errors as well as measurement uncertainty. The dotted line from bias to measurement uncertainty is to indicate that if bias can be estimated, it should be eliminated.

It is important to note that trueness and precision are performance characteristics (qualitative concepts) which express the qualitatively relative magnitude of the bias compared to a reference measurement result the average of two or more replicate measurements have regarding systematic error (trueness) and random error (precision), respectively.

A weakness in this concept diagram is that *accuracy* has a double meaning – a qualitative- and a quantitative meaning. The qualitative meaning expresses whether a *single* measurement result from measuring system A is likely to be more or less accurate that a measurement result from measuring system B. The accuracy is the difference between the sum of the random and systematic error minus a reference measurement result. The accuracy of a single measurement result in a quantitative meaning is the difference between the sum of the random and systematic error minus a reference measurement result.

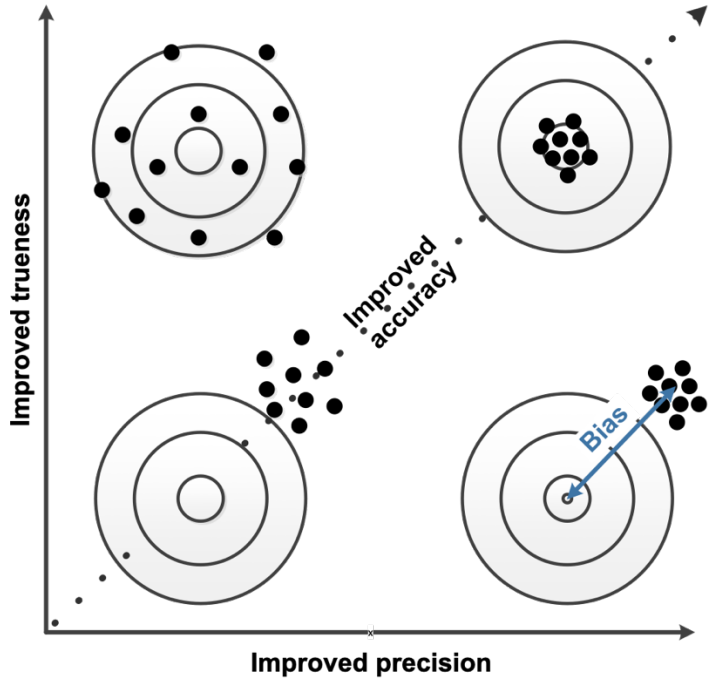

Figure 3: A graphical illustration of the meaning of trueness, precision, and their combination – accuracy.

Accuracy includes both random and systematic components which can be present in to any relative extent.

Measurement results are expressed on four "*measurement levels*"; *nominal*, *ordinal*, *interval* and *ratio*. Each level of measurement specifies how the numbers that are assigned to the measurands relate to the basic characteristics of the measurand determined by noting the presence or absence of four characteristics: 1) distinctiveness, 2) ordering in magnitude, 3) equal intervals, and the presence of a 4) absolute zero. A measurand has the characteristic of *distinctiveness* if measurement results can be expressed as different numbers, characters, or strings of characters. For example, persons have different names, blood groups have different characters or strings of characters making the distinctive regarding naming. Measurement results can also indicate an *ordering in magnitude*, with larger numbers representing more of the measurand being measured than smaller numbers. For example, in a urinary dipstic of +2 indicates a higher concentration of urine albumin than +1, even if this does not necessarily mean that +2 in this context means twice the concentration compared to +1. *Equal intervals* are obtained if equivalent differences between measurements represent the same quantity being measured. For example, if a two-point difference between the hemoglobin concentrations of 130 and 145 represents the same difference in concentrations as the two-point difference between the concentrations 115 and 130, the measurement has equal intervals. A measurement has an absolute zero when a

measurement of zero represents an absence of the property being measured. For example, a concentration of 0 means the absence of the molecules in question in the solution used for measuring. A very clear example are the Centigrade and Kelvin temperature scales. The Kelvin scale starts with 0 - the temperature when no molecules move. In contrast the Centrigrade (Celcius) scale does not mean the absence of temperature (movement of molecules). The four characteristics of measurement just described determine the four major levels of measurement: nominal, ordinal, interval, and ratio.

Characteristic	Nominal	Ordinal	Interval	Ratio
Distinctiveness	yes	yes	yes	yes
Ordering in magnitude	no	yes	yes	yes
Equal intervals	no	no	yes	yes
Absolute zero	no	no	no	yes

Table 1: Characteristics of the four levels of measurement (14).

References

1. 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.
2. Nordin G, Dybkaer R, Forsum U, Fuentes-Arderiu X, Pontet F. Vocabulary on nominal property, examination, and related concepts for clinical laboratory sciences (IFCC-IUPAC Recommendations 2017). Pure Appl Chem. 2018;90(5):913-35.
3. Shewhart WA. On the Measurement of a Physical Quantity Whose Magnitude is Influenced by Primary Causes beyond the Control of the Observer and on the Method of Determining the Relation between Two Such Quantities. P Natl Acad Sci USA. 1922;8:248-51.
4. Dybkaer R. ISO terminological analysis of the VIM3 concepts 'quantity' and 'kind-of-quantity'. Metrologia 2010;47:127-34.
5. Dybkaer R. Generic division of 'quantity' and related terms. Accredit Qual Assur. 2011;16(12):649-51.
6. Mundy B. The Metaphysics of Quantity. Philos Stud. 1987;51(1):29-54.
7. Cooper GA, Fisher WP. Continuous quantity and unit; their centrality to measurement Jena, Germany; 2011.
8. Michell J, Ernst C. The axioms of quantity and the theory of measurement (Reprinted from Mathematisch-Physikaliche Classe, vol 53, pg 1-64, 1901). J Math Psychol. 1996;40(3):235-52.
9. Lam D. Metaphysics of quantity and the limit of phenomenal concepts. Inquiry. 2019;62(3):256-75.
10. Wolff JE. The metaphysics of quantities. New product. ed. New York: Oxford University Press; 2020. pages cm p.
11. 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.
12. 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.
13. Menditto A, Patriarca M, Magnusson B. Understanding the meaning of accuracy, trueness and precision. Accred Qual Assur. 2007;12:45-7.
14. Stevens SS. On the theory of scales of measurement. Science. 1946;103:677-80.