OPEN ACCESS Freely available online

ISSN: 2375-4273
Health Care : Current Reviews

Research Article

The Relationship between Metrology Vocabulary and ISO or CLSI Vocabulary

Marco Pradella*

Quality and Accreditation Commission, Italian Society of Clinical Pathology and Laboratory Medicine, Via Ponchini 17/int.7, 31033 Castelfranco Veneto, TV, Italy

ABSTRACT

Medical laboratories are perhaps the largest measurement industry in the world. The metrology terminology is relevant for effective and efficient communication, particularly where metrology activities are carried out by operators with different metrology skills. WASPaLM and SIPMeL have had some opportunities to propose changes to the documents in preparation to Clinical & Laboratory Standards Institute (CLSI) and International Organization for Standardization (ISO) in order to harmonize the terminology with the Metrology Vocabulary (VIM). Many proposals have been accepted. Here we summarize some particularly critical points for metrological terms. The main terms discussed are: Measuring, measuring range, examination, pre-examination, post-examination, manufacturer, measuring instrument, quantitative, qualitative, semi-quantitative, processing, measurement error, maximum permissible error of measurement, total error of measurement, monitoring, variability, performance, reliability, influence, interference, selectivity, sensitivity, detection limit, reliability, comparability, compatibility and control material. Despite all efforts to coordinate terminologies, it is inevitable that overlapping and inconsistent terminologies will continue to be used because documents and policies are produced in different contexts. In some ISO and CLSI documents, phenomena of magnetic attraction towards common words (such as "analysis" and derivatives), without any consideration of the true metrological meaning, are noted. The ISO and CLSI working groups show, alongside moments of openness, phenomena of true self-referential conservatism.

Keywords: Medical laboratory; Metrology; Vocabulary; ISO; CLSI

INTRODUCTION

Medical laboratories are perhaps the largest measurement industry in the world, with staggering numbers and turnover and a extreme impact on health and our daily lives. The metrology terminology is relevant for effective and efficient communication, particularly in the laboratories of healthcare facilities, whose metrology activities are carried out by operators with different metrology skills a situation where communication problems can arise and where appropriate terminology is therefore important.

However, the diversity of operators involved in metrological activities, ranging from clinicians and lab technicians to manufacturers of diagnostic instruments, poses significant challenges. These operators often possess varying levels of metrological expertise, which can lead to discrepancies in understanding and applying metrology concepts. This makes clear, standardized vocabulary essential to minimizing miscommunication and improving overall reliability in lab measurements. The complexity of this task is further compounded by the fact that multiple regulatory bodies and standardization organizations, such as ISO and CLSI, are involved in drafting guidelines and standards for medical laboratories. While both ISO and CLSI aim to align terminology with the internationally recognized Metrology Vocabulary (VIM), achieving this harmonization has proven challenging. Despite efforts to streamline terms, divergence persists, reflecting the

Correspondence to: Marco Pradella, Quality and Accreditation Commission, Italian Society of Clinical Pathology and Laboratory Medicine, *Via* Ponchini 17/int.7, 31033 Castelfranco Veneto, TV, Italy, E-mail: labmedico@labmedico.it

Received: 09-Sep-2024, Manuscript No. HCCR-24-26885; **Editor assigned:** 11-Sep-2024, PreQC No. HCCR-24-26885 (PQ); **Reviewed:** 25-Sep-2024, QC No. HCCR-24-26885; **Revised:** 03-Oct-2024, Manuscript No. HCCR-24-26885 (R); **Published:** 11-Oct-2024, DOI: 10.35248/2375-4273.24.12.413

Citation: Pradella M (2024). The Relationship between Metrology Vocabulary and ISO or CLSI Vocabulary. Health Care Curr Rev. 12:413.

broader tension between scientific rigor and practical utility within these institutions.

MATERIALS AND METHODS

Regulatory work for medical laboratories takes place in different fora. A special ISO committee (ISO/TC 212) maintains the documents used worldwide for the accreditation of management systems for these activities (following the IAF, ILAC and EA mutual recognition chain), but also a number of documents that flank them for particular aspects or areas, such as sampling, near-patient examinations Point of Care Testing (POCT), risk management, anatomy and so on. Plus, a large number of documents dedicated to emerging areas of molecular methods. Clinical and Laboratory Standards Institute standard (CLSI), on the other hand, produces many guidelines, which are often given preferential consideration by entities such as the College of American Pathologists (CAP) and the United States Food and Drug Administration (U.S. FDA), but are also useful in many cases to support laboratories in responding to ISO requirements, especially when they are poorly detailed and prescriptive or even contradictory. The Società Italiana di Patologia Clinica e Medicina di Laboratorio (SIPMeL) is a regular member of CLSI, so it can propose new projects, volunteer or nominate candidates for standards development, vote on documents, participate in elections of officers and CLSI regulations. SIPMeL is a member of the World Association of societies of Pathology and Laboratory Medicine (WASPaLM), which participates in ISO work in TC 212 (Medical laboratories and *in vitro* diagnostic systems). Medical laboratories can take advantage of the recent revision of the ISO standard that concerns them. ISO 15189 introduces or confirms several innovative words in the chapter 'Terms and definitions', but others are present in the text of the document. Most of the terms defined by ISO 15189 concern practical aspects, but some (bias, reference interval, commutability, examination, examination procedure, IVD medical device, measurement accuracy, measurement uncertainty, measurement trueness, validation and verification) have obvious metrological connotations [1]. It would have been natural to expect that the ISO documents linked to ISO 15189 would follow the nomenclature, but this is not always happening, a bit like with the Metrology Vocabulary (VIM3) [2].

RESULTS

The following is a summary of some chapters where metrology terminology has played a significant role in the drafting of ISO and CLSI documents. In presenting the VIM and ISO 15189 definitions to the editors of the ISO and CLSI documents, when the opportunity arose, we experienced very varied reactions, ranging from welcome, to surprise, to perplexity, to rejection (Table 1).

Table 1: Summary of proposals for metrological or standard terms and outcome in the working groups.

Term	Instead of	Proposed to	Outcome
'Measurand' 'Measurement procedure' 'Measurement coefficient of variation'	'Analyte' 'Analysis' 'Analytical variability (CVA)' Analytical Coefficient of Variation (CVA)	ISO 22583 ISO 16766 ISO 18704 CLSI EP46 CLSI EP31	'Analyte' is a publicly known term, still used in the diagnostic community, including legislation
'Pre-examination'	'Pre-analytical'	ISO 18704	Partially accepted
'Measuring range'	'Analytical measuring interval'	CLSI EP31, EP21, EP32	Commonly used term
'Manufacturer of IVD'	'Manufacturer of analytical tests'	ISO 18704	Accepted
'Ordinal quantities'	'Semi-quantitative properties'	ISO 18704 ISO 22583	Accepted by 18704, "publicly known term" by 22583
'Measurement error' 'Systematic error' 'Random error'	'ANALYTICAL error' 'Total Allowable Error (ATE)'	CLSI EP21 CLSI EP46	Waiting
'Performance of a measuring system'	'Analytical Performance Specification(s) (APS)'	CLSI EP32	Commonly used and accepted term 2015 (!) IFCC recommendations
'Influence quantity' 'Selectivity'	'Interfering substance' 'Analytical specificity' 'specificity'	ISO 22583 ISO 18704 CLSI EP31 CLSI MM19	Too complicated for the target audience
'Limit of detection'	'Analytical sensitivity'	ISO 18704	Partially accepted

		CLSI EP47	
'Reliability'	'Clinical sensitivity' or 'Diagnostic sensitivity', 'Clinical specificity' or 'Diagnostic specificity' related to a reference method, 'Accuracy' (of qualitative examinations)	CLSI MM19	Partially accepted
'Metrological compatibility'	'Comparability'	CLSI EP31 CLSI EP39	Waiting
'Metrological comparability'	"Comparison" in IQC	CLSI EP31	Waiting
'Reference material' (non-commutable)	Commutable reference and control materials (surrogate samples)	CLSI EP31 CLSI EP35	Waiting
'Concordance' (qualitative)	'Trueness' (qualitative)	CLSI MM19	Waiting
'Accordance' (qualitative)	'Precision' (qualitative)	CLSI MM19	Waiting

Note: Sources of "term" Column: VIM3, VIM4+IUPAC, ISO 27877. IQC: Internal Quality Control.

Measuring, measuring range, examination, pre-examination, post-examination, manufacturer, measuring instrument, quantitative, qualitative, semi-quantitative and processing

We found in VIM3 2.3 'measurand' as the quantity that is to be measured, while the term 'analyte', i.e., the name of a substance or compound, must not be confused with the term 'measurand', because analytes are not quantities. Furthermore, VIM3 2.6 'measurement procedure' is a detailed description of a measurement according to a certain measurement method, based on a measurement model. IFCC-IUPAC proposes to supplement the Vocabulary with item 2.6 'examination of a nominal property', as the process of obtaining experimentally one or more values that can be reasonably attributed to a nominal property [3]. This new item is currently being discussed in the revision of VIM3. In some cases, an examination is performed through intermediate steps, which are measurements and the results of which are used to obtain the examination result. In ISO 15189 the term 'examining' is also used to refer to measuring, whereas in IFCC-IUPAC the term refers only to nominal properties. It is accompanied in IFCC-IUPAC by 'examining' (analogue of measurand) as a nominal property intended to be examined. ISO 15189 defines 'examination' as the set of operations aimed at determining the value or characteristics of a property. The lemma is taken up as such in many CLSI documents (QMS02, GP26, PRE04, QMS06, QMS11, QMS13, EP12, GP33, POCT07, POCT10, PRE01, GP23, QMS25, EP23, QMS22, QMS01, EP35). In discussion with the editors of ISO 22583 (pre-examination processes) and ISO 16766 (crisis IVD production), the proposal for metrological harmonization was rejected because 'analyte' is a publicly known term. The editors of ISO 18704 (collection of urine and other liquids) thanked the suggestion, but following the international discussion on the use of the term 'analyte' they noted that the term is currently still used in the diagnostic community, including legislation, e.g. EU-IVDR, FDA and CDC guidance documents. They therefore decide to continue using this term. However, they add 'measuring' in the definitions and in some notes to describe the correct metrological terms. In several sentences, the term 'analysis' will be replaced by 'measurement'. In ISO 18704 it is argued that the terms 'pre-analytical step' and 'pre-analytical workflow' are widely used, but 'pre-analytical variables' will be replaced by 'pre-examination variables'. Regardless, Section 3.23 will continue to have 'pre-analytical' as it is often used in practice as a synonym, citing an old 2015 reference. The use of a concept for a long time should not be the reason for rejecting innovations or technical corrections, regardless of the authority of the old sources. ISO in the medical laboratory field defines 'analyser' only in ISO/IEEE 11073-10422:2017 for informatics of urine instruments. VIM3 4.7 defines 'measuring range, working range' as the range of values of quantities (quantities) that can be measured by a given measuring system with a specified uncertainty. CLSI uses 'analytical measuring interval' in many documents, even quoting JCGM 200:2012 (VIM3) but with a mispronunciation of the original VIM3 lemma. Even when it does not provide a definition, it produces acronyms ('AMI' instead of 'MI') as in EP31 or EP32 (traceability). Evident is the unnecessary cacophonous tautology, which should be corrected in both lemma and acronym. The wording of EP32 ends in the justification of "commonly used and accepted term". VIM3 does not define 'manufacturer', but uses it in 2.44 ('verification' as 'manufacturer of measurement system'). In the discussion on ISO 18704 'manufacturer of analytical tests' is deleted, 'manufacturer of examinations' is changed to 'manufacturer of examination devices', with the definition updated accordingly: 'Entity that manufactures devices for *in vitro* diagnostic or research examinations, including measuring systems, instruments, reagents and instructions for use for a specific examination'. It should be noted that for a long time the concept of 'manufacturer' was superficially linked to the examination 'in toto', whereas today we know that the

examination or measurement is performed by the laboratory, while the manufacturer supplies devices or consumables. VIM3 1.26 states that properties that can be compared by order, but not by difference, are ordinal quantities, while (1.30) properties that do not have magnitude are nominal properties. In the relevant literature, ordinal properties are sometimes considered non-quantitative. The literature and practice, in order to handle the difficulty, have created a non-metrological category: Semi-quantitative properties. A solution that on the contrary worsens the difficulty. Ellison of EURACHEM has well described at least part of the confusing reality [4]. These can be ordinal scales (Absent/Low/Medium/High), interval scales (0-10, 10-100...), quantitative with great uncertainty. For Ellison, it is in fact difficult to find purely semi-quantitative examples, they are often quantitative expressed semi-quantitatively. ISO 18113-1:2022. 1.80 defines semi-quantitative examination as the set of operations that give results in an approximate range (e.g., trace, moderate). For ISO 18113 these are essentially qualitative examinations with an additional option for the response range (degree of positivity, dilution at which positive results are obtained or comparison with a colour chart). In the discussion with the editors of ISO 18704 (urine), it was admitted that 'semi-quantitative' is not a scientific term. The editors of ISO 22583 (POCT) justified themselves by stating that 'semi-quantitative' is a publicly known term. ISO 10993-18 (chemical characterization of medical devices) defines (3.29) qualitative analysis, (3.33) quantitative analysis and even (3.32) semi-quantitative analysis as 'an analytical approach that provides the concentration of an analyte using the response of one (or more) surrogate substance, specifically taking into account the relative responses of the analyte and the surrogate substance', whatever that means. And adds (3.4) analytical (screening) method-a method whose purpose is to discover, identify and semi-quantitatively estimate the concentration of all relevant analytes in a test sample above an established reporting threshold. 'Semi-quantitative' is used with different meanings in CLSI documents such as I/LA02, I/LA33, EP26, POCT04, H62, EP47. Semi-quantitative is used but not defined in EP39 (surrogate samples) and EP31 (method comparison). In discussion with the editors of EP47 (drag and drop), it emerged that 'semi-quantitative' is attributed to methods in which a threshold value divides the results between negative and positive. Ultimately, ISO documents, CLSI and practice introduce a non-metrological category to label procedures whose metrological models are not well understood. The practical result is the difficulty of applying the performance characteristics of qualitative properties (starting with uncertainty) as well as those of nominal properties to these procedures. In meetings with laboratories, it is noted that in these situations operators are as if paralyzed and accreditation inspectors are greatly embarrassed.

VIM defines 2.1 (measurement) as the activity of manipulating a sample to obtain a measurement as a treatment ("process"). The English "process" does not correspond exactly to the Italian meaning of "procedure, development, subsequent development, continuation", but rather to that of "processing".1,2 In ISO documents "analysis" is often used for statistics and risks, but in some cases it also means treatment of materials for measurement purposes. As in ISO 21043 (forensic sciences), where 3.2 analysis becomes part of the examination (3.18) which consists in detecting and/or measuring and/or comparing the properties of the elements (3.34) in order to obtain observations (3.37). The analysis can be instrumental, based on human perception or a combination of the two. It is accompanied by 3.3 analytical strategy, i.e., the choice of methods (3.36) and the sequence of analysis (3.2). We can perhaps ignore ISO/TS 5044:2023 (informatics for traditional Chinese medicine) which uses 3.2.1 "chemical analysis". CLSI EP35 (surrogate samples) also does not define "analysis", "analyze", "analyte" and "analytical", but uses the verbs "analyze" both for the treatment or processing of samples and for the statistical study of data. The same is done by CLSI EP31 (comparison of methods), where the recurrence of "analysis" in the statistical study is even prevalent. It can therefore be stated that the words derived from "analysis" and "analyte" are used in practice as multipurpose containers in the meaning of "processing, treatment". While the activities in statistics and even in philosophy (where they were born) are indisputable, VIM rightly points out that the uncontrolled dissemination in metrological fields can distract and distance from the necessary scientific rigor on the characteristics of the measurements.

Measurement error, maximum permissible error of measurement, total error of measurement, monitoring and variability

For VIM 2.16 'measurement error' or 'error' is the difference between the measured value and a reference value, having systematic and random components. For VIM 2.19 'random measurement error' is the component of measurement error that in replicate measurements varies unpredictably, with a distribution that can be summarized by its mean, generally assumed to be zero and its standard deviation. VIM 2.17 systematic measurement error or systematic error is the component of the measurement error that in replicate measurements remains constant or varies in a predictable manner, to which a compensatory correction can be applied, provided that the uncertainty of the correction is known or negligible. VIM 2.18 (measurement bias or distortion) is the estimate of a systematic error, which can be used for the correction of systematic error. Finally, VIM3 4.26 maximum permissible measurement error is the extreme measurement error, relative to a known reference value, allowed by the specifications or regulations for a given measurement, measuring instrument or measurement system. Only in ISO 14644-1:2015 (Clean rooms) do we find the entry as in VIM 4.26 in the health sector. 'Analytical error' appears in an older ISO document (ISO 18158:2016, linked to EN 1540:2011, Workplace Exposure-Terminology) concerning air in the workplace, even as a synonym for "uncertainty". In contrast, CLSI practices measurement error in several guides. CLSI EP35 and CLSI EP31 use the concept in the context of the comparison of methods, incorporated in the headword 'Total Allowable Error (ATE)'. CLSI EP21 and CLSI EP46 make it the centre around which recommendations are developed on how to estimate the measurement error and the identification of its limit of acceptability. In discussions with the committee developing the

two documents, numerous critical points and inconsistencies were noted on the operational level. From the metrological point of view, we can point out that in the EP21 definition itself, 'total error' includes what occurs in the preexamination processes (from sampling to receipt) and in the post-examination processes (from presentation of the result to its clinical use). Thus, all processing of the result, regardless of its correctness, only concerns the measurement. One should therefore speak of 'Total Measurement Error (TME)' and 'Total Acceptable Measurement Error (ATME)'. For errors in pre and post-examination processes, no guidelines are available at the moment.

But above all, one notices the approach of CLSI documents that adopt measurement error and the resulting statistics as an alternative to the measurement uncertainty demanded by ISO 15189 and driven by ISO 20914. An embarrassing situation for both laboratories and those involved in their ISO accreditation processes. Surveillance of performance characteristics seems at first sight to be an operational rather than a metrological issue. From another point of view, it can be placed alongside performance estimation activities. VIM3 2.22 (intermediate accuracy conditions) gives as an example measurement results obtained on quality control materials to monitor the quality of measurements. We can therefore ask which metrological concepts lend themselves well to monitoring and which less so. CLSI EP31 (comparability) uses the word 'monitoring' frequently, more than 40 times. The tendency to propose method comparison procedures with a fairly high frequency is evident. Despite conceptual weaknesses, i.e., the precarious relationship with calibration, the statistics of comparison data, the costs of procedures, the risks of false positives and false negatives, the weak connection with accreditation requirements. Clearly, the metrological concept VIM 5.26 (comparability based on calibration) does not lend itself to monitoring. But VIM 2.27 (compatibility on the basis of measurement pairs) also lends itself to monitoring in the form of inter-laboratory programmes, where costs and frequency are carefully managed, much less for in-house laboratory activities, as if it were a kind of alternative to in-house quality control. For CLSI EP21, 'total error' includes all random and systematic errors throughout the examination process and includes the combined effect of all precision and systematic errors that may affect the accuracy of a result. Total error incorporates sources of error from the pre-examination, examination and post-examination phases of a measurement procedure, but no CLSI guidelines are available for measuring total error. Neither are there any ISO guidelines for the error so defined. VIM does not have a definition of "variability", but uses the word "coefficient of variation under the specified conditions of measurement" in 'precision' (VIM3 2.15). CLSI only defines 'variability' in M23 (anti-biogram). CLSI EP21 (measurement error), EP46 (maximum acceptable error) and EP31 (method comparison) often use 'variability' in their abbreviations. They consider the concept of so called 'analytical' variability (CVA), i.e., that of measurement, alongside biological variability between subjects (CVG) and biological variability within the subject (CVI), as well as variability caused by personnel handling of samples, instruments, etc. They do not distinguish between variability and imprecision estimation, but speak of 'measurement system variability'. In fact, we find in EP21 that the only way to reduce variability (i.e., imprecision) is to increase the number of tests.

Performance, reliability, influence, interference, selectivity, sensitivity and detection limit

VIM does not define the lemma 'performance', but makes use of it in 4.10 (boundary operating conditions), 4.11 (reference operating conditions), 4.18 (detection limit), 2.44 (verification), as if the concept were taken for granted. For ISO 15189, on the other hand, definition 3.31 (Validation) lists the performance specifications: Accuracy of measurement, precision of measurement, including repeatability of measurement and intermediate precision of measurement, selectivity (analytical specificity), including influencing (interfering) substances, limit of detection and limit of quantification, range of measurement, clinical relevance, diagnostic specificity and diagnostic sensitivity. ISO 4307 (saliva) and ISO 20186 (cellular RNA) give for examination performance something generic, i.e., 'ability of an examination procedure to measure or detect'. Even ISO 17593 (anticoagulant self-examination) remains in the generic in 3.6 (control material) with 'verify the performance characteristics of an *in vitro* diagnostic (IVD) medical device'. ISO 16766 (IVD production) distinguishes measurement performance (3.2) and clinical performance (3.5). While the latter are declined into precise characteristics (sensitivity, specificity, predictive value), the former (under the name of 'analytical performance') are entrusted to the generic, perhaps tautological phrase 'the ability of an IVD medical device (3.10) to detect or measure', leaving the reader to imagine the content of 'capability'.

We find 'performance' in the term 'performance specification'. ISO in several standards, none of them health-related, understands it as 'document'. Instead, in the literature, the concept is linked to the measurement procedure, transferred into acronyms such as 'APS' and defined as 'Criteria that specify (in numerical terms) the required quality of analytical performance in order to provide laboratory test information that meets clinical needs to improve health outcomes' [5]. These are metrological performance targets (goals), collected in a database, estimated by precision and accuracy, to be used mainly in method validation and inter-laboratory exercises. Regardless of the technical aspects of the proposals on the subject, it is clear that a clean-up of terminology would promote clarity and harmonization. CLSI in EP32 (traceability), defends APS with the justification that it is a commonly used and accepted term. Still a descriptive approach, not prescriptive but unscientific. Referenced by key publications on the topic by CLSI in 2015, years before VIM metrological warnings with the start of the vocabulary revision process [6]. Metrological performance concepts are well defined for quantitative results [7,8]. Many difficulties arise with qualitative nominal results. IFCC-IUPAC proposed '3.12 examination precision' and '3.8 examination trueness'. ISO/TR 27877:2021 proposed '3.1.2 concordance' and '3.1.1 accordance' [9]. ISO/TR 27877:2021 '3.1.1 accordance' is the probability that two binary measured values be identical when they are taken from the same laboratory. The concept corresponds to the definition of 'repeatability' in ISO

5725. ISO/TR 27877:2021 '3.1.2 concordance' is the probability that two binary measured values be identical when they are taken from different laboratories. The concept corresponds to the definition of 'reproducibility' in ISO 5725. The issue is under discussion for the VIM revision, it seems for now to have landed in the VIM4 lemma '6.12 examination reliability'. We find the words 'sensitivity' and 'specificity' used for both laboratory measurements and the value of results from a diagnostic point of view. With the effect of introducing many ambiguities in meaning and the need for tortuous phrases in the texts to mitigate precisely the risks of ambiguity. VIM 2.52 defines 'influence quantity' as that which does not influence the quantity to be measured but the result of the measurement. VIM 4.13 puts the lemma 'selectivity' alongside it as a property of a measurement system that provides indications independent of quantities other than the measured quantity. ISO/IEC 17025 in 7.2.2.3 speaks of '... robustness against external influences or cross-sensitivity against matrix interferences'. In the discussion of ISO 22583 and ISO 18704, the objection was raised by the editors that the concepts of 'influence' and 'selectivity' are too complicated for the target audience of the documents, whereas 'interfering substance' is widely used by medical laboratories, IVD manufacturers and regulatory agencies such as the FDA. ISO is therefore represented as an entity that collects current usage and accepts the dialects of hypothetical categories of operators, right or wrong. An entity with a 'descriptive' rather than 'prescriptive' policy. Since it is clear that not only the behaviour of laboratories and IVD manufacturers, but also that of agencies such as the FDA, is derived from ISO documents, the phenomenon no longer appears as a progressive improvement on a scientific basis, but as a vicious circle. We find 'influence' with 'selectivity' in CLSI EP30 (Switchable Reference Materials), obtained through SIPMeL proposal. But also in CLSI EP32:2024 (Metrological traceability) and several other documents: C58, C51, EP07, C50, H62. CLSI EP31 maintains 'interference' between variations from special cause, i.e., from sources outside the examination process. The same is found in numerous CLSI documents. VIM is very clear when it defines 4.12 (sensitivity of a measurement system) as the quotient of the change in an indication of a measurement system and the corresponding change in a given value of the measured quantity, while in 4.18 (limit of detection) it refers to the term 'sensitivity' as not being recommended to refer to limits of detection. ISO 18113-1:2022 3.2.4 takes up the VIM dictate by even specifying that the sensitivity of a measuring system (3.2.40) is the slope of the calibration curve. Clearer than that it is not possible. The drafting of ISO 18704 (urine) maintains the term "sensitivity", adding the term "limit of detection" to the list of performances. Unfortunately, ISO 18704 inherits the definition of ISO 20184 1:2018 (RNA) 3.4 ('analytical test performance'), which is totally incompatible with VIM.

Instead, in CLSI documents, the exchange between LOD and sensitivity is often found, with notes such as "Sensitivity depends on the imprecision of the measurements" or "The lower limit of detection of a nucleic acid sequencing method", or in POCT04 "the lowest concentration reliably determined as non-zero with a minimum level reliably detectable". But also with uses in the text that conflict with the same definitions, such as in EP47 (drag-through). The uses of "sensitivity" for "in qualitative methods, the ability to obtain positive results in agreement with the reference method", clearly a linguistic cast from the concept of "diagnostic sensitivity", do not help. We found in MM19 that the term "clinical sensitivity" (United States) is equivalent to "diagnostic sensitivity" (Europe) and the term "clinical specificity" (United States) is equivalent to "diagnostic specificity" (Europe). Moreover, the term "clinical sensitivity" (United States) is equivalent to "diagnostic sensitivity" (Europe). The term "diagnostic specificity" (Europe) is equivalent to the US term "clinical specificity". These are big mistakes. US definitions are given by FDA: "The specificity of the test is estimated as the proportion of subjects without the target condition in whom the test is negative"; "The sensitivity of the new test is estimated as the proportion of subjects with the target condition in whom the test is positive" [10]. EU definitions are given by IVDR: (49) 'diagnostic specificity' means the ability of a device to recognize the absence of a target marker associated with a particular disease or condition; (50) 'diagnostic sensitivity' means the ability of a device to identify the presence of a target marker associated with a particular disease or condition [11]. Note that "condition" is very different from "marker". 'Specificity' is the percentage of true negatives out of all subjects who do not have a disease or condition. While 'sensitivity' is the proportion of true positives tests out of all patients with a condition [12]. Weak point of EU IVDR is the confusion between device reliability and marker characteristics.

Comparability, compatibility, control material VIM 2.46 provides the definition of 'metrological comparability' as a property of measurement results when they are metrologically traceable to the same reference. It complements VIM 2.47 with 'metrological compatibility' by attributing it to the difference of any pair of values from two different measurement results. VIM 2.46 states that metrological compatibility of measurement results replaces what was traditionally called "staying within error", as it represents the criterion for deciding whether or not two measurement results refer to the same measurand. If in a set of measurements of a measurand, assumed to be constant, one measurement result is not compatible with the others, it means that the measurement was not correct (for example, its measurement uncertainty was assessed as too small) or that the measured quantity has changed between the measurements. In CLSI documents, however, we still find the concept of "error" widely used. The two categories of VIM are taken as is from ISO 18113-1 under items '3.2.41 metrological comparability of measurement results' and '3.2.42 metrological compatibility of measurement results'. Unfortunately, it seems that the indications of VIM and ISO 18113 have fallen on deaf ears: No ISO document in the healthcare sector provides for the two terms. ISO 15189 does not provide definitions, but provides in point 7.3.7.4 the laboratory activities dedicated to the comparability of test results, with a description that does not correspond to the "comparability" of VIM and ISO 18113, but to "compatibility". It also slips dangerously by stating that the use of patient samples for the comparison of different test methods can avoid the difficulties related to the limited commutability of control materials, but when patient samples

are not available it calls upon all the options described for internal quality control and comparisons between laboratories. Of which there are many: Stored samples, calibrators, reference materials. ISO does not consider the operational difficulties and limitations of these solutions. All CLSI guides ignore the two definitions of VIM and ISO 18113. CLSI presents more than one guide on the topic of comparing laboratory methods, perhaps too many: EP09 2018 (Comparison with patient samples, considered the most robust), EP31 2012 Verification of comparability, interim revision but under review for approval today), EP31IG (Guide to the implementation of EP31), EP21 2016 (total measurement or "analytical" error, under review today), CLSI EP27 2022 (error grid). CLSI EP39 2021 (surrogate samples) is also added, where there are references to comparing methods. The topic of using materials other than those of patients is very slippery. Due to the difficulty of obtaining human samples, the temptation to turn to specially prepared materials is strong and the commercial proposals are incessant. VIM 5.13 defines reference material as usable for the evaluation of measurement accuracy and for quality control. But VIM 5.15 (commutability of a reference material) adds that the manufacturer's working calibrators and the end-user's calibrators do not have to be commutable. CLSI EP31 introduces the category of "Commutable reference and control materials", suitable for comparability tests. It is intended to complement EP35 (surrogate samples), which applies to both quantitative measurement procedures and qualitative tests and refers to EP14 for commutability. It should be noted, however, that EP31 with EP35 are intended for medical laboratories, while EP14 is intended for manufacturers of materials and suppliers of laboratory comparison programs. Although nothing prohibits its use in laboratories. Apart from H62 (flow cytometry), which requires the similarity of controls to patient samples, no guidance requires the commutability of control materials. The CLSI proposals, apart from EP09, present significant operational difficulties, outside the scope of this note, already identified in the 2016 SIPMeL Recommendations, in the recent ones and subject to critical comments on the ongoing revisions [13,14].

DISCUSSSION

Harmonization, meetings or clashes on the vocabulary of measures in medical laboratories

If medical laboratories are the largest measurement industry in the world, in their field the widest confrontation takes place between the principles of metrology, represented by words and the concreteness of reality, made of inertia, prejudices, misunderstandings, divergent interests. The scientific community may have a legitimate objective of harmonizing glossaries. ISO 860:2007 (Terminology Work-Harmonization of concepts and terms) tells us that despite all efforts to coordinate terminologies as they develop, it is inevitable that overlapping and inconsistent terminologies will continue to be used because documents and policies are produced in different contexts. Differences between concepts and misleading similarities in naming create obstacles to communication. Concepts and terms develop differently in individual languages and speech communities, depending on professional, technical, scientific, social, economic, linguistic, cultural or other factors. Harmonization is therefore desirable because the differences between concepts may not correspond to the names (the example case of "sensitivity"), similarity at the level of denomination does not necessarily mean that the concepts underlying the denominations are identical (the example case of "error" and "uncertainty"), errors occur when the same concept is designated by two synonyms that, by mistake, are considered as two different concepts (the example case of "sampling", "aliquot" and "sample" or "analysis", "examination" and "measurement"). The normative references of ISO 860 are ISO 704:2000 1), Terminological Work-Principles and methods, ISO 1087-1:2000 2), Terminology-Vocabulary-Part 1: Theory and application, ISO 10241:1992, International terminological standards-preparation and structuring. ISO 860 defines (3.4) harmonization of terms as the activity that leads to the selection of names for a harmonized concept both in different languages and within the same language. Terms harmonized between different languages are equivalent terms; terms harmonized within the same language are synonyms or variants of terms. The advantages of harmonization according to the Metrology Vocabulary would be numerous and significant. In our experience, harmonization of terms promotes the cleanliness of concepts understood by practitioners, reduces ambiguities, even improves the elegance and simplicity of sentences in standards and guidelines. In some ISO and CLSI documents, phenomena of magnetic attraction towards common words (such as "analysis" and derivatives) are noted without any consideration of the true metrological meaning, with the result of filling the texts with repetitions, redundancies, cacophonies, ambiguities. The phenomenon affects standards for medical laboratories but also those for forensic laboratories.

The distortion of vocabulary can have some negative consequences, effects unsuspected for simple "words", yet powerful. Like that of distancing from good practices, such as the estimation of measurement uncertainty according to the ISO standard. Or that of trapping medical laboratory operators in expensive practices, therefore reserved for a few, but scarcely effective. Like the obsessive pursuit of the comparison between methods in the same laboratory or in the articulations of the laboratory testing service in the same health company, in specialist laboratories or in POCT. Or the creation of "dialects" for the use and consumption of categories of operators, wrongly considered not educated enough, where traditional use is maintained, right or wrong as it may be. A "descriptivist" approach that contrasts with the function of "normative" type standards, therefore committed to changing practices, including linguistic ones, where necessary and possible, with the necessary caution and due respect.

CONCLUSION

The ISO and CLSI working groups show, alongside moments of openness, phenomena of true self-referential conservatism. Scientific progress is based on specialization: From time to time, chemistry, technology, medicine, engineering, metrology, glottology. Some specialists spend time and energy studying

etymologies, meanings and possible ambiguities. The linguistic proposals of the Vocabularies can be criticized and even rejected, but always with solid motivations based on objective elements: 'Measurement' and 'analysis' have different etymologies and very different meanings.

ACKNOWLEDGMENTS

The Author acknowledge the support given by SIPMeL and by WASPaLM in access discussions on ISO and CLSI documents.

REFERENCES

1. ISO 15189: 2012 Medical laboratories-requirements for quality and competence. 2020.
2. Balazs A. International vocabulary of metrology-basic and general concepts and associated terms. Chem Int. 2008.
3. 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):913-935.
4. Ellison S. Validation of qualitative and semi-quantitative methods. 2016.
5. Jones GRD. Using analytical performance specifications in a medical laboratory. Clin Chem Lab Med. 2024;62(8):1512-1519.
6. Sandberg S, Fraser CG, Horvath AR, Jansen R, Jones G, Oosterhuis W, et al. Defining analytical performance specifications: Consensus statement from the 1st Strategic Conference of the European Federation of Clinical Chemistry and Laboratory Medicine. Clin Chem Lab Med. 2015;53(6): 833-835.
7. Mari L, Pradella M, The terminology of metrology: Reflections, with particular reference to the case of medical laboratories. Tutto Misure. 2023;4:77-79.
8. Mari L. Evolution of 30 years of the International Vocabulary of Metrology (VIM). Metrologia. 2014;52(1):R1.
9. ISO/TR 27877:2021 Statistical analysis for evaluating the precision of binary measurement methods and their results. 2021.
10. FDA-2020-D-0957 FDA statistical guidance on reporting results from studies evaluating diagnostic tests. 2020.
11. Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on *in vitro* diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU (Text with EEA relevance). 2017.
12. Shreffler J, Huecker MR. Diagnostic testing accuracy: Sensitivity, specificity, predictive values and likelihood ratios. StatPearls. 2024.
13. Pradella, M., Cesana, B.M. Guidelines for the comparison of laboratory examination procedures: Use of indications of CLSI EP09-A3 and EP31-A-IR. Riv Ital Med Lab. 2016;12:26-35.
14. Pradella M. SIPMeL recommendations for ISO accreditation of the validity monitoring of exam results. Riv Ital Med Lab. 2024.