Eurachem
A focus for analytical chemistry in Europe

Terminology in Analytical Measurement

Introduction to VIM 3

Second Edition 2019

Eurachem

A focus for analytical chemistry in Europe

Eurachem Guide

Terminology in Analytical Measurement

Introduction to VIM 3

Second edition

Acknowledgements

This document has been produced by Members of the Eurachem Education and Training Working Group and others co-opted to the Project Group for this task. Those who have contributed to this version are listed below.

Project group

Fatma Akçdağ	TUBITAK Ulusal Metroloji Enstitusu (UME), TK
Vicki Barwick	LGC, UK
Mine Bilsel	TUBITAK Ulusal Metroloji Enstitusu (UME), TK
Oktay Cankur	TUBITAK Ulusal Metroloji Enstitusu (UME), TK
Michael Koch	University of Stuttgart, DE
Bertil Magnusson	Trollboken AB, SE
David Milde	Palacký University in Olomouc, CZ
Ulf Örnemark	Emendo Dokumentgranskning, SE
Marina Patriarca	Istituto Superiore di Sanità, IT
Elizabeth Prichard	UK
Eskil Sahlin	RISE Research Institutes of Sweden, SE
Lorens Sibbesen	Lab Quality, Denmark, DK
Michela Sega	Istituto Nazionale di Ricerca Metrologica, IT
Kyriacos Tsimillis	Pancyprian Union of Chemists, CY
Perihan Yolcı Ömeroglu	Uludag University, TK

Recommended citation

This publication should be cited* as: V. J. Barwick (Ed.), Eurachem Guide: Terminology in Analytical Measurement – Introduction to VIM 3 (2nd ed. 2019). ISBN 978-0-948926-36-5. Available from www.eurachem.org.

Subject to journal requirements

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64

78
79
80
81
82
83

65 Terminology in Analytical Measurement – Introduction to VIM 3

66
67 English edition

68
69 Second edition 2019

70
71 ISBN 978-0-948926-36-5

72
73 Copyright © 2019

74

Contents

Foreword to the second edition

In line with the Eurachem policy on the development and maintenance of guides, a review of the first edition of this Guide was initiated in 2016, five years after its original publication. The review involved a survey of users of the Guide and, while the feedback way very positive, the responses indicated that there was some scope for revision. In addition, since the publication of the first edition, a number of the key documents used in analytical measurement such as the ISO/IEC 17025 standard have been revised. In this second edition the scope and structure, and the terms and concepts discussed, remain unchanged from the first edition. However all sections have been reviewed and, where necessary, the text has been revised to improve clarity and ensure consistency with current guidance.

Abbreviations and symbols

The following abbreviations, acronyms and symbols occur in this Guide.

Abbreviations and acronyms

BIPM	International Bureau of Weights and Measures
CCQM	Consultative Committee for Amount of Substance – Metrology in Chemistry
CGPM	The General Conference on Weights and Measures
CITAC	Cooperation on International Traceability in Analytical Chemistry
CRM	certified reference material
ERM$^®$	European Reference Material
GC-FID	gas chromatography-flame ionisation detector
GC-MS	gas chromatography-mass spectrometry
IEC	International Electrotechnical Commission
IFCC	International Federation of Clinical Chemistry and Laboratory Medicine
ISO	International Organization for Standardization
IUPAC	International Union of Pure and Applied Chemistry
JCGM	Joint Committee for Guides in Metrology
JCTLM	Joint Committee for Traceability in Laboratory Medicine
LC-MS	liquid chromatography-mass spectrometry
LOD	limit of detection
LOQ	limit of quantification/quantitation
NIST	National Institute of Standards and Technology (USA)
NMI	National Metrology (or Measurement) Institute
p,p'-DDE	p,p'-dichlorodiphenyldichloroethylene
PT	proficiency testing
RM	reference material
SI	International System of Units
SOP	standard operating procedure
SRM$^®$	Standard Reference Material (NIST registered trademark)
VIM	International vocabulary of metrology – Basic and general concepts and associated terms (VIM)
VSMOW	Vienna Standard Mean Ocean Water
WHO	World Health Organization
XRF	x-ray fluorescence

Symbols

α	probability for a Type I error
β	probability for a Type II error
ρ	mass concentration
k	coverage factor used to calculate expanded (measurement) uncertainty
s	standard deviation
u	standard (measurement) uncertainty
u_c	combined standard (measurement) uncertainty
U	expanded (measurement) uncertainty

Introduction and scope

Introduction

In the world of metrology – the science of measurement and its application – there is a language which has to be learned. The International Vocabulary of Metrology (VIM) was produced to provide a common language, primarily for physical measurements. The third edition (International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (called VIM 3 in this document)) was produced by WG2 of the Joint Committee for Guides in Metrology (JCGM) and published as JCGM 200:2008 and as ISO/IEC Guide 99 [1]. A revised version of JCGM 200 was published in 2012 which is available free of charge from the BIPM website [2]. This revision included only minor corrections which do not affect the concepts described in this Guide, however ISO/IEC Guide 99 has not been updated. JCGM have also published an online annotated version of the VIM [3]. There are many differences between VIM 3 and earlier editions; one important change is captured in the title with the addition of the word 'concepts'. VIM 3 is a consistent set of concepts each described by a unique term – the 'label' of the concept. It is applicable across **all** scientific disciplines thus making it relevant to those involved in performing measurements in chemistry and biology. Consistent definitions of concepts with their associated terms and symbols are essential if analysts and customers across the globe are to understand each other.

Scientists from different sectors often attach different words to the same concept, which can make interdisciplinary conversations difficult. There needs to be a common language that is clear and unambiguous. The first thing one does when learning any new language is to acquire a vocabulary, which will grow with time. Anyone learning a new language will know that there are often peculiarities – from words that can have different meanings depending on the context, to words that sound the same but are spelled differently and obviously have a different meaning. In English the word 'standard' has often been cited as an example of a word with many meanings. A detailed knowledge of the language is required so as to avoid such words causing mistakes, especially for people who do not have English as a first language. Ambiguous terminology also becomes a problem for translators and can be an indirect barrier to trade.

So why do we need a guide to VIM 3 for analytical scientists? First, VIM is a normative reference in a number of International Standards and Guides that underpin accreditation, including ISO/IEC 17025 [4], ISO 15189 [5], ISO/IEC 17043 [6], ISO Guide 33 [7] and ISO Guide 35 [8]. It is also referenced in ISO 9000 [9] and ISO 17034 [10]. Second, those involved in education and training have realised that there is often confusion about both concepts and terminology. In addition the definitions are often written in a language that is difficult to understand. This is true for analysts even when the definitions are translated from English or French to the local language. Third, in VIM 3 there are some substantial changes to terminology in an attempt to accommodate chemical and biological measurements. Fourth, to make VIM 3 more accessible to analysts working in these sectors there is a need to provide context and additional examples which relate the concepts to chemical and biochemical measurements.

All languages use some words in several different ways which adds confusion when conversations are between different nationalities. As mentioned previously, the word 'standard' is one example in English but a more subtle case is the use of the word 'quantity'. In conversation we may say, e.g. 'the quantity of sample is 5 g'. This may be acceptable in daily life. However, the VIM 3 usage of the term is more specific. What we should say and write is, 'the mass of sample is 5 g'. In metrology quantity is not a synonym for amount. Quantity is a generic concept for things we measure, e.g. length, mass, time and amount-of-substance concentration. Validation and verification is another pair of words that have a changed definition in VIM 3 from what is generally used in analytical laboratories, although the actions in the laboratory to carry out these activities will be exactly the same.

This Eurachem Guide gives an explanation of selected concepts and provides examples over and above those in the Notes accompanying the definitions in VIM 3. The words that are defined in VIM 3 are highlighted and a VIM 3 reference number is provided for the concept. In VIM 3 the relations between concepts is displayed in 12 diagrams which have been used to help group concepts into families in this Guide. How the terms and definitions relating to these concepts are linked to each other, either within a family or between families, is illustrated in this Guide. The concepts which appear in this Guide are listed in Table A1 in the Appendix and in the text are organised into the following chapters; General Metrology, Metrological Traceability, Measurement Uncertainty, and Verification, Validation and Method Performance.

Scope

The scope of this Eurachem Guide is to cover a selection of the concepts in VIM 3, focusing on those most likely to be encountered in analytical laboratories. It aims to cover chemical, biological and clinical measurements. This Guide is intended for laboratory staff, accreditation bodies, for those commissioning measurements and for those using measurement results. Lecturers and trainers may also find this Guide useful when teaching aspects of metrology.

Notes for the reader

All concepts defined in VIM 3 appear in **bold** in the text. The terms discussed in this Guide are listed in Table A1 in the Appendix. If the full VIM 3 definition is included in this Guide, the VIM 3 reference number is given in the text box where the concept is defined and is not given each time the term is used in the text. When other VIM 3 terms are used in the text without a definition, the VIM 3 reference number is given the first time the term occurs in a section. VIM 3 permits multiple (often shorter) terms for the same concept; these are also included in the Appendix. If more than one term is given in VIM 3, the first term is the preferred one and it is used in this Guide as far as possible. However, a shorter alternative is used where it improves the readability of the text.

Single quotation marks (' ') are used both for emphasis and for quotations. The latter are always referenced. The decimal sign is the point on the line. For the word standard, an upper case S is used when it refers to a norm, e.g. the International Standard ISO/IEC 17025. When the word vocabulary refers to VIM 3 or previous editions, an upper case V is used.

The generic term 'concentration' is used on its own, i.e. unqualified, when a generality is required. It represents the family of quantities which includes, *mass* concentration, *amount* concentration, *number* concentration or *volume* concentration. Note that many other quantities used to express composition, such as mass fraction, substance content and mole fraction, can be directly related to concentration.

It is accepted that the metre is the SI base unit of length, and that volume should be expressed in m^3 and multiples or submultiples of this, i.e. 1 litre = 1 dm^3. Since litre is an accepted unit it is used in this Guide and is represented by L [11].

A key concept in this Guide is 'measurement'. However, the core activity of an analytical laboratory is often referred to using wording such as 'analysis', 'test', 'examination' or 'determination'. Unlike 'measurement' which is a quantitative feature, these terms are frequently used, and understood, to involve a qualitative and/or a quantitative aspect.

1 General metrology

1.1 Metrology

> science of **measurement** and its application (VIM 2.2)

Metrology covers all theoretical and practical aspects of **measurement** in all sectors, including routine **measurement**. It applies in analytical science, biological and clinical **measurement**, whatever the relative magnitude of the **measurement uncertainty** of the result.

1.2 Quantity

> property of a phenomenon, body, or substance, where the property has a magnitude that can be expressed as a number and a reference (VIM 1.1)

Quantity is a crucial concept in **metrology**, which applies across all disciplines involved with **measurement**, and is therefore the first term defined in VIM 3 [1]. The definition identifies a **quantity** as any property which has size (magnitude) that can be evaluated through **measurement**. Some of the terms related to **quantity** are shown in Figure 1.

There are many kinds of **quantity** including mass, volume, velocity (speed), electric current and flow. In everyday life, we are interested in specific examples of such **quantities** (formerly referred to as 'particular quantities') [12], e.g. the volume of gasoline dispensed into a vehicle, the speed at which my car was travelling when the police stopped me, or the number concentration of red cells in the blood sample taken yesterday from Mr. Smith.

The specification of the (particular) **quantity** we intend to measure (also called the **measurand**) is the first part of any **measurement**.

1.3 Nominal property

> property of a phenomenon, body, or substance, where the property has no magnitude (VIM 1.30)

The current definition of **quantity** clearly excludes properties that, although carrying valuable information, can only be described in words. Examples include the colour of a spot test in chemistry (e.g. a home pregnancy test), and molecular sequences (e.g. of amino acids in a polypeptide, of nucleotides in a DNA fragment). Such important properties, that have no magnitude, are however acknowledged in VIM 3 and described with the term **nominal property**. In analytical chemistry the term, *qualitative analysis* is often used to describe the examination of **nominal properties**.

It is possible to measure a **quantity** (see the definition of **measurement**), whereas obtaining information about a **nominal property** is not a **measurement**. The term *examination* is appropriate [13]. However, in ISO 15189 the term 'examination' is used for both the determination of **nominal property values** and for **measurement** [5].

1.4 Quantity value

> number and reference together expressing magnitude of a **quantity** (VIM 1.19)

The size (magnitude) of a **quantity** is expressed as a number accompanied by a **measurement unit** and – if appropriate – by additional reference to a **measurement procedure** or a **reference material (RM)** (VIM 5.13).
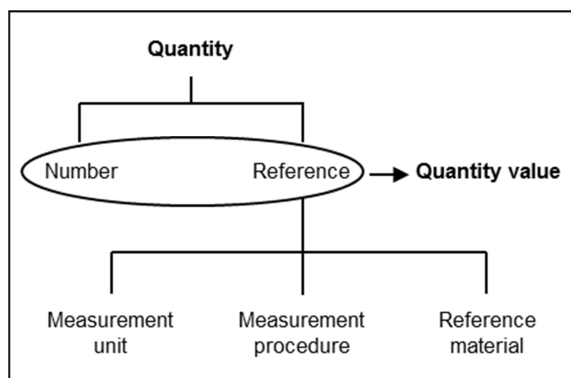
Figure 1 – Some terms related to Quantity.

Consider two examples from the field of chemical chemistry, both concerned with the **measurement** of lead in paint.

(i) A laboratory is required to determine the total mass concentration of lead in a paint sample. This is reported as $10\ mg\ L^{-1}$. In this case the specific **quantity** (the **measurand**) is the total mass concentration of lead. The **quantity value** is $10\ mg\ L^{-1}$, where 10 is the number and $mg\ L^{-1}$ (the **measurement unit**) is the reference.

(ii) A laboratory is required to determine the mass concentration of lead extracted from the paint on a toy following the **measurement procedure** described in European Standard EN 71-3 'Safety of Toys. Migration of certain elements' [14]. Since the amount of extracted lead is strictly dependent on the treatment applied (e.g. solvent, time, temperature), the **measurand** is defined by the **measurement procedure** applied (often referred to as an 'empirical method') and is therefore called an 'operationally defined **measurand**' (see section 1.11). A different **measurement procedure** would probably give a different result. Therefore, the 10 is still the number but the reference is both the **measurement unit** (mg L^{-1}) and the **measurement procedure** used.

The third case shown in Figure 1 is where the reference is the value of an **RM**. This can be illustrated using an example from laboratory medicine. The **quantity value** of the activity of a batch of an anticoagulant factor, Factor VIII, extracted from human blood, is referred to the value of the corresponding **RM**, periodically prepared and approved by WHO and known as a WHO International standard [15].

1.5 Nominal quantity value

> rounded or approximate **value** of a characterizing **quantity** of a **measuring instrument** or **measuring system** that provides guidance for its appropriate use (VIM 4.6)

In VIM 3 the word 'nominal' is used in a different sense in **nominal quantity value** and **nominal property** (see section 1.3).

A volumetric flask may be marked 100 mL, this is its **nominal quantity value** (or, more simply, the **nominal value**). The actual value of the volume of the flask may not be exactly 100.00 mL, but will fall in a range according to the glassware class. For example, if a class A 100 mL volumetric flask has a tolerance of 0.08 mL the actual volume will lie in the interval 99.92 mL to 100.08 mL.

1.6 Reference quantity value

> **quantity value** used as a basis for comparison with values of **quantities** of the same **kind** (VIM 5.18)

Many different types of materials and devices may have a **reference quantity value** and an associated **measurement uncertainty**. Some examples are:

- The **quantity value** given on the certificate of a **certified reference material (CRM)** (VIM 5.14) with its associated **measurement uncertainty** is a **reference quantity value** for the particular property to which it relates.

- The values of a set of solutions of known concentration, analysed to build a **calibration diagram** (VIM 4.30), are **reference quantity values** used for determining the value of the same **quantity** in other samples.

- In a calibration laboratory, mercury-in-glass thermometers are calibrated against a **measurement standard** (thermometer) reproducing specific temperature values (with associated **uncertainties**): these are **reference values** of the **quantity** 'temperature'.

- Analysts use the value assigned to a **CRM** as a **reference quantity value** for the assessment of the **trueness** of a **measurement procedure**.

- In order to assess the competence of staff and laboratories, the staff may be required to analyse samples which have assigned values. The value assigned to a sample may be a **quantity value** obtained either from previous analyses by more experienced staff/laboratories or from PT rounds, or from a certificate if the sample is a **CRM**. In this context, the value assigned to any of these materials is deemed to be a **reference quantity value**.

1.7 System of quantities

> set of **quantities** together with a set of non-contradictory equations relating those quantities (VIM 1.3)

In practice, it is useful to identify a set of **quantities** from which all other **quantities** can be derived. Such a set is a **system of quantities**.

In any **system of quantities** the **base quantities** (VIM 1.4) that constitute the set are, by definition, considered to be mutually independent – they cannot be described as a product of other **base quantities**. The choice of these **quantities** is by convention. Other choices are equally valid, provided that they satisfy the definition.

However, a specific **system of quantities** has been agreed and adopted. The Metre Convention established a permanent organisational structure for member governments to act in common accord on all matters relating to **units of measurement**. It led to the creation of the International Bureau of Weights and Measures (BIPM). The seven **base**

quantities, which are agreed and defined by the General Conference on Weights and Measures (CGPM), are shown in Table 1. This is called the **International System of Quantities** [11].

1.8 International System of Quantities

> **system of quantities** based on the seven **base quantities**: length, mass, time, electric current, thermodynamic temperature, amount of substance, and luminous intensity (VIM 1.6)

The definition of the units corresponding to the **base quantities** (VIM 1.4) is extremely important since they provide the foundation for the entire **system of units** (VIM 1.13). The **International System of Units** (VIM 1.16), the **SI**, is recognised by many countries (at the time of writing there are 59 members states of the Metre Convention) and is adopted as the only legal **system of units** within the European Union [11, 16]. The **base quantitie**s and their corresponding **base units** (VIM 1.10) are shown in Table 1.

In 2018 it was agreed that four of these **base units** would be redefined – the kilogram, the mole, the ampere and the kelvin. Following the adoption of the revised definitions in May 2019, all the **base units** are defined in terms of fundamental constants which are unchanged with respect to time and location. By definition, these fundamental constants have no **uncertainty**. The kilogram is defined in terms of the Planck constant $h = 6.626\,070\,15 \times 10^{-34}$ J s and the mole as the specific number of elementary entities (atoms, molecules, ions, etc) given by the Avogadro constant $N_A = 6.022\,140\,76 \times 10^{23}$ mol^{-1}. These changes do not affect the way routine **measurements** are made and **metrological traceability** is obtained exactly as before. It is only the definition of the **units** that has changed.

Many other **quantities** within the **SI** are expressed as relations between those shown in Table 1 and are called **derived quantities** (VIM 1.5). The definitions of the **derived units** (VIM 1.11) in terms of the **base units** follow from the equations defining the **derived quantities** in terms of the **base quantities**. For example, the **derived quantity** mass density is:

$$\text{mass density} = \frac{\text{mass}}{(\text{length})^3}$$

The **measurement unit** (**derived unit**) is obtained by applying the same formula to the **units**, i.e. $\frac{\text{kg}}{\text{m}^3}$ which is usually written as kg m^{-3} or kg/m^3.

Table 1 – Base quantities and base units.

Base quantity	Base unit (symbol)
length	metre (m)
mass	kilogram (kg)
time	second (s)
electric current	ampere (A)
thermodynamic temperature	kelvin (K)
amount of substance	mole (mol)
luminous intensity	candela (cd)

A common mistake is to confuse **quantities** and their **measurement units**. It should be kept in mind that whereas a **quantity** is a measurable property of a phenomenon, body or substance (e.g. mass), a **measurement unit** (e.g. the kilogram) is chosen by convention as the reference to which **measurements** of that property refer.

1.9 Measurement unit

> real scalar **quantity**, defined and adopted by convention, with which any other quantity of the same **kind** can be compared to express the ratio of the two quantities as a number (VIM 1.9)

We are all familiar with the concept of a **measurement unit**; the method of pricing many products is by showing the cost per agreed **unit**, e.g. food as cost per kg, price of gasoline quoted per litre (L). If we say the mass of an apple is 0.15 kg, this means that the mass of the apple is $0.15 \times$ mass of the kilogram, that is, the **measurement unit**. To obtain the number 0.15 you compare the value indicated for the apple with that indicated for a reference mass, i.e. the mass used to calibrate the balance. The reference mass in turn is compared with a practical realisation of the definition of the kilogram by national metrology institutes (NMIs), or by calibration or testing laboratories. The result of any such comparison is expressed as a ratio of the **indication** obtained to the value of a **quantity** of the same **kind** (VIM 1.2).

1.9.1 Quantities of the same kind

The classification of **quantities** of the same **kind** (VIM 1.2) is somewhat arbitrary, but the concept that only similar items can be compared is well understood. **Quantities** of the same **kind** will have

the same **unit** but two **quantity values** having the same **unit** do not have to be of the same **kind.** The **unit** of mass density and of mass concentration is kg m^{-3} but these are not **quantities** of the same **kind**. The **measurement unit** of both frequency and activity of radionuclides is s^{-1} but they are not **quantities** of the same **kind**. In this example the **unit** in each case is given a special name, namely hertz (Hz) and becquerel (Bq), respectively.

The example of the mass of the apple was easy because **quantities** of the same **kind** were compared. Sometimes it is not possible to obtain the **quantity value** by comparison with a **quantity** of the same **kind**, e.g. because of an incomplete understanding of the **measurand** or the complexity of the factors influencing the **measurement** process and its result. An example of such a **measurand** is the mass fraction of fibre in a food product (the 'fibre content'). However, it is still possible to compare results for such **measurands**, provided they are obtained using identical or proven equivalent **measurement procedures**. In such cases, reference must be made to which **measurement procedure** (including details such as reagent grade, **calibrator**, etc.) was used to obtain the **quantity value**.

1.10 Measurement

> process of experimentally obtaining one or more **quantity values** that can reasonably be attributed to a **quantity** (VIM 2.1)

A **measurement** is a series of actions (steps, stages) taking place in a defined manner (i.e. following a **measurement procedure**). Some **measurements** are a single step, others have many stages. There is potential for controversy as some regard **measurement** as the instrument response, often the last stage in a multi-stage process, e.g. for an aliquot of sample extract. What is clear is that **measurement** relates to the *whole* process of obtaining a **quantity value** and should not be used to refer to the numerical value obtained.

1.10.1 What is a 'measurement' and what is not?

In analytical sciences, a test sample submitted for analysis often undergoes a series of chemical and/or physical treatments in order to convert it to a form that can be presented to a **measuring instrument**. These steps are acknowledged to be part of the **measurement** process. In some cases there may be a particular sampling procedure included in the process.

A **quantity value** is expressed as a number and a reference, expressing the magnitude of the **quantity**. Does this mean that the procedure of counting items is a **measurement**? The answer is yes, because the result is quantitative and the reference is the counting procedure. However, visually inspecting a sample to note its colour is not a **measurement**, but an 'examination', as **measurement** does not apply to **nominal properties**. In contrast, using a spectrophotometer to record some property relating to the colour of a sample (e.g. absorbance at a particular wavelength) is a **measurement**.

1.10.2 Preliminary to making a measurement

Before making a **measurement**, the **quantity** must be clearly defined, bearing in mind the purpose for which the experimental result is required. The **quantity** referred to is the **measurand**. In addition, for **measurement results** to be fit for purpose, a validated **measurement procedure** needs to be available and must be applied using a calibrated **measuring system**. In this context 'fit for purpose' means that, when the **measurement procedure** is applied, the quantity intended to be measured is measured, and the **uncertainty** in the **measurement results** is acceptable. The concept of **target measurement uncertainty** (VIM 2.34) (see section 3.1.2) is used to describe the maximum **measurement uncertainty** that can be accepted by the customer for a specific application [17].

1.11 Measurand

> **quantity** intended to be measured (VIM 2.3)

There is a great deal behind this apparently simple definition. The **measurand** is a description of the specific **quantity** we intend to measure. The specification of the **measurand** should be sufficiently detailed to avoid any ambiguity. **Measurand** is not another name for analyte. Analyte is the component represented in the name of a measurable **quantity**, whereas **measurand** refers to a specific **quantity** to which **quantity values** are expected to be attributed by means of a **measurement**. Consider two examples of **quantities** which may be measured:

• mass of protein in a 24-hour urine collection;

1 • amount-of-substance concentration of glucose in
2 plasma.

3 In each case the complete statement represents the
4 **measurand**. The analytes are protein and glucose,
5 respectively.

6 The specification of the **measurand** is critical, in
7 order that the **measurement result** is suitable for its
8 intended use, and should include all important
9 parameters and conditions. For example, if the
10 volume of the liquid delivered by a pipette is to be
11 determined by weighing, the specification of the
12 **measurand** should at least include the type of liquid
13 to be used and the temperature at which the
14 **measurements** should be carried out. In chemical
15 and biological analysis the specification of the
16 **measurand** requires at least the description of the
17 **quantity** (e.g. mass fraction or amount-of-substance
18 concentration), the analyte and where relevant the
19 matrix, even if it is not possible to give a clear
20 chemical definition of the analyte, for example:

21 • mass fraction (e.g. mg kg^{-1}) of cadmium in soil;

22 • amount-of-substance concentration (e.g. mol L^{-1})
23 of total cholesterol in blood serum;

24 • mass fraction (e.g. mg kg^{-1}) of extractable fat in
25 a meat sample.

26 In the **measurement** of the mass fraction of
27 cadmium in a soil sample, the sample drying
28 conditions (e.g. dried to constant mass at
29 (105 ± 5) °C) should be included in the definition of
30 the **measurand** as they have an influence on the
31 basis for reporting results. It may be necessary to
32 specify the **measurement procedure** in even more
33 detail and define whether the **measurement result**
34 will be referring to the laboratory sample or the
35 whole bulk (e.g. a batch of animal feeding stuff,
36 whole lake). In other cases, the **measurand** can
37 only be defined with reference to an agreed
38 empirical **measurement procedure** (operationally
39 defined **measurand**. For example the **measurement**
40 of extractable fat in a sample of meat will depend
41 strongly on the solvent used and the conditions of
42 extraction. Such 'operationally defined'
43 **measurands** are still fit for the purpose of
44 comparing results and making decisions provided
45 that the agreed **measurement procedures** are
46 strictly followed.

47 ## 1.12 Measurement procedure

48
> detailed description of a **measurement** according
> to one or more **measurement principles** and to a
> given **measurement method,** based on a
> **measurement model** and including any calculation
> to obtain a **measurement result** (VIM 2.6)

49 The description of how **measurements** are
50 performed involves several levels of detail, with the
51 most comprehensive being the **measurement**
52 **procedure**, which encompasses all others.

53 Performing **measurements** requires an understanding
54 of the **measurement principle** (VIM 2.4), that is of
55 the phenomenon underlying the **measurement**. The
56 same **measurement principle** can be applied
57 according to different **measurement methods**
58 (VIM 2.5), e.g. using different techniques (such as
59 flame or electrothermal atomic absorption
60 spectrometry), or different **calibration** procedures
61 (external **calibration** or by the 'method of standard
62 additions'). The **measurement method** is a generic
63 description of the operations involved. The
64 following are examples of **measurement methods**,
65 with the **measurement principle** given in brackets:

66 • Determining by weighing the amount of a chemical
67 compound precipitated from a liquid test sample
68 using a defined chemical reaction (gravimetry).

69 • Determining the amount-of-substance
70 concentration of a compound in a given sample,
71 either directly, by measuring its absorbance at a
72 given wavelength, or indirectly, by measuring a so-
73 called 'surrogate quantity', such as the absorbance
74 of a complex formed as a result of a defined
75 chemical reaction (spectrophotometry).

76 • Determining the amount-of-substance
77 concentration of a compound by means of its
78 ability to become permanently linked to a specific
79 antibody carrying a tag (immunochemistry).

80 Although not defined in VIM 3, the term
81 'measurement process' is used in a number of
82 international Standards. It is defined in ISO
83 9000 [9] as 'a set of operations to determine the
84 value of a quantity'. It is the overall process of
85 planning, performing and evaluating **measurements**
86 and, as such, the **measurement** principle, method
87 and procedure form part of the **measurement**
88 process.

89 The most complete level of description of a
90 **measurement** is the **measurement procedure**
91 which should be sufficiently detailed to allow a
92 suitably trained person to perform the
93 **measurement**. In some laboratories the

measurement procedure may be documented in one or more standard operating procedures (SOP). Although ISO/IEC 17025 [4] refers to **measurement procedures** in some clauses, 'method' and 'test method' are also used and considered synonymous with **measurement procedure** as defined in VIM 3. However, it should be noted that the requirements of the Standard apply to both **measurements** and examinations. As mentioned earlier, ISO 15189 [5] uses the term 'examination' to refer both to the determination of **nominal properties** and to **measurement procedures**. In ISO/IEC 17025 the method includes, where relevant, aspects of 'sampling, handling, transport, storage and preparation of items to be tested and/or calibrated'. However, in ISO 15189, examination does not include sampling; the latter is part of the 'pre-examination' covered in clause 5.4 of that Standard.

A **measurement procedure** includes a description of how **measurement results** are obtained and reported, including any calculation. A **measurement result** is generally expressed as a single **measured quantity value** and a **measurement uncertainty**. The **measurement procedure** should therefore include an estimate of the **measurement uncertainty** to be used when reporting **measurement results** or information on how it should be calculated.

There are two types of **measurement procedure** that are included as two separate concepts; they are **reference measurement procedures** and **primary reference measurement procedures**.

1.13 Reference measurement procedure

> **measurement procedure** accepted as providing **measurement results** fit for their intended use in assessing **measurement trueness** of **measured quantity values** obtained from other measurement procedures for **quantities** of the same **kind,** in **calibration,** or in characterizing **reference materials** (VIM 2.7)

Reference measurement procedures are well characterised and will normally provide **measurement results** with a small **measurement uncertainty**. For example, in the clinical sector, to comply with the requirements of the *In vitro* Diagnostics Directive [18] manufacturers are required to use **reference measurement procedures** or **CRMs** (VIM 5.14) to establish the **metrological traceability** of values assigned to **calibrators**. The Joint Committee for Traceability in Laboratory Medicine (JCTLM) lists a number of **reference measurement procedures** [19], e.g. NIST LC-MS reference method for the determination of the mass concentration of cortisol in blood serum [20].

In the hierarchy of metrological order, the highest level is occupied by a **primary reference measurement procedure**.

1.14 Primary reference measurement procedure

> **reference measurement procedure** used to obtain a **measurement result** without relation to a **measurement standard** for a **quantity** of the same **kind** (VIM 2.8)

Primary reference measurement procedures (also known as primary methods of measurement or, more simply, 'primary methods') allow a **quantity value** to be determined with direct reference to the definition of its **measurement unit** or to fundamental constants. Because there are no other intermediate steps, they provide, under the stated conditions, **metrologically traceable measurement results** with the highest levels of **accuracy**. Examples of such procedures are the determination of the amount-of-substance concentration by coulometry, gravimetry or by isotope dilution mass spectrometry.

1.15 Measurement result

> set of **quantity values** being attributed to a **measurand** together with any other available relevant information (VIM 2.9)

The **measurement result** is the outcome of any **measurement** activity and is what is reported to the customer, be it a regulatory body, the accreditation body or a commercial client.

In the past, the term '**measurement result**' has been used to mean different things. A **measuring instrument** gives a number, i.e. an **indication**; the number can be converted into an uncorrected result using a **calibration curve** (VIM 4.31). In some cases, due to **measurement bias** (VIM 2.18), this value is corrected and the corrected result reported to the customer along with, e.g. a recovery factor with its **measurement uncertainty**. This is what constitutes a **measurement result** in the VIM 3 definition. Historically often a single number was all that was given to the customer. The VIM 3 definition aims to eliminate this lack of consistency, by clearly identifying a **measurement result** as the final

outcome of the process of determining the **quantity value(s)** of a **measurand**, i.e. providing an answer to the customer's request. In this context, all the relevant information relating to the **measurement** is also part of the **measurement result**.

A **measurement result** is generally expressed as a single **measured quantity value** and a **measurement uncertainty**. This can be interpreted as a 'set of **quantity values**', meaning that any value, within the interval defined by the **measurement uncertainty** is a possible value for the **measurand**. This provides the customer with information on the reliability of the **measurement result** which should be taken into account if, for example, it is compared with a stated limit.

The **measurement uncertainty** and the level of confidence associated with it are part of a **measurement result**. The **measurement uncertainty** may not always be explicitly reported if it is considered to be negligible in terms of interpreting the result, or if it is not relevant in the interpretation, or not required by the customer. Examples of where this normally holds true are a) the volume delivered by a petrol station pump, b) the mass of groceries weighed on a modern balance in a supermarket, and c) the examination results delivered to the physician by a hospital laboratory. However, the **uncertainty** is still taken into account, since the pump, the balance and the clinical tests must fulfil certain stated performance criteria before they can be put into use.

The requirements of ISO/IEC 17025 [4] are that information on **measurement uncertainty** should be presented in test reports when it is relevant to the validity or application of the test results, when the customer's instruction so requires or when the **uncertainty** affects conformity to a specification.

1.16 Measured quantity value

> **quantity value** representing a **measurement result** (VIM 2.10)

Measured quantity values are an essential part of a **measurement result**. In the simplest cases, e.g. when weighing bread or potatoes on a commercial weighing scale, the **measured quantity value** *is* the **measurement result** as the **measurement** is a simple single step process and no intermediate **measurements** or calculations are required. However, more often, in analytical sciences, a **measurement** involves different **quantities** and **indications**, to be combined according to the **measurement model** (VIM 2.48), to obtain the **measured quantity value**. The final value is

usually calculated as the average of the set of values obtained from repeated **measurements**, which will have a lower **measurement uncertainty** than the individual values. In many cases, the **measurand** requires more than one **measured quantity value**, obtained with separate **measurement procedures**. For example if the **measurand** is to be defined as the mass fraction of an analyte on a dry weight basis, the mass of the sample after drying and the mass fraction of the analyte of interest are both required.

1.17 Measurement error

> **measured quantity value** minus a **reference quantity value** (VIM 2.16)

No **measurement** is perfect; the very action of measuring introduces changes in the system subjected to **measurement**. It is convenient to describe this scenario in terms of a **measurement error**, affecting each individual **measurement**. In principle, the **measurement error** is represented by the difference between the **measured quantity value** and a **reference quantity value**. In practice, for a **measurement** on a test sample, the **measurement error** is unknowable. This is because, in this case, the **reference quantity value** is the unknown **true quantity value** (VIM 2.11) for the **measurand**. The **measurement error** consists of two components, **systematic measurement error** (VIM 2.17) and **random measurement error** (VIM 2.19), which represent respectively, the constant or predictable variation and the unpredictable variation in a series of replicate **measurements**. Well known parameters describing the performance of analytical methods are associated with the estimate of the random and systematic components of **measurement error** (see chapter 4).

1.18 Indication

> **quantity value** provided by a **measuring instrument** or a **measuring system** (VIM 4.1)

Most **measurements** are based on **indications** provided by **measuring instruments** or **measuring systems**. An **indication** (e.g. an instrument signal or response) and a corresponding value of the **quantity** being measured are not necessarily values of **quantities** of the same **kind** (VIM 1.2). In many cases, the **indication** provided by the **measuring instrument** or **measuring system** will be a value related to a **quantity** different from the **measurand**.

In most cases in analytical science, analysts rely on the **measurement** of physical **quantities** such as mass of precipitate, volume of titrant, or the change in electrical signal produced by the absorption of radiation of a particular wavelength. These **indications** are then converted, using well known stoichiometric relationships or a **calibration curve** (VIM 4.31), to **quantities** related to the amount of substance. In chemical analysis it is quite common to observe an instrument response for a series of reference solutions and then for the test sample, so in this context the term 'instrument response' is generally used for **indication**. Similarly **blank indication** (VIM 4.2) refers to the instrument response for a test material where the analyte of interest is believed to be absent (in VIM 3 the phrase 'supposed not to be present' is used).

1.19 Measuring instrument

> device used for making **measurements**, alone or in conjunction with one or more supplementary devices (VIM 3.1)

Measuring instrument is closely related to the concept '**measuring system**'.

1.20 Measuring system

> set of one or more **measuring instruments** and often other devices, including any reagent and supply, assembled and adapted to give information used to generate **measured quantity values** within specified intervals for **quantities** of specified **kinds** (VIM 3.2)

In some cases, when the **measuring instrument** can be used alone (e.g. mercury-in-glass thermometer) the **measuring system** consists of one **measuring instrument**. However, for the majority of analytical methods the **measuring system** consists of several **measuring instruments** and associated equipment and reagents.

VIM 3 defines three types of **measuring instruments**.

- **Indicating measuring instruments** (VIM 3.3) provide an output signal directly, for example a number (e.g. an electronic balance) or a visual/acoustic signal. The output may be transferred to another device, e.g. a computer with software providing data integration.

- **Displaying measuring instruments** (VIM 3.4), a particular type of **indicating instrument**, display the **result of a measurement** on a scale, e.g. mercury-in-glass thermometer, a spring balance. For instruments with analogue outputs, the **indication** is given by the position of a pointer on the display.

- **Material measures** (VIM 3.6) are intended to reproduce an assigned **quantity value**, e.g. volumetric flasks, standard weights or **CRMs** (VIM 5.14).

Distinguishing between these categories of **measuring instruments** is not normally a concern in analytical **measurement**.

1.21 Metrological comparability of measurement results

> comparability of **measurement results,** for **quantities** of a given **kind,** that are metrologically traceable to the same reference (VIM 2.46)

VIM 3 uses the word comparability in the sense of 'ability to compare' **not** in the sense of being 'similar in magnitude'. Hence, to be comparable the **measured quantity values** or the **measurement uncertainties** do not have to be of the same order of magnitude. For example, **measurement results** for the determination of the mass fraction of chromium in an alloy and in a sample of contaminated soil are **metrologically comparable** when they are traceable to the same **measurement unit**.

The purpose of making **measurements** is often to enable a comparison to be made between the **measurement result** obtained and another value for a **quantity** of the same **kind** (VIM 1.2), e.g. a legal limit or a reference interval. Examples of possible questions asked by a customer are, 'Is the mass fraction of lead in this sample of soil greater than the allowed limit?' or 'Is the mass fraction of lead in two samples of soil significantly different?' A question which often arises, for instance, in a legal context is, 'Are the results provided by the two laboratories different?' In order to be able to answer these questions, **measurement results** need to be metrologically comparable.

A comparison is only meaningful if the results are traceable to the same reference (preferably internationally accepted) which may be, e.g. the metre or the **quantity value** of a **CRM** (VIM 5.14).

The concept 'comparability' is associated with the concept 'compatibility'.

1.22 Metrological compatibility of measurement results

> property of a set of **measurement results** for a specified **measurand,** such that the absolute value of the difference of any pair of **measured quantity values** from two different measurement results is smaller than some chosen multiple of the **standard measurement uncertainty** of that difference (VIM 2.47)

For a set of two results to be **metrologically compatible** the difference between them should be smaller than the **expanded uncertainty** (VIM 2.35) of their difference. With knowledge of the value of the **measurement uncertainty** of the **quantity values** it is possible to calculate a permissible difference (d) between a pair of independent results which refer to the same **measurand**. A difference between x_1 and x_2 greater than d suggests a possible failure of the **measuring system**, a change in the **measurand** or that the **measurement uncertainty** of one or both results has been estimated incorrectly.

The **standard uncertainty** (VIM 2.30) of the difference $u(d)$ between two completely independent **measurement results** x_1 and x_2 (obtained, for example, from two different laboratories) is given by the equation:

$$u(d) = \sqrt{u(x_1)^2 + u(x_2)^2}$$

where $u(x_1)$ and $u(x_2)$ are the **standard uncertainties** associated with x_1 and x_2 respectively. Therefore, for two **measurement results** to be considered **metrologically compatible**, the difference d must be less than $k.u(d)$, where k is the **coverage factor** (VIM 2.38) appropriate for the required level of confidence.

For a larger set of results **metrological compatibility** cannot be determined so easily. The 'chosen multiple' in the definition would depend on the level of confidence required and the number of paired comparisons involved.

Correlation between the **measurements** influences **metrological compatibility** of **measurement results**. The **standard uncertainty** of the difference will be lower for positive correlation and higher for negative correlation.

2 Metrological traceability

This chapter describes the terminology relating to **metrological traceability**. Further information on how to establish the **traceability** of **measurement results** is given in a Eurachem/CITAC Guide [21] and an IUPAC Technical Report [22].

2.1 Metrological traceability

> property of a **measurement result** whereby the result can be related to a reference through a documented unbroken chain of **calibrations**, each contributing to the **measurement uncertainty** (VIM 2.41)

Most chemical analyses involve comparing a laboratory result with values produced at different times and locations, e.g. a value on a certificate, a legal limit, or a result obtained with a different **measurement procedure**.

Metrological traceability is essential for meaningful **measurement results** as it helps demonstrate that such comparisons are scientifically valid. It is meaningful to compare the length of a football field with the distance between street lights so long as they are both expressed in metres – the same **unit of measurement**. However, just because results are **metrologically traceable** does not mean they are fit for purpose as it does not ensure that the **measurement uncertainty** is adequate. For example, the **measurement result** obtained when weighing a certain mass of sodium chloride using a calibrated technical (2-figure) balance is **metrologically traceable** to the kilogram. This may be fit for purpose for preparing reagents such as buffers but may not be sufficiently accurate for the preparation of calibration solutions for the determination of low concentrations of sodium in water. In addition, to ensure that **measurement results** are fit for purpose, the **measurement procedures** used must be validated (see chapter 4) and adequate on-going quality control procedures must be in place.

2.1.1 Reference points

According to VIM 3 there are three types of reference (see Note 1 of the definition of **metrological traceability**).

- A **measurement unit**, e.g. $mol\,L^{-1}$, g, $mg\,kg^{-1}$, °C, $\mu kat\,L^{-1}$, through its practical realisation (see section 2.1.2).

- A **measurement procedure**, which is fully defined and internationally agreed upon, e.g. the procedures defined in the IFCC **primary reference procedure** for the **measurement** of catalytic activity concentration of alkaline phosphatase in human serum [23] or the procedures defined in the ISO Standard method for the determination of the fat content of dried milk and dried milk products [24].

- A **measurement standard**, e.g. the **CRM** (VIM 5.14) SRM® 2193a $CaCO_3$ pH standard which, when prepared according to the instructions given in the certificate, has a certified pH value of 12.645 at 20 °C with an **expanded uncertainty** (VIM 2.35) of 0.011 ($k = 2$).

For many **measurement results** the reference will be a **measurement unit** but in some cases additional metrological references, such as a **measurement procedure**, will also be required (see section 1.11). In such cases the references are used in combination.

The 'unbroken chain of calibrations' is a **traceability chain** (VIM 2.42), consisting of the sequence of **measurement standards** and **calibrations** used to relate a **measurement result** to a reference. A generic flow chart of **metrological traceability** is shown in Figure 2. The direction of increasing **measurement uncertainty** and the **calibration hierarchy** (VIM 2.40) are illustrated. The **calibration hierarchy** is a sequence of **calibrations** from the chosen reference to the final **measuring system** where the outcome of each **calibration** depends on the outcome of the previous **calibration**. The **traceability chain** is defined by the chosen **calibration hierarchy**.

The result should always be traceable to an appropriate reference point and accredited laboratories must be able to demonstrate this. The laboratory can draw its own **traceability chains** by studying the documentation for its routine procedures, equipment and **calibrators**. Examples of generic **traceability chains** can be found in the Standard ISO 17511 [25]. The IUPAC Technical Report on establishing traceability in chemistry contains the following seven illustrated examples of **traceability chains** [22]:

- Amount-of-substance concentration of an acid in a solution;

- pH of a solution;

- mass concentration of ethanol in breath;

- number-ratio of isotopes of an element in a material;

- mass fraction of glyphosate in an agricultural chemical;

- amount-of-substance concentration of creatininium[*] in blood plasma;

- mass fraction of protein in grain.

2.1.2 Practical realisation of a measurement unit

In the case of **metrological traceability to a measurement unit** (VIM 2.43), the reference is the definition of a **unit** through its practical realisation. What does this mean in practice? The realisation of the definition of a **unit** is the procedure by which the definition may be used to establish the **value**, and associated **measurement uncertainty**, of a **quantity** of the same **kind** (VIM 1.2) as the **unit**. Mass and amount of substance are two **base quantities** (VIM 1.4). They correspond to the **base units** (VIM 1.10) of the kilogram and the mole respectively. Although the definition of the kilogram is changed and there is no longer an international protype kilogram (see section 1.8) the realisation of the **unit** will still be made using well defined **measurement procedures** using, e.g. a Kibble balance [11]. The **measurement unit**, or multiples of it, are embodied in calibrated weights. The embodiment (realisation) is achieved by **measurement** using **a primary reference measurement procedure** and a **measuring system** to assign a **quantity value** and a **measurement uncertainty**.

The revised definition of the mole is discussed in section 1.8. A common realisation of the mole is achieved through weighing. The amount of substance n in a pure sample is measured by determining the mass m of the sample and dividing by the molar mass M according to: $n = \frac{m}{M}$. This approach is only possible when the chemical entity or entities, specified in a **measurand**, can be defined. If this is not the case then amount of substance cannot be measured. In such cases, other **quantities**, such as mass, which do not need entities to be specified, can be chosen. To report **measurement results** in the **SI** (VIM 1.16) **unit** mole, the embodiment of the definition of the mole

would require a **primary measurement standard** (VIM 5.4) for each of the millions of chemical compounds. To overcome this problem the Consultative Committee for Amount of Substance (CCQM) has selected **measurement principles** (VIM 2.4) and **measurement methods** (VIM 2.5) that have the potential to assign **quantity values** in mole, or its **derived units** (VIM 1.11), for the **quantities** carried by materials which then become primary **calibrators**, e.g. **CRMs** (VIM 5.14).

2.1.3 Route to establishing metrological traceability

Establishing and demonstrating **metrological traceability** in chemistry is often not straightforward. One reason is that there may be several ways of obtaining the **measurement result** for the same **measurand**. For example, the quantitative analysis of copper in a water sample can be carried out using various types of spectrometric instrumentation, with or without digestion, separation, and pre-concentration steps. In addition, the complexity of test materials means that extensive sample pre-treatment and clean-up is often required which makes straightforward comparisons between **measurement standards** and test samples difficult.

The **secondary measurement standard** (VIM 5.5) shown in Figure 2 serves to **calibrate** the **measuring system**, within the given **reference measurement procedure**. This **reference measurement procedure** is the one used to assign a **value** to the **calibrator**, e.g. a **CRM** (VIM 5.14), used in the laboratory during the analysis of routine samples. The choice of **calibrator** will depend on the **measurement procedure** and the purpose for which the **measurement** is being made. Analysts must assess the influence of the entire measurement process, and sampling if appropriate, on the **metrological traceability** of the **measurement result**.

Manufacturers normally offer various materials for the preparation of **working measurement standards** (VIM 5.7) for **calibration** of routine **measurements**. There are, e.g. pieces of copper metal with stated purity, and solutions with specified amount-of-substance concentration and matrix composition. The **uncertainty** in the **value** of the **calibrator** will directly influence the **measurement uncertainty** of the final result so here the analyst may have a choice.

Fewer standards qualify as **secondary measurement standards**, and there are an even smaller number of **primary measurement**

---

[*]'Creatininium' is the IFCC-IUPAC term for the sum of the species 'creatinine' and 'creatininium' ion.

1 **standards** (VIM 5.4) and primary procedures
2 available. So although the lower parts of the chain
3 shown in Figure 2 will differ, **measurements** of the
4 concentration of copper made in different
5 laboratories will be traceable to a stated endpoint
6 reference <u>via</u> the same primary **calibrator** or
7 procedure. Many of the **measurements** of the
8 protein transferrin in serum performed in medical
9 laboratories are traceable to the **SI unit** (VIM 1.16)
10 g L^{-1} via the **CRM** ERM® DA 470k/IFCC [26]. The
11 laboratory has to ensure the **metrological**
12 **traceability** of the steps shown below the dotted
13 line in Figure 2.

15 **Metrological traceability** is the property of a
16 result. In the generic example in Figure 2 the
17 **measurement result** is the **quantity value** and its
18 **measurement uncertainty**, together with any other
19 relevant information relating to the sample.

14
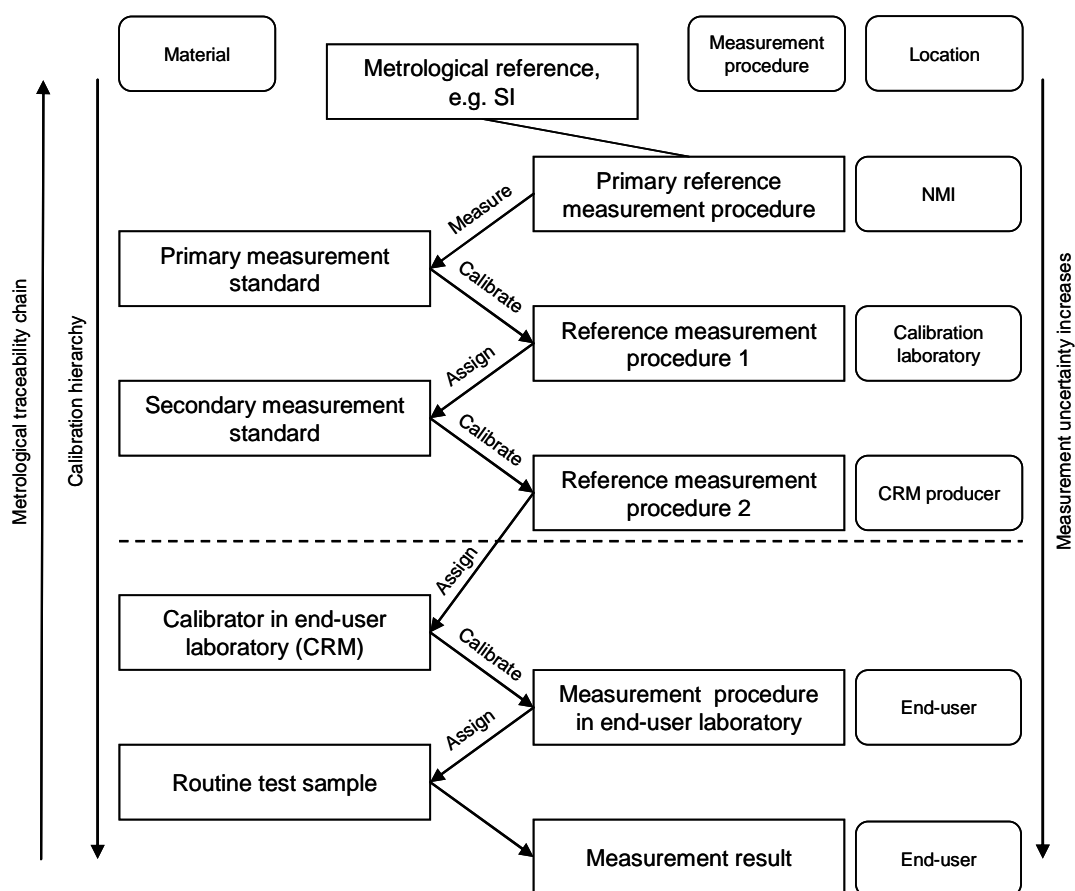
Figure 2 – Example of a generic traceability chain. The traceability chain relates the measurement result for a routine test sample to the reference point (here the SI) via a sequence of calibrations (the arrows). Uncertainties, present in all procedures and calibrators, are propagated to the final result. The arrows to the left illustrate the direction of the traceability chain (upwards) and the direction of the calibration hierarchy (downwards). The arrow on the right indicates the measurement uncertainty increasing from the metrological reference to the measurement result.

2.1.4 Practical demonstration of metrological traceability

Examples of how to achieve **metrological traceability** can be found in a number of guides [21, 22, 27]. For routine testing most of the information that the laboratory needs to establish and demonstrate **metrological traceability** is available in-house.

- Definition of the **measurand**: Includes at least the type of **quantity** (e.g. mass concentration), the analyte (e.g. methyl mercury), and type of samples (e.g. freshwater, dairy products ...), but may require additional specifications (e.g. dry mass and drying temperature), see also section 1.11 for details.

- A description of the **measurement procedure**: Includes details of all steps, equipment and materials required, the **measuring system** and a **model** (VIM 2.48) showing how the result is calculated.

- The **target measurement uncertainty** (VIM 2.34): The maximum **uncertainty** that is acceptable. This depends on the application (the intended use) and, ideally, the customer knows this, or can refer to specifications. In order to choose appropriate references knowledge of the **target measurement uncertainty** is required.

- **Reference**: The top of the **traceability chain** (VIM 2.42), (e.g. a **measurement unit**, a material with a specified **quantity value** (**CRM** (VIM 5.14)) or a **measurement procedure**).

- **Calibration hierarchy** (VIM 2.40) and **traceability chain**: The laboratory can perhaps choose between several working **calibrators**, each one provided with its own different, fixed **traceability chain**. The documentation accompanying the **calibrator** describes the sequence of steps (involving **measurement standards**, **measuring systems** and **measurement procedures**), each with increasing **uncertainty**, between the **calibrator** and its **reference** (**calibration hierarchy**) which define the **traceability chain** for the **calibrator** of choice. If the working **calibrators** are prepared in the laboratory, this step is added to the **traceability chain** and its **uncertainty** evaluated according to established procedures.

- Many **measurements** involve multiple **input quantities** (VIM 2.50) and **influence quantities** (VIM 2.52). These should all be **metrologically** **traceable** resulting in the **calibration hierarchy** having a branched structure. The **metrological traceability** of all relevant **input quantities** and **influence quantities** must be demonstrated by the laboratory by means of documented **calibrations**. The effort involved in establishing **metrological traceability** for each **quantity** should be commensurate with its relative contribution to the **measurement result**. Any **corrections** (VIM 2.53) applied before presenting the **measurement result** must also be traceable, for example when correcting results for **measurement bias** (VIM 2.18).

- Check that the relevant properties of the **calibrators** – **quantity values**, **uncertainties** and **metrological traceability** – are fit for purpose and fully documented.

2.1.5 Appropriate equipment and degree of control

Knowing the **target measurement uncertainty** (see section 2.1.4) the analyst can select appropriate equipment and **measurement standards**. It is important to identify the **input** and **influence quantities** (VIM 2.50 and 2.52) which have a significant effect on the **measurement result** so that the **uncertainty** associated with the **measurement** of these **quantities** can be controlled appropriately. For example, when measuring a volume of liquid there is a choice of apparatus available (measuring cylinder, volumetric flask, pipette, etc.). The **measurement uncertainty** associated with volumes measured using these devices will differ. When preparing a reagent, where the concentration is not critical to the **measurement result**, using a measuring cylinder may be acceptable. In contrast, the concentration of a calibration solution will have a direct influence on the **measurement result** so a higher degree of **accuracy** (smaller **measurement uncertainty**) in volume **measurements** is required. In addition, when preparing a calibration solution different grades of chemical substances are often available. The appropriate grade (quality) should be selected for a particular application. For example, two materials are available for preparation of a calibration solution to measure the mass fraction of pesticide p,p'-DDE in animal fat [27]:

- a commercial grade chemical with stated purity expressed as a mass fraction >95 %;

- a **CRM** (VIM 5.14) with certified purity expressed as a mass fraction of (99.6 ± 0.4) %.

1 The **measurement uncertainty** associated with the
2 stated purity of the commercial grade chemical may
3 be sufficient in a screening exercise to assess the
4 degree of contamination. However the **CRM**, which
5 has documented **metrological traceability** and a
6 smaller **measurement uncertainty** associated with
7 the stated purity, would be more appropriate to use if
8 the intention is to determine if a specific test sample
9 complies with a legal limit. As mentioned above, the
10 choice of **calibrator** fixes the **calibration hierarchy**
11 (VIM 2.40), and thereby the **traceability chain**
12 (VIM 2.42).

13 As a rule of thumb the **measurement uncertainty**
14 for those steps in the **measurement procedure** that
15 have a significant effect on the result should be ≤1/5
16 of the **target measurement uncertainty** for the
17 final result. When this condition is met the
18 individual steps concerned will make a negligible
19 contribution to the overall **measurement**
20 **uncertainty**.

21 When selecting **measurement standards**,
22 certificates of analysis and calibration certificates
23 should be considered in the light of the
24 accreditations or approvals held by the issuing body.
25 **Values** given on a certificate from a non-accredited
26 facility may not have the degree of **metrological**
27 **traceability** that the customer would anticipate.
28 Producers of **CRMs** should comply with the
29 requirements of ISO 17034 [10], as documented by
30 their accreditation or otherwise verified by external
31 assessment by customers or by self-assessment.

2.2 Calibration

operation that, under specified conditions, in a
first step, establishes a relation between the
quantity values with **measurement**
uncertainties provided by **measurement**
standards and corresponding **indications** with
associated measurement uncertainties and, in a
second step, uses this information to establish a
relation for obtaining a **measurement result**
from an indication (VIM 2.39)

34 **Calibration** in chemical analysis is frequently
35 associated with calibrating a **measuring instrument**
36 or **measuring system**. Typical features of these are:

37 • they contain chromatographic and/or
38   spectrometric equipment;

39 • they need frequent (daily, weekly, monthly)
40   **calibration**;

41 • the **indication**, i.e. the signal from the instrument
42   or system, corresponds to a **quantity** other than
43   that intended to be measured, e.g. an electric

44 charge or potential, rather than amount-of-
45 substance concentration or mass fraction.

46 The VIM 3 definition of **calibration** is divided into
47 two parts, Figure 3 illustrates the first part as a
48 **calibration diagram** (VIM 4.30).

50
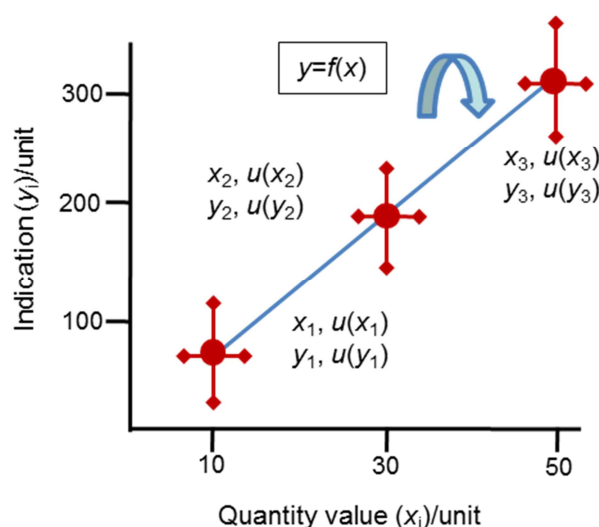
Figure 3 – **Schematic of the first clause of the**
definition of calibration. Indications ('signals' y_i)
from measurement standards (calibrators) with
quantity values x_i give the relation (the function)
$y = f(x)$. The vertical and horizontal arrows
indicate the standard uncertainties of the
indication and quantity values respectively (these
**are not to scale).**

60 In a **calibration** experiment, the analyst typically
61 prepares a set of calibration solutions (also known
62 as, '**calibrators**', 'calibrants', 'standard solutions' or
63 '**working standards**' (VIM 5.7)), i.e. a set of
64 **measurement standards**. When measured, each of
65 them gives rise to an **indication** ('signal',
66 'response'). The relation $y = f(x)$ between the
67 **indication** and the corresponding **quantity value** is
68 called a **calibration curve** (VIM 4.31). The
69 **uncertainty** of the **calibration** will include
70 contributions from the **uncertainty** of the
71 **measurement standards**, variation in **indications**,
72 and limitations in the mathematical model when
73 establishing the relation $y = f(x)$.

74 The analyst then analyses the unknown sample and
75 uses the **indication** (y_s) to calculate a corresponding
76 **quantity value** (x_s) from the **calibration curve**
77 using the function $x = f^{-1}(y)$. This second part of the
78 definition is illustrated in the **calibration diagram**
79 shown in Figure 4. If, for example, $f(x)$ is defined as
80 $a + bx$, where b is the slope of the curve and a is the
81 y-axis intercept when $x = 0$, then $f^{-1}(y)$ is $(y - a)/b$.

Uncertainties arising from the **indication**, the **calibration**, and other **corrections** (VIM 2.53), contribute to the **uncertainty** of the **measurement result**.
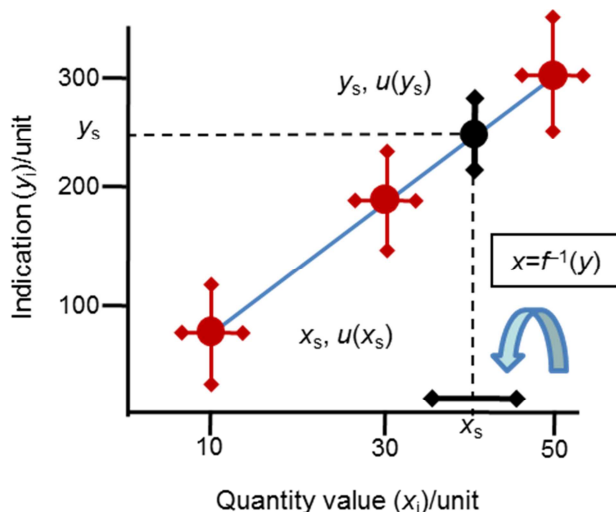
Figure 4 – **Schematic calibration diagram illustrating the second clause of the definition of calibration. The indication ('signal' y_s) from a sample corresponds to a quantity value x_s. The vertical and horizontal arrows indicate the standard uncertainty of the indication and of the quantity value respectively.**

2.3 Instrumental drift

> continuous incremental change over time in **indication**, due to changes in metrological properties of a **measuring instrument** (VIM 4.21)

Instrumental drift is the gradual change over time (in either direction) in the **indication** provided by an instrument. Drift will affect the **trueness** of results if the true **calibration** parameters have changed between the time the instrument was calibrated and the time test samples are analysed. The extent to which an **indication** drifts therefore determines the required frequency of recalibration of the instrument. In analytical chemistry, a 'drift correction standard' with a known **quantity value** can be measured regularly to monitor the state of **calibration** of an instrument and determine whether adjustment to the configuration of the instrument or recalibration is required.

2.4 Measurement standard

> realization of the definition of a given **quantity**, with stated **quantity value** and associated **measurement uncertainty**, used as a reference (VIM 5.1)

A commercial laboratory regularly monitors the level of cadmium in drinking water. Legislation has specified that the **quantity** measured should be the mass concentration. For instrument **calibration** the laboratory uses a **CRM** (VIM 5.14) for which, according to the certificate, the **value** of the mass concentration of cadmium is (1005 ± 3) mg L^{-1}. In this case the **quantity value** is 1005 mg L^{-1} and the **value** of the **measurement uncertainty** is 3 mg L^{-1}. The **CRM** is an example of a **measurement standard**.

Measurement standards are used in all scientific areas. **Material measures** (VIM 3.6), e.g. volumetric flasks and **CRMs**, as well as **measuring system** (e.g. a reference thermometer) can function as **measurement standards**. When analytical scientists talk about **calibrators** or calibrants they simply mean **measurement standards** used in **calibration**.

2.4.1 A hierarchy of measurement standards

Various terms are used to indicate properties or uses of **measurement standards**. Figure 5 shows the relationship between the types of standard.
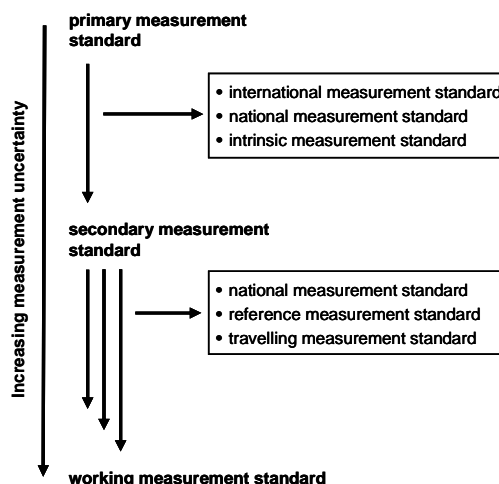
Figure 5 – **Illustration of hierarchy of measurement standards.**

The 'Vienna Standard Mean Ocean Water' (VSMOW2) is an **international measurement standard** (VIM 5.2) for differential stable isotope

ratio **measurements**. Many NMIs have a copy of the former international prototype kilogram serving as a **national measurement standard** (VIM 5.3) for mass. According to VIM 3, by accurately weighing and dissolving glucose of known purity to a known volume, a chemist can prepare a **primary standard** in the form of a solution of known concentration.

The standards mentioned above and **secondary measurement standards** (VIM 5.5) as well as other **reference measurement standards** (VIM 5.6) are needed by producers of **RMs** (VIM 5.13), instrument manufacturers and reference laboratories. In addition, for routine applications laboratories may use **working standards** (VIM 5.7) to check **measuring systems**.

Two other types of **measurement standard** are shown in Figure 5. An **intrinsic measurement standard** (VIM 5.10) is a **measurement standard** based on an inherent physical constant or inherent physical property. For example, the triple point of water cell is an **intrinsic standard** of thermodynamic temperature. A **travelling measurement standard** (VIM 5.8) is simply a **measurement standard** intended for transport between locations.

2.5 Calibrator

measurement standard used in **calibration** (VIM 5.12)

Analytical scientists frequently use the term calibrant, **calibrator** or standard when referring to **measurement standards** used in **calibration**. Many **measurements** are made using **measuring instruments** and **systems** that require regular **calibration**. An inherent part of the daily work for the analyst is, therefore, preparation and/or maintenance of this type of **measurement standard**.

Measurement standards are produced and used for different purposes. There are a number of documents which provide guidelines on choosing appropriate **measurement standards**, for example reference 28. Not all materials described by suppliers and producers as **measurement standards** can be used for **calibration**. The user needs to exercise caution when purchasing 'standards' as the producer may not interpret the requirements in the same way as in VIM 3.

Many routine **measuring systems** are designed to handle samples, without isolation or pre-concentration of the analyte. In such cases it is also necessary to demonstrate that the **calibrator** behaves in the same way as the routine samples.

This is done when the manufacturer or the user investigates the **commutability of a reference material** (see section 2.6).

Below are some examples of materials frequently used by analysts when carrying out **calibration**.

- Materials produced by manufacturers and intended to be used for **calibration** or **verification** of a commercial **measuring system**, e.g. a **working measurement standard** (VIM 5.7) with assigned **quantity value** and **measurement uncertainty** for amount-of-substance concentration of glucose in human serum, supplied as part of an *in vitro* diagnostic medical device.

- **RMs** (VIM 5.13) and **CRMs** (VIM 5.14).

- Materials produced by authoritative bodies, e.g. a national or international pharmacopoeia, and intended to be used within a limited specified scope.

- Materials produced and characterised by the laboratory in-house, e.g. in the absence of commercial products.

For practical purposes materials used as **calibrators** should have a statement of **measurement uncertainty** and **metrological traceability**.

2.5.1. Check the 'intended use'!

The content and layout of the documentation supplied with **RMs** (VIM 5.13) shows considerable variability. The headings may be adapted to meet the requirements of a particular sector. It is not always obvious to the user if the material can be used for **calibration**.

A description of the 'intended use' is an essential part of the certificate for a **CRM** (VIM 5.14) [29]. The primary purpose for which a **CRM** is issued by the producer should be stated. Many materials are not described as **CRMs** but still qualify as **calibrators**. Check the documentation and your own needs, e.g. in the case of medical laboratories the *In Vitro* Diagnostics Directive 98/79/EC applies [18].

Some materials could very well qualify as **calibrators** but the intention of the manufacturer of the material and/or **measuring system** is different, e.g. due to legislative restrictions. For instance, the manufacturer of a **measuring system** will only take responsibility if the prescribed **calibrator** is used. The laboratory may, however, wish to check the results by using other **measurement standards**, and therefore be looking, e.g. for an appropriate

1 'trueness control' to check for **measurement bias**
2 (VIM 2.18) during **verification**.

3 Below are three examples of statements from
4 certificates, indicating their intended use.

5 • 'The primary use of this material is for checking
6 the <u>**calibration**</u> of automatic density meters used
7 in industry to determine alcoholic strength ... '.

8 • 'The material is primarily intended to be used to
9 <u>**calibrate**</u> serum-based protein standards and
10 control products of organisations which offer
11 such preparations for the quantification of C-
12 reactive protein by immunoassay.'

13 • 'The material is primarily intended to be used to
14 <u>control</u> the performance of the IFCC reference
15 procedure ... . When the material is used as a
16 **calibrator** in a particular assay, the
17 **commutability** should be verified for the assay
18 concerned.'

19 2.6 Commutability of a reference
20 material

21
> property of a **reference material,**
> demonstrated by the closeness of agreement
> between the relation among the
> **measurement results** for a stated **quantity**
> in this material, obtained according to two
> given **measurement procedures**, and the
> relation obtained among the measurement
> results for other specified materials
> (VIM 5.15)

22 The wording of the definition is slightly different
23 from ones that appear in some ISO Standards and
24 Guides but the principle is the same. As mentioned
25 in section 2.5 it is important to check that the **RM**
26 (VIM 5.13) chosen as a calibrant behaves in the
27 same way as the samples. This is termed the
28 **commutability of a reference material**.
29 **Commutability** is of particular concern where
30 methods are very sensitive to the sample matrix or
31 'physical form' of the analyte of interest.

32

33 In such cases **calibration** with closely matching
34 materials is essential for accurate **measurement**
35 **results**. **Commutability of reference materials** is
36 also of concern where the **measurement procedure**
37 cannot be modified by the analyst but the **RMs**
38 available do not simulate the sample matrix. Medical
39 laboratories may encounter this problem when using
40 analysers with **calibrators** supplied by the
41 manufacturer. The issue of **commutability** in
42 relation to the analysis of clinical samples is
43 discussed in detail in recommendations published by
44 IFCC [30-32].

45 The 'other specified materials' mentioned in the
46 definition are usually samples analysed routinely in
47 a laboratory.

48 The concept of **commutability** is best described
49 diagrammatically as shown in Figure 6. Figure 6(a)
50 illustrates a case where the **reference material** M1
51 is commutable whereas in Figure 6(b) the **reference**
52 **material** M2 is not commutable. M1, M2 and S1
53 represent the **indication** for the **reference materials**
54 M1 and M2 and the sample S1, respectively. The
55 **indication** may be an instrument signal or a
56 particular **quantity value**.

57

1
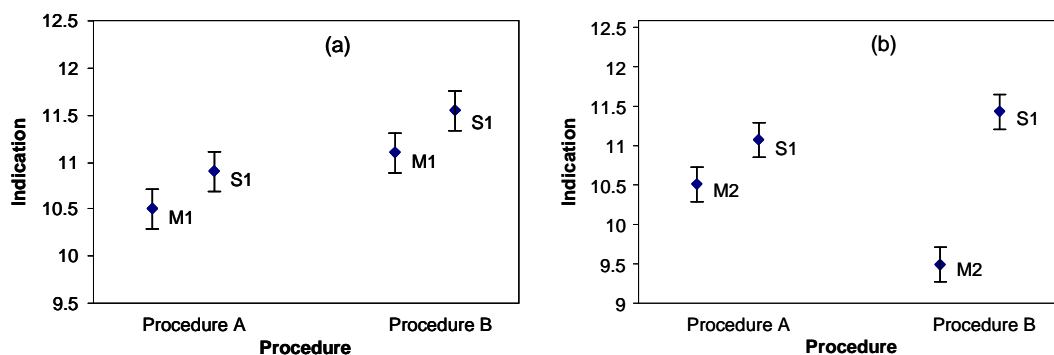
Figure 6 – Schematic to illustrate the commutability of a reference material, showing the measurement results and their associated confidence intervals. In case (a) the reference material is considered to be commutable – the relation between the indication obtained for the reference material (M1) and the indication obtained for the sample (S1) is independent of the measurement procedure. In case (b) the reference material is not commutable – the relation between the indication obtained for the reference material (M2) and the indication obtained for the sample (S1) is different for the two measurement procedures.

3 Measurement uncertainty

1 This chapter describes the terminology relating to
2 **measurement uncertainty**. The Guide to the
3 expression of uncertainty in measurement (GUM)
4 (published as JGCM100 [33] and ISO/IEC Guide
5 98-3 [34]), describes the principles of **uncertainty**
6 evaluation. Further information on how to evaluate
7 **measurement uncertainty** can be found in the
8 Eurachem/CITAC Guide, Quantifying uncertainty in
9 analytical measurement [35]. An overview of the
10 different approaches is given in a Eurolab
11 report [36]).

3.1 Measurement uncertainty

13
> non-negative parameter characterizing the
> dispersion of the **quantity values** being
> attributed to a **measurand**, based on the
> information used (VIM 2.26)

14 **Measurement uncertainty** provides a quantitative
15 indication of the quality of a **measurement result**.
16 Synonyms are 'uncertainty' and 'uncertainty of
17 measurement'.

18 This definition expresses the fact that parameters
19 used to describe the dispersion of distributions, e.g.
20 standard deviations, are usually positive. The
21 statement, 'based on the information used', explains
22 why it is necessary to declare what is included in the
23 estimate of **measurement uncertainty**. This does
24 not mean we can choose what to include and what to
25 leave out. There are many approaches to evaluating
26 **measurement uncertainty** and these are described
27 in the literature [35-39].

28 **Measurements** consist of many steps and require
29 various items of equipment. For example,
30 calculating the **measurement result** may involve
31 reagent concentrations and **values** from **measuring**
32 **instruments**, **calibrators** and **RMs** (VIM 5.13). All
33 of these **values** have some **uncertainty**; and their
34 **uncertainties** will make the calculated result
35 uncertain. Incompletely known properties of the
36 sample itself – such as possible interferents, matrix
37 effects and effects on analyte recovery – as well as
38 sampling and the manual operations carried out
39 during the **measurement**, also contribute to
40 **measurement uncertainty**. This means that, for a
41 specific calculated result, there is not one but a
42 whole range of **quantity values** (VIM 1.19) that
43 could reasonably have given rise to the **measured**
44 **quantity value**. **Measurement uncertainty**
45 describes the dispersion of these possible **values**.

46
47 The **result of a measurement** consists of two
48 quantitative parts: i) the **measured quantity value**
49 and ii) the **measurement uncertainty**. When the
50 result is reported with the **uncertainty**, it can be
51 presented in the format (value ± uncertainty) and
52 **unit**. For example, (5.5 ± 0.5) mL corresponds to the
53 interval $(5.0 - 6.0)$ mL (see Figure 7). The
54 **uncertainty** is interpreted as providing an interval
55 within which the **value** of the **measurand** is
56 believed to lie. The **uncertainty** is usually reported
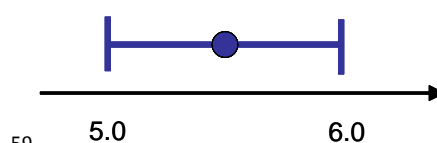57 as the **expanded uncertainty** (VIM 2.35) (see
58 section 3.1.1).

59

60 **Figure 7 – Illustration of the result (5.5 ± 0.5) mL.**
61

3.1.1 Expression of uncertainty

63 Estimates of **measurement uncertainty** can be
64 expressed in a number of different ways, e.g. as a
65 standard deviation or a confidence interval.
66 However, to be able to combine **uncertainty**
67 estimates they must be expressed in the same form,
68 so some conversion may be necessary. Following
69 ISO guidelines, **uncertainty** estimates should be
70 expressed as **standard uncertainties** (VIM 2.30)
71 (see below) before they are combined [34].

72 In certificates of analysis, test reports etc., the letter
73 u denotes **uncertainty**. However, there are different
74 forms of **uncertainty**:

75 • $u(x_i)$ – the **standard uncertainty** for **quantity** x_i
76   is an **uncertainty** expressed as a standard
77   deviation;

78 • $u(y,x_i)$ – the contribution to the **standard**
79   **uncertainty** of the **measurand** caused by the
80   quantity x_i;

81 • u_c – the **combined standard uncertainty**
82   (VIM 2.31) for the **measurand**, is a
83   mathematical combination of several individual
84   **standard uncertainties**;

85 • U – the **expanded uncertainty** (VIM 2.35) is
86   normally what the laboratory reports to the
87   customer. The **expanded uncertainty** provides
88   an interval within which the **value** of the
89   **measurand** is believed to lie with a higher level

of confidence. The value of U is obtained by multiplying the **combined standard uncertainty** u_c by a **coverage factor** (VIM 2.38) k, i.e. $U = k \cdot u_c$. The choice of the factor k is based on the level of confidence desired.

It follows that $u(y,x_i) < u_c < U$. These **uncertainties** are often expressed in relation to the **value**, e.g. as a **relative standard measurement uncertainty** (VIM 2.32).

3.1.2 Uncertainty evaluation

The overall performance of a **measurement procedure** is studied during method development and method **validation** (see chapter 4). Individual sources of **uncertainty** will be identified during this process and studied in detail if they are found to be significant compared to the overall requirements. The laboratory will aim to remove sources of significant **uncertainty** until the **measurement procedure** is deemed to be fit for purpose. This means that the laboratory should know the maximum **measurement uncertainty** that can be accepted by the customer for a specific application. This is called the **target measurement uncertainty** (VIM 2.34) . For example, the EU legislation regarding the official control for monitoring the status of surface waters and groundwater states that laboratories performing **measurements** should use **measurement procedures** capable of providing **results** with an '**uncertainty of measurement** of 50 % or below ($k = 2$) estimated at the level of the relevant environmental quality standard [40]. For example, the environmental quality standard for lead in surface waters is $7.2\,\mu g\,L^{-1}$ so the **target measurement uncertainty** is $3.6\,\mu g\,L^{-1}$ [41]. A Eurachem/CITAC Guide [17] provides an overview of the possible approaches to setting the **target measurement uncertainty** in various different scenarios.

During **validation/verification** of a **measurement procedure** the overall **measurement precision** of the method and the **uncertainty** of the **measurement bias** (VIM 2.18) are often evaluated. In many cases combining these two **uncertainty** components using the law of propagation gives a reasonable estimate of the **standard uncertainty** of the results obtained when using the **measurement procedure** [35-39]. Since **measurement uncertainty** can be estimated in different ways, the resulting **value** should be accompanied by an explanation, or by reference to available information, of how the **uncertainty** was evaluated. The customer is then in a position to interpret the **uncertainty** (see also section 3.2).

The minimum level of **uncertainty** associated with a given **measurement result** is implicit in the definition of the **measurand** (**definitional uncertainty**, VIM 2.27). For example, the **measurand**, 'volume of liquid contained in a volumetric flask' has a larger **definitional uncertainty** than the **measurand**, 'volume of water contained in a volumetric flask at 20 °C'. In the case of the former neither the nature of the liquid nor the temperature of liquid are specified. The **definitional uncertainty** depends on the analyst's ability to define the **measurand** adequately. No matter how much effort is put into the **measurement**, the **definitional uncertainty** cannot be reduced unless a new, more detailed definition of the **measurand** is given. For example, the **definitional uncertainty** associated with the **measurement** of the total amount of protein in a milk sample will be larger than the **definitional uncertainty** associated with the distribution of individual protein fractions in the milk sample. It is good practice to define the **measurand** in such a way that the **definitional uncertainty** is negligible for the purposes of the **measurement**.

3.2 Uncertainty budget

> statement of a **measurement uncertainty**, of the components of that measurement uncertainty, and of their calculation and combination (VIM 2.33)

All known sources of **measurement uncertainty** have to be evaluated and information about them summarised in an **uncertainty budget**. The word budget is used in a different sense from that in common usage; it is not an upper limit of **measurement uncertainty**; it is a statement of the sources of **uncertainty** and their values. The budget should also include the **measurement model** (VIM 2.48) and type of **uncertainty** evaluation. Two types of **uncertainty** evaluation are defined in VIM 3 and the ISO Guide to the expression of uncertainty in measurement [34]. An **uncertainty** contribution based on statistical analysis of **quantity values** obtained under defined conditions is a **Type A evaluation** (VIM 2.28). A contribution obtained by any other means, is a **Type B evaluation** (VIM 2.29). An example of **Type A evaluation** is the standard deviation of the mean of the results from ten replicate **measurements** performed under **repeatability conditions** (VIM 2.20). An **uncertainty** value taken from an **RM** (VIM 5.13) certificate is an example of a **Type B evaluation**. The **uncertainty budget** may also include the applied probability density function and degrees of

1 freedom for each **uncertainty** contribution, and the
2 **coverage factor** (VIM 2.38) used to calculate the
3 **expanded uncertainty** (VIM 2.35).

4 An example of an **uncertainty budget** for the mass
5 concentration of a cadmium calibration standard is
6 shown in Table 2. It is based on an example from the
7 Eurachem/CITAC Guide [35]. The mass
8 concentration of cadmium, ρ_{Cd} (mg L^{-1}) is given by:

9 $$\rho_{Cd} = (1000 \cdot m \cdot P)/V$$

10 where m is the mass in mg of cadmium, P its purity,
11 and V the volume of the flask in mL. Each of these
12 terms will introduce **uncertainty** in the calculated
13 concentration of the solution, as shown in the
14 **uncertainty budget** in Table 2. The **uncertainty** in
15 the mass is obtained from the calibration certificate
16 provided by an accredited calibration laboratory and
17 their recommendations on the estimation of the
18 **uncertainty** under the conditions of use. The purity

19 of the metal is obtained from the accompanying
20 information available from the producer and
21 converted to a **standard uncertainty** (VIM 2.30)
22 assuming a rectangular distribution. The
23 **uncertainty** in the volume of the flask consists of
24 three components – **calibration** (u_{cal}), **repeatability**
25 (VIM 2.21) of filling the flask (u_{rep}), and the
26 difference between the temperature at which the
27 **calibration** was made and the temperature when the
28 flask is used (u_{temp}).

29 An example of an **uncertainty budget** for a test
30 method for the determination of the mass
31 concentration of ammonium nitrogen NH$_4$-N in
32 water, validated in a single laboratory, is shown in
33 Table 3.

Table 2 – Uncertainty budget for the mass concentration of a cadmium calibration standard; values taken from the Eurachem/CITAC Guide [35]. The standard uncertainty in ρ_{Cd} was calculated by combining the relative standard measurement uncertainties and then multiplying by the value for ρ_{Cd}.

Quantity	Value	Standard uncertainty, $u(x_i)$	Unit	Relative standard uncertainty $u(x_i)/x_i$
m	100.28	0.050	mg	0.00050
P	0.9999	5.8×10^{-5}	g/g	5.8×10^{-5}
V*	100.00	0.066	mL	0.00066
ρ_{Cd}	1002.70		mg/L	
$\rho_{Cd,}$ combined standard uncertainty	0.84		mg/L	
Expanded uncertainty $k = 2$	1.7		mg/L	

*Volume contributions	Standard uncertainty, $u(x_i)$	Unit
u_{cal}	0.041	mL
u_{temp}	0.048	mL
u_{rep}	0.020	mL
Volume, combined standard uncertainty	0.066	mL

Table 3 – Uncertainty budget for the mass concentration of ammonium nitrogen NH$_4$-N in water. Values taken from the Nordtest Handbook [39]. The relative standard measurement uncertainty for the intermediate precision ($u(P)$) is combined with the relative standard measurement uncertainty associated with the bias estimate ($u(bias)$). The expanded uncertainty is given with a coverage factor of $k = 2$.

Concentration range (µg/l)	Estimate of intermediate precision	$u(P)$	Estimate of uncertainty associated with bias	$u(bias)$	Combined uncertainty	Expanded uncertainty
50-500	Control sample covering the whole analytical process	1.67 %	Proficiency tests	2.73 %	3.20 %	7.0 %

4 Verification, validation and method performance

1 This chapter describes the terminology relating to
2 **verification**, **validation** and method performance.
3 Further information on method **validation** can be
4 found in the Eurachem guide on the fitness for
5 purpose of analytical methods [42].

4.1 Verification and validation

> provision of objective evidence that a given item fulfils specified requirements (VIM 2.44)

> **verification**, where the specified requirements are adequate for an intended use (VIM 2.45)

9 Taking the concepts of **verification** and **validation**
10 together, the process of **validation** involves agreeing
11 with the customer 'specified requirements' for
12 performance characteristics such as **selectivity**,
13 **measuring interval**, **trueness** and **precision** that are
14 adequate for the intended use of the **measurement**
15 **procedure**, and then confirming, on the basis of
16 objective evidence, that they are fulfilled
17 (**verification**).

18 In order to illustrate the relationship between
19 **verification** and **validation**, consider an example
20 where a laboratory purchases an instrument. After
21 the instrument has been installed in the laboratory,
22 an analyst plans a series of experiments to check that
23 the instrument's performance meets that specified by
24 the manufacturer. This process is called
25 **verification** – the analyst will obtain objective
26 evidence (experimental data) which demonstrates
27 that the instrument meets the manufacturer's
28 specification. Once it has been confirmed that the
29 instrument performance is satisfactory it will be used
30 as part of a particular **measurement procedure**. The
31 performance requirements for the **measurement**
32 **procedure** are specified by the laboratory and
33 agreed with the customer as being fit for the
34 intended purpose, for example, to detect variations
35 greater than 1 % in the copper content of an alloy.
36 The analyst plans a new set of experiments to assess
37 the performance of the **measurement procedure**,
38 and checks that it meets the customer requirements.
39 This process is called **validation.**

4.1.1 Verification of a validated method

41 In the case where a method that has been validated
42 previously (e.g. a standard method) is being used,
43 the laboratory has to provide objective evidence that

44 the stated performance characteristics of the method
45 can be achieved. Consider the following example.

46 A laboratory is asked by a customer to perform a
47 **measurement** of the mass fraction of acrylamide in
48 bakery products, such as bread and biscuits. As a
49 first step, the laboratory agrees with the customer the
50 **measurement** requirements in terms of working
51 range and **expanded uncertainty** (VIM 2.35). Then,
52 it checks if a standard method exists, that has been
53 proven, by interlaboratory studies, to fulfil the
54 requirements for the intended use. It finds that the
55 method EN 16618:2015 'Determination of
56 acrylamide in food by liquid chromatography
57 tandem mass spectrometry (LC-ESI-MS/MS)'
58 covers bakery products in its scope, and that both the
59 working range and the **expanded uncertainty** fulfil
60 the agreed requirements for the intended use. The
61 laboratory must then verify, via an experimental
62 assessment, that the stated method performance can
63 be achieved.

64 The following sections describe the performance
65 characteristics defined in VIM 3 which are
66 commonly studied during **verification/validation**.

4.3 Selectivity of a measuring system

> property of a **measuring system,** used with a specified **measurement procedure,** whereby it provides measured **quantity values** for one or more **measurands** such that the values of each measurand are independent of other measurands or other **quantities** in the phenomenon, body, or substance being investigated (VIM 4.13)

69 The definition of **selectivity** in VIM 3 is consistent
70 with the more familiar definition proposed by
71 IUPAC: 'the extent to which the method can be used
72 to determine particular analytes in mixtures or
73 matrices without interferences from other
74 components of similar behaviour.' [43]. For
75 example, gas chromatography using a mass
76 spectrometer as the detector (GC-MS) would be
77 considered more selective than gas chromatography
78 using a flame ionisation detector (GC-FID), as the
79 mass spectrometer provides additional information
80 which assists with confirmation of identity. The use
81 of the term specificity is not recommended by
82 IUPAC and is not defined in VIM 3.
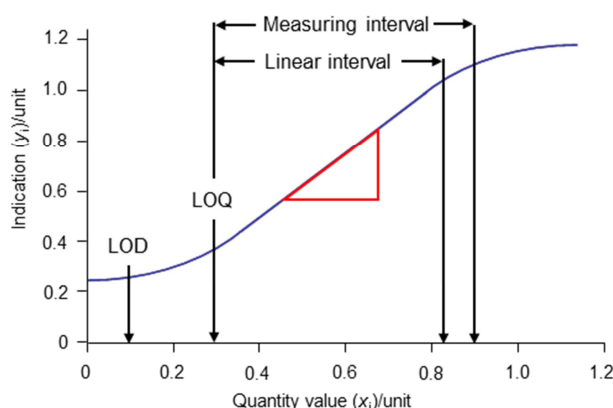
4.4 Measuring interval

set of **values** of **quantities** of the same **kind** that can be measured by a given **measuring instrument** or **measuring system** with specified **instrumental measurement uncertainty**, under defined conditions (VIM 4.7)

Within the **measuring interval**, **quantities** (e.g. mass concentration) can be measured with a specified **uncertainty** using a stated **measurement procedure**. The other phrases commonly used for this concept are: 'working range' and 'measurement range' (the latter is used in ISO/IEC 17025 [4]). The common usage of the term 'measurement range' or 'measuring range' to denote **measuring interval** is acknowledged in Note 1 of the VIM definition. However, it should be noted that in VIM 3, the term 'interval' denotes a set of numbers defined by its end values whereas the term 'range' or 'range of interval' is restricted to the difference between the highest and the lowest values of an interval. Following these conventions, in the example shown in Figure 8 the **measuring interval** is 0.3 to 0.9, written as [0.3, 0.9], and the range is 0.6.

The lower limit of the **measuring interval** is often considered to be the limit of quantification (LOQ) (a concept not defined in VIM 3). The upper limit is usually determined by the unacceptable change in **measurement uncertainty** or in the **sensitivity** (VIM 4.12), for example the plateauing effect observed at high absorbance values in UV/VIS spectroscopy. Figure 8 illustrates the relationship between some of the key terms related to '**measuring interval**'. The LOD is below the LOQ. The **measuring interval** should be compatible with the analytical requirement and, therefore, fit for the purpose. If, for example, the analyte level in samples is expected to be well above the LOQ, the laboratory may not need to cover the entire interval illustrated in Figure 8.

Many methods rely on the test sample received in the laboratory being processed (digested, extracted, diluted, for example) before it can be presented to the **measuring instrument** and a signal recorded. In such cases there are two **measuring intervals** to consider – the instrument **measuring interval** (described in the VIM definition) and the **measuring interval** for the **measurement procedure** as a whole (including any sample preparation steps). The evaluation of these different **measuring intervals** is discussed in detail in the Eurachem guide [42].

The **sensitivity of a measuring system** is, in the case of linear dependence, given by the slope of the **calibration curve** (VIM 4.31).

Figure 8 – A calibration diagram showing indications (y_i) versus quantity values (x_i) where the measuring interval, linear interval, LOQ and LOD are identified. The triangle illustrates the calculation of the sensitivity or the slope of the calibration curve (Δindication/Δquantity).

4.5 Detection limit

measured quantity value, obtained by a given **measurement procedure**, for which the probability of falsely claiming the absence of a component in a material is β, given a probability α of falsely claiming its presence (VIM 4.18)

VIM defines **detection limit** in terms of a **measured quantity value**.

This is not consistent with the IUPAC (and other) definitions currently used in analytical chemistry which refer to a **true quantity value** (VIM 2.11) rather than a **measured value**. It is not clear whether the difference is intentional or, if so, how it can be implemented. The description below therefore follows recommendations made by IUPAC for establishing detection capability for analytical methods [44].

Many analysts will be familiar with calculating the **limit of detection** (LOD) for a **measurement procedure** by multiplying a standard deviation, s (obtained from the results of the analysis of a blank sample or a sample containing a low level of the analyte) by an appropriate factor (typically between 3 and 5). The multiplying factor is based on statistical reasoning. The following text explains the background to the commonly used factor of 3.

The discussion deals with **LOD** in terms of concentration but it applies equally to other **quantities**, e.g. mass fraction. The aim when determining the **LOD** is typically to establish the lowest concentration of the analyte present in a

sample that can be detected, using a given **measurement procedure**, with a specified level of confidence. Defining the **LOD** is a two-step process. First a 'critical value' is established. This value is set so that the probability of obtaining a **measurement result** that exceeds the critical value is no greater than α, if a sample actually contains *none* of the analyte. The critical value sets a criterion for declaring a sample to be 'positive'. A false positive probability of $\alpha = 0.05$ is generally used; this leads to a critical value of approximately $1.65s$ (where s is the standard deviation of a large number of results for a blank sample or a sample containing a low concentration of the analyte, and 1.65 is the one-tailed Student t-value for infinite degrees of freedom at a significance level, $\alpha = 0.05$). The critical value is indicated on the vertical axis in Figure 9 to emphasise the fact that it is a **measured value**. The critical value is most conveniently expressed in terms of concentration, though in principle it may be any observation, such as peak area. Any result exceeding the critical value should be considered as indicating an analyte level that is significantly different from zero.

However, if the **true value** for the concentration in a sample were exactly equal to the critical value (expressed in terms of concentration), approximately half of the **measurement results** would be expected to fall below the critical value, giving a false negative rate of 50 %. This is illustrated by the distribution shown with the broken line in Figure 9. A false negative rate of 50 % is obviously too high to be of practical use; the method does not reliably give results above the critical value if the true value for the concentration is equal to the critical value. The **LOD** (also known as 'minimum detectable value') is intended to represent the true concentration for which the false negative rate is acceptable given the critical value. The false negative error, β, is usually set equal to the false positive error, largely for historical reasons (IUPAC recommends default values of $\alpha = \beta = 0.05$). Using $\alpha = \beta = 0.05$, the **LOD** is therefore located $1.65s$ above the value specified for the critical value. This is illustrated by the shaded distribution on the horizontal axis in Figure 9. The factor for calculating the LOD with $\alpha = \beta = 0.05$ is thus $1.65+1.65 = 3.30$, which is frequently rounded to 3.0. This is based on several approximations which are described in the literature [44].
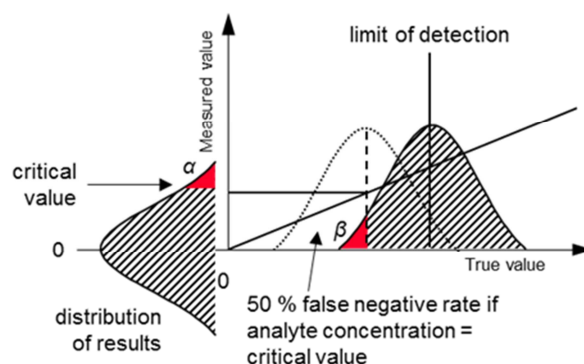

Figure 9 – Illustration of the statistical basis for detection limit calculations.

4.6 Measurement trueness

> closeness of agreement between the average of an infinite number of replicate **measured quantity values** and a **reference quantity value** (VIM 2.14)

Measurement trueness expresses the hypothetical ability of a **measurement procedure** to yield results close to expected **reference quantity values**, such as the **value** of a **CRM** (VIM 5.14). **Trueness** is not a **quantity** and therefore cannot be expressed numerically. However, **trueness** is inversely related to **systematic measurement error** (VIM 2.17) which may be estimated as **measurement bias** (VIM 2.18). An example of the estimation of **bias** as the difference between the mean value of several **measurement results** and a **reference quantity value** is shown in Figure 10. **Bias** can also be reported as the ratio of **measured** and **reference quantity values**.
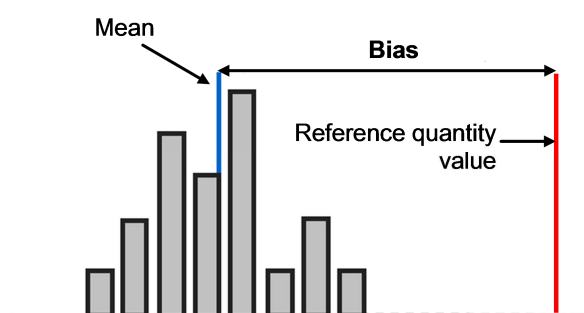
Figure 10 – Schematic illustration of the estimation of measurement bias. The mean of several measurement results is compared with a reference quantity value (note that the uncertainty in the reference value is not shown).

A **measurement bias** may be due, e.g. to inappropriate **calibration** or to lack of **selectivity** (see section 4.5). Where appropriate, the effect of known **systematic errors** on **measurement results** can be removed by introducing a **correction** (VIM 2.53) based on the estimated **bias**, e.g. the reading of a digital thermometer can be corrected on the basis of the **bias** observed during **calibration**. However, any factor used to make a **correction** will also have an associated **uncertainty**.

An estimate of the **bias** in **measurement results** produced by a laboratory can be obtained by measuring the **quantity value** of one or more **RMs** (VIM 5.13) several times under **repeatability conditions** (VIM 2.20) or under **intermediate precision conditions** (VIM 2.22), and calculating the mean value. The estimate of **bias** is then the difference between the mean value obtained and the **reference quantity value**. Note that there will be a **measurement uncertainty** associated with the **bias** value due to the **uncertainties** in the mean value and in the **reference quantity value**.

Example: The mean value of the mass fraction of CaO in a cement **CRM** calculated from 10 **measurement results** obtained over a six month period using XRF is 63.53 % with a standard deviation of the mean of 0.1 %. The certified **quantity value** is 63.23 % with an **expanded uncertainty** (VIM 2.35) of 0.21 % ($k = 2$). The **measurement bias** determined under **intermediate precision conditions** using this **CRM** is therefore estimated as $63.53 – 63.23 = 0.3$ %. The **bias** can also be expressed as a relative value (percent), i.e. $0.3/63.23 \times 100 = 0.47$ %.

4.7 Measurement precision

> closeness of agreement between **indications** or **measured quantity values** obtained by replicate **measurements** on the same or similar objects under specified conditions (VIM 2.15)

Measurement precision is related to **random measurement error** (VIM 2.19) and is a measure of how close results are to one another.

Measurement results cannot be corrected to remove the effect of **random error** but the size of the **random error** can be reduced by making replicate **measurements** and calculating the mean value.

Measurement precision is expressed numerically using measures of *imprecision* such as the standard deviation calculated from results obtained by carrying out replicate **measurements** on a suitable material under specified conditions. VIM 3 defines three **measurement** conditions: **repeatability condition** (VIM 2.20), **intermediate precision condition** (VIM 2.22) and **reproducibility condition** (VIM 2.24).

Estimates of **measurement repeatability** (VIM 2.21) and **intermediate measurement precision** (VIM 2.23) are obtained in a single laboratory. **Repeatability condition of measurement** refers to **measurements** being made on portions of the same material by a single analyst, using the same procedure, under the same operating conditions over a short time period. **Measurement repeatability** is often used to provide an estimate of within-run (also known as within-batch or intra-assay) variability in results. Under **intermediate precision conditions**, **measurements** are made on portions of the same material using the same procedure, but over an extended time period and, where possible, by different analysts, using different pieces of equipment, different batches of reagents, etc. **Intermediate measurement precision** is often used to provide an estimate of between-run (also known as between-batch or inter-assay) variability. **Intermediate precision** are user-defined and the conditions used should always be recorded (note that some laboratories use the term within-laboratory reproducibility for **intermediate measurement precision**).

Since **measurement repeatability** only reflects the variation in results over a short time period it is likely to underestimate the variability in results obtained when the **measurement procedure** is used routinely. Assuming appropriate **intermediate precision conditions** have been used during the **validation** study, the **intermediate measurement precision** provides a more realistic estimate of the long-term variability of **measurement results** in the laboratory.

Estimates of **measurement reproducibility** (VIM 2.25) are obtained from **measurement results** produced at different laboratories. **Reproducibility condition of measurement** refers to **measurements** being made on portions of the same material by different analysts working in different locations. In 'collaborative' method **validation** studies the same **measurement procedure** is used at all the participating laboratories. However, the term '**reproducibility condition**' also applies to interlaboratory comparisons where different **measurement procedures** may be used for the same **measurand** (VIM 2.24, Note 1), for example in a proficiency testing scheme. Therefore it is essential that the conditions under which **reproducibility** is evaluated are specified.

Figure 11 illustrates the relationship between **measurement repeatability**, **intermediate measurement precision** and **measurement reproducibility** in terms of the observed imprecision, which is estimated as a standard deviation, *s*. In the figure, 'between-injections' refers to replication of only the end **measurement** step of a multistage **measurement procedure** (e.g. repeat injections of portions of a test solution onto a gas chromatograph). Replicating this action would give the **measurement repeatability** of the final **measurement** stage, but would exclude the effect of **random errors** associated with any sample pre-treatment or clean-up steps. 'Within-run replicates' represents replication of the whole **measurement procedure** under **repeatability conditions**.

As the conditions of measurement become more variable (e.g. moving from replicating only part of the measurement procedure ('between-injections') to replicating the entire measurement procedure under **repeatability**, **intermediate precision** or **reproducibility conditions**) the observed imprecision of measurement results generally increases.
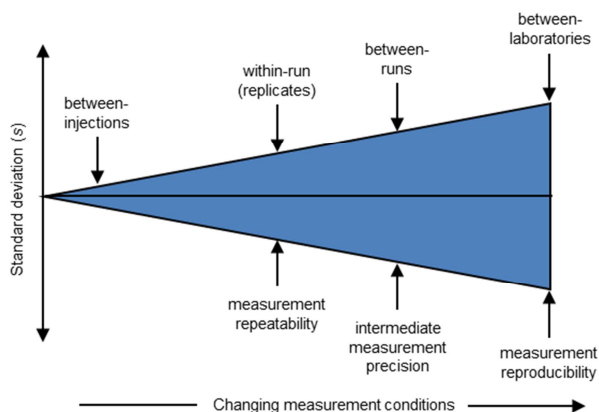

Figure 11 – Graphical representation of the expected relationship between precision estimates obtained under different measurement conditions, shown in terms of the magnitude of the observed imprecision.

4.8 Measurement accuracy

closeness of agreement between a **measured quantity value** and a **true quantity value** of a **measurand** (VIM 2.13)

Measurement accuracy describes how close a *single* **measurement result** is to the **true quantity value** (VIM 2.11). **Accuracy,** therefore, is influenced by both the random and systematic effects on the **measurement result**. **Accuracy** cannot be given a numerical value but **measurement results** are said to be 'more accurate' when the **measurement errors**, and therefore the **measurement uncertainty**, are reduced. Figure 12 illustrates this concept, using the example of shooting at a target.
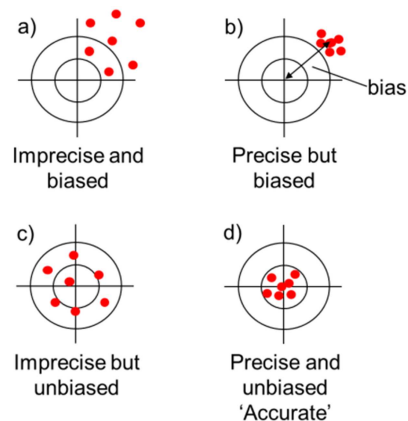
Figure 12 – Figurative representation of precision, bias and accuracy

The 'shots' on the target represent individual **measurement results**; the **reference quantity value** is the centre of the target. The best **accuracy** (lowest **measurement uncertainty**) is achieved in case d) where the individual results are all close to the **reference value**. In cases c) and d) there is no significant **bias** as the results are all clustered in the centre of the target. However, the **precision** is poorer in case c) as the results are more widely scattered. The **precision** in case b) is similar to that in case d). However, there is a significant **bias** in case b) as all the results are off-set from the centre in the same area of the target. The **accuracy** is poorest in case a) as the results are widely scattered and are off-set to the right of the target.

Measurement accuracy cannot be used to give a quantitative indication of the reliability of **measurement results**. Here an estimate of **measurement uncertainty** is required (see chapter 3).

Appendix

Table A 1 *Concepts discussed in this Guide, synonyms and the VIM reference. Concept in bold is the preferred term. The VIM reference is shown in bold for concepts where the full definition is given in this Guide.*

Concept	Synonym	VIM 3 Reference
base quantity		*1.4*
base unit		*1.10*
blank indication	background indication	*4.2*
calibration		**2.39**
calibration curve		*4.31*
calibration diagram		*4.30*
calibration hierarchy		*2.40*
calibrator		**5.12**
certified reference material	CRM	*5.14*
combined standard measurement uncertainty	combined standard uncertainty	*2.31*
commutability of a reference material		**5.15**
correction		*2.53*
coverage factor		*2.38*
definitional uncertainty		*2.27*
derived quantity		*1.5*
derived unit		*1.11*
detection limit	limit of detection	**4.18**
displaying measuring instrument		*3.4*
expanded measurement uncertainty	expanded uncertainty	*2.35*
indicating measuring instrument		*3.3*
indication		**4.1**
influence quantity		*2.52*
input quantity in a measurement model	input quantity	*2.50*
instrumental drift		**4.21**
intermediate measurement precision	intermediate precision	*2.23*
intermediate precision condition of measurement	intermediate precision condition	*2.22*
international measurement standard		*5.2*
International System of Quantities	ISQ	**1.6**
International System of Units	SI	*1.16*

Concept	Synonym	VIM 3 Reference
intrinsic measurement standard	intrinsic standard	*5.10*
kind of quantity	kind	*1.2*
material measure		*3.6*
measurand		**2.3**
measured quantity value	value of a measured quantity, measured value	**2.10**
measurement		**2.1**
measurement accuracy	accuracy of measurement, accuracy	**2.13**
measurement bias	bias	*2.18*
measurement error	error of measurement, error	**2.16**
measurement method	method of measurement	*2.5*
measurement model	model of measurement, model	*2.48*
measurement precision	precision	**2.15**
measurement principle	principle of measurement	*2.4*
measurement procedure		**2.6**
measurement repeatability	repeatability	*2.21*
measurement reproducibility	reproducibility	*2.25*
measurement result	result of measurement	**2.9**
measurement standard	etalon	**5.1**
measurement trueness	trueness of measurement, trueness	**2.14**
measurement uncertainty	uncertainty of measurement, uncertainty	**2.26**
measurement unit	unit of measurement, unit	**1.9**
measuring instrument		**3.1**
measuring interval	working interval	**4.7**
measuring system		**3.2**
metrological comparability of measurement results	metrological comparability	**2.46**
metrological compatibility of measurement results	metrological compatibility	**2.47**
metrological traceability		**2.41**
metrological traceability chain	traceability chain	*2.42*
metrological traceability to a measurement unit	metrological traceability to a unit	*2.43*
metrology		**2.2**
national measurement standard	national standard	*5.3*
nominal property		**1.30**
nominal quantity value	nominal value	**4.6**

Concept	Synonym	VIM 3 Reference
primary reference measurement procedure	primary reference procedure	**2.8**
primary measurement standard	primary standard	*5.4*
quantity		**1.1**
quantity value	value of a quantity, value	**1.19**
random measurement error	random error of measurement, random error	*2.19*
reference material	RM	*5.13*
reference measurement procedure		**2.7**
reference measurement standard	reference standard	*5.6*
reference quantity value	reference value	**5.18**
relative standard measurement uncertainty		*2.32*
repeatability condition of measurement	repeatability condition	*2.20*
reproducibility condition of measurement	reproducibility condition	*2.24*
secondary measurement standard	secondary standard	*5.5*
selectivity of a measuring system	selectivity	**4.13**
sensitivity of a measuring system	sensitivity	*4.12*
standard measurement uncertainty	standard uncertainty of measurement, standard uncertainty	*2.30*
system of quantities		**1.3**
system of units		*1.13*
systematic measurement error	systematic error of measurement, systematic error	*2.17*
target measurement uncertainty	target uncertainty	*2.34*
travelling measurement standard	travelling standard	*5.8*
true quantity value	true value of a quantity, true value	*2.11*
type A evaluation of measurement uncertainty	type A evaluation	*2.28*
type B evaluation of measurement uncertainty	type B evaluation	*2.29*
uncertainty budget		**2.33**
validation		**2.45**
verification		**2.44**
working measurement standard	working standard	*5.7*

Bibliography

1. ISO/IEC Guide 99:2007, International vocabulary of metrology – Basic and general concepts and associated terms (VIM), ISO/IEC, Geneva (2007).

2. JCGM 200:2012, International vocabulary of metrology – Basic and general concepts and associated terms (VIM). Available from www.bipm.org.

3. VIM definitions with informative annotations. Available at https://jcgm.bipm.org/vim/en/index.html.

4. ISO/IEC 17025:2017, General requirements for the competence of testing and calibration laboratories, ISO/IEC, Geneva (2017).

5. ISO 15189:2012, Medical laboratories – Requirements for quality and competence, ISO, Geneva (2012).

6. ISO/IEC 17043:2010, Conformity assessment – General requirements for proficiency testing, ISO/IEC, Geneva (2010).

7. ISO Guide 33:2015, Reference materials – Good practice in using reference materials, ISO, Geneva (2015).

8. ISO Guide 35:2017, Reference materials – Guidance for characterization and assessment of homogeneity and stability, ISO, Geneva (2017).

9. ISO 9000:2015, Quality management systems – Fundamentals and vocabulary, ISO, Geneva (2015).

10. ISO 17034:2016, General requirements for the competence of reference material producers, ISO, Geneva (2016).

11. The International System of Units (SI), Bureau International des Poids et Mesures, (9th ed, 2019). Available from www.bipm.org.

12. International vocabulary of basic and general terms in metrology, 2nd Edition, ISBN 92-67-10175, ISO, Geneva (1993).

13. G. Nordin, R. Dybkaer, U. Forsum, X. Fuentes-Arderiu, F. Pontet, Vocabulary on nominal property, examination, and related concepts for clinical laboratory sciences (IFCC-IUPAC Recommendations 2017), Pure Appl. Chem., **90**, 913–935 (2018). Available from www.degruyter.com/view/j/pac.

14. EN 71-3:2013+A1:2014, Safety of toys. Migration of certain elements, CEN, Brussels (2014).

15. S. Raut, S. Daniels, A. Heath, World Health Organization. Biologicals Unit & WHO Expert Committee on Biological Standardization, An international collaborative study to assign value to the 8th [eighth] WHO international standard for blood coagulation factor VIII concentrate (07/350), World Health Organization, Geneva (2009). Available from www.who.int/iris/handle/10665/70137.

16. Council Directive 80/181/EEC of 20 December 1979 on the approximation of the laws of the Member States relating to units of measurement and on the repeal of Directive 71/354/EEC, Official Journal L 039, 15/02/1980, p40-50.

17. R. Bettencourt da Silva and A. Williams (Eds.), Eurachem/CITAC Guide: Setting and using target uncertainty in chemical measurement (2015). Available from www.eurachem.org.

18. Directive 98/79/EC of the European Parliament and of the Council of 27 October 1998 on *in vitro* diagnostic medical devices, Official Journal L 331, 07/12/1998, p1-37.

19. JCTLM Database of higher-order reference materials, measurement methods/procedures and services. Available at www.bipm.org/jctlm/.

20. S. S-C. Tai and M. J. Welch, Development and evaluation of a candidate reference method for the determination of total cortisol in human serum using ID-LC/MS and LC/MS/MS, Anal. Chem., **76**, 1008-1014 (2004).

21. S. L. R. Ellison and A. Williams (Eds), Eurachem/CITAC Guide: Metrological traceability in analytical measurement (2nd ed. 2019). ISBN 978-0-948926-34-1. Available from www.eurachem.org.

22. P. De Bièvre, R. Dybkaer, A. Fajgelj, D. Brynn Hibbert, Metrological traceability of measurement results in chemistry: Concepts and implementation, (IUPAC Technical Report), Pure Appl. Chem., **83**, 1873-1935 (2011). Available from www.degruyter.com/view/j/pac.

23. G. Schumann et al., IFCC primary reference procedures for the measurement of catalytic activity concentrations of enzymes at 37 °C. Part 9: Reference procedure for the measurement of catalytic concentration of alkaline phosphatase, Clin. Chem. Lab. Med., **49**, 1439-1446 (2011).

24. ISO 1736:2008, Dried milk and dried milk products – Determination of fat content – Gravimetric method (Reference method), ISO, Geneva (2008).

25. ISO 17511:2003, In vitro diagnostic medical devices – Measurement of quantities in biological samples – Metrological traceability of values assigned to calibrators and control materials, ISO, Geneva (2003).

26. Certificate of analysis, ERM®- DA470k/IFCC, 2015, https://crm.jrc.ec.europa.eu/.

27. V. Barwick and S. Wood, Meeting the traceability requirements of ISO 17025, 3rd ed., ISBN 0-948926-23-6, LGC, Teddington (2005). Available from www.lgcgroup.com/nml.

28. B. King, The selection and use of reference materials: A basic guide for laboratories and accreditation bodies, EEE/RM/062rev3 (2002). Available from www.eurachem.org.

29. ISO Guide 31:2015, Reference materials – Contents of certificates, labels and accompanying documentation, ISO, Geneva (2015).

30. W. G. Miller et al, IFCC Working Group Recommendations for Assessing Commutability Part 1: General Experimental Design, Clin. Chem., **64**, 447-454 (2018). Available from http://clinchem.aaccjnls.org/.

31. G. Nilsson et al., IFCC Working Group Recommendations for Assessing Commutability Part 2: Using the Difference in Bias between a Reference Material and Clinical Samples, Clin. Chem., **64**, 455-464 (2018). Available from http://clinchem.aaccjnls.org/.

32. J. R. Budd et al, IFCC Working Group Recommendations for Assessing Commutability Part 3: Using the Calibration Effectiveness of a Reference Material, Clin. Chem., **64**, 465-474 (2018). Available from http://clinchem.aaccjnls.org/.

33. JCGM 100:2008, Evaluation of measurement data – Guide to the expression of uncertainty in measurement. Available from www.bipm.org.

34. ISO/IEC Guide 98-3 (JCGM/WG1/100), Uncertainty of measurement – Part 3: Guide to the expression of uncertainty in measurement (GUM:1995), ISO/IEC, Geneva (2008).

35. S. L. R. Ellison and A. Williams (Eds), Eurachem/CITAC Guide: Quantifying uncertainty in analytical measurement (3rd ed., 2012). ISBN 978-0-948926-30-3. Available from www.eurachem.org.

36. Eurolab Technical Report 2007/1, Measurement uncertainty revisited: Alternative approaches to uncertainty evaluation, Eurolab (2007). Available from www.eurolab.org.

37. ISO 21748:2017, Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation, ISO, Geneva (2017).

38. ISO 11352:2012, Water quality – Estimation of measurement uncertainty based on validation and quality control data. ISO, Geneva (2012).

39. B. Magnusson, T. Näykki, H. Hovind, M. Krysell, E. Sahlin, Handbook for calculation of measurement uncertainty in environmental laboratories, Nordtest Report TR 537 (ed. 4) (2017). Available from www.nordtest.info.

40. Commission Directive 2009/90/EC of 31 July 2009 laying down, pursuant to Directive 2000/60/EC of the European Parliament and of the Council, technical specifications for chemical analysis and monitoring of water status, Official Journal L 201, 01/08/2009, p36-38.

41. Directive 2008/105/EE of the European Parliament and of the Council of 16 December 2008 on environmental quality standards in the field of water policy, amending and subsequently repealing

Council Directives 82/176/EEC, 83/513/EEC, 84/156/EEC, 84/491/EEC, 86/280/EEC and amending Directive 2000/60/EC of the European Parliament and of the Council, Official Journal L 348, 24/12/08, p84-97.

42. B. Magnusson and U. Örnemark (Eds.), Eurachem Guide: The fitness for purpose of analytical methods: A laboratory guide to method validation and related topics (2nd ed., 2014). ISBN 78-91-87461-59-0. Available from www.eurachem.org.

43. J. Vessman, R. I. Stefan, J. F. van Staden, K. Danzer, W. Lindner, D. Thorburn Burns, A. Fajgelj, and H. Müller, Selectivity in analytical chemistry, (IUPAC Recommendations 2001), Pure Appl. Chem., **73**, 1381-1386 (2001). Available from www.degruyter.com/view/j/pac.

44. L. A. Currie, Nomenclature in evaluation of analytical methods including detection and quantification capabilities (IUPAC Recommendations 1995), Pure Appl. Chem., **67**, 1699-1723 (1995). Available from www.degruyter.com/view/j/pac.