

Measurement verification in the clinical laboratory:

A guide to assessing analytical performance during the acceptance testing of methods (quantitative examination procedures) and/or analysers

Zahra Khatami¹, Robert Hill², Catherine Sturgeon³, Edward Kearney⁴, Peter Breadon⁵ Anders Kallner⁶

¹Department of Clinical Biochemistry, Queen's Hospital Romford, ²Department of Clinical Biochemistry, King's Mill Hospital, Sutton-in-Ashfield, ³Department of Clinical Biochemistry, Edinburgh Royal Infirmary ⁴Department of Clinical Biochemistry, Queen Elizabeth The Queen Mother Hospital, Margate, ⁵Department of Clinical Biochemistry, Southport and Ormskirk Hospitals, ⁶Department of Clinical Chemistry Karolinska University Hospital, Stockholm, Sweden,

This paper and the associated spreadsheets, background statistical notes and worked examples were commissioned by the Scientific Committee of the Association for Clinical Biochemistry. Anders Kallner represented the Nordic Association for Clinical Chemistry. Peter Breadon represented the Institute of Biomedical Sciences.

Table of Contents

1. Aims	2
2. Scope	2
3. Definition of verification.....	2
4. Assumptions	2
5. Assessing imprecision	3
5.1 Choice of procedure and material	3
5.2 Procedure for assessing imprecision	3
6. Assessing bias	4
6.1 Specific procedure for assessing bias using patient samples	4
6.2 Specific procedures for assessing bias using reference material with an assigned value	5
6.3 Specific procedure for assessing bias using material from External Quality Assessment (EQA) Schemes.....	5
7. Instrument dilution check.....	5
7.1 Procedure.....	6
8. Glossary.....	7
9. Acknowledgements	8
10. References	9

Key words: Quality assessment, patient safety, procurement, measurement system, instrument, precision, imprecision, trueness, bias, dilution, clinical laboratory, guidance.

1. Aims

1. Provide practical recommendations on the minimum requirements for the verification of analytical performance when introducing new or modified methods and/or analysers into a clinical laboratory, i.e. to ensure that performance meets the performance characteristics claimed by the manufacturer of the measurement system.
2. Ensure patient safety by verifying analytical performance.

2. Scope

1. Experimental confirmation of the key analytical components of measurement uncertainty, i.e. precision, trueness {1} and linearity of dilution as claimed by the manufacturer.
2. Comparison of manufacturers' performance claims with data collected in the laboratory.
3. A consideration of the uncertainties associated with pre- and post-analytical procedures has not been included.
4. The procedures below cannot be used to verify ordinal (semi-quantitative) or dichotomous results.

3. Definition of verification

The generic definition of verification is: 'provision of objective evidence that a given item fulfils specified requirements' {1}.

In this document 'verification' is used to describe the experimental confirmation of performance characteristics by the user. This usually includes an assessment of precision (imprecision) and trueness (bias).

In Europe it is the statutory responsibility of the manufacturer to determine, describe and verify the performance of the measuring system (EU IVD Directive 98/79 {2}).

4. Assumptions

1. Verification is a responsibility of the laboratory director who ensures that it is completed in a timely manner within available resources (i.e. time and expertise).
2. There is an understanding of the measurement uncertainty {3, 4} (e.g. within and between measurement results) of laboratory results.
3. The performance specifications set by the user in the tendering specification are met by the manufacturer in their performance claims.
4. The performance claims made by the manufacturer were obtained on a comparable instrument, using appropriate reagents and calibrators.
5. Any modifications to the measurement system since the establishment of the performance claims are stated.
6. The manufacturer supplies, on demand, the results and the experimental design used in the studies to formulate their performance claims.
7. Where relevant, the following data have been made available: Analytical sensitivity, specificity, detection limit, measurement interval, linearity, effect of interferences including hooking and sample carry-over, and uncertainty profile.

8. Performance claims by the manufacturer have been determined using procedures recognised by appropriate regulatory bodies (e.g. ISO, CEN,) or professional organisations (e.g. IFCC, or CSLI). These procedures usually take more variables into consideration than verification by the laboratory. For example, if the CLSI EP5 guideline is followed by the manufacturer, it would be expected that the imprecision estimated in the laboratory during its verification would be smaller than that claimed by the manufacturer. If other procedures have been used for verification, the experimental design must be fully described.
9. Staff training and familiarisation with the new measurement system have been completed.
10. Laboratory infrastructure (e.g. interfacing, electrical connections, air conditioning, UPS support, etc.) is adequate.
11. Health and Safety requirements have been met.

The procedures described in this document may not be sufficient to establish that a measuring system is fit for purpose in the clinical context for which it is intended.

5. Assessing imprecision

5.1 Choice of procedure and material

An experimental design and statistical approach that is comparable to that used by the manufacturer should be chosen to allow a realistic verification of the manufacturer's claims. Either patient pooled samples or reference materials may be used. If reference materials are used commutability between the reference material and patient material should be demonstrated or documented by the manufacturer.

Where the medical decision point is close to the detection limit (e.g. Troponin, TSH or PSA) then pooled patient samples should be used.

Patient samples

Pooled samples should be prepared at two or more clinically relevant concentrations. All samples containing those interfering substances identified by the manufacturer e.g. drug metabolites, haemolysis, should be excluded. For stored samples, if the manufacturer's instructions and the stability of the measurand allow, refrigeration is the preferred method of storage to avoid artefacts introduced in the freeze-thaw cycle such as particulate matter.

Reference material

Reference material e.g. internal quality control material should be chosen at two or more clinically relevant concentrations.

5.2 Procedure for assessing imprecision

Before commencing verification ensure satisfactory calibration of the measurement system using the manufacturer's recommended procedure and ensure that reagents are within their expiry date.

1. Measure the concentration of one series of the samples per day for a minimum of five days. The series must consist of five replicate samples at two or more concentrations and the same material must be used in all five series. It is preferable to divide the original sample into five aliquots containing the same material, one for each day (series). If a series must be rejected because of operating difficulties or assumed outliers, record the circumstances, discard the data, and conduct

an additional series. **NOTE** It is advised that sufficient pooled material is prepared and stored under conditions that minimise sample deterioration.

2. Estimate the within and between series imprecision using analysis of variance components and the ANOVA statistic. [Spreadsheet A](#) is designed to calculate the imprecision, compare it with claims of the manufacturer and determine whether they are statistically comparable. **Note** that the manufacturer's claims may be based on more extensive studies e.g. CLSI EP-5 guideline.
3. The analytical conditions (e.g. reagent or calibrator lot) may be held constant or varied between series to simulate normal operating procedures. When imprecision is estimated between series under varying analytical conditions it is referred to as "intermediate imprecision" for the purpose of this paper.

6. Assessing bias

Bias should be estimated as part of the verification process and assessed with patient-based material. Bias should also be determined using a reference material with a traceable assigned value or against a peer group of laboratories using the same method or group of methods. It is important to know whether the two quantities measured are identical before drawing conclusions about bias estimates. Reference intervals should be reassessed considering any demonstrated significant bias.

6.1 Specific procedure for assessing bias using patient samples

Measurement of the bias against the previous or alternative laboratory method should be undertaken using paired patient comparisons. This approach is only comparative and provides no information about whether either method is biased against the appropriate reference method. A patient sample comparison is particularly useful when reference materials with assigned values do not exist or when peer groups in EQA schemes are small.

1. Obtain at least 20 patient samples of clinically relevant concentrations within the measuring interval of the method. All samples containing those known interfering substances identified by the manufacturer (e.g. drug metabolites, haemolysis) should be excluded.
2. Measure the concentrations of the samples, preferably in duplicate, using both the test and comparative procedures. Enter the results into [Spreadsheet B](#). Measurement of paired samples should be performed within a short time interval taking into account the stability of the measurand under the storage conditions used. It is preferable to use separate sample cups/tubes for each replicate rather than allowing the instrument to sample twice from the same cup/tube.
3. Where duplicates have been measured, their mean should be used for comparison purposes.
4. A scatter plot and difference plots are displayed in the spreadsheet B which also calculates the mean bias (difference) between the measurements, and estimates the statistical significance between the results using Student' *t*-test and Wilcoxon sign rank test for parametric and non parametric distributions respectively.
5. The spreadsheet allows partitioning of the results to further detail the bias in up to three concentration intervals.
6. A clinically significant bias should be confirmed by an extended study of 40 samples measured in duplicate {5}.

Information from the above procedures can be used to decide whether amendments of reference intervals will be necessary. The methodology for establishing reference intervals is outside the scope of this paper.

6.2 Specific procedures for assessing bias using reference material with an assigned value

For each selected concentration of reference material, the results of the laboratory's measurements must be compared with the assigned value and its uncertainty. At least two different concentrations should be studied and preferably chosen close to medical decision points.

1. Measure the concentration of the analyte (component) under investigation in duplicate on three to five different occasions and enter the results in [Spreadsheet C](#) {6}.
2. A comparison with the value of the reference material is performed by the spreadsheet C, using Student's *t*-statistics.

Practical points

Commutable certified reference materials are available for a limited number of analytes (e.g.). The Joint Committee for Traceability in Laboratory Medicine (JCTLM) {7} maintains a list of materials that are available and may be obtained from NIBSC {8}, IRMM {9} NIST Some companies may provide similar materials that are not certified but have a nominal assigned value and uncertainty. Where no reference material exists (e.g. free thyroid hormones) bias assessment should be carried out using the procedures described in 6.3.

6.3 Specific procedure for assessing bias using material from External Quality Assessment (EQA) Schemes

Bias can be estimated by measuring the concentration of relevant analytes present in EQA materials. **NOTE** EQA samples may not always be commutable with clinical samples (e.g. matrix effects). The EQA target value used is the peer group consensus value. Its uncertainty is expressed as the SEM of the results of the peer group

1. Measure the concentration of the analytes under investigation in 7 to 10 EQA samples of different concentrations in duplicate.
2. Estimate the significance of the bias by calculating the difference between the peer group method mean of the EQA material and the mean of the results as measured in the laboratory. ([Spreadsheet D](#))

Practical points

Previously circulated EQA samples and their associated reports may be obtained from EQA providers. A representative mean, the standard error of the mean and a description of the methods used should be supplied.

7. Instrument dilution check

Confirmation of the measurement linearity of a method is not a requirement of verification. However, it is advisable to confirm that any built-in dilution procedure on an instrument provides an accurate dilution of samples that are above the measurement interval i.e. beyond the concentration at which the manufacturer recommends measurement without dilution. This is particularly important for measurands such as hormones, tumour markers and enzymes where occasional samples require dilution prior to measurement.

Reference materials and EQA samples with such high concentrations may be difficult to obtain but are an invaluable resource for measurement verification.

7.1 Procedure

1. Identify a reference material or EQA sample with a known or assigned high concentration of the measurand i.e. higher than the threshold above which the manufacturer recommends dilution but below the threshold above which a possible 'hook effect' is known to occur.
2. Conduct at least five dilution experiments measuring the concentration of the diluted sample five times. Enter the results to [Spreadsheet A](#) and estimate the average, within- between- and total imprecision (section 5.2). Alternatively, make three to six dilutions and measure the diluted material in duplicates (section 6.2). Analyze the results using [Spreadsheet C](#)

8. Glossary

analyte, substance or chemical whose concentration is measured in an analytical procedure. {10}

NOTE This term has been superseded by ‘component’ (defined below) but is used in this document to simplify the language of metrology.

analytical sensitivity, quotient of the change in an indication of a measuring system and the corresponding change in a value of a quantity being measured {1} § 4.12

analytical specificity, ability of a measuring system, using a specified measurement procedure to determine solely the measurand.

between series precision – see intermediate precision

certified reference material (CRM), reference material accompanied by documentation issued by an authoritative body and providing one or more specified property values with associated uncertainties and traceability, using valid procedures {1} § 5.14

commutability of a reference material, 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 {1} § 5.15

component, definable part of a system {11}

definitional uncertainty, component of measurement uncertainty resulting from the finite amount of detail in the definition of a measurand {1} § 2.27

detection limit, the minimum single result with a stated probability that can be distinguished from a suitable blank value. The limit defines the point at which analysis becomes possible and this may be different from the lower limit of determinable analytical interval. {1} § 4.18

hook effect, falsely low values on an immunoassay when an overwhelming amount of antigen affects the binding capacity of the added antibody; especially when testing for thyroglobulin in management of thyroid cancer

<http://www.medilexicon.com/medicaldictionary.php?t=28026> (accessed 090603)

intermediate precision, (intermediate measurement precision//intermediate precision) measurement precision under a set of intermediate precision conditions of measurement. {1} § 2.23

intermediate precision conditions of measurement, (intermediate precision condition), condition of measurement, out of a set of conditions that includes the same measurement procedure, same location, and replicate measurements on the same or similar objects over an extended period of time, but may include other conditions involving changes; {1} § 2.22

matrix effect, the combined effect of all components of a sample other than the analyte on the measurement of the quantity of interest. If a specific component can be identified as causing an effect then this is referred to as interference. {11}

measurand, quantity intended to be measured {1} § 2.3

measurement bias, – estimate of a systematic measurement error. {1} § 2.18

measurement procedure, 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; {1} § 2.6

measurement uncertainty, non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used; {1} § 2.26

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 uncertainty, under defined conditions; {1} § 4.7

measuring system, set of one or more measuring instruments and often other devices, including any reagent and used to generate measured quantity values within specified intervals for quantities of specified kinds.

NOTE A measuring system may consist of only one measuring instrument {1} § 3.2

precision, closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions; {1} § 2.15

reference material, material, sufficiently homogeneous and stable with reference to specified properties, which has been established to be fit for its intended use in measurement or in examination of nominal properties {1} § 5.13

repeatability, (measurement repeatability, repeatability), measurement precision under a set of repeatability conditions of measurement {1} § 2.21

sample carry-over, carry over from a preceding sample probe into a following specimen cup which will influence the result {11}

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; {1} § 2.41

trueness, closeness of agreement between the average of an infinite number of replicate measured quantity values and a reference quantity value; {1} § 2.14

uncertainty profile, absolute or relative uncertainty estimated at different concentrations within a measuring interval

verification, provision of objective evidence that a given item fulfils specified requirements. In the context of this document confirmation that performance characteristics of a measuring system are achieved. {1} § 2.44

VIM, International vocabulary of metrology — Basic and general concepts and associated terms – see reference 1

within series precision, – see repeatability

9. Acknowledgements

The encouragement and help in defining the scope for this paper from Gordon Avery (Abbott Diagnostics) and Poly Gerondaes (Ortho Clinical Diagnostics) is acknowledged. They also provided constructive criticism of the drafts.

The authors are also grateful to Graham White (Adelaide, Australia) and Sophie Barnes (London UK) for comments on early drafts of this document.

10. References

1. International vocabulary of metrology — Basic and general concepts and associated terms (VIM) JCGM 200:2008.VIM 07, Geneva 2008.
http://www.bipm.org/utis/common/documents/jcgm/JCGM_200_2008.pdf (accessed 100603).
2. Directive 98/79/EC of the European Parliament and of the Council of 27 October 1998 on in vitro diagnostic medical devices Official Journal 1998;41:L331.
<http://eur-lex.europa.eu> (accessed 100603).
3. Guide to the expression of Uncertainty in Measurements. ISO, Geneva 1993.
<http://www.bipm.org/en/publications/guides/gum.html> (accessed 100603).
4. Quantifying Uncertainty in analytical measurement, Eurachem/Citac
www.eurachem.org (accessed 100603).
5. Clinical and Laboratory Standards Institute (CLSI). Method Comparison and Bias Estimation Using Patient Samples; Approved Guideline—Second Edition. CLSI document EP9-A2 (ISBN 1-56238-472-4). Wayne, Pennsylvania 19087-1898 USA, 2002.
6. Clinical and Laboratory Standards Institute (CLSI). User Verification of Performance for Precision and Trueness; Approved Guideline—Second Edition. CLSI document EP15-A2 (ISBN 1-56238-574-7). Wayne, Pennsylvania 19087-1898 USA, 2005.
7. Joint Committee for Traceability in Laboratory Medicine (JCTLM).
<http://www.bipm.org/en/committees/jc/jctlm/> (accessed 100603)
8. National Institute for Biological Standards and Control (NIBSC)
http://www.nibsc.ac.uk/products/biological_reference_materials.aspx (accessed 101108)
9. Institute for Reference Materials and Measurements (IRMM).
<http://irmm.jrc.ec.europa.eu/html/homepage.htm> (accessed 100603).
10. Compendium of Chemical Terminology. IUPAC 2006.
<http://old.iupac.org/publications/compendium/> (accessed 101108)
11. Dybkaer R. Vocabulary for use in measurement procedures and description of reference materials in laboratory medicine. Clin Chem Clin Biochem 1997;35(2):141-173.

Version 9.27 Document date 22/10/10 Review date 22/10/11