NICOLAUS COPERNICUS UNIVERSITY IN TORUŃ
Faculty of Pharmacy
Collegium Medicum in Bydgoszcz

INTERNATIONAL SCIENTIFIC MEETING "EVALUATING AND MONITORING ANALYTICAL QUALITY IN THE TRACEABILITY ERA"

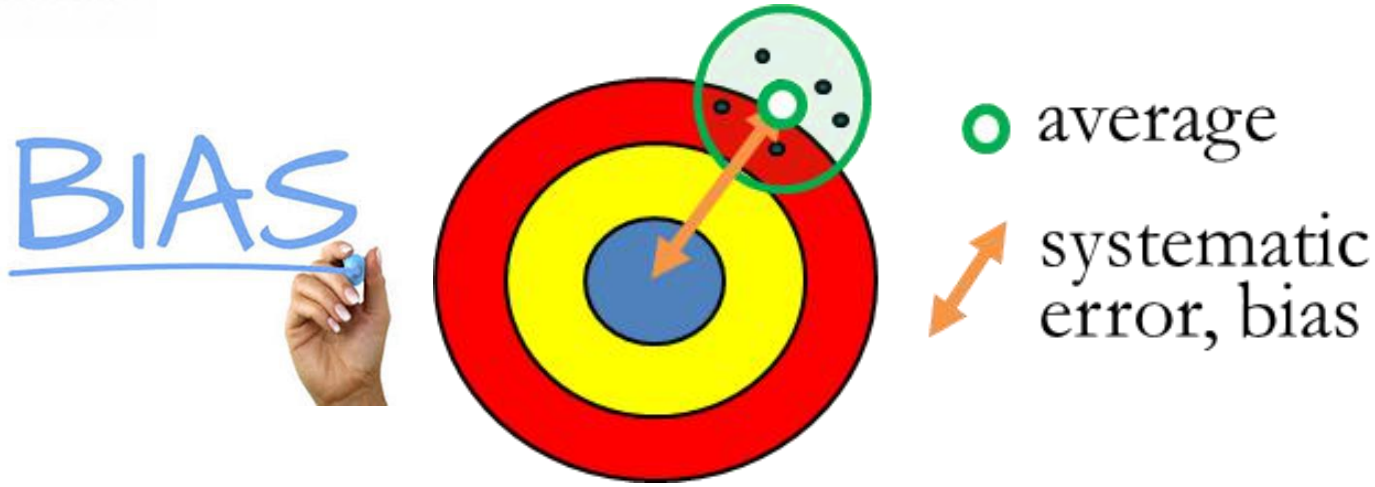
29 NOVEMBER 2024

Not all biases are created equal: how to deal with bias on laboratory measurements

Mauro Panteghini

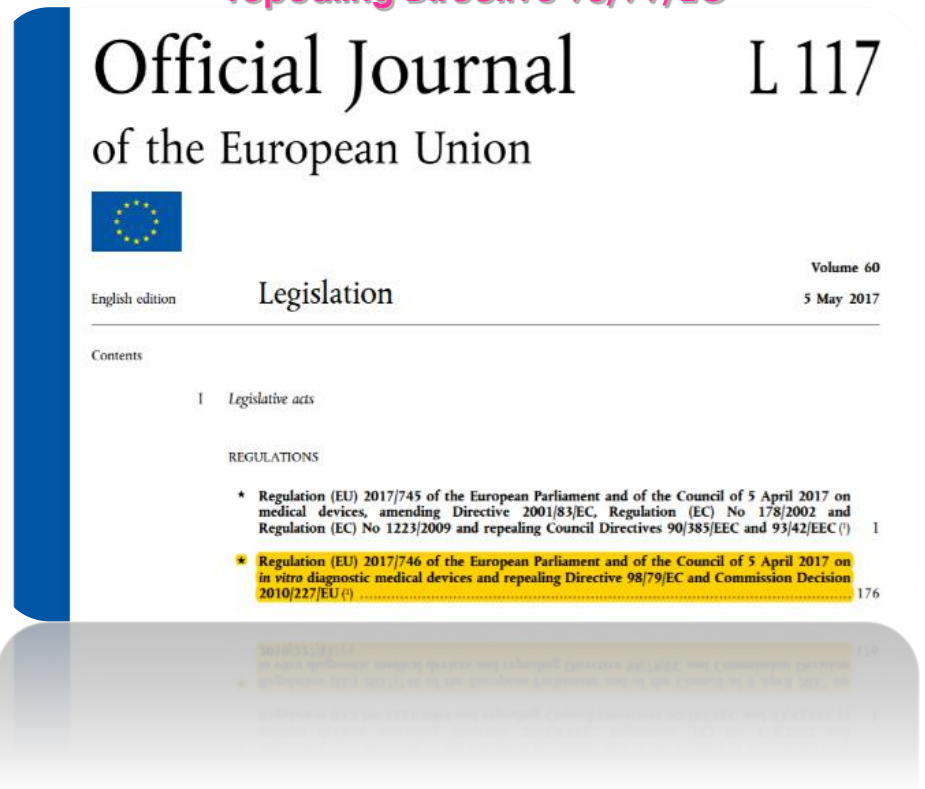
Definition International Vocabulary of Metrology, JCGM 200:2012

MEASUREMENT BIAS = estimate of a systematic measurement error = component of measurement error that in replicate measurements remains constant or varies in a predictable manner



EU 98/79/EC-IVD Directive

REGULATION (EU) 2017/746 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 5 April 2017 on *In Vitro* Diagnostic Medical Devices and repealing Directive 98/79/EC



Official Journal L 117

of the European Union

English edition — Legislation — Volume 60 — 5 May 2017

Contents

I *Legislative acts*

REGULATIONS

* Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC (1) 1

* Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on *in vitro* diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU (1) 176

Both require IVD manufacturers to ensure metrological traceability of their IVD-MDs to higher-order references

INTERNATIONAL STANDARD **ISO/FDIS 17511**

This standard is no longer just an informative document for the IVD manufacturers, but it has become a normative reference to the EU IVD Regulation, obliging the industry to implement metrological traceability as described in the document

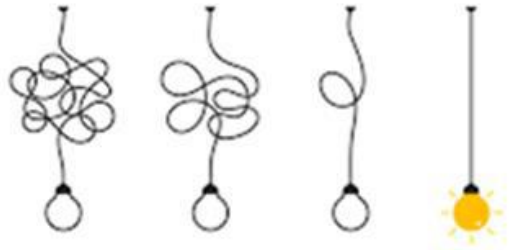
Safety
Quality
Traceability

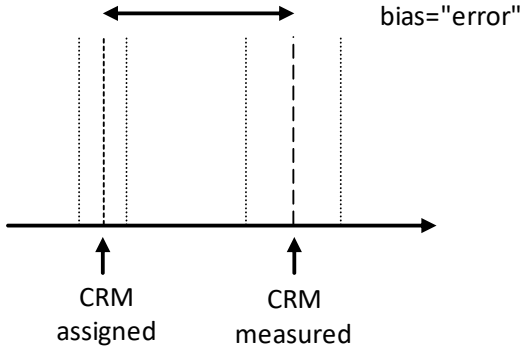
Key responsibilities of IVD manufacturers

- Identification of appropriate higher-order metrological references
- Definition of a calibration hierarchy to assign traceable values to measuring system calibrators and bias correction during the trueness transfer process
- Estimation of combined measurement uncertainty of calibrators
- Fulfillment of measurement uncertainty specifications for calibrators, which represent a proportion of the uncertainty budget allowed for medical laboratory results

Braga F & Panteghini M, Clin Chim Acta 2014;432:55
Braga F et al. CCLM 2021;59:291

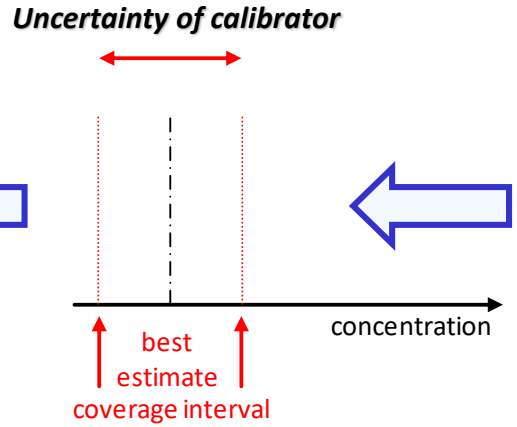
BIAS CORRECTION by IVD manufacturers during the trueness transfer process





Bias, **systematic measurement error due to, e.g., inappropriate model for the calibration curve, incorrect values assigned to the calibrators, matrix-related bias, etc.**

Bias correction, ***realignment of IVD-MD by adjusting the value assigned to the calibrator***



Provide unbiased clinical results

CE

Panteghini M. Clin Chem Lab Med 2024 Nov 26. doi: 10.1515/cclm-2024-1208. Online ahead of print

How to deal with bias on laboratory measurements

Which bias?

1. Bias arising from an insufficient correction during the implementation of traceability to higher-order references, because of manufacturing limitations, such as, e.g., the use of too large validation criteria for traceability and value assignment of different IVD calibrator lots.

The manufacturers' internal specifications to validate the calibrator traceability to higher-order references (and its measurement uncertainty) may not be based on suitable performance specifications

Abbott - γ-Glutamyltransferase 2^ gen. -

→Manufacturer's calibrator release specification: ±14% from the target

BUT: Desirable maximum allowable uncertainty on clinical samples: ±4.5%

BECKMAN COULTER - Serum folate -

→Manufacturer's calibrator release specification: ±10% from the target

BUT: Desirable maximum allowable uncertainty on clinical samples: ±8.0%

How to deal with bias on laboratory measurements

Which bias?

1. IVD-MD BIAS → when medically unacceptable, this source of bias in the individual laboratory's stable performance can be discovered by the ongoing IVD-MD surveillance by traceability-based EQA and should be confirmed by an *ad-hoc* experiment

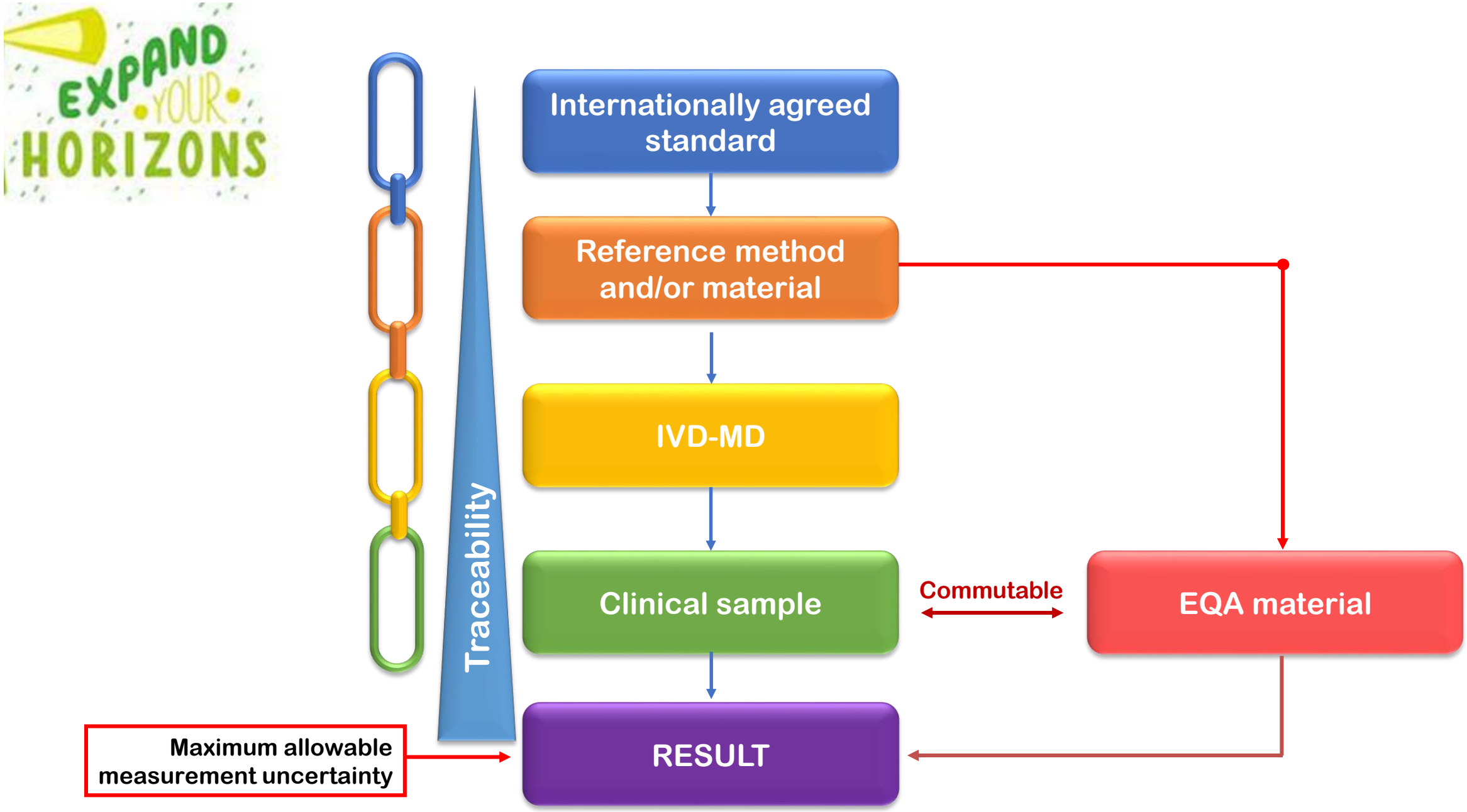
Requirements for a traceability-based EQA

Feature	Aim
EQA material value-assigned with reference measurement procedures or strictly controlled procedures, if the reference procedure is lacking	To check the traceability of employed IVD-MD to reference measurement systems and the performance of participating laboratories against higher-order references
Proved commutability of EQA materials	To allow transferability of partici-pating laboratory performance to the measurements of patient samples
Use of objectively defined analytical performance specifications	To verify the suitability of laboratory measurements in clinical setting

Panteghini M. Clin Chem Lab Med 2023;61:759

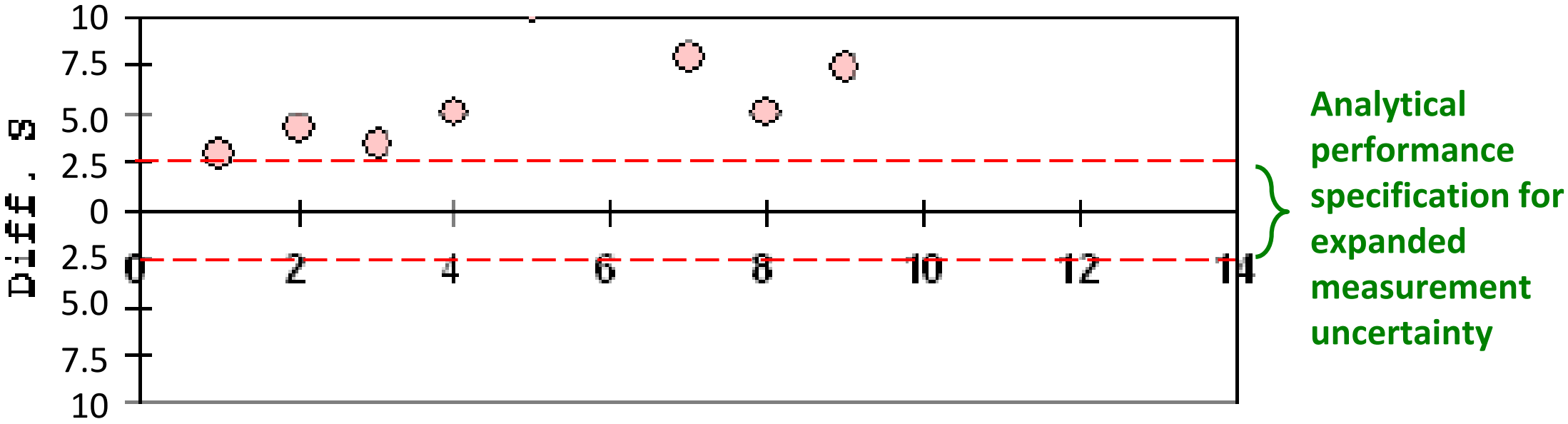
EQA evaluation of the performance of participating laboratories in terms of traceability of their measurements and harmonization of their results

- AIM: Confirm (or not) the analytical quality required to deliver laboratory test information that would satisfy clinical needs, including the traceability of the calibration and the test result equivalence among laboratories (i.e., result harmonization).
- PREMISE: EQA programs should meet requirements for evaluation of the performance of participating laboratories in terms of traceability of their measurements and harmonization of their results.
- CRITERIUM: The deviation of a laboratory measurement from the value assigned to the EQA material should stay within the expanded MAU for that measurand.



[Panteghini M & Krintus M. Crit Rev Clin Lab Sci 2024]

If EQA results in following exercises are all above or below the maximum allowable uncertainty limits, a medically unacceptable measurement bias can be suspected*

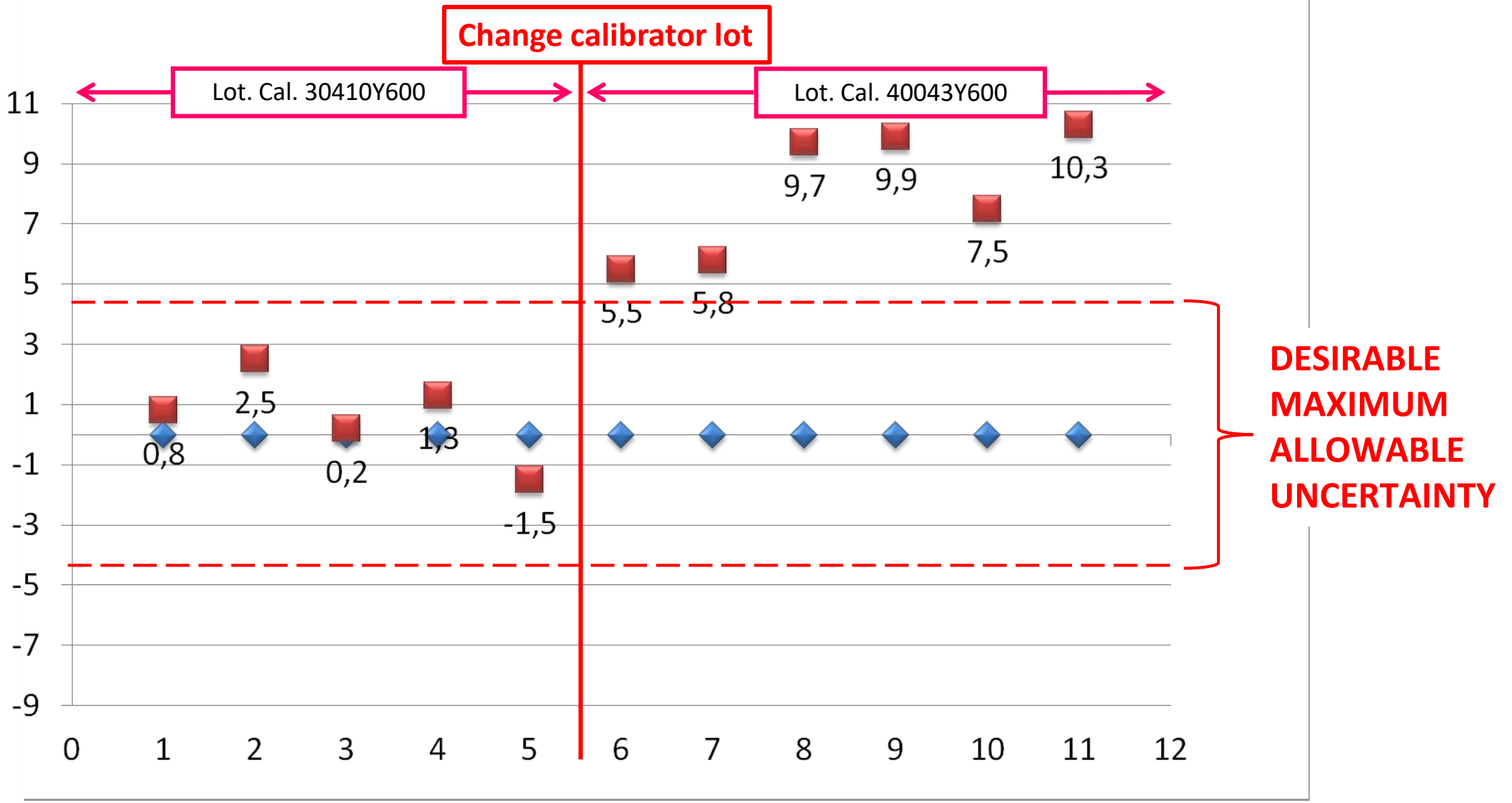


* The drawback with EQA is that the laboratory result is based on one measurement for exercise, which results in an increased uncertainty compared to a mean value. This is why a suspicion of bias needs to be confirmed by at least 3 following results all above or below the limits.

Example: Creatinine measurement by Abbott Architect

NICOLAUS COPERNICUS UNIVERSITY IN TORUŃ
Faculty of Pharmacy
Collegium Medicum in Bydgoszcz

Pasqualetti S et al. Clin Chim Acta. 2015;450:125



How to confirm an IVD-MD bias by a*d-hoc* experiments

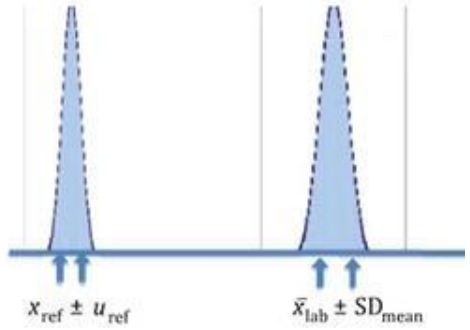
Note 1 to entry: Difference between the accepted value of a commutable reference material and the mean value of replicate measurements produced under repeatability conditions by a medical laboratory measurement procedure

Note 2 to entry: Difference between the mean value of replicate measurements produced by a reference measurement procedure and the mean value of replicate measurements produced under repeatability conditions by a medical laboratory measurement procedure.

ISO/TS 20914:2019
MEDICAL LABORATORIES -- PRACTICAL GUIDANCE FOR THE ESTIMATION OF MEASUREMENT UNCERTAINTY

How to confirm an IVD-MD bias by a*d-hoc* experiment (I)

Use of a commutable certified reference material (CRM) as trueness control



- Perform triplicate measurements of CRM for four consecutive days (n=12)
- Calculate the mean of means of triplicate measurements
- Calculate bias by using the mean of means of triplicate measurements and target value of CRM
- The estimated bias is considered significant if: $(x_{ref} - x_{IVD\text{-}MD}) > 2 * u_{bias}$, where x_{ref} is the certified CRM value, $x_{IVD\text{-}MD}$ is the mean of mean of means of triplicate measurements, and u_{bias} is equal to $\sqrt{(u_{ref}^2 + SD_{mean}^2)}$, where u_{ref} is the standard MU of the CRM and SD_{mean} is the standard deviation of the mean value of CRM obtained by the evaluated IVD-MD, calculated as $SD/\sqrt{n}$, where SD is the standard deviation of the replicate measurements of the CRM and n=12.

Example: Creatinine measurement by Abbott Architect

NICOLAUS COPERNICUS UNIVERSITY IN TORUŃ
Faculty of Pharmacy
Collegium Medicum in Bydgoszcz

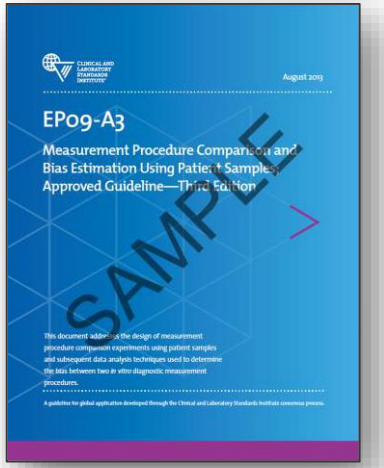
Pasqualetti S et al. Clin Chim Acta. 2015;450:125

enzymatic assay on Architect *c*16000 platform after calibration with two different lot of system calibrator. Data obtained by measurements of NIST SRM 967a reference material (certified value ± expanded uncertainty: L1, 0.847 mg/dL ± 0.018 mg/dL and L2, 3.877 mg/dL ± 0.082 mg/dL).

	SRM 967a level 1	SRM 967a level 2
Multigent Clin Chem Calibrator lot no. 40043Y600		
Imprecision (u_{Rw})	0.47%	0.40%
Bias	3.57%	7.05%

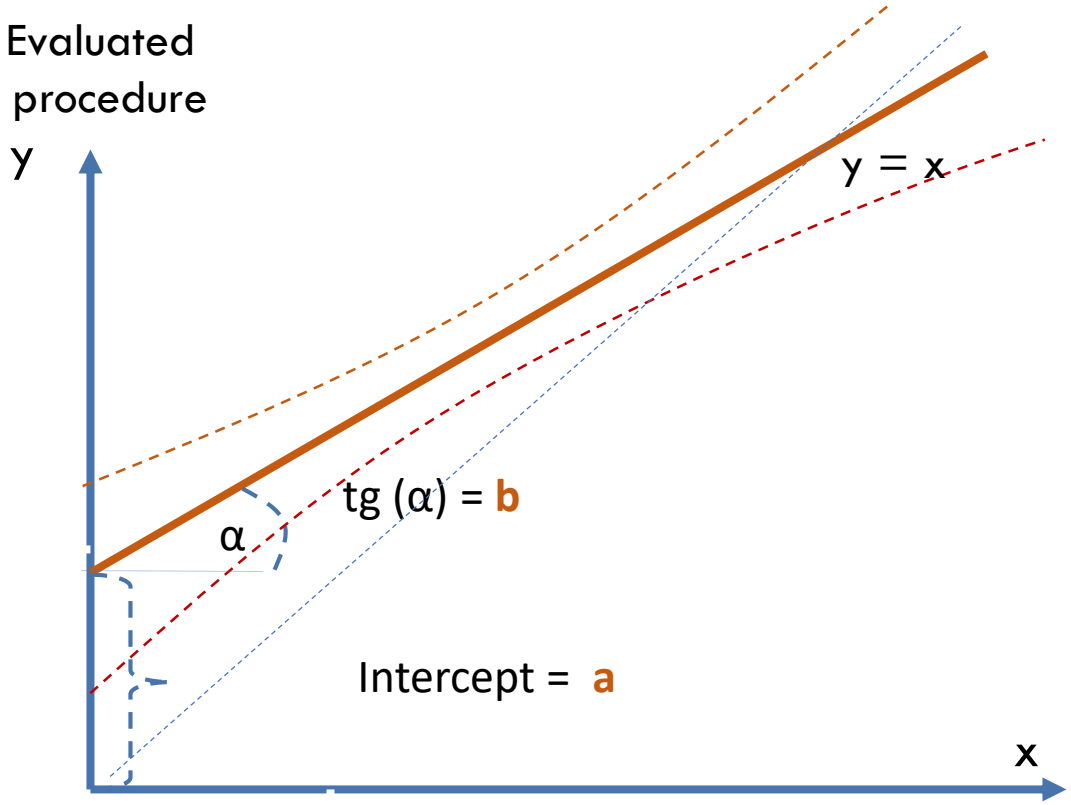
How to confirm an IVD-MD bias by a*d-hoc* experiment (II)

Method comparison study on a set of clinical samples with a mentor procedure (reference measurement procedure, if available)



1. Minimum 40 clinical samples
2. Cover 90% of the method measurement range
3. Replicate measurements with both methods

Significant constant and proportional bias



Regression equation

y = a (95% CI) + b (95% CI) x

Excluding 0
Constant error

Excluding 1
Proportional error

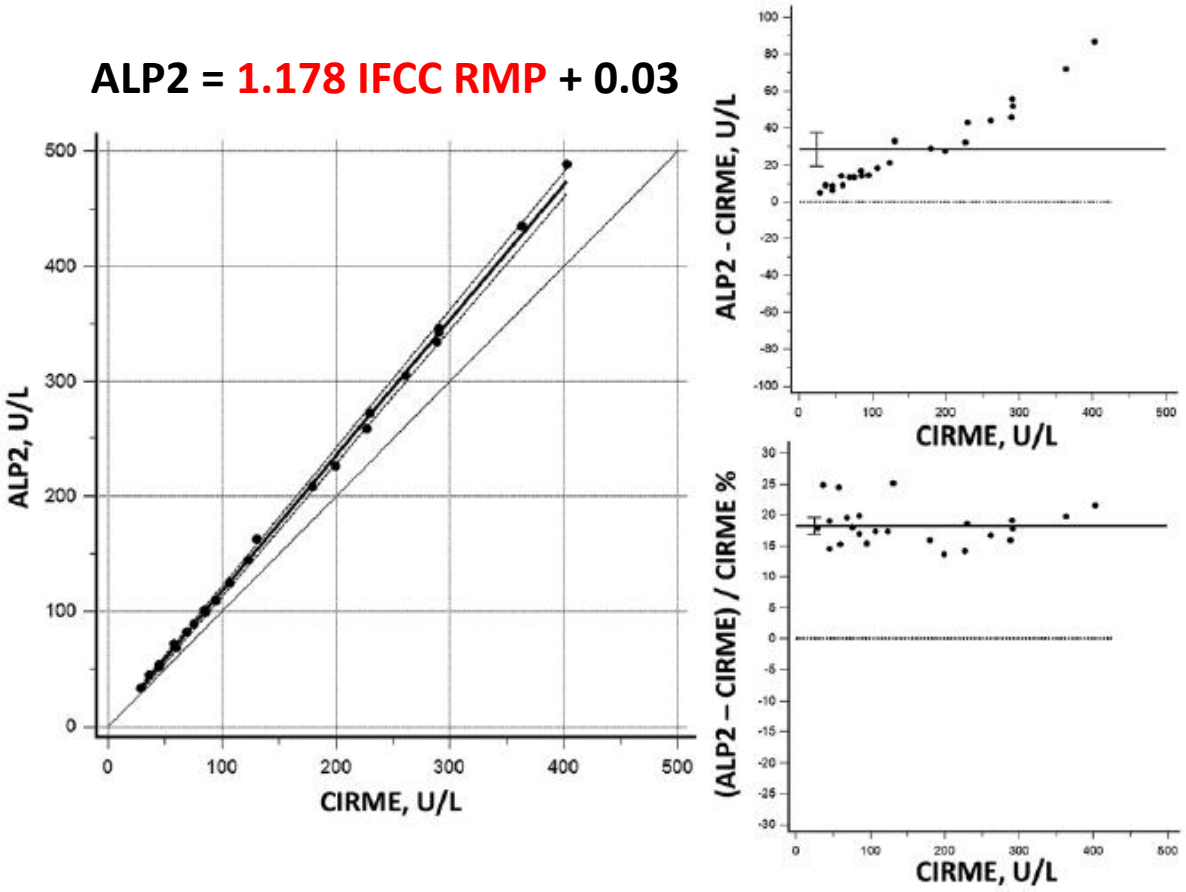
Example: ALP measurement by Abbott Alinity

NICOLAUS COPERNICUS UNIVERSITY IN TORUŃ
Faculty of Pharmacy
Collegium Medicum in Bydgoszcz

Bianchi G et al. Clin Chem Lab Med. 2024;62:280

Comparison between Abbott Alinity ALP 2nd gen. *vs.* IFCC reference measurement procedure

ALP2 = 1.178 IFCC RMP + 0.03



Even if a measurement bias is statistically significant, the final assessment of significance should be based on its impact on clinical interpretation of measurement results

PLEASE NOTE

Steps related to how to deal with major bias on clinical measurements

1. Discover a medically unacceptable measurement bias during the external quality assessment (EQA) program (only schemes fulfilling category I/IIA criteria are however usable to this scope)
2. If a medically significant bias (meaning a bias that does not fulfill the corresponding performance specifications) is suspected during the ongoing EQA surveillance, the bias against a reference (material or procedure) for that measurand should be estimated and the presence of a significant systematic error confirmed. Note that as reference may act any material or procedure positioned at the top of the corresponding traceability chain, even in the absence of high-order options
3. The obtained bias value should be included in the estimate of measurement uncertainty (MU) of clinical samples

Braga F & Panteghini M. Clin Chem Lab Med 2020;58:1407

Milan model allocation and proposed analytical performance specifications (APS) for standard MU on clinical samples of 56 common laboratory measurands

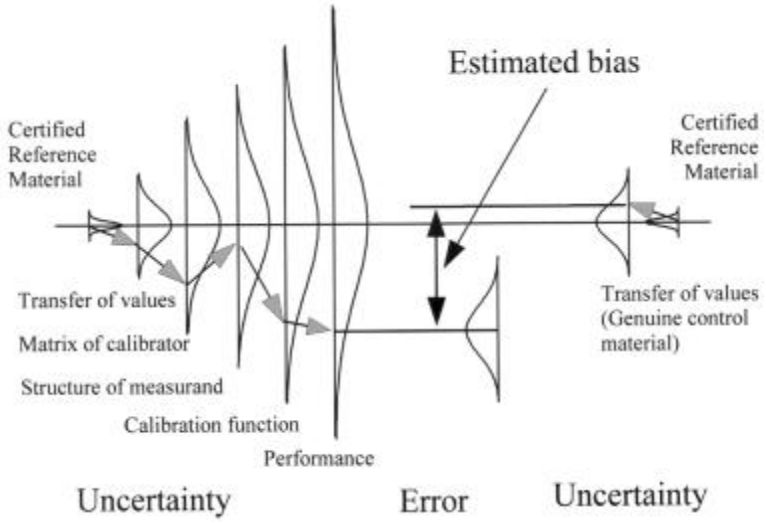
Measurand	u_{result} APS, %	
	Desirable	Minimum
Outcome-based model		
Fasting plasma glucose	2.00	3.00
Blood HbA_{1c}	3.00	3.70
Blood total hemoglobin	5.60	8.50
Serum total cholesterol	3.00	7.00
Serum HDL cholesterol	2.90	5.60
Serum triglycerides	6.10	12.4
Serum cardiac troponin	9.40	13.0
Urine albumin	9.00	17.0
Serum total folate	8.00	12.0
Serum 25-hydroxyvitamin D_3	10.0	15.0
Serum transferrin saturation	10.0	15.0
State-of-the-art model		
Serum C-reactive protein	3.76	5.64
Serum intact human chorionic gonadotropin	4.55	6.83
Temporarily belonging to state-of-the-art model[d]		
Serum ferritin	4.31	6.47
Serum thyroid stimulating hormone	2.89	4.34
Urine total protein	4.97	7.46
Model 1&2[e]		
Serum digoxin	6.00	9.00
Blood cyclosporine	11.5	17.2
Blood everolimus	4.03	6.05
Blood sirolimus	3.44	5.17
Blood tacrolimus (immediate release)	8.27	12.4
Temporarily belonging to biological variation model[c]		
Serum albumin	1.25	1.88
Plasma D-dimer	10.6	15.9
Blood platelets	4.85	7.28
Serum alanine aminotransferase	4.65	6.98
Serum creatine kinase	7.25	10.9
Serum pancreatic lipase	3.85	5.78
Serum pancreatic amylase	3.15	4.73
Biological variation model		
Serum sodium	0.27	0.40
Serum potassium	1.96	2.94
Serum chloride	0.49	0.74
Serum total carbon dioxide	2.10	3.15
Serum total calcium	0.91	1.36
Serum inorganic phosphate	3.84	5.75
Serum magnesium	1.44	2.16
Serum creatinine	2.20	3.30
Serum urea	7.05	10.6
Serum urate	4.16	6.24
Plasma lactate	13.6	20.4
Serum total bilirubin	10.5	15.7
Serum alkaline phosphatase	2.65	3.98
Serum aspartate aminotransferase	4.75	7.13
Serum γ-glutamyltransferase	4.45	6.68
Serum lactate dehydrogenase	2.60	3.90
Serum cholinesterase	2.10	3.15
Serum total proteins	1.30	1.95
Serum IgG	2.20	3.30
Serum IgA	2.50	3.75
Serum IgM	2.95	4.43
Serum prostate-specific antigen	3.40	5.10
Plasma homocysteine	3.52	5.27
Red blood cells	1.55	2.33
White blood cells	5.65	8.48
Serum free T3	2.35	3.53
Serum free T4	2.80	4.20
Serum parathyroid hormone	7.85	11.8

[Panteghini M. Clin Chem Lab Med 2024;62:1497]

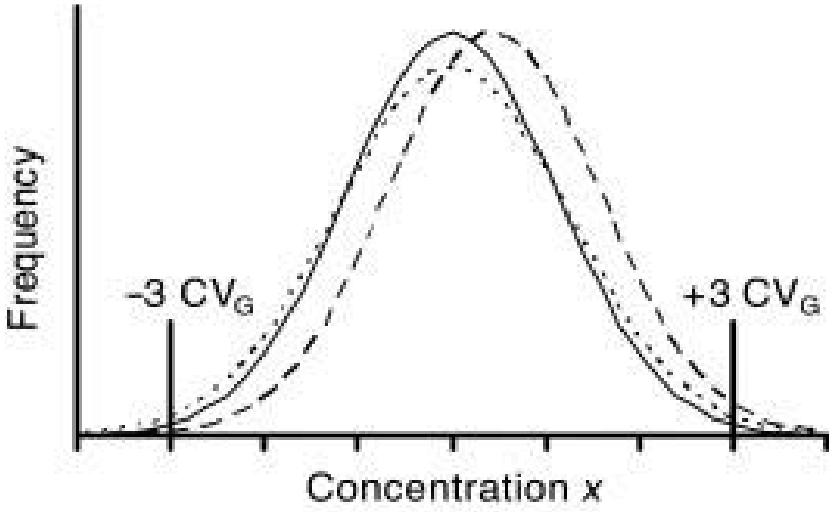
Including uncorrected IVD-MD bias in uncertainty calculations: a breathtaking twist of logic?

There has been a long debate about the correctness of the "variance approach" treating bias (asymmetric) as a standard MU component (symmetric), which is added in the usual root-sum-of-square (RSSu) manner:

$\rightarrow (MU_{cal}^2 + MU_{end\text{-}user}^2 + \text{IVD-MD-related bias}^2)^{½}$



[Hyltoft Petersen P et al. Clin Chem Lab Med 2001;39:589]



[Thienpont LM. Clin Chem 2008;54:1587]

The contentious nature of the “enlargement approach” is demonstrated by the numerous articles describing procedures for inclusion of uncorrected bias into the overall MU budget

Volume 102, Number 5, September–October 1997
Journal of Research of the National Institute of Standards and Technology
[J. Res. Natl. Inst. Stand. Technol. 102, 577 (1997)]

Guidelines for Expressing the Uncertainty of Measurement Results Containing Uncorrected Bias

Clinica Chimica Acta 495 (2019) 129–138
Contents lists available at ScienceDirect
Clinica Chimica Acta
journal homepage: www.elsevier.com/locate/cca

Review

Bias in analytical chemistry: A review of selected procedures for incorporating uncorrected bias into the expanded uncertainty of analytical measurements and a graphical method for evaluating the concordance of reference and test procedures

Robert Frenkel[a,*,1], Ian Farrance[b], Tony Badrick[c]

Clin Chem Lab Med 2001; 39(7):589–595 © 2001 by Walter de Gruyter · Berlin · New York

Meas. Sci. Technol. 9 (1998) 1010–1011. Printed in the UK
PII: S0957-0233(98)90780-9

DESIGN NOTE

Evaluation of the uncertainty associated with a measurement result not corrected for systematic effects

Ignacio H Lira† and Wolfgang Wöger‡

Accred Qual Assur 3:418–422
© Springer-Verlag 1998

Werner Hässelbarth

Measurement uncertainty procedures revisited: Direct determination of uncertainty and bias handling

Anal Bioanal Chem (2008) 390:201–213
DOI 10.1007/s00216-007-1693-1

REVIEW

Treatment of uncorrected measurement bias in uncertainty estimation for chemical measurements

Bertil Magnusson · Stephen L. R. Ellison

Opinion Paper

Models for Combining Random and Systematic Errors. Assumptions and Consequences for different Models

Per Hyltoft Petersen[1,2], Dietmar Stöckl[3], James O. Westgard[4], Sverre Sandberg[2], Kristian Linnet[5] and Linda Thienpont[3]

Talanta 65 (2005) 829–837
www.elsevier.com/locate/talanta

Attempts to include uncorrected bias in the measurement uncertainty

Václav Synek*

A variety of modeling methods using the “variance approach” have been proposed, but no one is perfect.

Whether statistics should primarily be used for making decisions or for investigating factors causing variation continues to be amongst the most compelling causes of disagreements in statistics.

“The analysis of variance is not a mathematical theorem but a simple method of arranging arithmetical facts to isolate and display the essential features of a body of data with the utmost simplicity”.

Sir Ronald Aylmer Fisher
1890-1962

“In medical laboratories attempting to fulfil APS for measurement uncertainty, there is a need to identify the factors that primarily cause variation in different circumstances to make effective decisions aimed at decreasing measurement uncertainty in the interest of fulfilling APS.”

Elvar Theodorsson, Clin Chem Lab Med 2024;62:1520

The choice of a specific “MU enlarging model” is therefore not compelling, particularly if a laboratory sets limits to acceptable bias.

The proposed model should be simple to apply in daily lab practice, *not for reporting MU with given results but just to establish if the estimated IVD-MD bias can influence the fulfillment of APS for MU.*

The “enlargement approach” combining the systematic error as a standard uncertainty with the other but genuine standard MU components by using RSS model should be considered the most pragmatic and simplest compromise.

RSSu curve at bias >1.8 x significant bias ratio start to overestimate the uncertainty interval slightly (“consistent expansion of the coverage properties”)

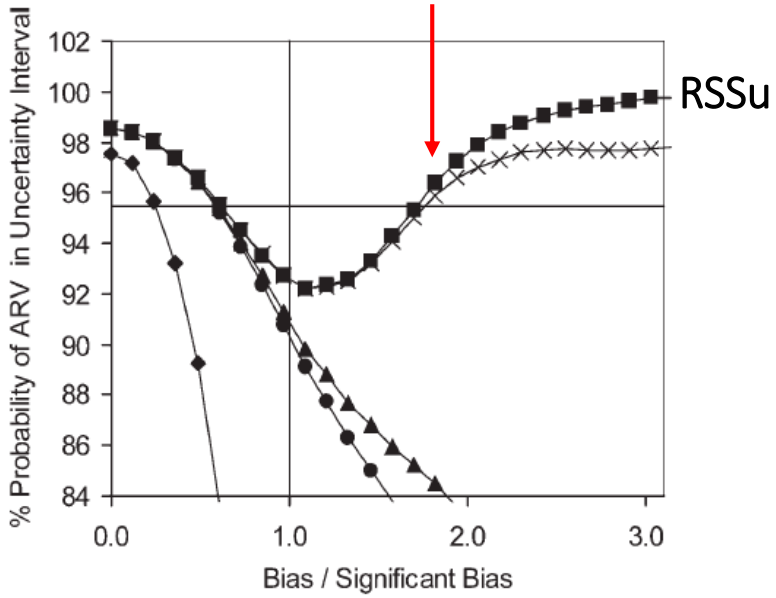


Fig. 3 Enlarge when significant. This graph shows the probability of the true value occurring within the uncertainty interval when the uncertainty interval is increased by an enlargement method when bias is significant.

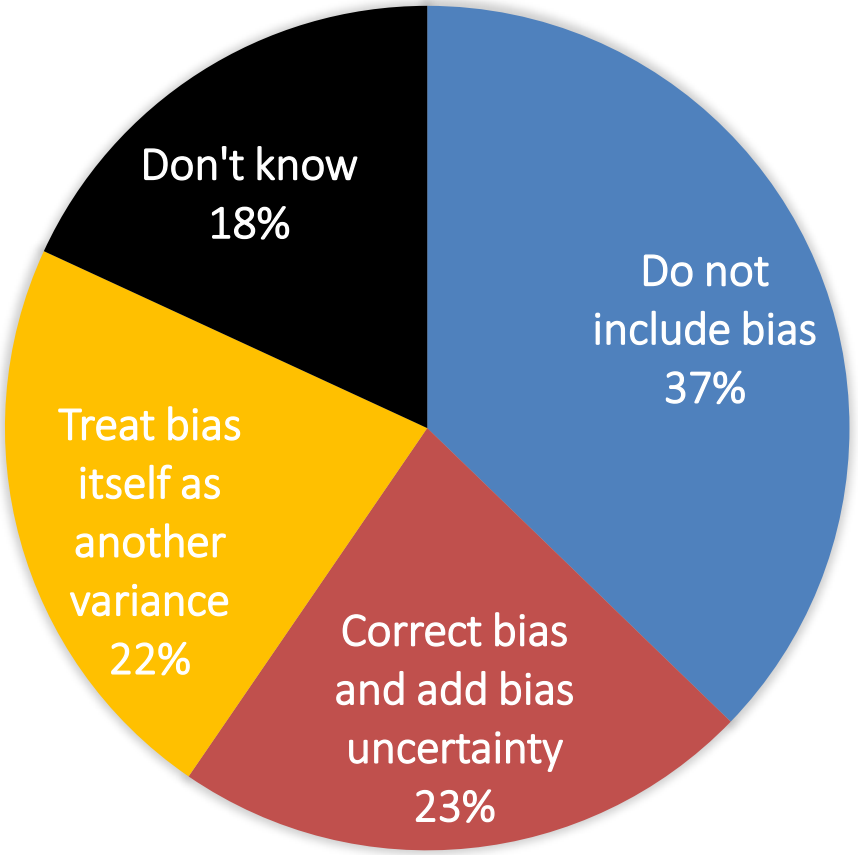
O'Donnell GE, Hibbert DB. Analyst 2005;130:721

WESTGARD QC

Measurement Uncertainty Survey Results

August 2024

How do European labs handle bias in the calculation of MU?



Whatever you decide:

1. Considering bias separately and use specific APS for bias

2. Including bias itself in MU and use specific APS for MU

in cases of “IVD-MD bias” the results must be inescapably corrected

Where an IVD-MD-related bias is not corrected for, it is clearly misleading to report only the uncorrected result and an uncertainty that does not take account the known and significant systematic error.

Where do the obligations of laboratory professionals end and IVD manufacturers' obligations begin when an IVD-MD bias is detected and must be corrected?

SCHOOLS OF THOUGHT

Once detected, the laboratory should immediately introduce a factor for the bias correction

It is the responsibility of the manufacturer appropriately warned to take an immediate investigation and eventually fix the problem with a corrective action

Individual lab

IVD manufacturer

ISO/TS 20914:2019
MEDICAL LABORATORIES -- PRACTICAL GUIDANCE FOR THE ESTIMATION OF MEASUREMENT UNCERTAINTY

→ It is the responsibility of the IVD manufacturer, or the laboratory in the case of LDT, to take immediate corrective action.

→ If the manufacturer is unable to rectify an unacceptable bias, the laboratory may, if local regulations permit, manage such measurement bias by applying a correction factor to the results or by re-assigning a calibrator value.

→ When the laboratory implements a correction for a medically significant measurement bias, the laboratory should estimate and account for u_{bias} in the estimate of u(y) [Annex C.5].

The introduction of correction factors by individual laboratories requires the - in many cases unrealistic - responsibility to carefully monitor the persistent need and stability of such factors, as the manufacturer may take its responsibility to correct the source of IVD-MD-related bias in the reagent production stage without any explicit communication.

Because performed by different laboratories on a voluntary basis, the jeopardized use of correction factors may create a situation worsening the harmonization status, where the same IVD-MD is corrected or not depending on the individual end-user's decision.

The use of bias correction factors by individual laboratories represents a *de facto* alteration of the IVD-MD status, depriving the system (and, consequently, the produced results) of the certification originally provided through CE ('Conformité Européen') marking by the manufacturer.

Involved laboratories should primarily insist in order that the providing manufacturer quickly solves the issue. Progress in the definition of responsibilities is hopefully expected as quickly as possible.

How to deal with bias on laboratory measurements

Which bias?

1. Bias arising from an insufficient correction during the implementation of traceability to higher-order references
2. Systematic variation due to sudden changes in the alignment of the IVD-MD, arising from poor calibrations or changes of reagent or calibrator lots → LABORATORY COMPONENT OF BIAS

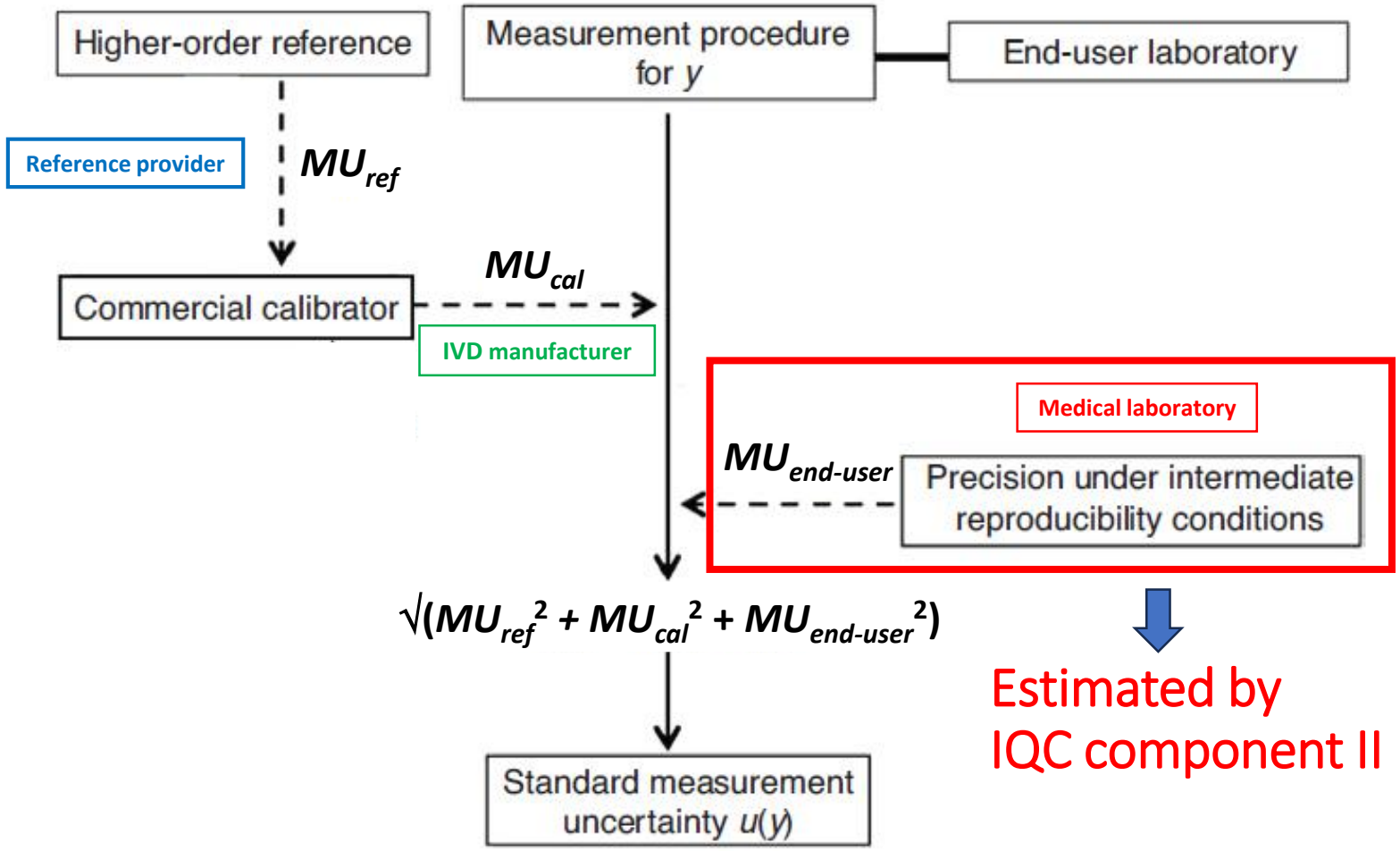
→ Medical laboratories should rely on IVD manufacturers, which entirely assume the responsibility of ensuring traceability of their products to the highest available reference and to provide unbiased results.

→ However, even if the IVD-MD is correctly aligned in the validation steps and the bias is correctly eliminated, during daily use the system may undergo some systematic changes, such as those caused by poor recalibrations and lot changes.

→This “implementation-dependent” sources of randomly occurring bias are incorporated in the estimate of MU on clinical samples through $MU_{end\text{-}user}$ and can be tolerated until the former fulfills the predefined maximum allowable uncertainty (MAU).

35

Sources of measurement uncertainty (MU) across the entire metrological traceability chain contributing to the estimate of MU of clinical sample [*u(y)*]



Estimated by IQC component II

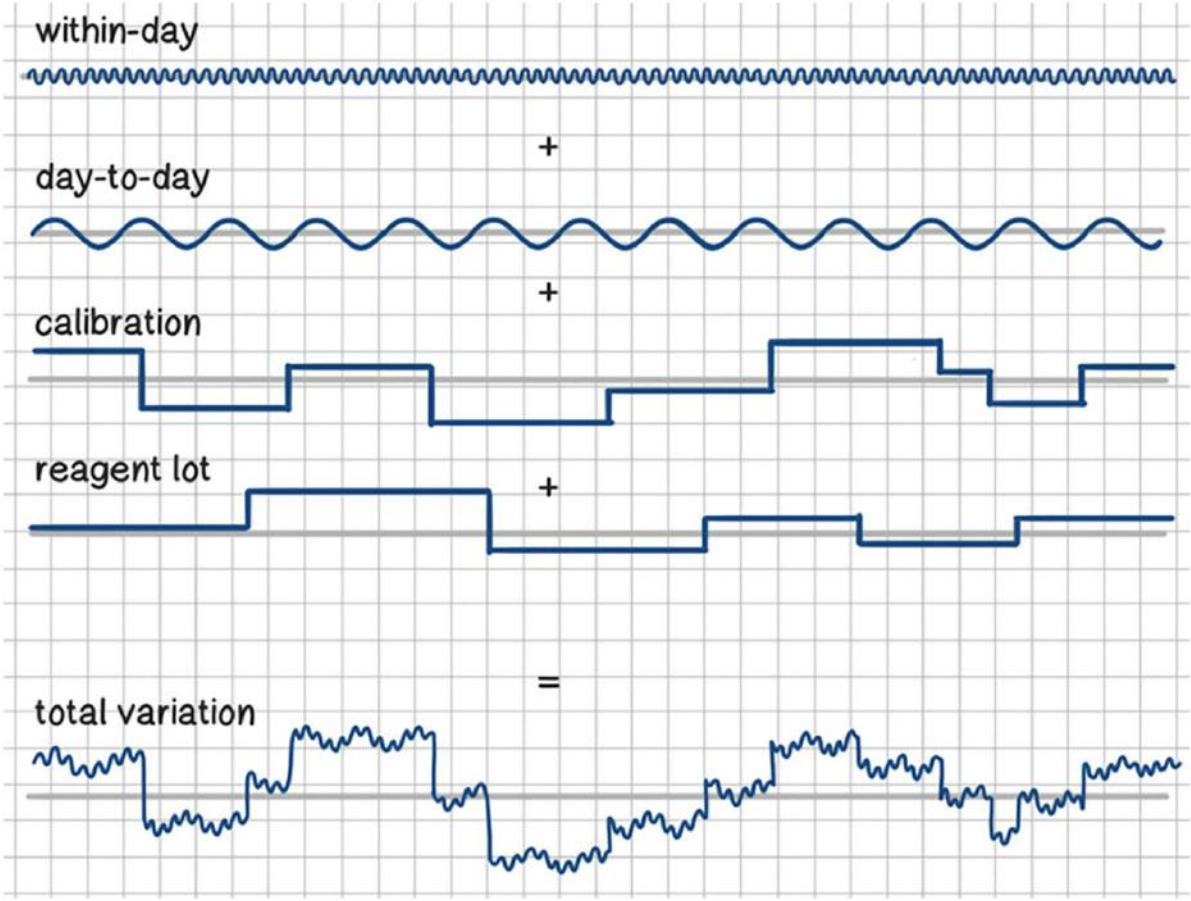
Defining and deriving $MU_{end\text{-}user}$ as 'precision under intermediate reproducibility conditions' according to the ISO/TS 20914:2019 guidance

What is it: Within-laboratory imprecision for a period sufficient to include most changes to measuring conditions.

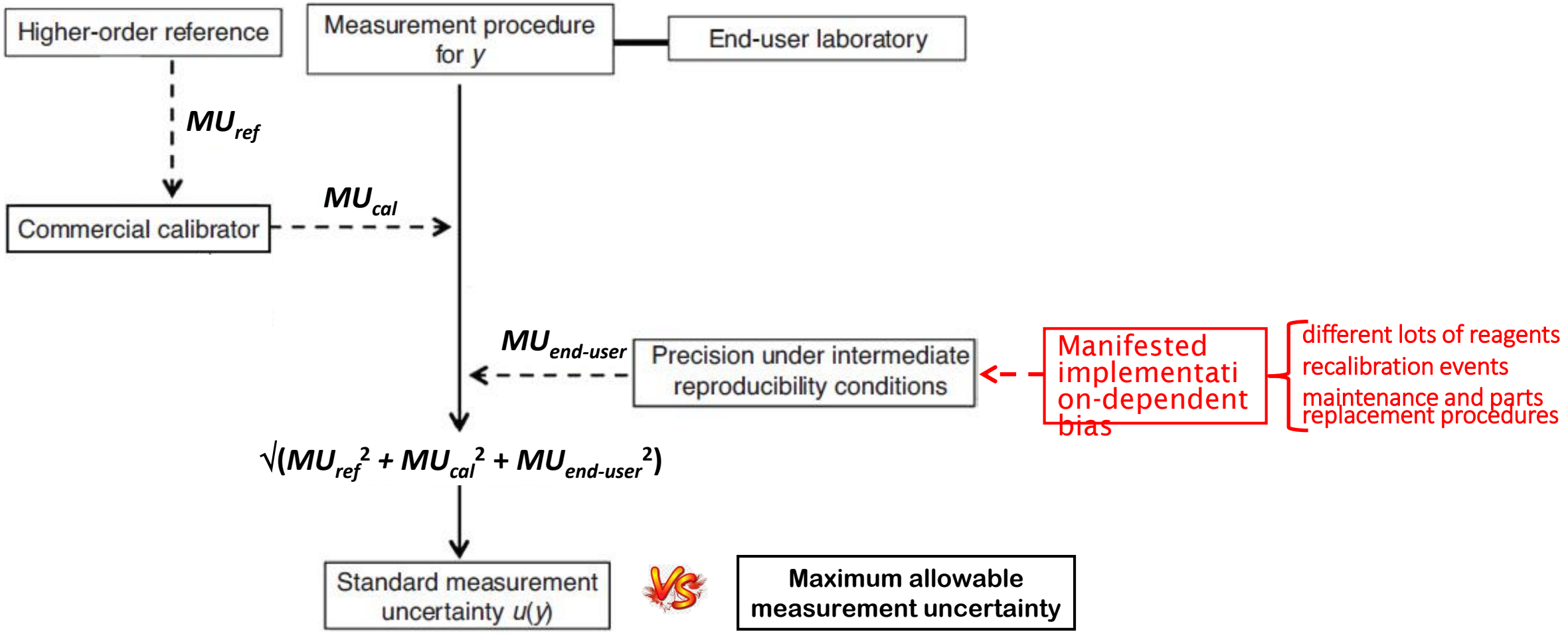
How to get it: The intermediate reproducibility should be estimated from six consecutive months Internal Quality Control daily data to capture systematic sources of uncertainty, such as those caused by different lots of reagents, different calibrations, different environmental conditions, etc.

ISO/TS 20914:2019
MEDICAL LABORATORIES -- PRACTICAL GUIDANCE FOR THE ESTIMATION OF MEASUREMENT UNCERTAINTY

Main sources of 'within-laboratory precision under intermediate reproducibility conditions'



van Schrojenstein Lantman M et al. Clin Chem Lab Med 2022;60:681


Higher-order reference
Measurement procedure for y
End-user laboratory
MU_ref
MU_cal
Commercial calibrator
MU_end-user
Precision under intermediate reproducibility conditions
Manifested implementation-dependent bias
different lots of reagents
recalibration events
maintenance and parts replacement procedures
√(MU_ref² + MU_cal² + MU_end-user²)
Standard measurement uncertainty u(y)
VS
Maximum allowable measurement uncertainty

MUTUAL EXCLUSION MODEL OF IMPLEMENTATION-DEPENDENT BIAS VALIDATION

in which the laboratory component of bias is interrelated with random sources of variability in $MU_{end\text{-}user}$ to be compared with allowable MU

Consider the average u(y) for a given IVD-MD as
$\sqrt{(1.0^2 + 1.9^2 + 4.5^2)} = 5.0\%$, with a MAU of 10.0%.
In this situation, the max allowable laboratory component of bias that is permitted "to enlarge" u(y) will be
$= \sqrt{(10.0^2 - 1.0^2 - 1.9^2 - 4.5^2)} = 8.7\%$

In this mutual relationship, smaller bias is better because, becoming part of the MU, it eats MU budget

Panteghini M. Clin Chem Lab Med 2024 Nov 26. doi: 10.1515/cclm-2024-1208. Online ahead of print

If u(y) is >MAU

It would be necessary by medical laboratory to verify all analytical conditions that may affect $MU_{end\text{-}user}$ including systematic changes.

- How is the $MU_{end\text{-}user}$ of other analytes determined on the same material or on the same instrument?
- What are IQC-I and EQA data like?
- Is the reagent on board stable?
- Has full instrumental maintenance been carried out?
- Has the reagent been reconstituted and stored properly?
- Were calibrations performed that significantly changed the results?
- Has the lot of reagent been changed?

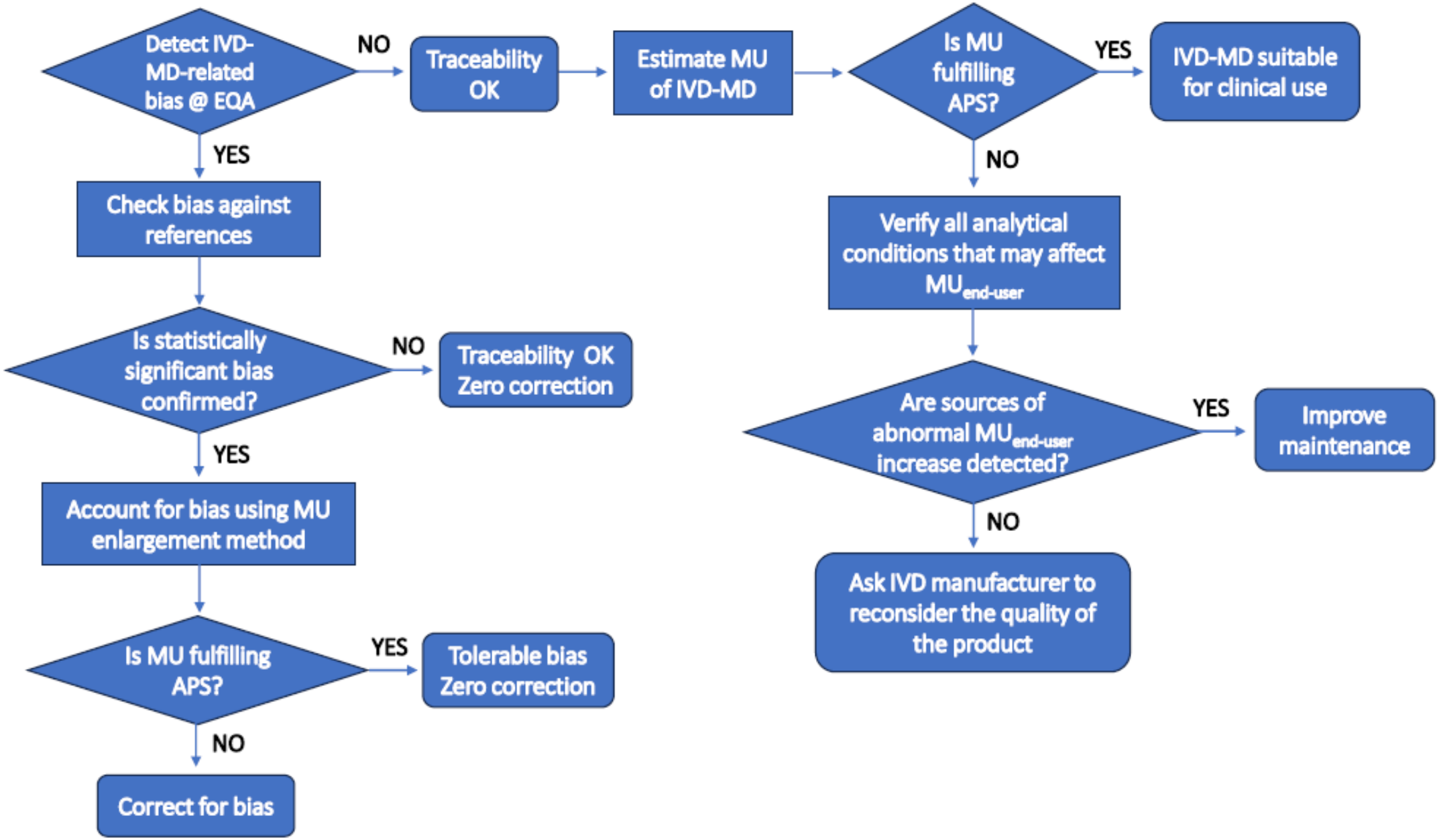
If u(y) is not fulfilling MAU because of too wide systematic deviations, a readjustment of the measuring system by the end-user must be undertaken to try to correct the bias that became significant., e.g., by a recalibration step.

If the significant bias remains and the calculated u(y) is still not meeting the predefined MAU, the IVD manufacturer should be requested to take immediate investigation and corrective action to rectify the problem.

In particular, the IVD manufacturer should be asked to reconsider:

1. the quality of the product in terms of performance stability
2. the internal specifications for accepting MU_{cal}
3. the metrological chain selected for implementing traceability.

The ultimate option can be the replacement of the IVD-MD if the MU of the measuring system is stably exceeding MAU and other marketed IVD-MD are performing better.



Panteghini M. Clin Chem Lab Med 2024 Nov 26. doi: 10.1515/cclm-2024-1208. Online ahead of print

→ The bias should not be considered in isolation.

→ Its magnitude may, or may not, be significant only if compared with the required measurement uncertainty.

→ A unifying approach to bias in the name of measurement uncertainty is therefore recommended.

Thank you for listening