



IFCC Working Group Recommendations for Correction of Bias Caused by Noncommutability of a Certified Reference Material Used in the Calibration Hierarchy of an End-User Measurement Procedure

W. Greg Miller,^{a,*} Jeffrey Budd,^b Neil Greenberg,^c Cas Weykamp,^d Harald Althaus,^e Heinz Schimmel,^f Mauro Panteghini,^g Vincent Delatour,^h Ferruccio Ceriotti,ⁱ Thomas Keller,^j Douglas Hawkins,^k Chris Burns,^l Robert Rej,^m Johanna E. Camara,ⁿ Finlay MacKenzie,^o Eline van der Hagen,^d Hubert Vesper,^p for the IFCC Working Group on Commutability

Establishing metrological traceability to an assigned value of a matrix-based certified reference material (CRM) that has been validated to be commutable among available end-user measurement procedures (MPs) is central to producing equivalent results for the measurand in clinical samples (CSs) irrespective of the clinical laboratory MPs used. When a CRM is not commutable with CSs, the bias due to noncommutability will be propagated to the CS results causing incorrect metrological traceability to the CRM and nonequivalent CS results among different MPs. In a commutability assessment, a conclusion that a CRM is commutable or noncommutable for use with a specific MP is made when the difference in bias between the CRM and CSs meets or does not meet a criterion for that specific MP when compared to other MPs. A conclusion regarding commutability or noncommutability requires that the magnitude of the difference in bias observed in the commutability assessment remains unchanged over time. This conclusion requires the CRM to be stable and no substantive changes in the MPs. These conditions should be periodically reverified. If an available CRM is determined to be noncommutable for a specific MP, that CRM can be used in the calibration hierarchy for that MP when an appropriately validated MP-specific correction for the noncommutability bias is included. We describe with examples how a MP-specific correction and its uncertainty can be developed and applied in a calibration hierarchy to achieve metrological traceability of results for CSs to the CRM's assigned value.

Background

TERMINOLOGY FOR MEASUREMENT PROCEDURE, MEASURING SYSTEM, AND IN VITRO DIAGNOSTIC MEDICAL DEVICE

As explained in part 1 of this series (1), we use the term measurement procedure (MP) when referring to a written specification for equipment, reagents, calibrators, and other components for making a measurement. We use the term in vitro diagnostic medical device (IVD-MD) for a physical measuring system manufactured according to the MP specifications and used to make measurements on clinical samples (CSs). An IVD-MD can be manufactured by a commercial company or by a clinical laboratory for its own use as a laboratory developed test.

INTRODUCTION

Standardized results among different MPs for the same measurand are essential for the application of clinical practice guidelines that direct medical decisions based on specific values of the measurand determined in CSs. Establishing metrological traceability of reported values for CSs to a certified reference material (CRM) for all available IVD-MDs for the measurand as described in ISO 17511 (2) is an accepted approach to achieve equivalent reported values.

Matrix-based CRMs are required to be commutable with CSs to be suitable for use in the calibration hierarchies of IVD-MDs (1, 2). Figure 1, A shows results from a difference in bias commutability assessment (3)

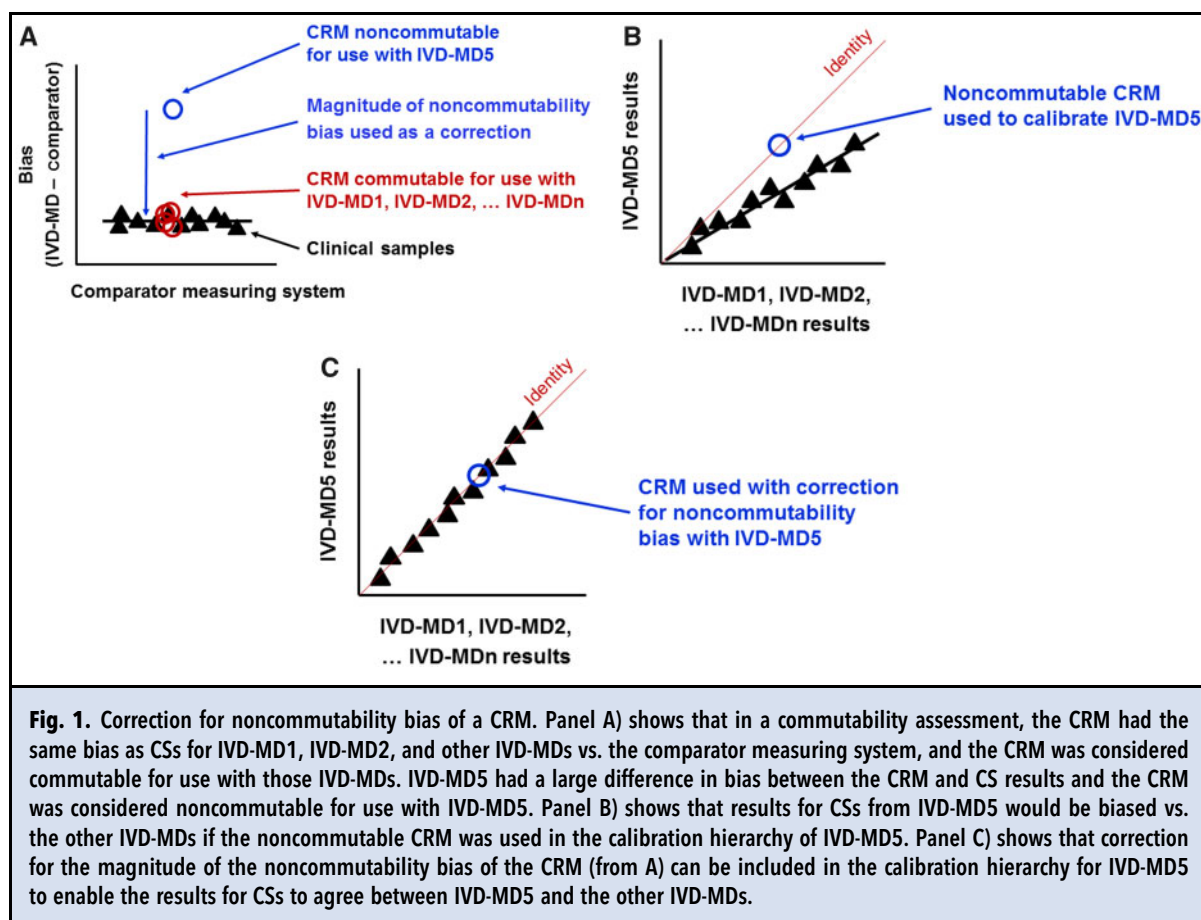
^a Virginia Commonwealth University, Richmond, VA; ^b Jeff Budd Consulting, St. Paul, MN; ^c Neil Greenberg Consulting, LLC, Rochester, NY; ^d Queen Beatrix Hospital, Winterswijk, the Netherlands; ^e Siemens Healthcare Diagnostics Products GmbH, Marburg, Germany; ^f European Commission, Joint Research Centre, Directorate F, Geel, Belgium; ^g Research Centre for Metrological Traceability in Laboratory Medicine (CIRME), University of Milan, Milan, Italy; ^h Laboratoire national de métrologie et d'essais (LNE), Paris, France; ⁱ Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy; ^j ACOMED Statistic, Leipzig, Germany; ^k School of Statistics, University of Minnesota, Minneapolis, MN; ^l National Institute for Biological Standards and Control, A Centre of the MHRA,

Hertfordshire, UK; ^m Wadsworth Center for Laboratories and Research, New York State Department of Health, and School of Public Health State University of New York at Albany, Albany, NY; ⁿ National Institute of Standards and Technology, Gaithersburg, MD; ^o Birmingham Quality/UK NEQAS, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK; ^p Centers for Disease Control and Prevention, Atlanta, GA.

* Address correspondence to this author at: Virginia Commonwealth University, PO Box 980286, Richmond, VA 23298-0286. E-mail: gmiller@vcu.edu.

Received October 2, 2019; accepted February 6, 2020.

DOI: 10.1093/clinchem/hvaa048



where most of the IVD-MDs in the assessment had very small biases between the CRM and CSs, and the CRM was considered to be commutable with the CSs for use with those IVD-MDs. **Figure 1, A** shows that the difference in bias between the CRM and CSs for IVD-MD5 was large, and the CRM was considered to be noncommutable with the CSs for use with IVD-MD5. **Figure 1, B** shows that if the CRM is used as a calibrator for IVD-MD5, the results for CSs are biased low compared to any of the other IVD-MDs in proportion to the bias (referred to as noncommutability bias) seen in **Fig. 1, A** for the CRM vs. CSs with IVD-MD5. The magnitude of the noncommutability bias of the CRM for IVD-MD5, shown in **Fig. 1, A**, can be used as a correction for noncommutability bias in the calibration hierarchy of IVD-MD5 such that results for CSs will agree with results from the other IVD-MDs as shown in **Fig. 1, C**.

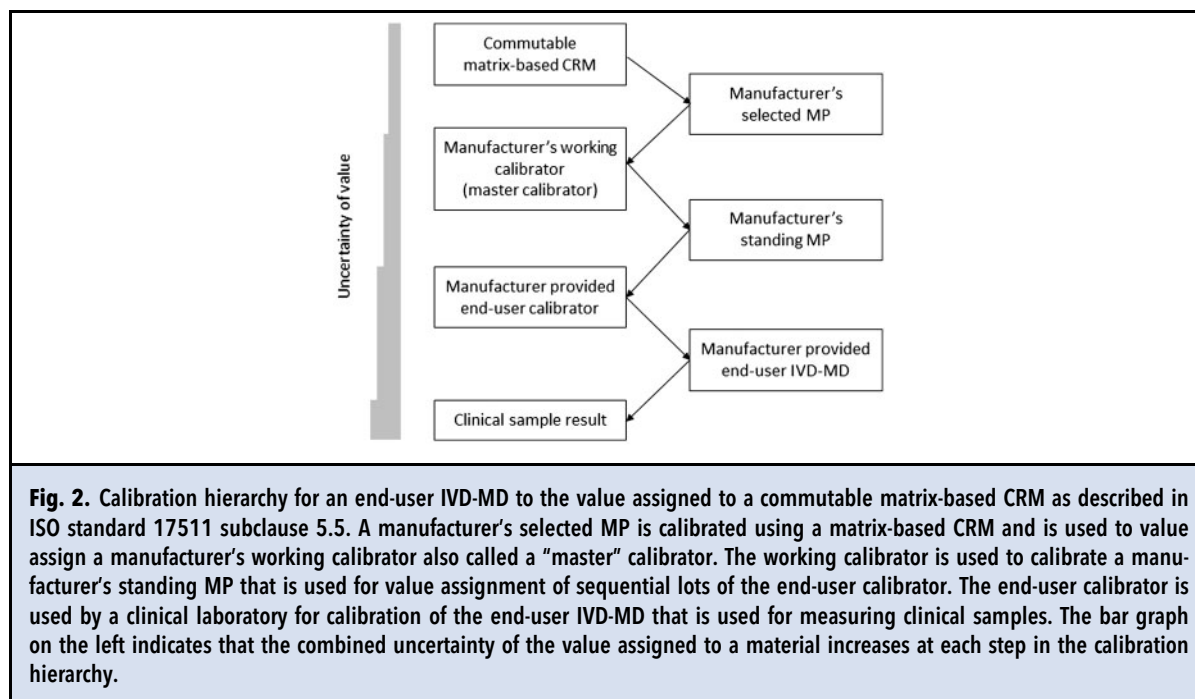
Unfortunately, there are matrix-based CRMs in use that are either noncommutable for use in the calibration hierarchies of some IVD-MDs, or have not been validated for commutability and are functionally noncommutable.

In these situations, results for CSs do not agree among different end-user IVD-MDs even though the same CRM was used in each IVD-MD's calibration hierarchy (4–16). For example, results for parathyroid hormone varied 4-fold (12) and results for ceruloplasmin varied 80% (13) among end-user IVD-MDs.

This report provides scientific rationale and technical procedures with examples to correct for noncommutability bias in a matrix-based CRM. Using a matrix-based CRM with correction for noncommutability bias, when necessary, will improve standardization of laboratory test results for patient care decisions.

Matrix-Based Certified Reference Materials in the Calibration Hierarchy

Matrix-based CRMs from metrology institutes, the World Health Organization, and other producers are widely used as higher-order references in the calibration hierarchies of end-user IVD-MDs in clinical laboratories as shown in **Fig. 2** based on the ISO 17511 standard for



metrological traceability (2, 17). Establishing metrological traceability to a matrix-based CRM (also called a secondary CRM) value that has been validated to be commutable among a group of end-user IVD-MDs, or to dilutions of that CRM that have been validated to be commutable, is central to producing equivalent results for a measurand in CSs irrespective of the IVD-MD used for measurements. Especially in cases of measurands for which there are no available reference MPs, matrix-based CRMs are frequently used as calibrators for a manufacturer's selected MP in a calibration hierarchy for an end-user IVD-MD that is used in a clinical laboratory to produce values for CSs. Primary CRMs are pure substances that are used to prepare calibrators for reference MPs. This report does not address primary CRMs. This report applies to matrix-based CRMs and to matrix-based calibrators prepared from a pure substance primary CRM or prepared from a high concentration matrix-based CRM by dilution into a biological matrix to prepare calibrators for use with a manufacturer's selected MP in the calibration hierarchy of an end-user IVD-MD. In addition, this report applies to situations when a matrix-based CRM is used as a calibrator for a clinical laboratory developed IVD-MD.

Because it is typically neither logistically nor financially reasonable to directly use a higher-order matrix-based CRM for calibration of an end-user IVD-MD, manufacturers produce working calibrators to facilitate the frequent value assignment of sequential manufactured lots of end-user calibrators. Working calibrator(s)

and end-user calibrator(s) are prepared by IVD manufacturers to be stable and reproducible over prolonged time intervals to support the manufacturing process. The manufacturer's working calibrator and end-user calibrator are not required to be commutable but are required to have values assigned such that the results for CSs are metrologically traceable to the matrix-based CRM value (1). IVD manufacturers assign values to their proprietary working calibrator(s) and end-user calibrator(s) that correct or compensate for any noncommutability bias that may be present in those calibrators used at those positions in the calibration hierarchy. Value assignment of these manufacturer's calibrators is not addressed in this report except to the extent that the value assigned by the manufacturer may be altered based on a correction added to the calibration hierarchy as described here. However, approaches similar to those described here for CRMs can also be applied by IVD manufacturers to correct for noncommutability of materials used in other positions of their calibration hierarchies.

Limitations of Noncommutable CRMs

Using a noncommutable matrix-based CRM in the calibration hierarchy of an end-user IVD-MD, without correction for noncommutability bias, is an incorrect practice because the bias due to noncommutability will be propagated to the final CS results measured using that end-user IVD-MD as shown in Fig. 3, A. In this

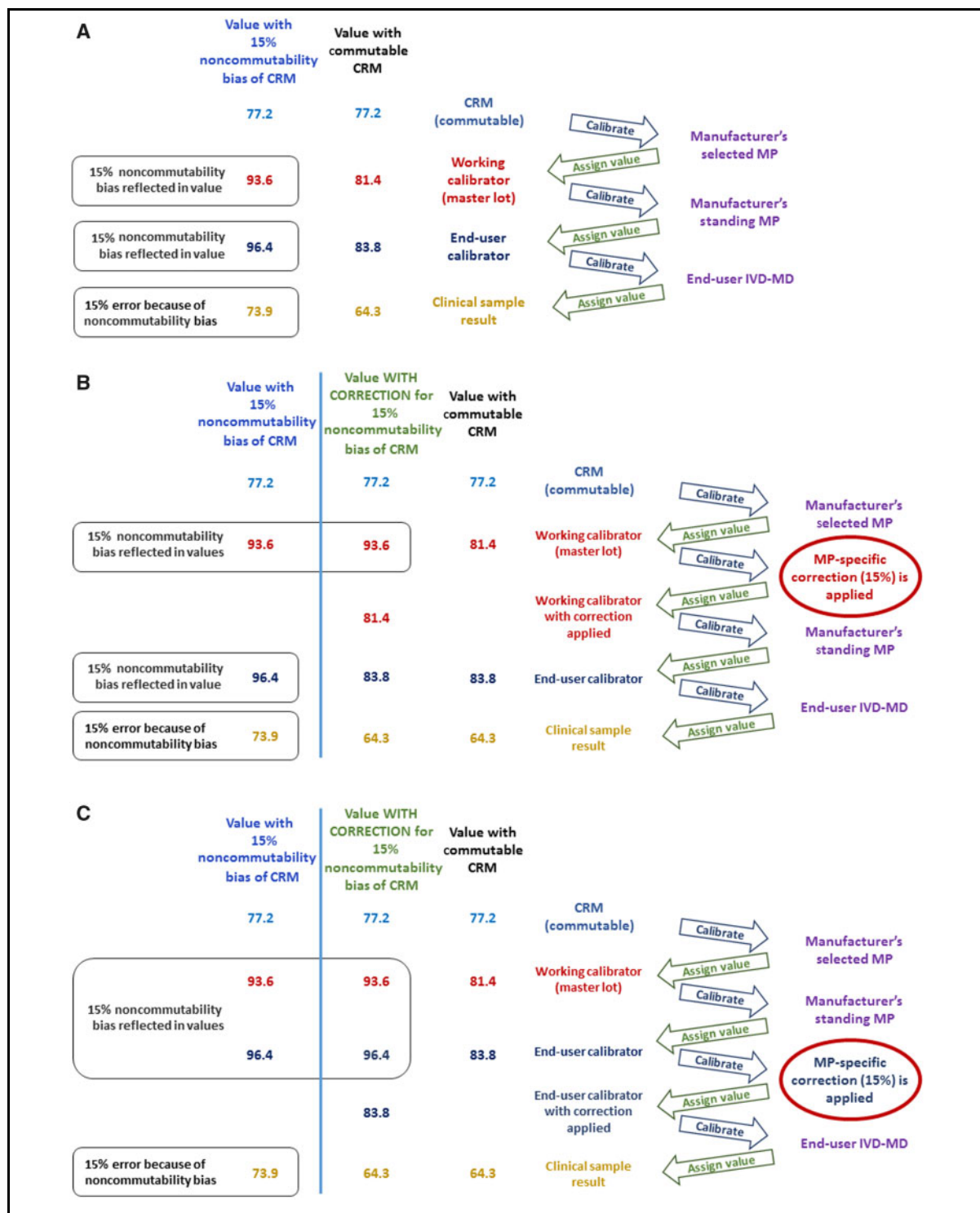


Fig. 3. Calibration hierarchies showing values assigned to calibrators and clinical samples when a commutable and a noncommutable CRM are used. In panel A), the column "Value with commutable CRM" shows values assigned when a CRM is commutable with clinical samples for use with an MP. The column "Value with 15% noncommutability bias of CRM" shows values assigned when a CRM is not commutable for use causing incorrect results assigned to the working calibrator and propagated through the calibration hierarchy to produce a 15% bias in the clinical samples. Panels B) and C) show that a MP-specific correction for noncommutability bias can be applied at the working calibrator step (panel B) or at the end-user calibrator step (panel C) in the calibration hierarchy. The column "Value WITH CORRECTION for 15% noncommutability bias of CRM" shows that adding a step in the calibration hierarchy to correct for the noncommutability bias of a CRM with a specific MP will produce results for clinical samples that are equivalent to those from a different MP for which the CRM was commutable for use.

situation, results from the IVD-MD for which the CRM is noncommutable for use are biased compared to other IVD-MDs for which the CRM is commutable for use. Consequently, errors in diagnosing, treating, or monitoring patients will occur when the nonequivalent results from that IVD-MD are used in medical decisions. Some regulations require metrological traceability to a higher-order reference measurement system component when available (18, 19) without specifying that, if the higher-order reference system component is a matrix-based CRM, the CRM is required to be commutable with CSs. Consequently, to fulfill regulations, IVD manufacturers have in some cases claimed “traceability” to a CRM that is noncommutable for use with a particular IVD-MD. In this situation, such a calibration hierarchy is not technically valid. Applying a correction for noncommutability bias allows an appropriate calibration hierarchy to be implemented.

As mentioned in the introduction, there are matrix-based CRMs in use that are either noncommutable for use in the calibration hierarchies of some IVD-MDs, or have not been validated for commutability and are functionally noncommutable. Using CRMs that are noncommutable, or have unknown commutability when that CRM is in fact noncommutable, in the calibration hierarchies of end-user IVD-MDs is an inappropriate practice because nonequivalent results for CSs will be produced among different end-user IVD-MDs. Bias between results from different end-user IVD-MDs can also be due to issues such as procedures applied for calibration (20), to inadequate calibrator value assignment protocols (21), or to nonselectivity for the measurand (1).

This report addresses the situation when a matrix-based CRM’s commutability is suitable for use with a large fraction of IVD-MDs but is noncommutable for use with a particular measurand for one or a few IVD-MDs. In this situation, a correction for bias caused by noncommutability enables that CRM to be correctly used in the calibration hierarchies for such IVD-MDs.

Correction for Bias Caused by Noncommutability of a CRM

Metrological traceability is defined by the International Vocabulary of Metrology as a 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 (22). Note 6 in the definition states that a comparison between 2 measurement standards may be viewed as a calibration if the comparison is used to check and, if necessary, correct the quantity value and measurement uncertainty attributed to one of the measurement standards. A matrix-based CRM is a measurement standard in the context of the International

Vocabulary of Metrology language. In addition, subclause 4.5.7 in ISO 17511 (ed 2) “*Requirements for establishing metrological traceability of values assigned to calibrators, trueness control materials and human samples*” states that in cases where a CRM demonstrates noncommutability with human samples for some IVD-MDs, the noncommutable CRM may still be used as a calibrator within the calibration hierarchy of a relevant IVD-MD by application of a correction factor, as long as the documentation provided for the IVD-MD discloses details for derivation and validation of the correction factor, including any incremental uncertainty (2). Based on the preceding statements, we conclude that adding a step to a calibration hierarchy to correct for noncommutability bias of a CRM is consistent with international standards.

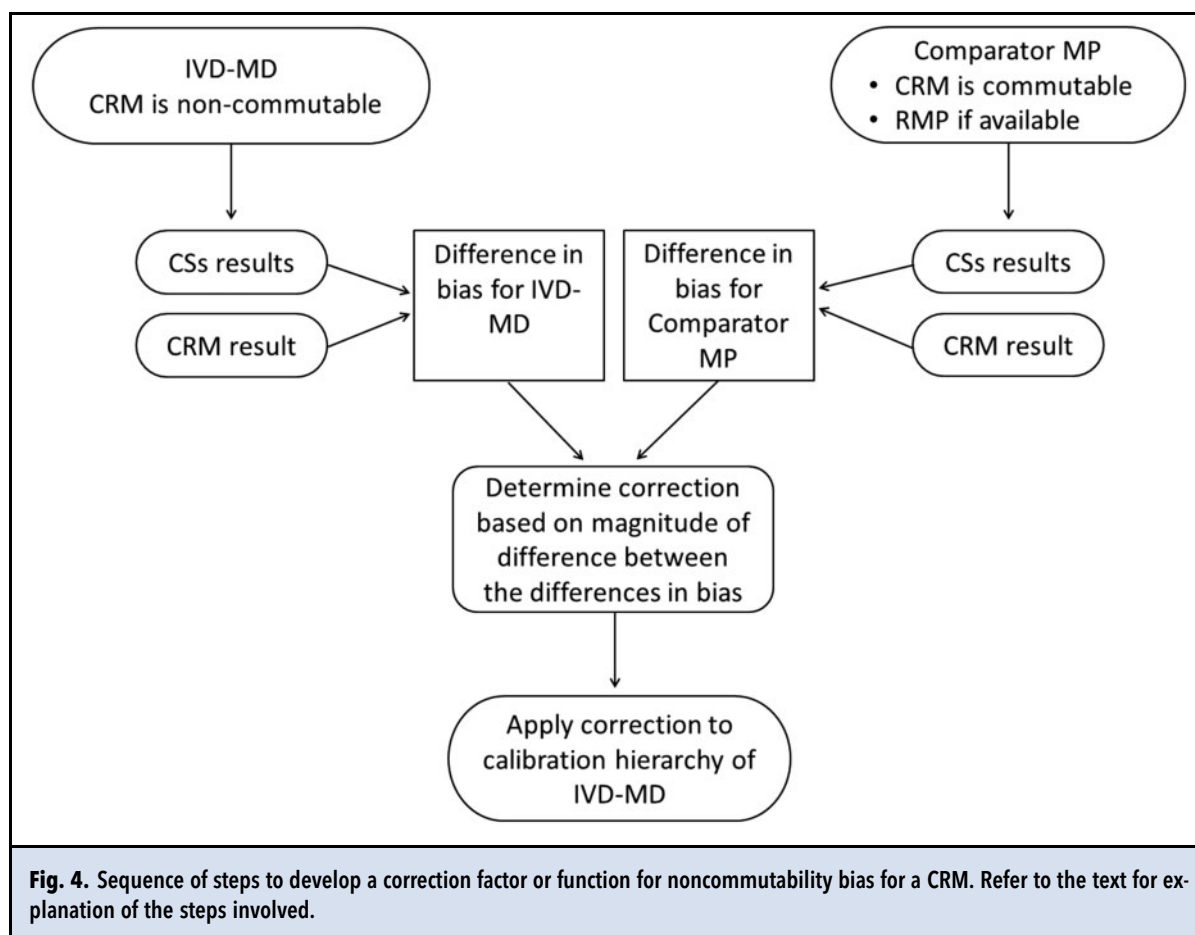
The correction step for noncommutability bias can be added in any suitable position in the calibration hierarchy as shown in Fig. 3, B and C. The CS results are metrologically traceable to the CRM value because the noncommutability bias is corrected such that the trueness of the assigned value of the CRM is correctly transferred to the calibration of the end-user IVD-MD. The correction step in the calibration hierarchy will add additional uncertainty to the final CS result.

There is no predefined magnitude of noncommutability bias that precludes developing a correction factor or function. A difference in bias between a CRM and CSs for an IVD-MD and for a comparator measuring system is considered commutable if within a criterion or noncommutable if outside a criterion. The magnitude of this difference in bias is what is corrected, if necessary, to use a CRM with excessive difference in bias (i.e. noncommutable) in the calibration hierarchy for an end-user IVD-MD. Applying a correction assumes the magnitude of the noncommutability bias remains constant over time.

Commutability assessment is typically performed using IVD-MDs. When the manufacturer’s selected MP is the same MP as the end-user IVD-MD but operated with more stringent specifications for calibration and measurement replication to reduce uncertainty, a correction for noncommutability bias determined for the end-user IVD-MD will be applicable to the manufacturer’s selected MP. When the manufacturer’s selected MP is a different MP than the end-user IVD-MD, then a correction for noncommutability bias at the manufacturer’s selected MP must be determined. In some cases, a manufacturer may use a CRM to calibrate their standing MP in which case the same considerations apply.

REQUIREMENTS TO APPLY A CORRECTION FOR NONCOMMUTABILITY BIAS OF A MATRIX-BASED CRM

One important requirement for the suitability of a correction is that the noncommutability bias of the CRM



and the performance characteristics of an IVD-MD for which the correction is applied must remain stable over time; e.g., across reagent lots. The second key requirement is that the uncertainty of the correction step in the calibration hierarchy allows the final combined uncertainty of the results for CSs to be within the allowable uncertainty budget specification for the measurand (23–25). Refer to the worked examples and explanation of the examples in the [Supplemental Data](#) that accompanies the online version of this paper for experimental designs that can be realistically implemented and achieve small uncertainties for the corrections. If the preceding conditions are not satisfied, then a correction for noncommutability bias is not acceptable and should not be applied.

Determination of a Correction Factor or Function for Noncommutability Bias

The user of a CRM (typically an IVD manufacturer or a clinical laboratory) who develops a calibration hierarchy

for an end-user IVD-MD is responsible for establishing a correction factor or function for noncommutability bias of that CRM, if such a correction is needed. The data available from an initial commutability assessment (e.g., 1, 3) could be sufficient for developing a correction if the incremental increase in uncertainty due to the correction step is acceptable. If a lower uncertainty in the correction step for noncommutability bias is needed, then additional data from a new experiment are required. The experimental design regarding replication of measurements, and related details (see examples in [Supplemental Data](#)) are established by the manufacturer of the IVD-MD to meet the uncertainty requirements.

Figure 4 shows a flow diagram of the general approach for developing a correction factor or function. The CSs selected to quantitate the noncommutability bias are chosen to represent patients with common medical conditions for which an MP is intended to be used, applying the same selection criteria described in part 1 of this series to avoid using CSs with unusual pathological forms of a measurand or known interfering

substances (1). Since the manufacturer of an MP will have information on the selectivity of the MP, suitable CS selection criteria can be specified to avoid known sample specific effects that could compromise the experimental design to quantitate the noncommutability bias of the CRM when used with that specific MP.

A manufacturer, or a clinical laboratory in the case of a laboratory developed MP, requires a comparator MP, for which a CRM has been shown to be commutable, to assign values to the CSs that will be used to quantitate the noncommutability bias of the CRM for the MP under investigation. A single comparator MP is adequate for development of a correction factor or function. A higher-order reference MP should be used as the comparator MP when available (26). Otherwise, a comparator MP is another end-user IVD-MD chosen based on its performance characteristics in the commutability assessment for the CRM (3, 20). Collaborating with a laboratory that uses the comparator MP may be necessary.

The key attribute for a comparator MP is that the CRM was shown to be commutable for use with that MP. The CRM provider can be contacted for advice regarding end-user MPs that had suitable commutability with CSs to be considered as comparator MPs. Alternatively, external quality assessment (proficiency testing) results from commutable samples can be used to identify a comparator MP that used the CRM in its calibration hierarchy and for which there was negligible bias observed vs. the target values for those samples. Negligible bias infers that the CRM is commutable for use in the calibration hierarchy of that MP selected as a comparator MP. When considering external quality assessment data, caution should be used when a small number of participants use an MP in the survey. In addition to the CRM being commutable for use with a comparator MP, its precision and selectivity for the measurand must be adequate.

The key requirement for a comparator MP to be used to develop a correction for another MP is that the CRM has been shown to be commutable for use with the comparator MP. The experimental design for developing a correction factor for a single level CRM described in example 1 in the [Supplemental Data](#) that accompanies the online version of this article includes measuring the CRM and the CSs by both the IVD-MD for which a correction is being developed and by the comparator MP. This experimental design eliminates the influence of calibration error of either MP on determination of the correction factor for noncommutability bias of the CRM for a given IVD-MD. In addition, the experimental design eliminates variance components between-run; such as day-to-day, instrument-to-instrument, reagent lot-to-lot,

and calibration event-to-calibration event from influencing the correction factor. In example 1, sources of variability other than within-run and sample specific effects are eliminated from influencing the correction factor. This experimental design minimizes the uncertainty associated with the correction step in the calibration hierarchy.

The reason for noncommutability of a CRM should be investigated as part of the process to develop a correction for the noncommutability bias. This information can be useful to assess the risk that a correction will be consistent and stable over time and to monitor the stability of the contributor(s) to the noncommutability bias. In addition, a cause for noncommutability may be possible to eliminate in the next batch of a CRM. However, the influence quantity or quantities may not be identified reinforcing the need for periodic reverification of the correction factor.

Considerations When a CRM Is Provided as a Multilevel Set of Materials

Some CRMs are provided as a set of 2 or more concentration levels intended to be used together for calibration of MPs. The underlying reason for providing a multilevel CRM is that MPs may not produce measurement responses for CSs that are linearly proportional to the amount of measurand in those samples over the measuring interval. One approach to develop a correction for noncommutability bias is to treat each CRM level separately with measured values of CSs clustered near the value of each CRM level. Another approach is to use CSs that cover the values of the CRM levels and fit a suitable mathematical function to develop correction factors for each CRM level. Because the measurement response cannot be assumed to be proportional to concentration over the measuring interval, a matching design is not possible as used in example 1 to eliminate influence of calibration drift in the comparator MP. Consequently, the experimental design must be appropriate to determine the correction factors for noncommutability bias for each level of the CRM that enable the final combined uncertainty requirement for CSs to remain within that required for medical decisions. Example 2 in the online [Supplemental Data](#) that accompanies the online version of this article illustrates one approach for determining and using correction factors for a three-level CRM.

Considerations When a CRM is Intended to Be Diluted to Obtain One or More Calibrators

Some CRMs are provided as a single high-concentration material that is intended to be diluted, or otherwise

reduced in concentration, in an appropriate matrix to prepare one or more calibrators for the manufacturer's selected MP or for an end-user IVD-MD in the case of a clinical laboratory developed MP. Dilution in an appropriate matrix is intended to create a calibrator that is commutable with CSs and has a concentration within the measuring interval. Each concentration level of diluted CRM must be validated to be commutable with CSs. The CRM provider should work collaboratively with IVD manufacturers to validate commutability of dilutions of a CRM as part of the initial validation of the CRM. It may be possible that a common diluent can be used by all MPs in which case commutability assessment as described in parts 2 or 3 of this series can be used (3, 20). If dilutions of a CRM are commutable for use by most MPs but shown to be noncommutable for one or more specific MPs, then a MP-specific correction for noncommutability bias as described in this report can be included in the calibration hierarchy of the end-user IVD-MD. A separate correction factor is needed for each dilution of a CRM.

Alternatively, some MPs may require unique diluents and possibly different concentrations of calibrators suitable for an MP's calibration model or measuring interval. In this case, only commutability assessment as described in part 3 (20) of this series can be used because the CRM is diluted with different matrices and/or to different concentrations and thus the dilutions are not suitable for measurement by all MPs included in the commutability assessment.

Assumptions and Limitations When Assessing Commutability and Applying a Correction for Noncommutability Bias

A commutability assessment is an experiment performed at a point in time including representative CSs of the type intended to be measured and IVD-MDs representing all or a substantial fraction of the MPs in use in clinical laboratories at the time of the assessment. A commutability decision for a CRM is in principle only applicable to the measurement conditions (e.g., calibrator, reagent lots, and IVD-MD performance characteristics) represented in the assessment experiment.

A conclusion that a CRM is commutable for use in the calibration hierarchies of a stated group of MPs requires that the assessment results meet a specified criterion. Similarly, a conclusion that a CRM is noncommutable for use with a specific MP is made when the assessment results do not meet the criterion. These commutability conclusions remain valid when the CRM is stable regarding its commutability property and no substantive changes, e.g., reagent lot-to lot influences,

occur in the end-user IVD-MDs or the other MPs used in the calibration hierarchies.

It is unavoidable that IVD-MDs will use different lots of calibrators and reagents, and that other measurement conditions will vary over time within the MP manufacturer's specifications. The changes in measuring conditions that occur over time must be small enough that the differences in biases between CRM and CSs observed in the commutability assessment remain essentially the same over time for the CRM to remain suitable for use in the calibration hierarchies of the end-user IVD-MDs when future measurements are made.

Alternative approaches to establishing metrological traceability to a CRM that is noncommutable for a particular MP can also be considered as stated in ISO 17511 (2). For example, CSs with values assigned by a suitable comparator measuring system that is calibrated with a commutable CRM can be used directly in the calibration hierarchy as calibrators for the manufacturer's selected MP. Uncertainty for alternative approaches needs to be determined and added to the calibration hierarchy.

Periodic Reverification of the Commutability of a CRM and of a Correction for Noncommutability Bias

The commutability of a CRM as well as a correction factor for noncommutability bias might change over time. A commutable matrix-based CRM used as a calibrator is a critical step in the calibration hierarchy, as is a correction step for noncommutability bias. ISO/IEC 17025:2017 (27) and a report from the National Institute for Standards and Technology (28) recommend to periodically verify metrological traceability of the calibration hierarchy. The commutability of a matrix-based CRM and a correction for noncommutability bias when included in the calibration hierarchy should be considered when reverifying metrological traceability. CRM providers are required (29, 30) to monitor the stability of the value assigned to a CRM but are not currently required to monitor its commutability. However, the outcome of a commutability assessment at a point in time may not be valid if there has been an aging effect on the CRM matrix, a structural change in the measurand in the CRM, a change in an MP, or an unintended change in manufacturing IVD-MDs based on an MP. In particular, replacement lots of reagents can influence the stability of a commutability decision.

Assessing the stability of the assigned value of a measurand in a CRM with some higher-order MPs, e.g., mass spectrometry, might not assess changes in the molecular form because the higher-order MP can measure a quantity derived from the measurand; e.g., a tryptic-digested peptide derived from an intact protein.

Similarly, a higher-order MP may not identify changes in the matrix of a CRM because higher-order MPs typically include steps intended to make them highly selective for the measurand and insensitive to matrix influences. Reassessing commutability could, in many cases, be more suitable using one or more manufacturer's selected, standing, or end-user MPs because such MPs may be more sensitive to an aging effect on the CRM matrix or a structural change in the measurand.

The frequency to perform reverification of commutability or of a correction for noncommutability bias uses a risk assessment approach. Critical factors to consider include: changes to a MP such as reformulation of reagents or other changes likely to affect selectivity for a measurand or other performance characteristics of the MP; evidence from a manufacturer's quality system that monitors the stability of working calibrator(s) and standing and/or selected MPs in a manufacturer's calibration hierarchy; evidence from stability assessment of the value assigned to the CRM; results from external quality assessment (proficiency testing) using commutable samples; and evidence from a comparison of results for CSs among MPs. Evidence of change from any of these sources should initiate an investigation of root cause that can include reverification of the value assigned to a CRM, its commutability, and, when applicable, a correction for noncommutability bias. A risk assessment will be limited by available knowledge and state of the art but is the best approach to determine the frequency to perform a reverification activity.

Some stability monitoring approaches cannot distinguish between a change in the CRM or a change in an MP over time. Consequently, cooperation is needed between CRM producers and IVD manufacturers to review stability assessment data and determine the appropriate approach when stability is unacceptable.

Responsibility for a Correction for Noncommutability Bias

Development of a correction factor or function for noncommutability bias and where to apply the correction step in the calibration hierarchy is the responsibility of the manufacturer of an MP, including a clinical laboratory that develops an IVD-MD for its own use. If MP performance characteristics for random variability, selectivity for the measurand, or the calibration function of a MP are not suitable, the MP might need to be improved before a correction factor or function can be successfully developed. These performance characteristics are discussed in more detail in parts 1 and 3 of this series of reports (1, 20).

When a suitable correction for noncommutability bias of a CRM is incorporated in the calibration hierarchy for an end-user IVD-MD, the manufacturer can claim traceability to the value of the CRM with full disclosure and description of the correction step including

the comparator MP provided to the end-user of the IVD-MD.

Documentation of Commutability and of a Correction Factor for Noncommutability

The documentation that accompanies a CRM should include the MPs and diluents, when applicable, for which the CRM, or dilutions of the CRM, were validated to be commutable.

The documentation of a correction for noncommutability bias for a particular MP is maintained by the manufacturer as part of the technical file for metrological traceability.

Disclaimer

The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the Centers for Disease Control and Prevention/the Agency for Toxic Substances and Disease Registry. Use of trade names is for identification only and does not imply endorsement by the Centers for Disease Control and Prevention, the Public Health Service, and the US Department of Health and Human Services.

Supplemental Material

Supplemental material is available at *Clinical Chemistry* online.

Nonstandard Abbreviations: CRM, certified reference material, CS, clinical sample, IEC, International Electrotechnical Commission, ISO, International Organization for Standardization, IVD, in vitro diagnostic, MD medical device, MP measurement procedure.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 4 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; (c) final approval of the published article; and (d) agreement to be accountable for all aspects of the article thus ensuring that questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved.

J.R. Budd, statistical analysis; T. Keller, statistical analysis.

Authors' Disclosures or Potential Conflicts of Interest: Upon manuscript submission, all authors completed the author disclosure form. Disclosures and/or potential conflicts of interest:

Employment or Leadership: W.G. Miller, *Clinical Chemistry*, AACC; R.Rej, *Clinical Chemistry*, AACC.

Consultant or Advisory Role: None declared.

Stock Ownership: None declared.

Honoraria: None declared.

Research Funding: None declared.

Expert Testimony: None declared.

Patents: None declared.

Role of Sponsor: No sponsor was declared.

References

1. Miller WG, Schimmel H, Rej R, Greenberg N, Ceriotti F, Burns C, et al. IFCC Working Group recommendations for assessing commutability part 1: general experimental design. *Clin Chem* 2018;64:447-54.
2. ISO 17511:2020 (2nd edition). In vitro diagnostic medical devices—requirements for establishing metrological traceability of values assigned to calibrators, trueness control materials, and human samples. Geneva: International Organization for Standardization; 2020.
3. Nilsson G, Budd JR, Greenberg N, Delatour V, Rej R, Panteghini M, 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* 2018;64:455-64.
4. Rose MP. Follicle stimulating hormone international standards and reference preparations for the calibration of immunoassays and bioassays. *Clin Chim Acta* 1998; 273:103-17.
5. Panteghini M, Pagani F. AACC creatine kinase MB (CK-MB) standardization material used as manufacturer's working calibrator is unable to harmonize CK-MB results between two commercial immunoassays. *Clin Chem* 2004;50:1711-2.
6. Steele BW, Wang W, Klee G. Analytical bias of thyroid function tests: analysis of a College of American Pathologists fresh frozen serum pool by 3900 clinical laboratories. *Arch Pathol Lab Med* 2005;129:310-7.
7. Miller WG, Myers GL, Rej R. Why commutability matters. *Clin Chem* 2006;52:553-4.
8. Miller WG, Thienpont LM, Van Uytanghe K, Clark PM, Lindstedt P, Nilsson G, Steffes MW. Toward standardization of insulin immunoassays. *Clin Chem* 2009;55: 1011-8.
9. Infusino I, Valente C, Dolci A, Panteghini M. Standardization of ceruloplasmin measurements is still an issue despite the availability of a common reference material. *Anal Bioanal Chem* 2010;397:521-5.
10. Thienpont LM, Van Uytanghe K, Beestall G, Faix JD, Ieiri T, Miller WG, et al. Report of the IFCC Working Group for Standardization of Thyroid Function Tests; part 1: thyroid-stimulating hormone. *Clin Chem* 2010;56: 902-11.
11. Sokoll LJ, Rosenwald S, Lyons J, Elliott DJ, Chan DW. Is the WHO 90:10 prostate-specific antigen (PSA) first international reference standard really 90% α 1-antichymotrypsin-bound PSA and 10% free PSA?. *Clin Chem* 2011; 57:1776-7.
12. Almond AA, Ellis AR, Walker SW. Current parathyroid hormone immunoassays do not adequately meet the needs of patients with chronic kidney disease. *Ann Clin Biochem* 2012;49:63-7.
13. Zegers I, Beetham R, Keller T, Sheldon J, Bullock D, MacKenzie F, et al. The importance of commutability of reference materials used as calibrators: the example of ceruloplasmin. *Clin Chem* 2013;59:1322-9.
14. Ross HA, Lentjes EW, Menheere PM, Sweep CG. Harmonization of growth hormone measurement results: the empirical approach. *Clin Chim Acta* 2014; 432:72-6.
15. Sturgeon CM, Sprague S, Almond A, Cavalier E, Fraser WD, Algeciras-Schimmich A, et al. Perspective and priorities for improvement of parathyroid hormone (PTH) measurement—a view from the IFCC Working Group for PTH. *Clin Chim Acta* 2017;467:42-7.
16. Bachmann LM, Yu M, Boyd JC, Bruns DE, Miller WG. State of harmonization of 24 serum albumin measurement procedures and implications for medical decisions. *Clin Chem* 2017;63:770-9.
17. Vesper HW, Thienpont LM. Traceability in laboratory medicine. *Clin Chem* 2009;55:1067-75.
18. Directive 98/79/EC of the European Parliament and of the Council of 27 October 1998 on in vitro diagnostic medical devices OJ L 331 of 7 December 1998.
19. Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on in vitro diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU (Text with EEA relevance.) <http://www.ce-mark.com/IVD%20Regulation.pdf> (Accessed March 2020).
20. Budd JR, Weykamp C, Rej R, MacKenzie F, Ceriotti F, Greenberg N, et al. IFCC Working Group recommendations for assessing commutability part 3: using the calibration effectiveness of a reference material. *Clin Chem* 2018;64:465-74.
21. Braga F, Pasqualetti S, Panteghini M. The role of external quality assessment in the verification of in vitro medical diagnostics in the traceability era. *Clin Biochem* 2018; 57:23-8.
22. JCGM. International Vocabulary of Metrology—Basic and General Concepts and Associated Terms (VIM), 3rd edition. JCGM 200:2012.
23. Sandberg S, Fraser CG, Horvath AR, Jansen R, Jones G, Oosterhuis W, et al. Defining analytical performance specifications: consensus statement from the 1st Strategic Conference of the European Federation of Clinical Chemistry and Laboratory Medicine. *Clin Chem Lab Med* 2015;53:833-5.
24. Braga F, Infusino I, Panteghini M. Performance criteria for combined uncertainty budget in the implementation of metrological traceability. *Clin Chem Lab Med* 2015; 53:905-12.
25. Braga F, Panteghini M. Defining permissible limits for the combined uncertainty budget in the implementation of metrological traceability. *Clin Biochem* 2018;57: 7-11.
26. Joint Committee for Traceability in Laboratory Medicine. Database of higher-order reference materials, measurement methods/procedures and services. <https://www.bipm.org/jctlm/> (accessed March 2020).
27. ISO/IEC 17025:2017. General requirements for the competence of testing and calibration laboratories. Geneva: International Organization for Standardization; 2017.
28. Ehrlich CD, Raspberry SD. Metrological timelines in traceability. *J Res Natl Inst Stand Technol* 1998;103: 93-105.
29. ISO 15194:2009. In vitro diagnostic medical devices—Measurement of quantities in samples of biological origin—Requirements for certified reference materials and the content of supporting documentation. Geneva: International Organization for Standardization; 2009.
30. ISO 17034:2016. General requirements for the competence of reference material producers. Geneva: International Organization for Standardization; 2016.