Guidelines for CCPR and RMO Bilateral Key Comparisons

CCPR Working Group on Key Comparison
CCPR-G5 October 10th, 2014

These guidelines are prepared by CCPR WG-KC and RMO P&R representatives, and approved by CCPR, to ensure that bilateral Key Comparisons in CCPR or RMOs are prepared and performed in a fair and uniform manner and that the results be linked to the CCPR KCs appropriately. This document is to supplement the CIPM guide, *Measurement comparisons in the CIPM MRA*, CIPM-MRA-D-05 (December 2011) [1].

1. Purposes and initiation of a bilateral comparison

1.1 A CCPR bilateral KC is carried out only right after a CCPR KC is completed and when there was a problem in the reported results with a participant and the participant wants to correct the results. All other bilateral KCs are done as RMO comparisons.

1.2 RMO bilateral KCs are carried out when the CCPR KC or an RMO KC of the quantity of interest is not available in the required time frame, and a supporting evidence for a CMC claim is urgently needed.

1.3 With the aim of reducing the number of bilateral comparisons, it is recommended that two or more bilateral comparisons of the same quantity planned at a similar time at different RMOs be combined, or an NMI seeking for a bilateral comparison can be invited to an RMO comparison of the same quantity in other RMOs planned at a similar time.

1.3.1 An NMI needing a comparison to underpin its CMCs shall contact its RMO TC chair before looking for a bilateral comparison.

1.3.2 When a need for a bilateral comparison is identified, the RMO TC chair should ask the WG-CMC chair and its members whether there are any comparisons of the same quantity planned in any other RMO, in which the NMI can participate or with which the NMI can coordinate.

1.4 These bilateral comparisons are normally conducted between two laboratories. However, there can be cases where two or three bilateral comparisons with the same pilot laboratory are combined as one comparison. In this case, the comparison may be registered as one comparison in KCDB, and there should be one comparison report.

1.4.1 If there are significant differences in measurement time or measurement conditions for each participant, the comparison can be registered separately for each pair of participants, in which case, report should be prepared separately.

1.5 If the participants are members of different RMOs, the RMO to which the lab requesting the comparison belongs (laboratory to be linked) should organize the comparison, unless the other direction is agreed by both RMOs.

2. Pilot laboratory and Link laboratory

2.1 The pilot laboratory is responsible for developing the comparison protocol, preparing and distributing transfer standards (transfer standards may be prepared by the other lab), conducting measurements of all transfer standards, and preparing the comparison report.

2.2 A link laboratory in a bilateral CCPR KC is a participant of the bilateral KC and also is a participant of the previous CCPR KC and provides the link of results between the bilateral KC and CCPR KC.

2.3 A link laboratory in a bilateral RMO KC is a participant of the bilateral KC and also is a participant of the previous or current RMO or CCPR KC and provides the link of results between the bilateral KC and the RMO or CCPR KC.

2.4 The participant whose result is to be linked to the CCPR or RMO KC is called "non-link laboratory".

2.5 The link laboratory normally serves as the pilot laboratory, but it is possible for the non-link laboratory to serve as the pilot lab. In this situation the link laboratory will be responsible for pre-draft A, the non-link laboratory will act as pilot for comparison registration, preparation of Draft A and subsequent work.

3. Development of Technical Protocol

3.1 Follow the CCPR *G6: Guidelines for RMO PR Key Comparisons*. The procedures for submitting the results to a third party lab (5.2, 5.3) should be included in the protocol.

4. Registration to KCDB and approval of the technical protocol

4.1 In case of an RMO bilateral KC, upon completion of the technical protocol, the pilot lab sends it to the RMO PR TC Chair, who will submit the technical protocol to WG-KC Secretary (copy to WG-KC Chair) to request for approval.

4.2 After approval of the protocol, the RMO PR TC Chair sends the KCDB entry form and the protocol document to CCPR Executive Secretary. The RMO PR TC Chair should obtain confirmation of the receipt of the form by BIPM.

4.3 In case of a bilateral CCPR KC, the pilot laboratory sends the technical protocol to the WG-KC Secretary for approval and informs the KDCB manager.

4.4 Bilateral comparisons are registered in KCDB with an ID number ending with a dot and a serial number (e.g., EURAMET PR-K4.1). The number before the dot is identical with that of the main CCPR or RMO comparison.

5. Measurements

5.1 Follow the CCPR *G6: Guidelines for RMO PR Key Comparisons* in general, with exceptions for the points given here (5.2-5.4).

5.2 A third party (WG-KC Secretary) is designated for the comparison, and all the measurement results, both from the non-link laboratory and the link laboratory are

submitted to the third party upon completion of each measurement, to ensure blindness of the comparison.

5.3 At completion of all measurements, the third party sends all the data received to the link laboratory, so that the link laboratory can start Pre-Draft A process.

5.4 The third party sends to both participants all the raw data received after Draft A is issued.

6. Pre-Draft A process

6.1 The link laboratory always takes responsibility for pre-draft A, even if the non-link laboratory is acting as pilot for other stages.

6.2 Sections 1, 2, and 3 of CCPR G2: *Guidelines for CCPR Comparison Report Preparation* shall be followed by the link laboratory, but simplified as below.

6.3 In the verification of results, the link laboratory sends the other participant's reported results (as received from the third party) to ask them to verify.

6.4 In the review of uncertainty budget, only the non-link laboratory's uncertainty budget is reviewed.

6.5 Relative data should be prepared by the link laboratory. In the review of Relative Data, the link laboratory prepares the Relative Data and sends to the other participant, and discusses any need for removing data of unstable artifacts, before the non-link laboratory has seen any absolute results.

6.6 If the non-link laboratory is acting as pilot, at the end of the pre-draft A stage, the link laboratory passes the results and the calculations of the pre-draft A phase to the pilot for the Draft A report preparation.

7. Preparation of Draft A (for Key Comparisons)

7.1 After the Pre-Draft A processes are complete, the pilot laboratory prepares and distributes Draft A to both participants, which discloses the absolute results of the comparison. The Draft A should tabulate all the results as well as present them in graphical form as necessary. It is recommended that the pilot laboratory also distribute the data of the analyses in a spreadsheet file. The Draft A should be distributed within six months after completion of all the measurements of the comparison.

7.2 The results of the bilateral KC are linked to the results of the most recent CCPR KC of the same quantity. The unilateral Degree of Equivalence (DoE) of the non-link laboratory should be calculated using all appropriate information available from the Key Comparison and Bilateral Comparison. The analysis process should be as simple as possible and transparent. In order to reduce the checking requirements it should follow, where possible, previously published approaches. An example approach is given in Appendix A.

7.3 Bilateral DoEs are not required.

8. Review of Draft A and preparation of Draft B

- Follow the procedure given in CCPR G2: *Guidelines for CCPR Comparison Report Preparation.*

References

1. CIPM MRA-D-05, *Measurement comparisons in the CIPM MR*A, Version 1.1 of December 2011.

Appendix A: An example analysis approach

This example analysis is based on the work of Ojanen *et al* [1]. It assumes that a single artifact was measured during each comparison. When multiple artifacts are used, the 'effective single artifact' can be determined using a simple mean of different artifact values of $(y_\alpha - y_i)$ in Equation (1) or a simple mean of different artifact values of $(y_\alpha / y_i - 1)$ in Equation (2). Here, subscript α denotes the non-link laboratory, and subscript *i* denotes the link laboratory (*i* is the designation of the link laboratory as a participant in the CCPR KC to be linked). The uncertainty associated with this effective single artifact is given by the uncertainty declared by the participant for a single artifact. This ensures that the uncertainties associated with random effects are not reduced by taking a mean, depending on the number of measurements made by the participant – in this way the participants are treated more equitably.

The abbreviation KC represents here the CCPR Key Comparison that the bilateral comparison is linked to. The abbreviation BC represents the bilateral comparison.

It is important to understand whether the comparison model is based on absolute differences or relative differences. In an absolute-difference model:

- The KCRV has units of the key comparison quantity.
- All uncertainty components have the same units.
- The DoE of the BC participant will be the best estimate of the systematic offset of that participant's measurements. So if the participant measured the CCPR KC 'virtual artifact', this would be the offset in that participant's measured value from the KCRV in the same units.

In a relative-difference model:

- The KCRV is the value 1. It is the 'average' ratio of participant's measurements of the 'virtual artifact' to the KC value of the 'virtual artifact'.
- All uncertainty components are expressed relatively, in percentages.
- The DoE of the BC participant will be the best estimate of the systematic ratio (minus one) of that participant's measurements to the KCRV. So if the participant measured the CCPR KC 'virtual artifact', this would be the ratio of that participant's measured value and the KCRV minus one. It could be expressed as 'percentage error'.

The value component of the unilateral Degree of Equivalence (DoE) of the non-link laboratory α is calculated from the unilateral DoE of the link laboratory *i* in the KC and the difference between the measurement results of the participants of the BC. Where multiple artifacts have been used, the average (simple mean) difference can be used instead. Where at least one participant has made two measurements, the two measurements should be averaged (simple mean).

For an absolute-difference model:

$$D_{\alpha(i)} = \underbrace{x_i - x_{\text{ref}}}_{D_i} + \underbrace{\left(y_\alpha - y_i\right)}_{\substack{\text{from average of} \\ \text{artefacts}}}, \tag{1}$$

where

- $D_{\alpha(i)}$ is the unilateral DoE of the non-link laboratory α
- $D_i = x_i - x_{\text{ref}}$ is the unilateral DoE for the link laboratory, calculated during the KC from the measurement result x_i and KCRV x_{ref}.
- $y_\alpha - y_i$ is the "average value" (of multiple artifacts) of the difference between the non-link laboratory's measurement result (or average of the non-link laboratory's measurements) and the link-laboratory's measurement result (or average result) for each artifact in the BC.

For a relative-difference model:

$$D_{\alpha(i)} = \underbrace{x_i / x_{\text{ref}} - 1}_{D_i} + \underbrace{\left(y_\alpha / y_i - 1\right)}_{\substack{\text{from average of} \\ \text{artefacts}}} \tag{2}$$

where

- $D_{\alpha(i)}$ is the unilateral DoE of the non-link laboratory α
- $D_i = x_i / x_{\text{ref}} - 1$ is the unilateral DoE for the link laboratory, calculated during the KC from the measurement result x_i and KCRV x_{ref}.
- $y_\alpha / y_i - 1$ is the "average value" (of multiple artifacts) of the ratio between the non-link laboratory's measurement result (or average of the non-link laboratory's measurements) and the link-laboratory's measurement result (or average result) for each artifact in the BC, subtracting unity to obtain a DoE.

The uncertainty component of the unilateral DoE is given as an expanded uncertainty

$$U\left(D_{\alpha(i)}\right) = 2u\left(D_{\alpha(i)}\right) \tag{3}$$

where the standard uncertainty is calculated using

$$u^2\left(D_{\alpha(i)}\right) = \underbrace{u_\alpha^2}_{\substack{\alpha \text{ declared uncert.}}} + \underbrace{u^2\left(x_{\text{ref}}\right) + \left(1 - 2w_i\right)s_{\text{KC}}^2}_{\text{KC effects}} + \underbrace{u_{i,\text{r,KC}}^2 + u_{i,\text{st}}^2 + u_{i,\text{r,BC}}^2}_{i \text{ linking quality}} + \underbrace{s_{\text{BC}}^2}_{\text{BC effects}} \tag{4}$$

There are four contributions in this uncertainty calculation. These are exactly the same for absolute-difference and relative-difference models except that for absolute-difference models the uncertainties are absolute uncertainties (expressed in the unit of the KC quantity) and for relative-difference models the uncertainties are relative uncertainties (expressed, for example, as a percentage).

Uncertainties associated with non-link laboratory measurement results

- u_α is the declared total standard uncertainty of the non-link laboratory for a single artifact. This includes uncertainties due to both correlated and uncorrelated effects.

Key comparison effects

- $u\left(x_{\text{ref}}\right)$ is the standard uncertainty associated with the Key Comparison Reference Value. This value is available from the KC report.
- s_{KC} is the transfer uncertainty for the KC. This may be an artifact instability factor calculated from known effects, or it may be the additional s term added during a Mandel-Paule approach in obtaining consistency of the KC results. In either case this term is available from the KC report. If it is not used in the KC report, $s_{\text{KC}} = 0$.
- The multiplier $\left(1-2w_i\right)$ is calculated from the weight w_i assigned to the linking laboratory's measurements in the calculation of the KCRV. This value is provided in the KC report; however, where it is not available, then $w_i = 0$ will give a conservative estimate of the uncertainties.

Linking quality

- This term considers the quality of the link provided by the link laboratory. It includes the standard uncertainty associated with uncorrelated effects (random uncertainty) of the link laboratory during the KC, $u_{i,\text{r,KC}}$,
- the standard uncertainty associated with reproducibility of the link laboratory's scale between the KC and BC, $u_{i,\text{st}}$.
- the standard uncertainty associated with uncorrelated effects (random uncertainty) of the link laboratory during the BC, $u_{i,\text{r,BC}}$

Bilateral comparison effects

- This term considers uncertainties from the bilateral comparison. This includes the standard transfer uncertainty of the bilateral comparison, s_{BC}. The transfer uncertainty may come from known effects (e.g. known artifact instability), or from the similar term during the key comparison (to account for artifact instability that is not 'visible' in the BC). Alternatively it may be appropriate to consider $s_{\text{BC}} = 0$.

Further notes on linking quality for older KCs

For some of the older KCs, the declared uncertainty associated with the link laboratory's measurements during the KC may not be available broken down into uncertainties associated with correlated (systematic) and uncorrelated (random) effects. In this case the term $u_{i,\text{r,KC}}$ is not readily available. Then its value during the KC may be estimated and/or the cases (a) and (b) below may be considered:

(a) If the time interval between KC and BC is short or if the link laboratory can use additional evidence to confirm the stability of its scale, then it may be assumed that $u_{i,\mathrm{st}} = 0$ and/or $u_{i,\mathrm{r,KC}} = u_{i,\mathrm{r,BC}}$ (stable scale).

(b) Alternatively for some older KCs, it may be more appropriate to assume that the quality of the link comes from the combined standard uncertainty associated with the link laboratory's measurements during the KC. In this case $u_{i,\mathrm{r,KC}}^2 + u_{i,\mathrm{st}}^2 = u_i^2$ (scale stability unknown), where u_i is the declared total standard uncertainty of the link laboratory during the KC.

[1] Ojanen, M, Shpak, M, Kärhä, P, Leecharoen, R and Ikonen, E 2009 Uncertainty evaluation for linking a bilateral key comparison with the corresponding CIPM key comparison *Metrologia* **46**(**5**) 379-403