Key Comparison Report Template

Consultative Committee for Mass and Related Quantities

The Consultative Committee on Mass and Related Quantities (CCM) and its Working Groups (WGs) organize international key comparisons to measure the degree of equivalence of national standards for mass and related quantities. General rules about the conduct of comparisons can be found in CIPM document [1]. The key comparisons (KCs) are usually multi-year efforts. A pilot laboratory finds a suitable transfer standard (TS), circulates it between participants, processes data, and writes a comparison report that is reviewed by the participants, the responsible WG, the CCM Executive Secretary, and the CCM WG on Strategy and MRA coordination, and posted on the BIPM key comparison data base (KCDB).

The goal of this document is to provide a template for the comparison report that will lead to more uniform formatting, brevity, readability, accessibility of the results, completeness, and assist pilot labs that are new to the process.

Summary comparison results should be located early in the report so that they are accessible. For most readers, the abstract and the graphical summary of results at the beginning of the report will tell them everything that they want to know. The body of the report should be 15 pages or less and should tell a more interested reader details unique to the particular comparison regarding the transfer standard, its performance, the protocol, and comparison calculations. Commonly used calculation methods should use references and not reproduce the equations. Using a suggested set of variable names given in this document will reduce the need to reproduce equations for calculating the key comparison reference value (KCRV) and other quantities. Appendices should be used to archive all of the information necessary to reproduce the comparison calculations.

In this report template, recommended section titles are followed by general comments in grey text. Following the general comments, examples of text and figures from various CCM comparisons are given. In some cases, the information or text has been altered from the original.

Title Page *The comparison identifying number, measurand, and its range should appear in the title. When the report is approved and submitted to the Executive Secretary for publication on the KCDB, the label should be updated to "Final Report". All participants should be listed as authors in order to receive appropriate credit for their contributions. The pilot may differentiate the roles (pilot, co-pilot, participant) if they like, as shown in this example.*

Example 1

CCM.FF-K6.2011:

CIPM Key Comparison of Low-Pressure Gas Flow, 2 m^3/h to 1000 m^3/h

Final Report

Pilot
Miroslava Benková – CMI, Czech Republic
Participants
Bodo Mickan – PTB, Germany
Stefan Makovnik – SMU, Slovakia
Roberto Arias – CENAM, Mexico
Khaled Chahine – NMI, Australia
Tatsuya Funaki – NMIJ AIST, Japan
Chunhui Li – NIM, China
Hae Man Choi – KRISS, Korea
Denys Seredyuk – GP, Ukraine
Chun-Min Su – CMS, Chinese Taipei
Christophe Windenberg – LNE-LADG, France
John Wright – NIST, USA

February, 2014

Abstract *An abstract is required for the KCDB. Details about linkage to other comparisons can also be included here.*

Example 1 based on CCM.FF-K6.2011

The CCM.FF-K6.2011 comparison was organised for the purpose of determination of the degree of equivalence of the national standards for low-pressure gas flow measurement over the range 2 m^3/h to 100 m^3/h. A rotary gas meter was used as a transfer standard. Eleven laboratories from four RMOs participated between August 2010 and December 2012 – EURAMET: PTB, Germany; SMU, Slovakia; LNE-LADG, France; SIM: NIST, USA; CENAM, Mexico; APMP: NMIJ AIST Japan; KRISS, Korea; NMI, Australia; NIM, China; CMS, Chinese Taipei; COOMET: GP GP Ivano-Frankivs'kstandart-metrologia, Republic of Ukraine and all participants reported independent traceability chains to the SI. The measurements were provided at prescribed reference pressure and temperature conditions. All results were used in the determination of the key comparison reference value (KCRV) and the uncertainty of the KCRV. The reference value was determined at each flow separately following "procedure A" presented by M. G. Cox [3]. The degree of equivalence with the KCRV was calculated for each flow and laboratory. All reported results were consistent with the KCRV. This KCRV can now be used in the further regional comparisons.

Graphical Summary of Results *A figure illustrating the degree of equivalence of the participants is required for the KCDB. The recommended format of the figure is shown in the following examples, i.e., the degree of equivalence (difference between the participants' results and the KCRV) with 95 % confidence level error bars for the uncertainty of the degree of equivalence. The x-axis represents the KCRV and error bars touching the x-axis are a good indicator of participants meeting their uncertainty claims. Note that it is sometimes necessary to use multiple graphs to cover the entire range of the comparison.*

Example 1 based on CCM.P-K4 (1998)

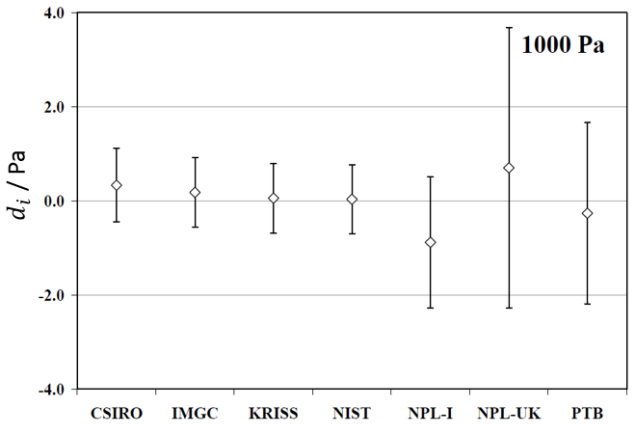

Figure 1. Degrees of equivalence expressed as the deviation of corrected mean gauge readings from the key comparison reference value at 1000 Pa. The error bars refer to expanded uncertainties of the deviations at a 95 % level of confidence.

Example 2 based on CCM.D-K4 (2010) *This example illustrates how multiple measurand set points (or transfer standards) should be plotted in the same figure to save space.*

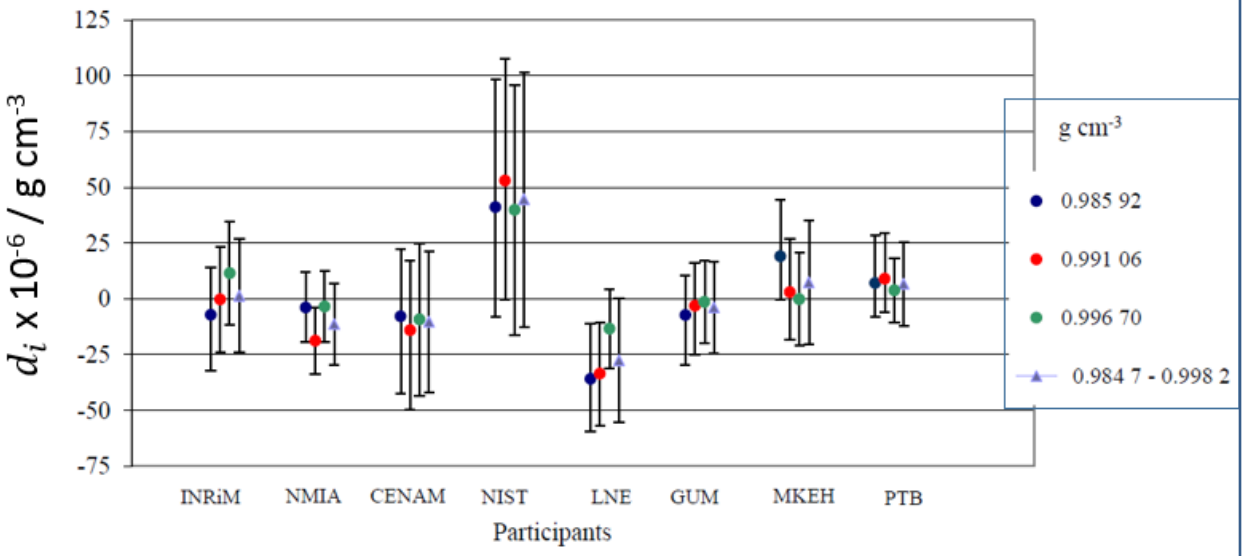

Figure 2. Degrees of equivalence with respect to the KCRV of each laboratory for hydrometer 9340171. The symbols represent the three set points used and the average of the three set points in g/cm^3. The error bars show the expanded uncertainty of the degree of equivalence for each calibrated value.

Contents *A table of contents is recommended to assist the reader in finding particular information.*

1. Introduction 5
2. List of Participants, Facilities Used, Circulation Scheme 5
3. Transfer Standard: 6
4. Comparison Protocol 7
5. Methods of Measurement and Range of Conditions 8
6. Transfer Standard 10
6.1. Corrections to the Transfer Standard 10
6.2. Uncertainty due to the Transfer Standard 10
7. Data Processing and Computation of the KCRV 14
8. Results 15
9. The Key Comparison Reference Value and Its Uncertainty 15
10. Degrees of Equivalence between individual participants and the KCRV 16

11. Summary and Conclusions 17
13. Explanations for Discrepant Reported Values 18
14. Nomenclature 18
15. Appendices 19
16. References: 19

1. Introduction *The introduction should not repeat information that can be found elsewhere in the KC report, for instance in the title or the abstract. It is not necessary to describe the structure of the relevant CC, WG, the general purpose of a comparison, etcetera. Information about prior comparisons for the same measurand, regional comparisons, and how the range of the measurand used in the comparison was selected are appropriate here.*

Example 1 based on CCM.FF-K4

During the 10th WGFF meeting, held in Taiwan in October 2010, it was agreed to perform the second round of CCM.FF-K4 for Volume of Liquids at 20 L and 100 mL. CENAM offered to act as the pilot laboratory; and started re-manufacturing the transfer standards for the intended purpose. Based on comments from volume technical experts that participated in the first version of this KC, some improvements were implemented on the transfer standards, so that better repeatability and reproducibility were expected for CCM.FF-K4.1.2011.

2. List of Participants, Facilities Used, Circulation Scheme *A table is a compact way to present information about the participants, including the country, the acronym of the lab, the uncertainty of the reference standards, the date that the testing occurred in each lab, and whether or not the reference standard has a traceability chain that is independent from the other participants. If the lab's traceability is dependent, the report can state the source, e.g., "No, NEL". The type of reference standard or facility name is particularly useful if it matches a name that is given in the NMI's calibration and measurement capabilities (CMCs). Other information that can be included in this section:*

- *Descriptions or diagrams explaining the circulation scheme (single loop, multiple loop [petal], star, or other),*
- *Failures or damage to the transfer standard and actions taken,*
- *Explanation of unexpected delays,*
- *Participation by non-Mutual Recognition Arrangement (MRA) signatories.)*

Example 1 from CCM.FF-K6 (2005)

Table 1. KC Participants, facilities used, reference standard uncertainty, dates of test, and independence of the participant's traceability from other participants.

Participant	Type of reference standard	Reference standard uncertainty (k=2, %)	Date of test	Independent traceability?
NIST (United States)	34 L and 677 L *PVTt*	0.025	Mar 2005	yes
PTB (Germany)	Piston provers	0.075	May 2005	yes
	CFV working standards	0.04		yes
	Positive displacement working standard	0.059 to 0.076		yes
NEL (United Kingdom)	Piston provers	0.08	June 2005	yes
KRISS (Korea)	12 L piston prover	0.065	Aug 2005	yes
	150 L and 600 L bell provers	0.055		yes
NMIJ (Japan)	Gravimetric	0.05	Oct 2005	yes
	PVTt	0.075 to 0.1		yes
NMIA (Australia)	Piston provers	0.05	Jan 2006	yes
	Bell prover	0.08		yes
CENAM (Mexico)	Piston provers	0.045	April 2006	yes
	Bell prover	0.075		yes

3. Transfer Standard: *A description of the transfer standard with relevant technical information should be included. Pictures of the transfer standard are encouraged.*

Example 1 from CCM.FF-K6.2011

The transfer standard was a rotary gas meter, a new model of S-Flow meter inside the body Actaris Delta 2050. The transfer standard, a pulse transmitter connector and a filter were shipped in one transfer box.

Figure 3. Rotary gas meter Actaris Delta S-Flow

Basic technical specification:

Type:	Delta 2050 S-Flow
Manufacturer:	ActarisGaszählerbau GmbH, Germany
Size:	G65
Serial number:	GN-HD-001
Flow range:	2 m^3/h to 100 m^3/h
P_{max}:	4 MPa
Inside diameter:	50 mm

4. Comparison Protocol *May include: a brief description of the protocol (including a description of how to select the method to be used to determine the KCRV for CCM KCs), any special instructions about conducting the measurements, such as a warm-up times or required ambient conditions, deviations from the protocol by any participants, information learned about the transfer standard and the protocol that might improve future similar comparisons. Reproducing the entire protocol here is not recommended, however, it can be included as an appendix. If the transfer standard was damaged, repaired, or replaced, that information can be included in this section.*

Example 1 from CCM.FF-K6.2011

The measured range was 2 m^3/h to 100 m^3/h. If the laboratory was not able to cover the whole flow range they could make measurements in one part of the flow range.

- The transfer standard was tested in the horizontal position using air.
- The reference pressure from the transfer standard was measured from the output "P_m" (pressure tap located at the outlet of the meter).
- The second pressure point to determine the pressure loss of the transfer standard was defined at the inlet of the meter.
- The reference temperature from transfer standard was measured upstream of the transfer standard (Figure 4).

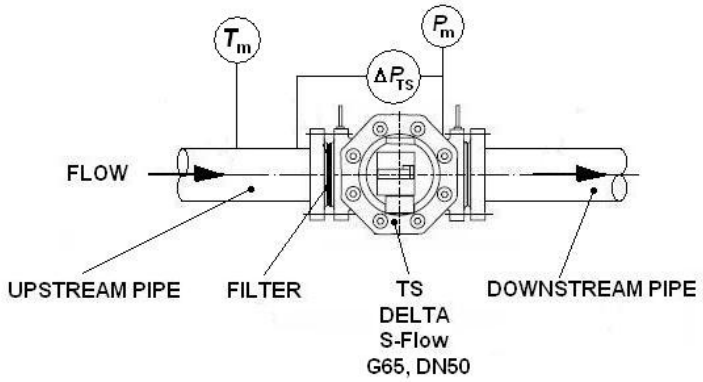

Figure 4. Recommended installation of the meter

- It was necessary to use the pulse transmitter.
- There was no lubrication of the meter.
- Operating conditions:
  - the calibration medium was air,
  - air temperature: (20 ± 5)°C,
  - ambient relative humidity range: 25 % to 75 %,
  - ambient atmospheric pressure range: 86 kPa to 106 kPa.
- The flow rate had to be within ± 3 % of the required value.
- Flow set points: (2; 4.5; 6.6; 9.1; 13.1; 16; 24; 32; 40; 50; 60; 70; 80; 90; 100) m^3/h.

5. Methods of Measurement and Range of Conditions *A description of the methods and equipment used in each participant's reference standard may be given. If the operating conditions in the participants' laboratories are relevant to the performance of the transfer standard, the range of conditions can be given in this section.*

Example 1 from CCM.FF-K2.2015

A summary of the calibration methods used by the participants is shown in Table 2. Details are given in Appendix B.

Table 2. Calibration methods.

NMI	Calibration method	Reference standard
BEV	Volumetric method with flying start and stop	Volume tank
CENAM	Volumetric method with flying start and stop	Unidirectional Pipe prover
CMS	Static and gravimetric method with standing start and stop	Weighing scale
LNE-TRAIL	Volumetric method with flying start and stop	Unidirectional Pipe prover
NEL	Static and gravimetric method with standing start and stop	Weighing scale
NMIA	Volumetric method with flying start and stop	Small volume prover
NMIJ	Static and gravimetric method with flying start and stop	Weighing scale

Example 2 from CCM.FF-K4.1.2011

Table 3. Summary of the experimental procedure employed at the different NMIs

	Weighing*		Water**	De-aerated water?	Density formula
	20 L	100 mL			
CENAM	DS	DR	IE + O	No	Tanaka et al
NIST	DR		O	No	Patterson & Morris
IPQ	SS	SS	IE + O	No	Tanaka et al
VSL	DS	DS	DM+2D	No	Bettin & Spieweck
SP	DS	SS	IE	Yes	Bettin & Spieweck
INRIM	SS	SS	IE + 2D	No	Tanaka et al
NIM	ABA	SS	IE	No	Tanaka et al
INMETRO	ABA	DR	DI	No	measured

**Weighing: DS*: Double substitution; *DR*: direct reading; *SS*: single substitution; *ABA*: substitution weighing

***Water:* IE: Ion exchange; O: Inverse osmosis; 1D: single distillation; 2D: double distillation, DM: demineralized

Example 3 from CCM.FF-K6.2011

The conditions during measurements were described by all participants. The values are given in Table 4.

Table 4. Temperature (°C) in participating laboratories during measurements

NMI	SlovakiaSMU	GermanyPTB	Ukraine GP Ivano-Frankivs'kstandart-metrologia	AustraliaNMI	USA NIST	Mexico CENAM	Korea KRISS	China NIM	Chinese Taipei CMS	Japan NMIJ/AIST	France LNE-LADG
Max	20.71	22.71	18.89	21.45	23.89	20.40	21.33	21.49	22.83	24.45	18.43
Min	19.84	21.42	18.76	21.30	23.13	19.80	18.67	20.26	22.26	23.91	17.65
Max-Min	0.88	1.29	0.13	0.16	0.76	0.60	2.66	1.23	0.57	0.54	0.78
Mean	20.19	21.77	18.82	21.38	23.66	20.16	20.07	20.89	22.63	24.14	18.15

6. Transfer Standard *Transfer standards often drift in time and are sensitive to the conditions under which they are used, for example the pressure, temperature, and humidity conditions in a participant's laboratory.*

6.1. Corrections to the Transfer Standard *In some comparisons, the transfer standard's drift and its sensitivity to measurement conditions are well understood and corrections are made for them by the pilot lab. In these cases, the corrections should be explained. In CCM.P-K12, leak flows from permeation tube transfer standards were corrected by a function of time [2].*

Example 1, from CCM.P-K4 (2002)

6 REDUCTION AND ANALYSIS OF THE REPORTED DATA

The reduction and analysis of the key comparison data required that several factors be addressed. These included zero-pressure offsets (Section 6.1), thermal transpiration effects (Section 6.2), deviations of the actual pressures realized from the target pressures (Section 6.3), relatively large calibration shifts in the capacitance diaphragm gauges (Section 6.4), and normalization of the data from two different transfer standard packages (Section 6.5). Methods for estimating uncertainties (Sections 6.6 and 6.7) and for evaluating degrees of equivalence (Section 6.8) are also described.

6.1 CORRECTIONS FOR ZERO-PRESSURE OFFSETS

The first step in reducing the comparison data was to correct the reading of each gauge i for its zero-pressure offset. The index i is equal to either 1 or 2 and refers to either, CDG1 and CDG2 or RSG1 and RSG2 (see Figure 1). At a given target pressure during calibration run k, the corrected reading of gauge i is given by:

$$p_{\mathrm{ikl}} = p_{\mathrm{Gikl}} - \langle p_{\mathrm{Gik0}} \rangle_{10} + p_{\mathrm{REFkl}} \quad \text{for liquid column manometer data} \qquad (1)$$

Where p_{Gikl} is the uncorrected gauge reading, $\langle p_{\mathrm{Gik0}} \rangle_{10}$ is the mean of 10 zero-pressure readings taken just prior to the start of calibration run k, and p_{REFkl} is the reference pressure reading during repeat set l.

6.2. Uncertainty due to the Transfer Standard *The uncertainty of the transfer standard generally includes the uncertainty due to its sensitivity to measurement conditions and the uncertainty due to its instability in time. In cases where transfer standard uncertainty may be significant relative to any of the participating labs' uncertainties, the transfer standard uncertainty and how it was estimated should be included in the report. If transfer standard uncertainty is not a significant contributor to the uncertainty of the values reported by participants, this should be stated in the report. If the transfer standard was tested by the pilot*

lab more than once to assess stability, state which of the Pilot's multiple data sets was used in calculating the key comparison reference value and degree of equivalence.

Example 1 from CCM.FF-K6.2011 *In this example, calibration stability and environmental temperature sensitivity were deemed to be the most significant sources of transfer standard uncertainty. Note that in some cases, a transfer standard drifts in a predictable manner with respect to time and the drift can be corrected.*

The stability of the transfer standard was checked before starting the comparison by the assisting lab LNE-LADG France and seven times before and during the comparison by the pilot laboratory (Figure 5). The range of error $(\varepsilon_{\max} - \varepsilon_{\min})$ of these seven calibrations was 0.103 %. A rectangular distribution was applied to the range of the calibration changes observed by the pilot lab, giving a standard uncertainty due to transfer standard calibration stability of $u_{\text{drift}} = \frac{(\varepsilon_{\max} - \varepsilon_{\min})}{2\sqrt{3}}$ = 0.030 %.

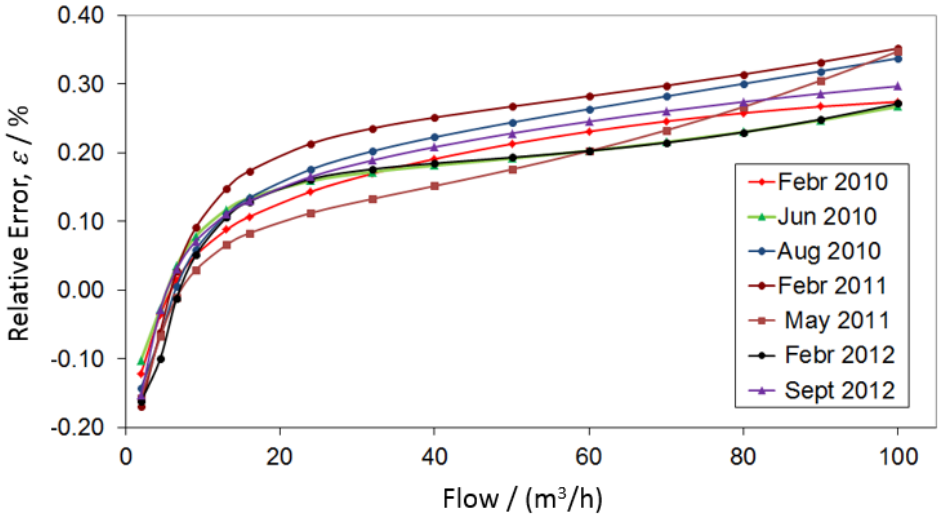

Figure 5. Stability of the transfer standard.

Based on preliminary testing in the pilot laboratory, temperature is the only significant sensitivity of the transfer standard to the testing conditions. The temperature sensitivity of the transfer standard was checked by PTB Germany (Figure 6) and measured to be 0.0035 %/°C. The range of temperatures the transfer standard was exposed to in the participants' labs was 6.8 °C. Assuming a rectangular distribution leads to a standard uncertainty due to temperature effects of $\boldsymbol{u_T} = \frac{\mathbf{0.0035\ \%/°C\ (6.8\ °C)}}{\mathbf{2\sqrt{3}}}$ = 0.007 %.

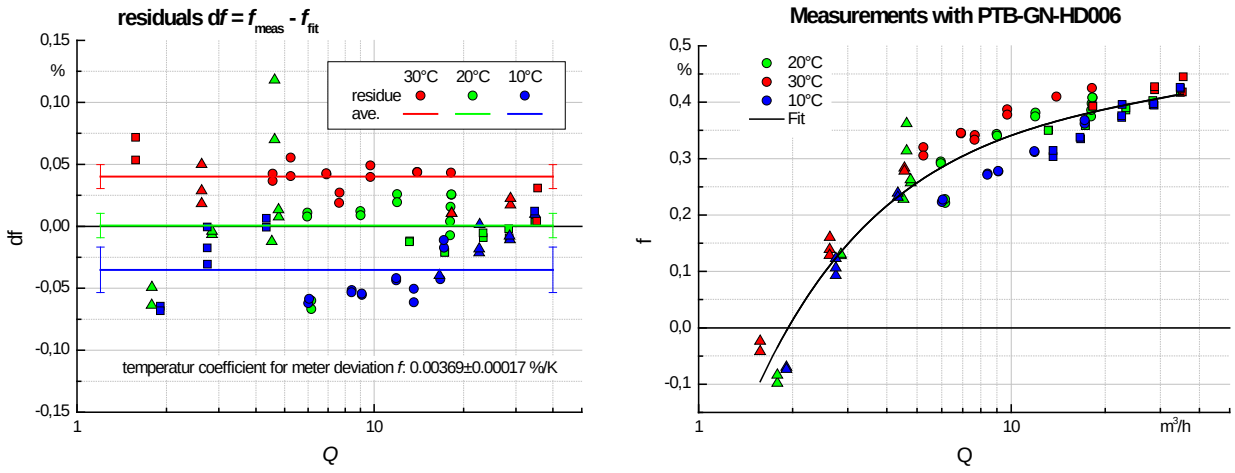

Figure 6. Temperature stability.

Combining the uncertainties by root-sum-of-squares:

$$u_{\mathrm{TS}} = \sqrt{u_{\mathrm{drift}}{}^2 + u_{\mathrm{T}}{}^2} = 0.031\ \% \qquad (2)$$

leads to a standard uncertainty due to the transfer standard of 0.031 %.

This transfer standard uncertainty component was combined by root-sum-of-squares with the standard uncertainty provided by each participating laboratory and the standard deviation of the mean for the repeated measurements at each set point (Type A). The ratio of the transfer standard uncertainty to any participant's flow standard uncertainty is ≤ 1.24. The data set collected by the pilot lab in May 2011 was used in the KCRV and degree of equivalence calculations.

Example 2 from CCM.FF-K2.2011 *In this example, the influence of internal pressure on a liquid flow transfer standard is quantified by experiments performed by the pilot lab. Other sections of the report (not shown here) quantified transfer standard sensitivities to other variables.*

Pressure Effect: The relative *K* factor of the transfer standard at different pressures measured using the pilot laboratory's primary standard is shown in Fig. 7. The calibration liquid was light oil and the liquid temperature was 35 ºC. The pressure effect on the relative *K* factors is less than 0.003 4 %/MPa, and the difference of liquid pressure between each pair of the participants is estimated to be less than ± 0.25 MPa. Therefore, the standard uncertainty due to the difference of the pressure between each pair of the participants is estimated to be 0.0009 %.

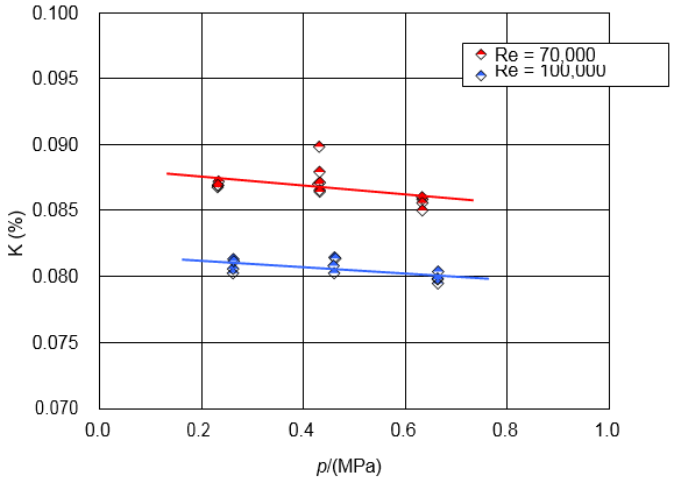

Figure 7. Relative *K* factor of transfer standard at different pressure.

Example 3 from CCM.FF-K2.2011 *In this section of the K2.2011 report, the transfer standard uncertainty contributions are listed and combined by root-sum-of-squares.*

The standard uncertainty of the calibration results, that is the relative *K* factor at each participating laboratory, is expressed by Equation 2.

$$u_{x_i}^{\ 2}(K_i) = u_{\text{base }i}^{\ 2} + u_{\text{DUT},i}^{\ 2} + u_{TS}^{\ 2} = u_{\text{base }i}^{\ 2} + u_{\text{DUT},i}^{\ 2} + u_{\text{drift}}^{\ 2} + u_{vis}^{\ 2} + u_{\text{Re}}^{\ 2} + u_{p}^{\ 2} + u_{upst}^{\ 2} \quad (3)$$

$u_{\text{base }i}$: Standard uncertainty due to calibration facility in the laboratory i , Base uncertainty

$u_{\text{DUT},i}$: Standard uncertainty due to repeatability of transfer standard at calibration, DUT uncertainty

u_{TS} : Standard uncertainty due to the transfer standard

u_{drift} : Standard uncertainty due to reproducibility of the transfer standard

u_{vis} : Standard uncertainty due to the effect of viscosity and temperature differences between each pair of the participants on the transfer standard.

u_{Re} : Standard uncertainty due to the effect of differences of Re between each pair of the participants on the transfer standard.

u_{p} : Standard uncertainty due to the effect of pressure differences between each pair of the participants on the transfer standard.

u_{upst} : Standard uncertainty due to the effect of differences of upstream condition between each pair of the participants on the transfer standard.

The standard uncertainty due to the transfer standard μ_{TS} is estimated to be 0.008 %.

7. Data Processing and Computation of the KCRV *A survey of CCM comparisons conducted in 2016 showed that most were using the methods documented by Cox [3] to calculate the key comparison reference value (KCRV) using Procedure A (uncertainty weighted mean and χ-squared consistency test) or Procedure B (median). These methods are now sufficiently well known that the equations for calculating the KCRV and related quantities should not be reproduced or explained in comparison reports. Covariance due to labs with traceability to other participants or participants that share a common source of traceability to a third party must be taken into account and listed so that a reader could duplicate the comparison calculations (see Section 11, Degrees of Equivalence). While Cox's approach is widely used, new methods are constantly proposed and may be more appropriate for particular data sets, for example the random effects model used in the NIST Consensus Builder [4]. KCRV calculation methods that are not widely known or where the citations are difficult to obtain should be explained in sufficient detail that they can be duplicated by a reader. If unusual methods are applied, the reasons for doing so should be explained.*

Example 1, based on CCM.FF.K4.1.2001

The KCRV for volume of liquids at 20 L was calculated by applying the "weighted mean" method as described by Cox [3]. The reported values were found to be consistent.

Table 5. Consistency check and computation of KCRV for TS 710-05.

TS 710-05	x_i/mL	$u(x_i)$/mL	$x_i/u(x_i)^2$	$1/u(x_i)^2$	$(x_i - x_{\mathrm{KCRV}})^2/u(x_i)^2$
CENAM	19 993.50	0.40	124 959.401	6.25	0.005
NIST	19 993.39	0.58	59 433.371	2.972 651 61	0.064
IPQ	19 992.97	0.69	41 993.209	2.100 399 08	0.672
VSL	19 993.25	0.34	172 951.948	8.650 519 03	0.714
SP	19 993.45	0.25	319 895.179	16	0.112
INRIM	19 993.55	0.19	553 837.95	27.700 831	0.009
NIM	19 993.14	0.3	222 146.033	11.111 111 1	1.685
INMETRO	19 993.81	0.17	691 827.2	34.602 076 1	2.590
		Σ	2187044.29	109.387588	5.851
			x_{KCRV}/mL	19 993.53	$\chi^2_{0.05,7}$ = 14.07
			$u_{x\ \mathrm{KCRV}}$/mL	0.096	pass

8. Results *Summary results (averages and statistics from repeated measurements) from all participants are required. For each participant, reported values of the measurand, the type B uncertainty for the reference standard used to calibrate the transfer standard, and the type A uncertainty of the reported measurements should be listed. If listing the data requires more than 2 pages, please put them in an appendix. Also refer to an appendix if complete uncertainty budgets for each participant are shown.*

9. The Key Comparison Reference Value and Its Uncertainty *Final uncertainties should have 95% confidence limits and this should be clearly stated. Note that only MRA signatories' and designated institutes' reported values can be used in calculating the KCRV and its uncertainty. Also, "the graphs and tables of equivalence explicitly shown include results only from MRA signatories and designated institutes" [5]. However, the results of non-signatory participants can be included in other portions of the report, e.g. an appendix, so that these results are available as evidence to support future CMC submissions.*

Example 1 based on CCM.FF-K6.2011

Table 6. Key comparison reference values (KCRVs).

Flow/ (m3/h)	2	4.5	6.6	9.1	13.1	16	24	32	40	50
x_{KCRV} (%)	-0.134	0.017	0.070	0.107	0.139	0.165	0.189	0.214	0.233	0.250
$U_{\mathrm{x\ KCRV}}$ (%)	0.025	0.025	0.024	0.023	0.022	0.022	0.022	0.022	0.022	0.022

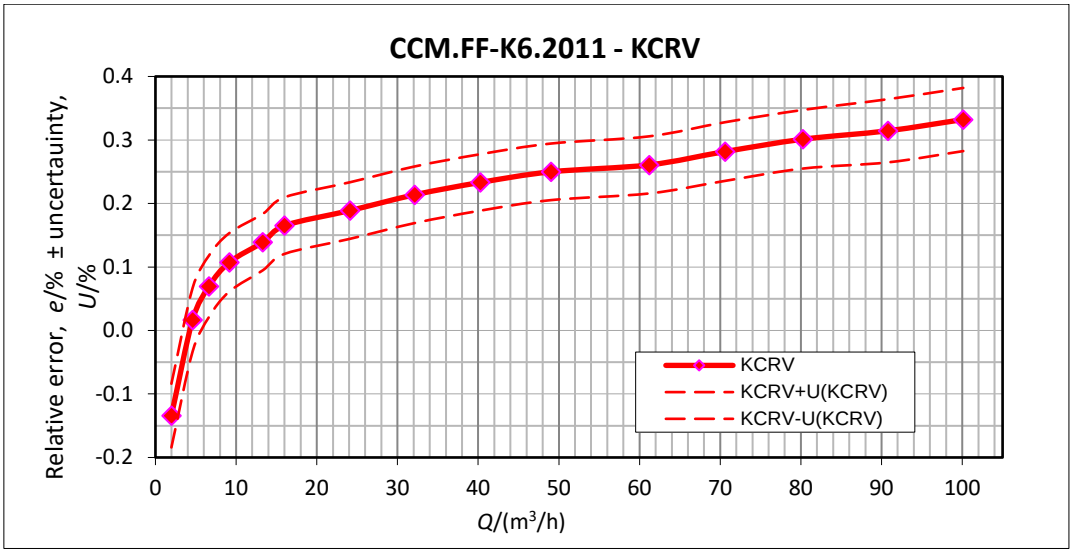

Figure 8. Key comparison reference value.

10. Degrees of Equivalence between individual participants and the KCRV *Tables and figures giving the degree of equivalence and the uncertainty of the degree of equivalence are required in a KC report. Pairwise degrees of equivalence should not be reported for space reasons.*

The uncertainty of degree of equivalence $u(d_i)$ is normally calculated by one of two formulas that depend on whether participant i's results 1) were included in the calculation of the KCRV and 2) are strongly correlated with another participant via their traceability path.

The most common situation is that participant i has traceability that is independent from the other participants and their results were included in the calculation of the KCRV. In this case, Equation 5 in Cox 2002 applies:

$$u(d_i) = \sqrt{{u_{x_i}}^2 - {u_{x\ KCRV}}^2}\,. \qquad (4)$$

This equation also applies to a participant reporting results that are highly correlated with another participant's results. For example, if Lab B produced their comparison results using a reference standard that was calibrated by Lab A (Lab B is traceable to Lab A) and if we conservatively estimate that the correlation coefficient between their results is 1, Equation 4 should be applied.

If participant i's results were not included in the KCRV calculations then their uncertainty for the degree of equivalence is,

$$u(d_i) = \sqrt{{u_{x_i}}^2 + {u_{x\ KCRV}}^2}\,. \qquad (5)$$

Possible reasons for not including a participant's results in the KCRV calculations are: 1) their results were determined to be outliers by a statistical test or 2) the participant is not a signatory of the Mutual Recognition Arrangement or a Designated Institute.

Note that Equations 4 and 5 apply to methods that use a weighted mean of the participants' results as the KCRV. If the median of the participants' results is used as the KCRV (Cox's Procedure B), a Mont Carlo simulation is used to determine $u(d_i)$.

Example 1, from CCM.FF-K4.2.2011 *Text of a different color is often used to highlight results with degree of equivalence greater than the uncertainty of the degree of equivalence.*

Table 7. Degree of equivalence with KCRV for micropipette 354868Z.

NMI	d_i/µL	$U(d_i)$/µL
IPQ	0.05	0.26
CENAM	-0.07	0.27
LNE	0.26	0.32

UME	0.14	0.29
NMISA	0.09	0.27
NIM	-0.15	0.25
INMETRO	**0.56**	0.37
KEBS	-0.18	0.25

11. Summary and Conclusions *The summary and conclusions should include:*

- *A synopsis of the measurand, number of participants, duration of the comparison,*
- *Quantitative information about the performance of the transfer standard,*
- *Discussion of the degrees of equivalence and uncertainty of the participants,*
- *Information on ALL results where d_i exceeds $U(d_i)$. In the case that the performance or stability of the transfer standard is not adequate to support the CMCs of some participating institutes, this should also be noted in the summary/conclusion.*
- *A statement about "how far the light shines", i.e. how far beyond the range of the values tested in the comparison the NMIs and DIs should be allowed to claims CMCs (for example, "provided the results agree with the KCRV, the comparison can be taken as confirmation that the NMI in question has demonstrated equivalence for the calibration of mass standards within an order of magnitude of the nominal value of the transfer standard(s)").*

Example 1, from CCM.FF-K4.1.2011

CONCLUSIONS

CCM.FF-K4.1.2011 was piloted by CENAM. Eight NMIs tested the two 20 L transfer standards, whereas 7 tested the three 100 mL pycnometers

1. CCM.FF-K4.1.2011 for Volume of Liquids at 20 L and 100 mL was conducted during 2012 – 2014. The execution of the CCM.FF-K4.1.2011 was affected by the fact that the transfer package remained at the Brazilian Customs for nearly 8 months; despite this fact, the artifacts did not change their metrological properties, and the KC was completed successfully.

2. No discrepant measurements were distinguished on the 20 L artifacts. The largest difference between two NMIs was 0.004 2 %; whereas the average degree of equivalence for artifacts 710-04 and 710-05 resulted in 0.000 1 % and 0.000 5 %, respectively.

3. Only one participant produced anomalous results for 100 mL measurements; NIM´s result for TS 03.01.12 was inconsistent with IPQ, VSL, SP, INRIM and INMETRO. However, results for artifacts 03.01.16 and 03.01.17 were all fully consistent with each other. The average degree of equivalence for artifacts 03.01.16 and 03.01.17 resulted in 0.000 17 % and 0.001 1 %, respectively.

4. Subsequent linkage to CCM.FF-K4.1.2011 will be based on the results for artifacts 710-04 and 03.01.17, for 20 L and 100 mL, respectively.

Example 2, from CCM.FF-K4.1.2011

In order to assess the support of CMC entries provided by this comparison, it is necessary to compare d_i against declared uncertainty values from the CMC tables. It is expected that d_i values are smaller than U_{CMCs} for supporting purposes.

Table 8. Consistency check for CMC entries for volume of liquids at 20 L.

NMI	U_{CMCs}	$U_{\mathrm{K4.1.2011}}$	$\frac{d_i}{x_{\mathrm{KCRV}}}$	Are the CMCs supported by CCM.FF-K4.1.2011
	%	%	%	
CENAM	0.004 0	0.004 0	0.000 016	yes
NIST	0.015 + 1.2/V, V in L	0.005 8	0.000 15	yes
IPQ	0.01	0.008 5	–0.001 5	yes
VSL	0.01	0.003 4	–0.001 2	yes
SP	0.003	0.002 5	–0.000 44	yes
INRIM	0.005	0.001 9	0.000 082	yes
NIM	n/a	0.003 0	–0.001 6	n/a
INMETRO	0.010	0.002 0	0.001 5	yes

13. Explanations for Discrepant Reported Values *If explanations for discrepant results are submitted to the pilot by a participant between Draft A and Draft B stages of the report, these may be included in the KC report.*

In the event of a dispute over the calculation or interpretation of results, the KC report must "contain a section that briefly describes the points where agreement among the participants could not be achieved and how the final decision was made" [5].

14. Nomenclature *Recommended variable names for quantities commonly used in comparisons are given below.*

k	Coverage factor associated with a specified confidence level.
n	Number of measurements made at a set point.
s	The standard deviation of a set of measurements, sample standard deviation.
U	Expanded, approximately 95 % confidence level uncertainty.
u_{drift}	Long term reproducibility (calibration drift) of the transfer standard.
u_{T}, u_{P}, u_{prop}	Standard uncertainties due to temperature, pressure, and property sensitivities of the transfer standard.
u_{TS}	Standard uncertainty of the transfer standard, accounting for uncertainty due to

	transfer standard drift during the comparison, temperature sensitivities, pressure sensitivities, property sensitivities, etc.
i	Participating lab index.
x_i	Reported value of the measurand by the participating laboratory i.
ε_i	Difference between the transfer standard and the participants reference standard measurements. May have the dimensions of the measurand, or in percent, parts in 10^6, etc.
$u_{\text{base } i}$	Type B standard uncertainty of the participating laboratory's reference standard obtained by using the law of propagation of uncertainty as described in the ISO Guide to Uncertainty in Measurement [6].
u_{x_i}	Standard uncertainty of the reported value from the participating laboratory, accounting for uncertainty due to base reference standard uncertainty, transfer standard uncertainty, and standard deviation of the mean of *n* measurements at each set point.
u_{d_i}	Standard uncertainty of the difference between a participant's reported result and the KCRV.
$u_{\text{CMC } i}$	Standard uncertainty for a lab's calibration and measurement capabilities (CMCs).
d_i	Degree of equivalence $= x_i - x_{\text{KCRV}}$.
En_i	Standardized degree of equivalence between a lab i and the key comparison reference value, $= d_i/2u_{d_i}$.
x_{KCRV}	The comparison reference value.
$u_{x\ \text{KCRV}}$	Standard uncertainty of the comparison reference value (KCRV).

15. Appendices *Data sets longer than 2 pages, detailed uncertainty budgets, descriptions of the reference standards used by participants, detailed explanations of unusual KCRV calculations, the full comparison test protocol, etc. are best placed in appendices, not in the body of the report.*

16. References:

[1] *Measurement comparisons in the CIPM MRA*, CIPM-MRA-G-11.
[2] Jousten, K, Arai, K., Becker, U., Bodnar, O., Boineau, F., Fedchak, J., Gorobey, V., Jian, W., Mari, D., Mohan, P., Setina, J., Toman, B., Vicar, M., and Yan, Y., *Final report of key comparison CCM.P-K12 for very low helium flow rates (leak rates)*, Metrologia Technical Supplement, Volume 50.
[3] Cox, M. G., *The Evaluation of Key Comparison Data*, Metrologia, 2002, **39**, 589-595.
[4] Koepke, A., LaFarge, T, Possolo, A., and Toman, B., *NIST Consensus Builder User's Manual*, May 13, 2017.
[5] *CCM Guidelines for approval and publication of the final reports of key and supplementary comparisons*, CCM-Working Group on Strategy and MRA coordination.
[6] Guide to the expression of uncertainty in measurement, Joint Committee for Guides in Metrology.

Document history

Version	Date	Remarks
1.2	22.05.2018	First approved document
1.3	24.05.2019	Addition of a statement about “how far the light shines”
1.4	24/02/2022	Adapt references to updated CIPM MRA guidance documents and the new BIPM website