CCM.M-K8.2024

Key comparison of kilogram realizations

Final report

8 December 2025

M. Stock[1], P. Conceição[1], H. Fang[1], F. Bielsa[1], A. Kiss[1], Y.-H. Wu[2], S.-J. Shen[2], L. Tsao[2], S.-J. Chen[2], L. Nielsen[3], M. Thomas[4], F. Beaudoux[4], K. Dougdag[4], P. Espel[4], D. Ziane[4], H. Baumann[5], K. Marti[5], D. Zahnd[5], Z. Li[6], J. Wang[6], Y. Bai[6], Y. Wang[6], R. Zhong[6], D. Wu[6], D. Haddad[7], J. Graybill[7], Z. Kubarych[7], N. Kuramoto[8], Y. Kano[8], Y. Ota[8], K. Fujita[8], H. Inaba[8], S. Okubo[8], L. Zhang[8], S. Davidson[9], R.G. Green[10], N. Murnaghan[10], H. Parks[10], D. A. Castro[10], K. Kato[10], M. Borys[11], D. Knopf[11], D. Eppers[11], S. Quabis[11], E. Kuhn[11], A. Peter[11], M. Müller[11], F. Scholz[11], A. Schofeld[11], M. Hämpke[11], H. Ahmedov[12], R. Orhan[12], H. Mutaf[12]

[1] BIPM
[2] CMS/ITRI, Chinese Taipei
[3] DFM, Denmark
[4] LNE, France
[5] METAS, Switzerland
[6] NIM, China
[7] NIST, USA
[8] NMIJ, Japan
[9] NPL, UK
[10] NRC, Canada
[11] PTB, Germany
[12] UME, Türkiye

Bureau International des Poids et Mesures

Table of Contents

Abstract

This report describes the third CCM key comparison of realizations of the kilogram definition based on the fixed numerical value of the Planck constant, which came into force on 20 May 2019. As for the previous comparisons, the objectives were to determine the level of agreement between realizations of the kilogram using Kibble and joule balances and the X-ray crystal density (XRCD) method and to provide input for the calculation of the third "consensus value" of the kilogram. The consensus value serves as the basis for an internationally coordinated dissemination of the kilogram and is updated after each new key comparison. Its use will continue until the CCM decides that satisfactory agreement between realization experiments has been achieved. Another objective was the determination of the reproducibility of the realization experiments by comparing the new results with those of the previous key comparisons of kilogram realizations, CCM.M-K8.2019 and CCM.M-K8.2021.

The comparison was organized by the BIPM and had ten participants. The BIPM, LNE, METAS, NIST, NRC and UME operated Kibble balances, the NIM used a joule balance and the CMS/ITRI, NMIJ and the PTB participated using ^{28}Si spheres, the masses of which were determined with the XRCD method. These realization methods were used to calibrate 1 kg mass standards under vacuum or in air. The standards were sent (in containers filled with air) to the BIPM where they were compared with each other (in vacuum or in air, as at the participants' laboratories) and with BIPM Pt-Ir working standards. The latter were traceable to the International Prototype of the Kilogram (IPK), the mass of which served as the definition of the kilogram until 20 May 2019.

The results of the weighings at the BIPM together with the measurement results communicated by the participants allowed comparison of the values attributed to 1 kg mass standards using the realization experiments of the participants. The level of agreement between mass determinations with the realization experiments and the BIPM as-maintained mass unit, traceable to the Planck constant through the mass of the International Prototype of the Kilogram, could also be deduced.

1 Introduction

On 20 May 2019, a revision of the International System of Units, the SI, came into force [1]. Since then, the kilogram has been defined based on a fixed numerical value of the Planck constant [2]. This leads to the fundamentally new situation in mass metrology whereby the mass unit can, in principle, be realized individually by any National Metrology Institute (NMI) which has developed an experiment allowing the realization of the new definition. At present, the *mise en pratique* of the kilogram [3] recognizes two independent primary methods that are capable of realizing the kilogram with relative uncertainties of a few parts in 10^8, corresponding to a few tens of micrograms at the level of 1 kg: the Kibble balance [4] (a special realization of which has been described as a joule balance [5]) and the XRCD method [6].

In 2017, the numerical values of the defining constants for the four new SI definitions were determined by a least squares adjustment of all available data by the CODATA Task Group on Fundamental Constants [7]. The set of eight results for the Planck constant was not statistically consistent. The Consultative Committee for Mass and Related Quantities (CCM) discussed this situation at its meeting in 2017. It decided that until the dispersion between values became compatible with the individual realization uncertainties, NMIs should base their dissemination on an agreed "consensus value". The details of this international coordination of the kilogram dissemination are described in the "CCM detailed note on the dissemination process after the redefinition of the kilogram" [8]. The consensus value was implemented for the first time in February 2021 after the completion of the first key comparison of kilogram realization, CCM.M-K8.2019 [9, 10]. It was updated after completion of the second key comparison, CCM.M-K8.2021 [11], in March 2023.

Following the completion of the present comparison, a new consensus value will be calculated. It will be determined from the following three data sets, following the rules established by the CCM [8]:

- the key comparison reference value (KCRV) of the first key comparison of kilogram realizations, CCM.M-K8.2019;
- the KCRV of the second key comparison of kilogram realizations, CCM.M-K8.2021.
- the KCRV of the present, third, key comparison of kilogram realizations, CCM.M-K8.2024

The reference values of the first two comparisons are maintained by the BIPM working standards. The three data sets from 2019, 2021 and 2024 can be tied together through the quantifiable stability of the BIPM working standards.

This comparison has the following three objectives: (1) to study the present level of agreement between realization experiments, (2) to provide input for the calculation of the third consensus value and (3) to determine the reproducibility of the realization experiments by comparing the results of the three key comparisons.

The new consensus value will be determined by the *CCM Task Group on the Phases for the Dissemination of the kilogram following redefinition* (TGPfD-kg). The result will be published in a CCM document.

2 Organization of the key comparison

The comparison followed the same scheme as the two previous key comparisons of kilogram realizations. It was organized in the form of a star-comparison, in which each participating institute was requested to send one or two of its own 1 kg mass standards to the BIPM, which acted as the pilot laboratory. The participants could choose between Pt-Ir prototypes, stainless steel standards, Si-spheres or any other standards of their choice. The standards should be well characterized with regard to their mass stability, in particular under repeated air-to-vacuum transfers.

The participants determined the mass of their travelling standards under vacuum or in air, using their realization experiment, before sending them to the BIPM. The mass values attributed to the standards were calculated using the numerical value of the Planck constant that is now fixed in the SI:

$$h = 6.626\,070\,15 \times 10^{-34}\ \text{J s}.$$

While at the BIPM, the travelling standards were all compared under the same environment as in the participant's laboratory to BIPM working standards which served as reference mass standards. The comparison of the mass standards at the BIPM, together with the mass values attributed by the participants, allowed a determination of the differences between the participants' realizations of the kilogram. The BIPM reference mass standards are traceable to the International Prototype of the Kilogram (IPK) but also effectively maintain the reference values of the KCRVs of the 2019 and 2021 key comparisons.

The participants were asked to verify the mass stability of their travelling standards over the period of the comparison by comparing them before and after the measurements at the BIPM with other, stable mass standards, either in air or in vacuum. These measurements were to be made with the lowest possible uncertainty, in particular by avoiding large air buoyancy corrections.

3 Participants and travelling standards

All NMIs working on realization experiments had been invited to participate in this comparison, under the following conditions:

- standard uncertainty of the mass of the 1 kg travelling standards under vacuum (or in air if measurements under vacuum were not possible) based on the realization experiment < 200 µg, that is 200 parts in 10^9;
- a peer reviewed publication of the realization experiment, including a detailed uncertainty budget for a combined uncertainty close to the one claimed in the comparison, and some evidence of the reproducibility of the results over time;
- availability to perform the required measurements within the schedule of the comparison.

Ten institutes participated (table 1): the BIPM, LNE (France), METAS (Switzerland), NIST (USA), NRC (Canada) and UME (Türkiye) used Kibble balances, the NIM (China) used a joule balance and the CMS/ITRI (Chinese Taipei), NMIJ (Japan) and the PTB (Germany) used the XRCD method as the basis for their mass determinations.

Table 2 lists the travelling standards sent by the participants.

Table 1: Comparison coordinator, support group members and contact persons of the participating institutes.

Comparison coordinator		
Institute	**Contact person**	
BIPM	Michael Stock	
Support group members		
Institute	**Contact person**	
DFM	Lars Nielsen	
NPL	Stuart Davidson	
Participants		
Institute	**Contact person(s)**	**Realization method**
BIPM	Hao Fang	Kibble balance
CMS/ITRI	Yu-Hsin Wu	XRCD method
LNE	Matthieu Thomas	Kibble balance
METAS	Henri Baumann	Kibble balance
NIM	Zhengkun Li Jian Wang	Joule balance
NIST	Darine Haddad	Kibble balance
NMIJ	Naoki Kuramoto	XRCD method
NRC	Richard Green	Kibble balance
PTB	Dorothea Knopf Michael Borys	XRCD method
UME	Haci Ahmedov	Kibble balance

Table 2: Travelling standards and their properties, as communicated by the participants.

Institute	Identification of standard	Manufacturer	Type (dia x ht)/mm	Estimated air-vacuum surface sorption / µg	Magnetic susceptibility	Magnetic polarization / µT
BIPM	691	BIPM	Pt-Ir prototype (39 x 39)	4	24×10^{-5}	< 0.02
CMS/ITRI	H0	Häfner	stainless steel cylinder (54 x 54)	10	< 0.02	< 2.5
	H1	Häfner	stainless steel cylinder (54 x 54)	9	< 0.02	< 2.5
LNE	JM15	Johnson-Matthey and CNAM	Pt-Ir standard (39 x 39)	6.5	24×10^{-5}	0
METAS	WB-1	Häfner	Stainless steel cylindrical (54 x 54.3), E0	8.5		
	WB-3	Häfner	Stainless steel cylindrical (54 x 54.3), E0	weighed in air		
NIM	110	BIPM	Pt-Ir prototype (39 x 39)	unknown	24×10^{-5}	0
	B22	Changzhou Fuyue	stainless steel cylinder (54.2 x 54.2)	5.4	3.4×10^{-4}	< 0.02
NIST	K85	BIPM	Pt-Ir prototype (39 x 39)	11.9	24×10^{-5}	
	K104	BIPM	Pt-Ir prototype (39 x 39)	8.3	24×10^{-5}	
NMIJ	94	BIPM	Pt-Ir prototype (39 x 39)	5.7 (3.3)	3×10^{-4}	< 0.02
	E59	Stanton Instruments	Pt-Ir cylinder (39 x 39)	11.6 (3.3)	2×10^{-4}	< 0.02
NRC	S38	Troemner	Stainless steel cylindrical (54 x 54)	14.3	4.5×10^{-3}	0.02
	H1000W1	Häfner	Tungsten cylinder (38 x 47)	14.5	$< 1 \times 10^{-4}$	<0.1
PTB	109	BIPM	Pt-Ir prototype (39 x 39)	2	< 0.001	< 0.1
	Si14-02	PTB	Si sphere (diam. 93.6)	20	-2.6×10^{-7}	0
UME	2950120	Häfner	Stainless steel cylinder (54 x 54)	weighed in air	1.05×10^{-3}	0.04
	01	Häfner	Stainless steel cylinder (54 x 54)	weighed in air	3.27×10^{-3}	0.03

4 Measurements made by the participants

The following are short summaries of the measurement reports provided by the participants. The complete reports, including the detailed uncertainty budgets, are available as an annex to this report.

4.1 BIPM

The mass of the travelling standard was directly measured under vacuum in the Kibble balance using the fixed numerical value for the Planck constant. The main changes to the BIPM apparatus since the last comparison in 2021 are:

- use of a new 2 V PJVS array operated in a cryocooler system for voltage measurement
- multiple acquisition of interferometer signals rather than one for velocity measurement
- improvement of the detection program for the interferometer beam vertical alignment.

The period of operation of the Kibble balance relevant for the key comparison was 29 November to 10 December 2024.

The detailed traceability chain was that described in 2020 in Metrologia [12].

4.2 CMS/ITRI

CMS/ITRI used the sphere Si28kg03a made of isotopically enriched ^{28}Si as primary mass standard. The core volume, mean molar mass, lattice parameter and mass deficit are traceable to the PTB. CMS/ITRI only measured the surface layer mass using a combined X-ray fluorescence (XRF)/X-ray photoelectron spectroscopy (XPS) surface-analysis system. The reference samples for the XRF analysis were calibrated at the PTB. The surface analysis system of CMS/ITRI is described in [13].

The two travelling standards were calibrated against the Si-sphere in a M_one mass comparator in vacuum.

4.3 LNE

The details of the LNE Kibble balance are described in [14]. The paper describes the operation of the LNE Kibble balance in air. The measurements for the present comparison were made in vacuum.

The determination of mass of the travelling standard JM15 under vacuum was done in three steps. In the first step, a 500 g Pt-Ir mass standard, W1, was calibrated on the LNE Kibble balance. In the second step, the 500 g Pt-Ir mass W2 was calibrated against W1 in a M_one mass comparator. In the last step the travelling standard JM15 was calibrated against W1+W2 in the M_one mass comparator. The three steps were carried out between September 2024 and February 2025.

4.4 METAS

The value given for the mass WB1 is a primary realization with the METAS Kibble balance in vacuum. The measurement campaign lasted from end of October to the end of November 2024. The mass value given for WB3 is based on a comparison with the primary realization in the METAS M_one mass comparator in air in December 2024.

The METAS Kibble balance is described in [15].

4.5 NIM

Since the last key comparison NIM has carried out a series of improvements on the alignment, flux linkage measurement, and force measurement. A vertical laser alignment apparatus has been built to monitor the verticality of the laser in vacuum. A new current source with higher resolution and higher stability than the formerly used one has been built to improve the force measurement and the statistical uncertainty.

For the electrical measurement, a 100 Ω resistor used for the current measurement in the weighing phase was calibrated by the QHR standard of NIM on Oct.12, 2024. This resistor has been calibrated by QHR at least once a year since 2019 and has a very good long and short term stability. A PJVS system is used in situ to calibrate the digital voltmeters in the NIM-2 joule balance for the voltage and flux linkage measurement.

An iodine stabilized laser with vacuum wavelength stability of 1E-11 is used to calibrate the interferometers for the length measurement.

Based on the absolute gravity reference point in the Changping campus of NIM and a value transfer method, the gravitational acceleration has been measured during Feb. 12-14, 2025 by the gravimetry group of NIM and the uncertainty for the gravity measurement has been reevaluated.

Two travelling standards with the numbers 110 and B22 were measured separately in the vacuum chamber of the NIM-2 joule balance from October to December of 2024.

Recent improvements of the NIM-2 Joule balance are described in [16].

4.6 NIST

NIST measured two Pt-Ir masses K85 and K104 directly on the fourth generation Kibble Balance, NIST-4. Each mass was determined under vacuum in NIST-4, and then transferred under vacuum to a vacuum mass comparator to check the stability of masses before and after the mass determination. K85 was measured in NIST-4 from 17 to 27 January 2025. K104 was measured in NIST-4 from 28 January to 3 February 2025. As a result of detailed investigations (described in the measurement report), the dark uncertainty has been reduced from 24.3 ppb in CCM.M-K8.2019 and CCM.M-K8.2021 to 8.8 ppb in the present comparison.

The NIST Kibble balance NIST-4 is described in [17].

4.7 NMIJ

The reproducibility of the realization of the kilogram with the XRCD method at the NMIJ is described in [18, 19]. The NMIJ submitted a very comprehensive description of their application of the XRCD method which can be found in the annex to this report.

For the primary realization of the kilogram, the ^{28}Si-enriched crystal sphere AVO28-S5c was used. The values of its lattice constant, the relative atomic mass of Si and the influence of point defects on the core mass were already determined by the IAC project. There is no known mechanism that changes the values of these properties of Si crystals with respect to time. The mass of the sphere was therefore redetermined from the value of the Planck constant by measuring the volume of the core (by optical interferometry) and the mass of the surface layers (by XPS and ellipsometry). The mass of the travelling standards in vacuum was determined based on the mass of the sphere AVO28-S5c using a mass comparator, on 2 and 3 February 2025.

4.8 NRC

The travelling standards are traceable to two realizations (2024, 2025) using the NRC Kibble Balance and the values aggregated using the methods described by B M Wood et al [20]. The value of the 2024 realization in which the mass of transfer standard AuCu500B (500 g gold plated copper cylinder) was determined, was transferred via PtIr sorption artefacts to be stored in K74, the primary national standard of Canada, whose physical drift rate is known to low uncertainty.

In advance of the 2025 realization the value of the physical drift of AuCu500B was determined with respect to K74 and this drift value was used to link the 2024 and 2025 realizations in order to perform an inverse covariance weighted mean of the two realizations on AuCu500B. The 2024 and 2025 realizations showed AuCu500B changed by 7 µg, which could largely be accounted for by the physical drift of AuCu500B as determined with respect to K74 which was 6.4 µg ± 2.2 µg, demonstrating very good agreement. It should be highlighted that the two realization measurements are completely independent of the absolute value of K74, only dependent upon its physical drift rate.

Due to a broken QHR system the resistance value used for the 2024 realization was based on the long term fit of the resistors in time [21], and verified by comparing to a stable and calibrated 10 kΩ reference resistor via a CCC. The fit provided a resistance uncertainty of approximately 14 ppb, for a combined uncertainty of 17 ppb on the 2024 realization. The QHR system was repaired in advance of the 2025 realization and calibrations of the Kibble Balance resistor were performed before and after the realization measurements with lower uncertainty, the combined uncertainty for the 2025 realization is 12.3 ppb. The main uncorrelated uncertainty components between the 2024 and 2025 realizations lie within the balance alignment, stability of the transfer standard, interferometer laser stability (laser was replaced between realizations), and some components of gravity.

Each realization was performed in a manner similar to those reported by NRC in the 2021 CCM.M-K8 report by M Stock *et al* [11]. Before each realization the AuCu500B was cycled between air and vacuum as described previously [11], additionally it was also transferred back and forth to the KB laboratory once to simulate the route taken for realization, but without performing the realization. In this way the transfer uncertainty has been reduced reproducibly to about 2 ppb between the Mone vacuum balance and the Kibble Balance.

The masses were hand carried to and from the BIPM while contained within hermetically sealed travel enclosures. The stability measurements of the travelling standards were performed in vacuum, with reference to artefacts that remained under vacuum at NRC over the course of comparison.

4.9 PTB

The status of the kilogram realization using the XRCD method at the PTB is described in [22, 23].

The PTB used two spheres made from isotopically enriched ^{28}Si as references: AVO28-S8c from the earlier work of the IAC and the more recently made sphere Si28kg01a. The travelling standards were calibrated against these references in a CCL1007 mass comparator during the period from September to December 2024. The measurement report showed the masses with respect to each of the two references and the weighted mean. The latter is used for the analysis of the present comparison. The results based on the two reference spheres differ by about 10 µg.

Only the volume of the ^{28}Si spheres and the surface layers were measured anew. The measurements of the volumes of the spheres were performed in June/September 2023. The surface layers were measured in January/March 2024. In contrast to [24], the surface layers were measured only by means of XRF/XPS-methods. As a result of an extensive investigation of the XRF system used for the calibration of the XRF reference samples, it was discovered that the fitting procedure of the XRF spectra was sometimes unstable and systematically overestimated the background caused by inelastic scattering. This resulted in scattered results and underestimated the mass depositions of oxygen. The model for the spectral distribution of the XRF spectra in the data analysis, particularly the part related to the spectral background caused by X-ray scattering, was revised. In consequence, the mass of the surface layers on both ^{28}Si spheres determined in the course of the measurements increased by about 14 µg (18 %). One additional microgram has to be considered due to the change of the surface correction of the volume measurements. From an analysis of historical data, it can be concluded that PTB results reported in the CCM.M-K8.2019 and CCM.M-K8.2021 comparisons were impaired in a similar way [23].

4.10 UME

In comparison with KB measurements conducted for the key comparison CCM.M-K8.2021 the following improvements have been made for the ongoing key comparison:

1) The inner frame of the UME KB 3 has been modified by adding three linear motors which produce rotations and horizontal motions of the magnet assembly. The new setup was used to reduce the alignment uncertainty from 40 ppb to 22 ppb. For details see [25].
2) A Cryogenic Current Comparator was used to reduce the uncertainty for the resistance calibration from 26 ppb to 20 ppb.
3) A more stable current source was used to reduce the magnetic flux measurement uncertainty from 16 ppb to 12 ppb.

5 Measurements at the BIPM

The objective of the weighings at the BIPM was to determine the mass differences between the participants' travelling standards and to compare them with BIPM working standards. The masses of the latter are traceable to the mass of the IPK from measurements made in 2014 [26] and maintain the reference values of the previous key comparisons [10,11].

All measurements were made with the M_one mass comparator, equipped with a six-place mass exchanger. The influence of the mass handler position is determined once a year, most recently in September 2024. Any position dependency has always been found to be close to the statistical uncertainty of the comparator. No correction is applied, but an uncertainty of 0.001 mg is taken into account. The sensitivity of the mass comparator is determined at the beginning and at the end of each weighing set using a stainless steel sensitivity weight of 95 mg, which was recalibrated in April 2022 with an uncertainty of 0.0018 mg. Since the masses of all travelling standards except the BIPM standard are within a few mg, the uncertainty contribution of the sensitivity is generally negligible. For standard 691, the mass of which is close to 1 kg – 70 mg, a sensitivity uncertainty of 0.0013 mg is taken into account. The residual pressure during the vacuum weighings was between 5×10^{-4} Pa and 8×10^{-4} Pa.

The masses of the BIPM, CMS-ITRI, LNE, NIM, NIST, NMIJ, NRC, PTB and the mass WB1 of METAS were weighed under vacuum. The mass WB3 of METAS and the masses of UME were weighed in air, as in the UME and METAS laboratories. The mass comparator stayed under vacuum during the whole duration of the vacuum weighings. The cylindrical 1 kg Pt-Ir standard A0 of the BIPM stayed inside the comparator under vacuum during the duration of the vacuum weighings and served as the reference with which all travelling standards were compared. A0 forms together with A18, a 1 kg Pt-Ir stack consisting of eight disks, a set of sorption standards. The surface of A18 is 3.6 times larger than the surface of A0. The change of the mass difference between A18 and A0 when going from air to vacuum or vice versa allows to determine the differences between the mass values in air and in vacuum for both standards. The evolution of the mass difference between A18 and A0 while they were kept in vacuum for about one month allows to determine small changes of the individual masses.

During most of February 2025 both A0 and A18 underwent 7 cycles between air and vacuum, staying each time about one day in each environment. Past experience indicated that this should lead to an improved mass stability during an extended stay in vacuum. On 23 February A0 and A18 were calibrated in air against the BIPM working standards 63 and 77. The mass of A0 was determined as 1 kg + 0.8584 mg and the mass difference of *m*(A18) - *m*(A0) as -0.0498 mg. On 1 March measurement of the mass difference in vacuum resulted in – 0.0506 mg. From this the change of mass of A0 related to the air-to-vacuum transfer can be calculated as -0.0003 mg. The mass of A0 in vacuum was therefore 1 kg + 0.8581 mg. This value was used in the least-squares adjustment to determine the masses of the travelling standards based on the BIPM mass scale.

During the period of the vacuum weighings, about one month, *m*(A18) - *m*(A0) changed very little, indicating that the mass of A0 had increased by 0.0009 mg and that of A18 by 0.0034 mg. After completion of the vacuum weighings, A0 and A18 were brought again to air. The mass difference *m*(A18) - *m*(A0) changed from -0.0482 mg in vacuum to -0.0361 mg in air, indicating a mass increase of A0 of 0.0047 mg. Adding to the initial vacuum mass of A0 (1 kg + 0.8581 mg) the mass increases during vacuum (+0.0009 mg) and related to the transfer to air (+0.0047 mg) leads to the expected mass value in air of 1 kg + 0.8637 mg. This value was confirmed within less than 1 µg by a new comparison in air with working standards 63 and 77.

Seven weighing sets were carried out under vacuum, two sets in air.

- Weighings in vacuum

Set 1 (01-02/03/2025):	A0	A18	H1 (CMS)	H0 (CMS)	B22 (NIM)	110 (NIM)
Set 2 (05-06/03/2025):	A0	A18	K85(NIST)	K104 (NIST)	B22 (NIM)	110 (NIM)
Set 3 (08-09/03/2025):	A0	A18	E59 (NMIJ)	94 (NMIJ)	B22 (NIM)	110 (NIM)
Set 4 (16-17/03/2025):	A0	A18	E59 (NMIJ)	94 (NMIJ)	109 (PTB)	Si14-02 (PTB)
Set 5 (20-21/03/2025):	A0	A18	H1000W1 (NRC)	S38 (NRC)	109 (PTB)	Si14-02 (PTB)
Set 6 (27-28/03/2025):	A0	A18	H1000W1 (NRC)	S38 (NRC)	691 (BIPM)	WB1 (METAS)
Set 7 (02-03/04/2025):	A0	A18	H1000W1 (NRC)	JM15 (LNE)	691 (BIPM)	WB1 (METAS)

- Weighings in air

Set 8 (05-06/04/2025):	A0	A18	01 (UME)	2950120 (UME)	Cp2	Cc2
Set 9 (09-10/04/2025):	A0	A18	01 (UME)	WB3 (METAS)	Cp2	Cc2

The 1 kg masses Cp2 and Cc2, used during the weighings in air, have different volumes and form a set of standards for the determination of the air density used to calculate buoyancy corrections.

Standards which were included in two consecutive weighing sets stayed under vacuum in-between. An exception was the Si sphere Si14-02 which was cleaned before each weighing, as requested by the PTB, by applying the recommended procedure [27] and using the cleaning products provided by the PTB. None of the other mass standards were cleaned, only dust was removed with a soft brush and by gently blowing some air.

For each set, four full weighing schemes were carried out. In each vacuum weighing scheme all 15 pairwise mass-differences between the standards were determined. In the weighing schemes in air, six pairwise mass-differences between the standards were determined. This led to 60 mass differences for each of the vacuum sets and to 24 mass differences for the air sets and to 468 mass differences in total. Each individual mass difference and its uncertainty were obtained from an A-B-A-B-A-B-A scheme.

All 468 mass differences were used to carry out a generalized least-squares adjustment to obtain the masses of the travelling standards. The weights of the individual weighing results were taken as the inverse variances. The initial vacuum mass of A0 served as the constraint in the adjustment. The small increase of mass during its stay in vacuum (+0.0009 mg) and the mass increase after the transition to air (+0.0047 mg) were taken into account in a drift correction. The statistical uncertainty of the adjusted masses was at the level of 0.0005 mg.

The uncertainty budget for the vacuum weighings is shown in table 3. The table shows the combined standard uncertainties of the masses of the travelling standards with respect to two different references:

- with respect to the mass of A0 under vacuum: 0.0015 mg
- with respect to the Planck constant (via the IPK): 0.012 mg

The uncertainty of the first part of the table, 0.0015 mg, representing the uncertainty of the mass of the travelling standards weighed under vacuum with respect to the mass of A0 in vacuum, is considered to be uncorrelated for all travelling standards.

For the uncertainty of the masses of the UME standards and of WB3 from METAS, weighed in air, with respect to the *vacuum* mass of A0 (which served as the reference for all other travelling standards) we consider in addition the uncertainty related to the vacuum-to-air transfer of A0 (0.001 mg) and to the buoyancy correction (0.0026 mg) between Pt-Ir and stainless steel masses, leading to a combined uncertainty of 0.0032 mg. As already mentioned above, for standard 691, the mass of which is close to 1 kg – 70 mg, an uncertainty of 0.0013 mg for the sensitivity of the mass comparator is taken into account.

The uncertainty contributions of the second part of the table, representing the uncertainty of the vacuum mass of A0 with respect to the Planck constant, are completely correlated for all travelling standards, and have no influence on the comparison of the realization experiments.

Table 3: Uncertainty budget for the mass of the travelling standards under vacuum with respect to the vacuum mass of the reference A0 and with respect to the Planck constant (via the IPK). The uncertainties of the second part of the table are totally correlated for all travelling standards.

Source of uncertainty	Uncertainty	Sensitivity coefficient	Unc. contribution / mg
position error of M_one	0.001 mg	1	0.001
centre of gravity of standards	0.5 mm	0.0003 mg/mm	0.0002
statistical uncertainty	0.0005 mg	1	0.0005
mass stability of A0 during vacuum weighings	0.001 mg	1	0.001
Standard uncertainty of mass of travelling standards with respect to mass of A0 in vacuum			**0.0015**
air to vacuum transfer of A0	0.001 mg	1	0.001
mass of A0 in air with respect to IPK	0.006 mg	1	0.006
mass of IPK with respect to Planck constant	0.010 mg	1	0.010
Standard uncertainty of mass of travelling standards with respect to Planck constant			**0.012**

6 Results of the comparison

6.1 Mass stability of the travelling standards

Each participant determined the masses of the travelling standards against another mass standard or a set of mass standards before and after the weighings at the BIPM, to verify the mass stability:

$m_{\mathrm{before},i,j}^{\mathrm{NMI}}$: The mass of the travelling standard *j* (*j*=1,2) determined by participant *i* (*i*=1,..,10) against a stable reference mass before sending it to the BIPM.

$m_{\mathrm{after},i,j}^{\mathrm{NMI}}$: The mass of the travelling standard *j* (*j*=1,2) determined by participant *i* (*i*=1,..,10) against the same stable reference mass after the return from the BIPM.

Some participants made these measurements in air (LNE, METAS, NIM, NIST, NMIJ, PTB), others in vacuum (BIPM, NRC, UME). Only the Si-sphere Si14-02 was cleaned before these weighings at the PTB. The observed mass difference $\Delta m_{i,j}^{\mathrm{NMI}} = m_{\mathrm{after},i,j}^{\mathrm{NMI}} - m_{\mathrm{before},i,j}^{\mathrm{NMI}}$ and its uncertainty $u(\Delta m_{i,j}^{\mathrm{NMI}})$ were communicated to the BIPM. Table 4 and Figure 1 show the results of these measurements.

The mean value of the observed mass changes is -1.6 µg with a standard deviation of 7 µg. Eleven of the 18 standards show changes within ±5 µg, 15 lie within ±10 µg. In general, the uncertainty related to the mass changes is negligible compared to the realization uncertainties.

Table 4: Changes of the mass of the travelling standards during the comparison $\Delta m_{i,j}^{\mathrm{NMI}}$, and related uncertainty $u(\Delta m_{i,j}^{\mathrm{NMI}})$, as determined by the participants. Also shown are the corrections $\delta m_{\mathrm{stab},i,j}^{NMI}$ for the mass changes, applied by the BIPM to the masses of the travelling standards determined with the NMIs realization experiments, and their uncertainties $u_{\mathrm{stab},ij}^{\mathrm{NMI}}$.

NMI	mass standard	mass change (1) $\Delta m_{i,j}^{\mathrm{NMI}}$/ mg	uncertainty of mass change (1) $u(\Delta m_{i,j}^{\mathrm{NMI}})$/ mg	correction for mass change $\delta m_{\mathrm{stab},i,j}^{\mathrm{NMI}}$ / mg	uncertainty of correction $u_{\mathrm{stab},i,j}^{\mathrm{NMI}}$ / mg
BIPM	691	+0.001	0.0019	+0.0005	0.0019
CMS/ITRI	H0	-0.0051	0.0025	-0.0026	0.0029
	H1	-0.0146	0.0025	-0.0073	0.0049
LNE	JM15	+0.0032	0.0050	+0.0016	0.0051
METAS	WB 1	-0.007	0.010	-0.0035	0.0102
	WB3	+0.001	0.010	+0.0005	0.0100
NIM	110	-0.008	0.0048	-0.0040	0.0053
	B22	-0.015	0.012	-0.0075	0.0128
NIST	K85	-0.0031	0.0030	-0.0016	0.0031
	K104	-0.0012	0.0025	-0.0006	0.0025
NMIJ	94	+0.0027	0.0026	+0.0014	0.0027
	E59	+0.0089	0.0026	+0.0044(2)	0.0037
NRC	H1000W1	+0.0013	0.0010	+0.0006	0.0011
	S38	+0.0010	0.0012	+0.0005	0.0012
PTB	109	+0.0029	0.0019	+0.0014(2)	0.0021
	Si14-02	-0.0031	0.0048	-0.0016	0.0049
UME	01	-0.005	0.009	-0.0025	0.0091
	2950120	+0.011	0.009	+0.0055	0.0095

(1) Number of decimals as given by the participants.
(2) Rounded to the nearest even number.

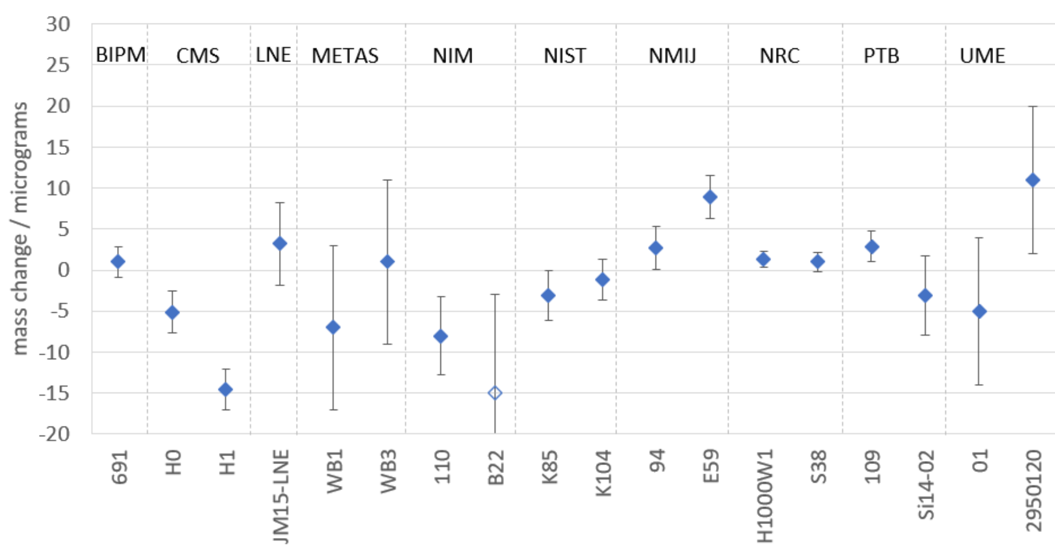

Fig. 1: Changes of the mass of the travelling standards during the comparison, $\Delta m_{ij}^{\mathrm{NMI}}$, as determined by the participants. The standard B22 shown with open symbol was discarded from the comparison (see 6.2).

The masses of the travelling standards determined by the participants using their realization experiments were corrected by the BIPM for the observed mass changes, to represent the masses of the standards while at the BIPM. In the absence of detailed knowledge about when and how the mass changes occurred (before or/and after the comparison at the BIPM, stepwise or continuously, …), the best estimate for the correction is half of the mass change observed by the participants:

$\delta m_{\mathrm{stab},i,j}^{\mathrm{NMI}} = \frac{1}{2}\Delta m_{i,j}^{\mathrm{NMI}}$: correction to be added to the mass of travelling standard *j* (*j*=1,2) determined by participant *i* (*i*=1,..,10), to predict the mass of the standard while at the BIPM.

The uncertainty of the stability correction $u_{\mathrm{stab},i,j}^{\mathrm{NMI}}$ has two components. The first is the uncertainty of the determination of the mass change made by the participant $u\left(\Delta m_{i,j}^{\mathrm{NMI}}\right)$, shown in the fourth column of table 4. The second component is related to the choice of the value chosen for the correction. Assuming that the mass while at the BIPM lies between the two NMI stability measurements, a rectangular probability distribution, centered at half of the observed change (used for the correction), and with width of half the observed change can be assumed. Adding both components quadratically leads to:

$$\left(u_{\mathrm{stab},i,j}^{\mathrm{NMI}}\right)^2 = u\left(\Delta m_{i,j}^{\mathrm{NMI}}\right)^2 + \left(\frac{1}{2\sqrt{3}}\Delta m_{i,j}^{\mathrm{NMI}}\right)^2$$

The corrections for mass stability and their uncertainties are shown in the two rightmost columns of table 4.

6.2 Results of the mass determinations by the participants and by the BIPM

The results of the mass determinations by the participants using their realization experiments and the BIPM, based on the mass unit maintained on the BIPM working standards, are shown in Table 5. The columns have the following meaning:

$m_{i,j}^{\mathrm{NMI}}$: the mass of the travelling standard *j* (*j*=1,2) determined by participant *i* (*i* = 1,..,10) using its realization experiment, presented as the deviation from 1 kg.

$u_{i,j}^{\mathrm{NMI}}$: the standard uncertainty associated with the mass value $m_{i,j}^{\mathrm{NMI}}$, as communicated by the NMI.

$m_{\mathrm{corr},i,j}^{\mathrm{NMI}}$: the predicted mass of the travelling standard *j* of participant *i* while at the BIPM, obtained by applying the correction for mass change (table 4) to $m_{i,j}^{\mathrm{NMI}}$, presented as the deviation from 1 kg:

$$m_{\mathrm{corr},i,j}^{\mathrm{NMI}} = m_{i,j}^{\mathrm{NMI}} + \delta m_{\mathrm{stab},ij}^{\mathrm{NMI}}$$

$u_{\mathrm{total},i,j}^{\mathrm{NMI}}$: uncertainty communicated by the participant, increased by the contribution for the mass stability correction (table 4):

$$\left(u_{\mathrm{total},i,j}^{\mathrm{NMI}}\right)^2 = \left(u_{i,j}^{\mathrm{NMI}}\right)^2 + \left(u_{\mathrm{stab},i,j}^{\mathrm{NMI}}\right)^2$$

For the sphere Si14-02 an additional uncertainty of 0.002 mg for the potentially different cleaning efficiency at the PTB and the BIPM was included.

$m_{i,j}^{\mathrm{BIPM}}$: The mass value obtained by the BIPM for travelling standard *j* of NMI *i*, based on the mass unit maintained on the BIPM working standard (traceable to the IPK), presented as deviation from 1 kg.

$u_{i,j}^{\mathrm{BIPM}}$: The standard uncertainty associated with the BIPM results. Only the uncertainty components which are uncorrelated amongst the travelling standards are included (see table 3) and related text.

$\Delta m_{i,j}$: The difference between the mass values obtained by the NMI, corrected for the mass change, and the BIPM:

$$\Delta m_{i,j} = m_{\mathrm{corr},i,j}^{\mathrm{NMI}} - m_{i,j}^{\mathrm{BIPM}}$$

$u_{i,j}$: The standard uncertainty associated with the mass difference:

$$u_{i,j}^2 = \left(u_{i,j}^{\mathrm{BIPM}}\right)^2 + \left(u_{\mathrm{total},i,j}^{\mathrm{NMI}}\right)^2$$

r_j: correlation coefficient between the results obtained for two standards of an NMI, calculated from the detailed uncertainty budgets submitted by the participants, after including $u_{\mathrm{stab},i,j}^{\mathrm{NMI}}$, $u_{i,j}^{\mathrm{BIPM}}$ and the additional uncertainty of 0.002 mg for sphere Si14-02. The correlation coefficient is required for the calculation of the mean value of the results for both standards.

Table 5: Masses (deviations from 1 kg) and standard uncertainties of the travelling standards determined by the NMIs using their realization experiments and by the BIPM using its working standards (traceable to the IPK). On the right, the mass differences $\Delta m_{i,j} = m^{\mathrm{NMI}}_{\mathrm{corr},i,j}$ - $m^{\mathrm{BIPM}}_{i,j}$ are shown.

Institute	Identification of standard	NMI realization results				BIPM results		Mass difference, NMI-BIPM		
		$m^{\mathrm{NMI}}_{i,j}$ / mg	$u^{\mathrm{NMI}}_{i,j}$/ mg	$m^{\mathrm{NMI}}_{\mathrm{corr},i,j}$/ mg	$u^{\mathrm{NMI}}_{\mathrm{total},i,j}$ / mg	$m^{\mathrm{BIPM}}_{i,j}$ / mg	$u^{\mathrm{BIPM}}_{i,j}$ / mg	$\Delta m_{i,j}$ / mg	$u_{i,j}$ / mg	corr. coeff. r_i
BIPM[(1)]	691	-70.211	0.036	-70.2105	0.0361	-70.1927	0.0020	-0.0178	0.0361	NA
CMS/ITRI	H0	-0.053	0.037	-0.0556	0.0371	-0.0607	0.0015	0.0052	0.0371	0.99
	H1	-0.058	0.037	-0.0653	0.0373	-0.0664	0.0015	0.0011	0.0374	
LNE	JM15	-0.788	0.0364	-0.7864	0.0368	-0.7856	0.0015	-0.0008	0.0368	NA
METAS	WB 1	-0.0144	0.0533	-0.0179	0.0543	0.0294	0.0015	-0.0473	0.0543	0.90
	WB 3	-0.1434	0.0569	-0.1429	0.0578	-0.0739	0.0032	-0.0690	0.0579	
NIM	110	0.025	0.0288	0.0210	0.0293	0.0532	0.0015	-0.0322	0.0293	0.73
	B22	-0.128	0.0338	-0.1355	0.0361	-0.2984	0.0015	0.1629	0.0362	
NIST	K85	-0.7796	0.0138	-0.7812	0.0142	-0.7662	0.0015	-0.0149	0.0142	0.88
	K104	0.398	0.0133	0.3974	0.0135	0.4124	0.0015	-0.0150	0.0136	
NMIJ	94	0.3167	0.0210	0.3181	0.0212	0.3394	0.0015	-0.0214	0.0212	0.95
	E59	4.9017	0.0210	4.9062	0.0213	4.9277	0.0015	-0.0216	0.0214	
NRC	H1000W1	-7.0043	0.0119	-7.0037	0.0119	-7.0149	0.0015	0.0112	0.0120	0.93
	S38	-0.1535	0.0118	-0.1530	0.0119	-0.1611	0.0015	0.0081	0.0120	
PTB	109	0.1714	0.0135	0.1729	0.0137	0.1885	0.0015	-0.0157	0.0137	0.88
	Si14-02	-4.2302	0.0136	-4.2318	0.0146[(2)]	-4.2016	0.0025	-0.0302	0.0148	
UME	01	-0.351	0.036	-0.3535	0.0371	-0.3271	0.0032	-0.0264	0.0373	0.94
	2950120	0.039	0.036	0.0445	0.0372	0.0908	0.0032	-0.0463	0.0374	

(1) For the BIPM, the "NMI result" is the mass of standard 691 obtained with the BIPM Kibble balance, and the "BIPM result" is its mass based on the BIPM working standards.
(2) For the sphere Si14-02 an additional uncertainty of 0.002 mg for the reproducibility of the washing between the PTB and the BIPM is included.

The differences $\Delta m_{i,j}$ between the mass values attributed by the NMIs, corrected for the mass changes, and by the BIPM and their uncertainties $u_{i,j}$ are shown in Figure 2.

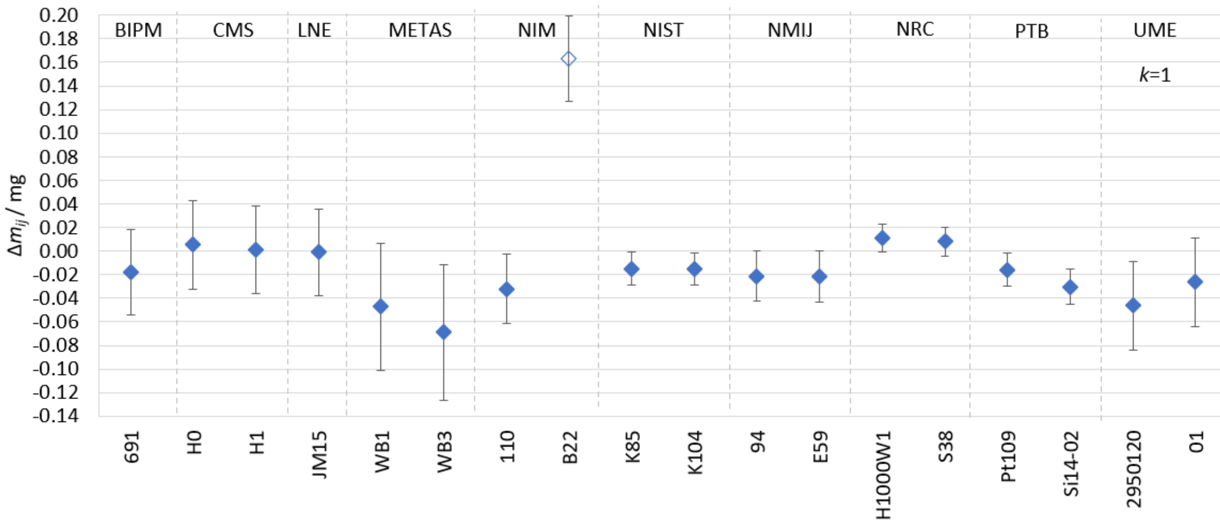

Fig. 2: Differences $\Delta m_{i,j} = m^{\mathrm{NMI}}_{\mathrm{corr},i,j}$ - $m^{\mathrm{BIPM}}_{i,j}$ between the mass determinations of each travelling standard with the NMI's realization experiment and the BIPM working standards and associated standard uncertainties $u_{i,j}$. The standard B22 shown with open symbol was discarded from the comparison.

At the present stage of the data analysis, the mass in vacuum of BIPM standard A0 serves only as a common reference for the comparison, with respect to which the masses of all travelling standards are determined. Therefore the BIPM comparison uncertainty $u^{\mathrm{BIPM}}_{i,j}$ includes only uncertainty components related to the weighings against this reference, but not components related to the absolute mass value of A0. For the travelling standards weighed in vacuum, the weighing uncertainty of 0.0015 mg shown in table 5 includes only the uncertainty components of the first part of the uncertainty budget in table 3, which are uncorrelated between the travelling standards. For the three standards measured in air, the uncertainty of the vacuum-air transfer of A0 (0.001 mg) and of the buoyancy correction (0.0026 mg) are also included, resulting in 0.0032 mg. For standard 691, the mass of which is close to 1 kg – 70 mg, an uncertainty of 0.0013 mg for the sensitivity of the mass comparator is taken into account. Based on experience of the PTB, an uncertainty of 0.002 mg for the repeatability of the cleaning of the sphere before each weighing at the BIPM was included in the weighing uncertainty of the Si-sphere Si14-02 (in addition to an uncertainty of 0.002 mg for an eventual systematic difference between cleaning of the sphere at the PTB and the BIPM, mentioned in footnote 2 of table 5). The BIPM comparison uncertainty does not include the remaining components of table 3 which are related to the uncertainty of the vacuum mass of A0 with respect to the Planck constant. These components are totally correlated for the travelling standards and can be omitted here, because for the purpose of the comparison of the realization experiments with each other, the BIPM mass unit has to be stable, but not accurate.

The result for B22 of NIM is a clear outlier. It deviates from the result for n° 110 by 195 µg. NIM was approached to verify the communicated results for their standards but they did not find any anomaly. Their measurements for stability of B22 before and after BIPM resulted in a mass change of -15 µg with an uncertainty of 12 µg. In discussion with NIM we could not identify a clear reason for this anomaly.

Hypothetically, the mass of B22 changed by about -195 µg between the weighing in the Joule balance and the initial stability check. It was agreed with NIM to withdraw B22 from the calculation of the comparison results.

Ideally, for NMI *i* participating with two travelling standards, the differences between the NMI and the BIPM results for both standards, $\Delta m_{i,1}$ and $\Delta m_{i,2}$, should be the same. In practice they can differ due to the uncertainty contributions which are uncorrelated between the two results. Uncertainty contributions which are correlated between the results for the two standards have no influence on the consistency of the results. It was verified that for each NMI *i* which participated with two standards, the two results $\Delta m_{i,1}$ and $\Delta m_{i,2}$ agreed within the expanded uncertainty (*k*=2) of $\Delta m_{i,1}$ - $\Delta m_{i,2}$. This uncertainty was calculated from those components of the uncertainty budgets which are uncorrelated between the two standards.

It should be noted that the mass difference between the travelling standards sent by the PTB measured at the BIPM differs from the mass difference measured at PTB by 0.0115 mg. After applying the corrections for mass changes this leads to the difference of 0.0145 mg between the results for 109 and Si14-02, as shown in Fig. 2 and Table 5. The expanded uncertainty of this difference (calculated from the uncorrelated uncertainty components) is of about the same size. Similar differences were observed during the two previous comparisons (0.024 mg in 2019 and 0.014 mg in 2021), which indicates that this is a systematic effect. Attempts made in 2023 and 2024 to discover the origin of this systematic difference did not lead to a clear conclusion. The difference between mass values assigned by the PTB to any of the two travelling standards (Pt109 or Si14-02) by the two realizations (AVO28-S8c and Si28kg01a sphere) amounts to 0.010 mg.

To obtain one single result, representative for each NMI, the results for the two travelling standards were averaged. The correlation coefficients r_i between the results of the two standards, shown in table 5, were taken into account. The weighted mean Δm_i of the two results $\Delta m_{i,1}$ and $\Delta m_{i,2}$ with uncertainties $u_{i,1}$ and $u_{i,2}$ and covariance $u_{i,12} = r_i u_{i,1} u_{i,2}$ and its variance were calculated as [28]:

$$\Delta m_i = \frac{(u_{i,1}^2 - u_{i,12})(u_{i,2}^2 - u_{i,12})}{u_{i,1}^2 + u_{i,2}^2 - 2u_{i,12}} \left(\frac{\Delta m_{i,1}}{u_{i,1}^2 - u_{i,12}} + \frac{\Delta m_{i,2}}{u_{i,2}^2 - u_{i,12}} \right) \qquad \text{(eq. 1)}$$

$$\mathrm{var}\,(\Delta m_i) = \frac{u_{i,1}^2 u_{i,2}^2 - u_{i,12}^2}{u_{i,1}^2 + u_{i,2}^2 - 2u_{i,12}} \qquad \text{(eq. 2)}$$

This corresponds to the weighted mean in which only the uncorrelated uncertainty is used in the determination of the weights. For totally uncorrelated results, the equations become identical to those for the normal weighted mean. Table 6 shows the averaged results Δm_i for all participants and figure 3 shows the results in graphical form.

Table 6: Differences Δm_i between mass values attributed to 1 kg mass standards using the realization experiment of the participants and by the working standards of the BIPM, and associated standard uncertainty.

Institute	$\boldsymbol{\Delta m_i}$ **/ mg**	$\boldsymbol{u(\Delta m_i)}$ **/ mg**
BIPM	-0.0178	0.0361
CMS/ITRI	0.0040	0.0371
LNE	-0.0008	0.0368
METAS	-0.0511	0.0541
NIM	-0.0322	0.0293
NIST	-0.0150	0.0135
NMIJ	-0.0215	0.0210
NRC	0.0095	0.0118
PTB	-0.0185	0.0137
UME	-0.0359	0.0367

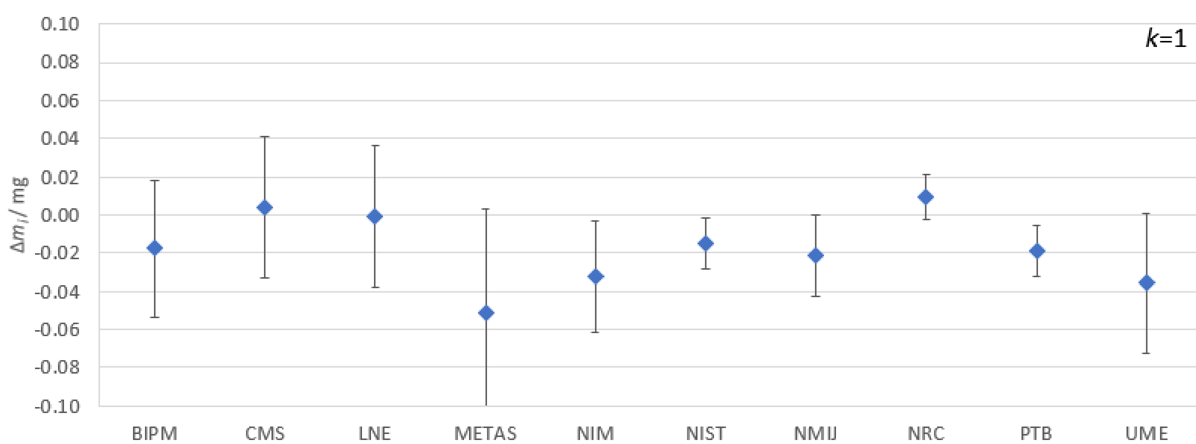

Fig. 3: Differences Δm_i between mass values attributed to 1 kg mass standards using the realization experiment of the participants and by the working standards of the BIPM, and associated standard uncertainty.

6.3 Calculation of the Key Comparison Reference Value

Following the rules in [29], the result of CMS/ITRI is not included in the calculation because Chinese Taipei is not a member of the BIPM, but an Associate of the CGPM. The key comparison reference value (KCRV) is calculated as the inverse-variance-weighted mean of the Δm_i of the realization experiments:

$$\overline{\Delta m} = \frac{\sum_i w_i \Delta m_i}{\sum_i w_i} \quad \text{with} \quad w_i = \frac{1}{u(\Delta m_i)^2} \quad \text{and} \qquad \text{(eq. 3)}$$

$$u(\overline{\Delta m}) = \sqrt{(\sum w_i)^{-1}} \qquad \text{(eq. 4)}$$

The result is $\overline{\Delta m}$ = -0.0107 mg with a standard uncertainty of 0.0064 mg. The largest statistical weights, between 20 % and 30 % are attributed to the NRC, NIST and PTB results. This value of the KCRV (with respect to the mass unit maintained on the BIPM working standards) will be the input of this comparison to the calculation of the third consensus value.

In the last step of the analysis we calculate the deviation of the participants' results from the key comparison reference value $\overline{\Delta m}$:

$$\Delta m_i' = \Delta m_i - \overline{\Delta m} \qquad \text{(eq.5)}$$

For the participants which contributed to the calculation of the KCRV, the uncertainty is calculated as

$$u(\Delta m_i') = \sqrt{u(\Delta m_i)^2 - u(\overline{\Delta m})^2} \qquad \text{(eq.6)}$$

For CMS/ITRI, the result of which does not contribute to the KCRV, the uncertainty is calculated as

$$u(\Delta m_{CMS}') = \sqrt{u(\Delta m_{CMS})^2 + u(\overline{\Delta m})^2} \qquad \text{(eq. 7)}$$

These deviations and their uncertainties are shown in table 7 and figures 4 and 5. The BIPM working standards are traceable to the Planck constant through its known relationship with the IPK. The difference between mass values based on the BIPM as-maintained mass unit and the KCRV can be determined as:

$$\Delta m_{h(IPK)}' = \Delta m_{h(IPK)} - \overline{\Delta m} = -\overline{\Delta m} \qquad \text{(eq.7)}$$

$\Delta m_{h(IPK)} = 0$ because the BIPM working standards (traceable to the IPK) served as the reference for the comparison. Since the BIPM as-maintained mass unit was not used in the calculation of the KCRV, the uncertainty is calculated as:

$$u\left(\Delta m_{h(IPK)}'\right) = \sqrt{u\left(\Delta m_{h(IPK)}\right)^2 + u(\overline{\Delta m})^2} \qquad \text{(eq.8)}$$

The uncertainty $u\left(\Delta m_{h(IPK)}\right)$ includes all uncertainty components of table 3 and is dominated by the uncertainty of the mass of the IPK in terms of the Planck constant, 0.010 mg.

The difference between the results of NRC and PTB is significantly reduced, compared to the last comparison. In CCM.M-K8.2021 the difference was 0.050 mg with a combined uncertainty of 0.018 mg. In the present comparison the difference is 0.028 mg with a combined uncertainty of 0.018 mg.

Table 7: Deviations $\Delta m_i'$ of the NMIs' results from the KCRV, related standard uncertainties $u(\Delta m_i')$ and expanded uncertainties for *k* = 2, $U(\Delta m_i')$. The difference between mass values based on the BIPM working standards, traceable to the Planck constant through the IPK, and those based on the reference value is also shown.

Institute	Deviation from KCRV $\Delta m_i'$ / mg	$u(\Delta m_i')$ / mg	$U(\Delta m_i')$ / mg
BIPM	-0.0070	0.0355	0.0711
CMS/ITRI	0.0147	0.0377	0.0753
LNE	0.0099	0.0362	0.0724
METAS	-0.0404	0.0537	0.1075
NIM	-0.0215	0.0286	0.0572
NIST	-0.0043	0.0118	0.0236
NMIJ	-0.0107	0.0200	0.0401
NRC	0.0202	0.0099	0.0198
PTB	-0.0078	0.0121	0.0241
UME	-0.0252	0.0361	0.0723
BIPM (*h*(IPK))	0.0107	0.0136	0.0272

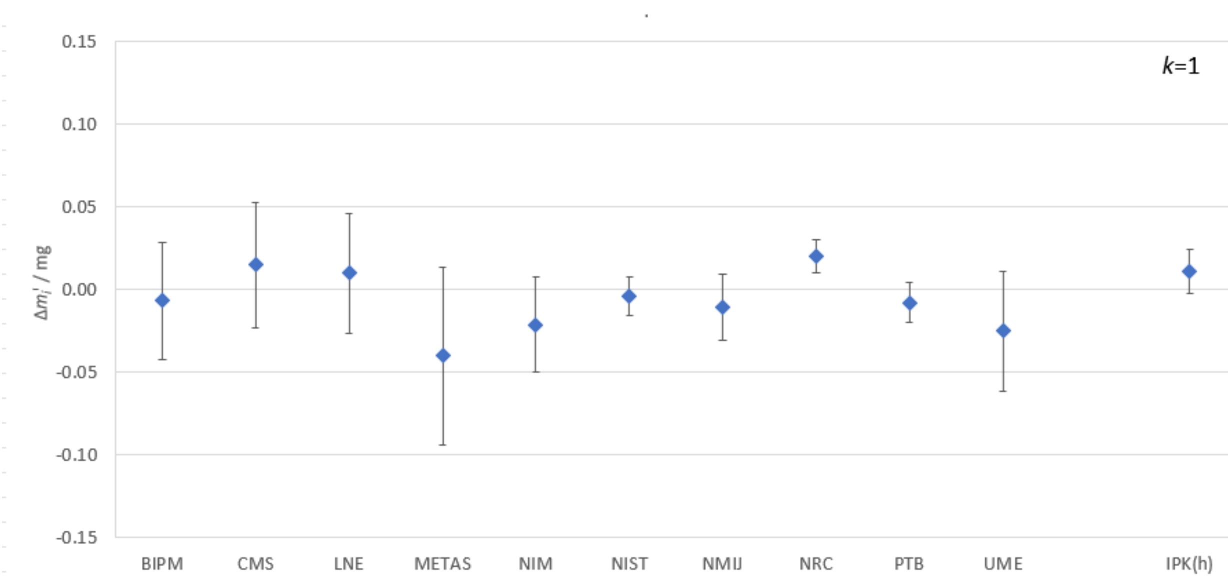

Fig 4: Differences $\Delta m_i'$ between mass values attributed to 1 kg mass standards using the realization experiment of the participants and the KCRV, and associated standard uncertainty. The difference between mass values based on the BIPM working standards, traceable to the Planck constant through the IPK, and those based on the reference value is also indicated.

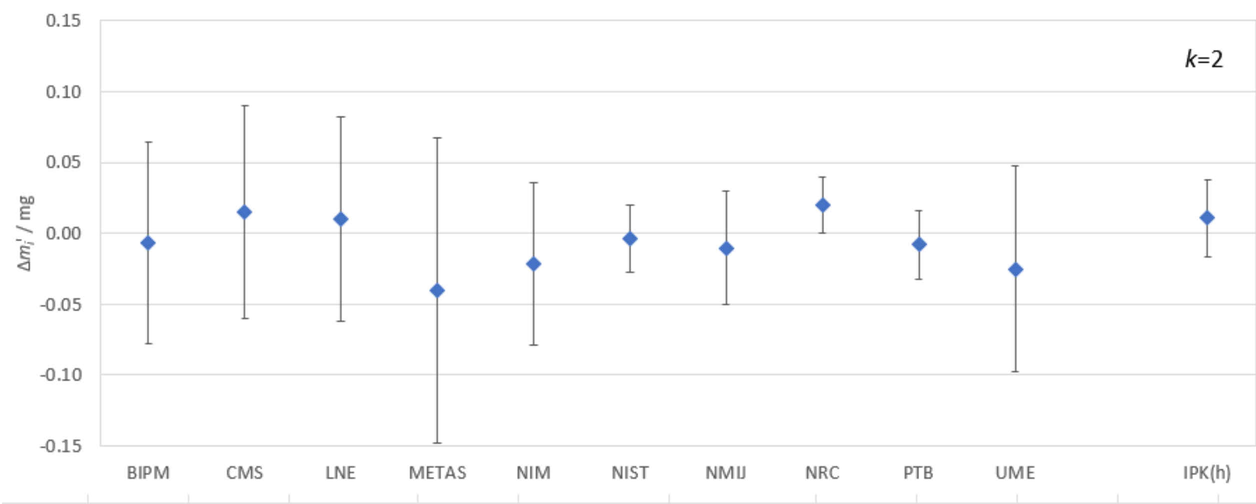

Fig 5: Same as figure 4, with uncertainty bars showing the expanded standard uncertainty for *k*=2.

To test the consistency of the data set, the chi-squared value of the data set was calculated from the data in table 7 according to

$$\chi_{obs}^2 = \sum_{i=1}^{10} \frac{\Delta m_i'^2}{u^2(\Delta m_i)} \qquad \text{(eq. 9)}$$

The experimental chi-squared value is 5.5. The chi-squared value for nine degrees of freedom at the 95 % cut-off of the distribution is 16.9. A more conservative criterion is to require that the observed chi-squared lies within the expectation value of the chi-squared distribution plus its standard deviation. For *ν* degrees of freedom this value is $\nu + \sqrt{2\,\nu}$, for 9 degrees of freedom one obtains 13.2. Both tests are passed.

6.4 Comparison with the results of previous comparisons

The CCM has established a list of criteria for a decision on the transition from the dissemination from the consensus value to dissemination from local, independent realizations [8]. One of these criteria requires that at least five realization experiments demonstrate stability by producing consistent results for two consecutive key comparisons. It is therefore interesting to compare the results obtained in the present comparison with those obtained in the Pilot Study of 2016, CCM.M-K8.2019 and in CCM.M-K8.2021. Such a comparison is made possible by the quantified stability of the BIPM mass unit, as maintained by the BIPM Pt-Ir standards, and the link of this mass unit to the different comparison reference values.

Figure 6 shows the differences Δm_i between mass values attributed to 1 kg mass standards using the realization experiment of the participants who participated in more than one comparison and by the working standards of the BIPM, for the four comparisons. The results of the pilot study were corrected by +0.013 mg to account for the difference between the value of the Planck constant used for this study and the value fixed in 2019 for the new definition of the kilogram, and for a retrospective correction of the mass values of the BIPM working standards.

The figure shows that for all realization experiments the results of the three key comparisons are consistent within the uncertainties.

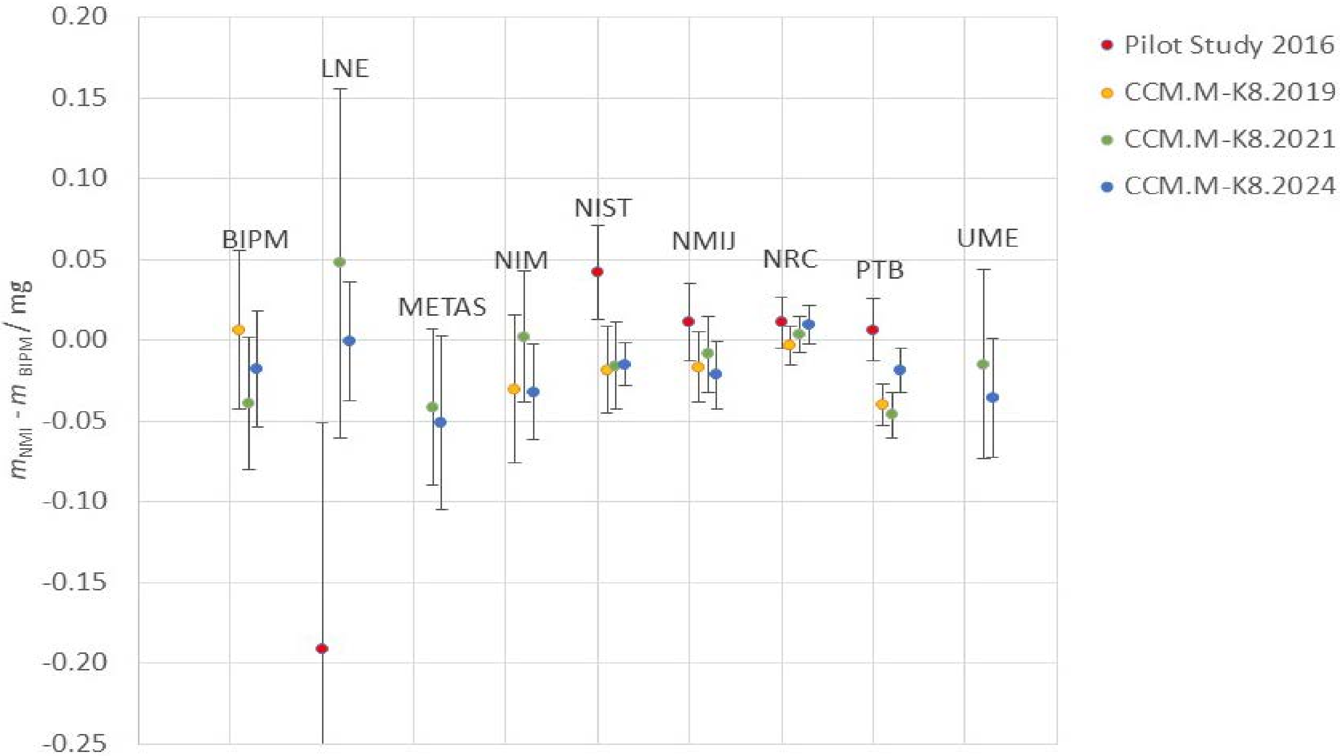

Fig. 6: Differences Δm_i between mass values attributed to 1 kg mass standards using the realization experiment of the participants who participated in more than one comparison and by the working standards of the BIPM (traceable to the IPK), for four comparisons and associated standard uncertainties (*k*=1).

7 Summary

This key comparison had the objective of comparing realizations of the kilogram based on six Kibble balances, a joule balance and three applications of the XRCD method. The participants determined the mass of one or two 1 kg mass standards traceable to their primary methods and sent them to the BIPM for comparison. At the BIPM all these travelling standards were compared with each other using a BIPM mass standard as the reference.

Four types of travelling standards were used: 8 Pt-Ir standards, 8 stainless steel standards, a tungsten weight and a Si-sphere. In general, the uncertainty related to the stability of the travelling standards was negligible in comparison to the realization uncertainty. In most cases the observed mass changes were within or close to the uncertainty of measurement of this change.

The key comparison reference value was calculated as the weighted mean of the results. It has a deviation of -0.0107 mg with respect to the mass unit maintained by the BIPM working standards, with a standard uncertainty of 0.0064 mg. This reference value will be the contribution of the present comparison to the calculation of the next CCM consensus value. The chi-squared test for consistency using the 95 % cut-off criterion is passed. The difference between the results of NRC and PTB is significantly reduced compared to the last comparison.

References

[1] M. Stock, R. Davis, E. de Mirandés, M. Milton, "The revision of the SI – the result of three decades of progress in metrology", *Metrologia* **56** (2019) 022001 and Corrigendum in *Metrologia* **56** (2019) 049502, https://doi.org/10.1088/1681-7575/ab0013 and https://doi.org/10.1088/1681-7575/ab28a8

[2] 9th edition of the SI Brochure, available on the BIPM web site: www.bipm.org, https://doi.org/10.59161/AUEZ1291

[3] *Mise en pratique* for the definition of the kilogram in the SI, available in Appendix 2 of the 9th edition of the SI brochure on the BIPM web site: www.bipm.org

[4] I. A. Robinson, S. Schlamminger, "The watt or Kibble balance: a technique for implementing the new SI definition of the unit of mass", *Metrologia* **53** (2016) A46, http://dx.doi.org/10.1088/0026-1394/53/5/A46

[5] Z. Li et al., "The first determination of the Planck constant with the joule balance NIM-2", *Metrologia* **54** (2017) 763, https://doi.org/10.1088/1681-7575/aa7a65

[6] K. Fujii et al., "Realization of the kilogram by the XRCD method", *Metrologia* **53** (2016) A19, http://dx.doi.org/10.1088/0026-1394/53/5/A19

[7] D. B. Newell et al., "The CODATA 2017 values for h, e, k, and N_A for the revision of the SI", *Metrologia* **55** (2018) L13, https://doi.org/10.1088/1681-7575/aa950a

[8] CCM, 2019, "CCM detailed note on the dissemination process after the redefinition of the kilogram", available on the BIPM web site: www.bipm.org

[9] S. Davidson, M. Stock, "Beginning of a new phase of the dissemination of the kilogram", *Metrologia* **58** (2021) 033002, https://doi.org/10.1088/1681-7575/abef9f

[10] M. Stock et al., "Report on the CCM key comparison of kilogram realizations CCM.M-K8.2019", *Metrologia* **57** (2020) 07030, https://doi.org/10.1088/0026-1394/57/1A/07030

[11] M. Stock et al., "Final report on the CCM key comparison of kilogram realizations CCM.M-K8.2021", *Metrologia* **60** (2023) 07003, https://doi.org/10.1088/0026-1394/60/1A/07003

[12] H. Fang et al., "The BIPM Kibble balance for realizing the kilogram definition", *Metrologia* **57** (2020) 045009, https://doi.org/10.1088/1681-7575/ab860c

[13] Y.-H. Wu, S.-J. Shen, L. Tsao, S.-J. Chen, "Minimal System for Dissemination and Realization of the Kilogram via the XRCD Method at CMS/ITRI", *IEEE Instrum. Meas*. **73** (2024) 1007511, https://doi.org/10.1109/TIM.2024.3453348

[14] M. Thomas et al., "A determination of the Planck constant using the LNE Kibble balance in air", *Metrologia* **54** (2017) 468, https://doi.org/10.1088/1681-7575/aa7882

[15] A. Eichenberger et al., "First realisation of the kilogram with the METAS Kibble balance", *Metrologia* **59** (2022) 025008, https://doi.org/10.1088/1681-7575/ac566f

[16] Z. Li et al., "Improvements of the NIM-2 Joule balance since 2020", *IEEE Trans. Instrum. Meas.,* 72 (2023) 1004907, https://doi.org/10.1109/TIM.2023.3276012

[17] D. Haddad et al., "Measurement of the Planck constant at the National Institute of Standards and Technology from 2015 to 2017", *Metrologia* **54** (2017) 633, https://doi.org/10.1088/1681-7575/aa7bf2

[18] N. Kuramoto et al., “Reproducibility of the Realization of the Kilogram Based on the Planck Constant by the XRCD Method at NMIJ”, *IEEE Trans. Instr. Meas*. **70** (2021) 1005609, https://doi.org/10.1109/TIM.2021.3061805

[19] N. Kuramoto et al., “Evaluation of the long-term reproducibility of a realization of the kilogram by the XRCD method”, *Metrologia* **62** (2025) 035001, https://doi.org/10.1088/1681-7575/adc817

[20] B. M. Wood et al., “A summary of the Planck constant determinations using the NRC Kibble balance”, *Metrologia* **54** (2017) 399, https://doi.org/10.1088/1681-7575/aa70bf

[21] R. G. Green et al., “Realisation and dissemination of mass through the consensus value of the kilogram and preparation for the final phase of traceability to independent realisations” *Measurement Sensors* **38** (2025) 101352, https://doi.org/10.1016/j.measen.2024.101352

[22] D. Eppers et al., “Status of the kilogram realization using the XRCD method at the PTB”, *Measurement: Sensors* 38 (2025), 101353, https://doi.org/10.1016/j.measen.2024.101353

[23] D. Eppers et al., “Realisation of the kilogram with the XRCD-method – improved XRF evaluation and its impact”, *Metrologia* **62** (2025) 065001, https://doi.org/10.1088/1681-7575/ae158d

[24] G. Bartl et al., “A new ^{28}Si single crystal: counting the atoms for the new kilogram definition”, *Metrologia* **54** (2017) 693, https://doi.org/10.1088/1681-7575/aa7820

[25] H. Ahmedov, “Alignment Procedure for air operating UME Kibble balance”, Poster for BIPM 150th Anniversary, https://thebipm150.org/posters-online/

[26] M. Stock *et al*., “Calibration campaign against the international prototype of the kilogram in anticipation of the redefinition of the kilogram part I: comparison of the international prototype with its official copies”, *Metrologia* **52** (2015) 310, http://dx.doi.org/10.1088/0026-1394/52/2/310

[27] J. W. Chung, V. Görlitz, M. Vogtmann, E. Beyer, F. Härtig, “The PTB Cleaning Procedure for Silicon Spheres”, 2016, https://www.ptb.de/cms/fileadmin/internet/presse_aktuelles/messen_events/si-kg_workshop/Documents/PTB_Si_bestPracticeCleaningProcedure_PTB.pdf

[28] M. Cox, C. Eiø, G. Mana, F. Pennecchi, “The generalized weighted mean of correlated quantities” *Metrologia* **43** (2006) S268, http://dx.doi.org/10.1088/0026-1394/43/4/S14

[29] “Measurement comparisons in the CIPM MRA”, CIPM MRA-G-11, Version 1.1, 2021, available on the BIPM web site: www.bipm.org