Journal of Chemical Metrology
www.acgpubs.org
Open Access
This article is licensed under CC-BY-NC-ND 4.0
ORIGINAL ARTICLE

Metrological framework for the development of plant-based certified reference materials. Part II: Multilaboratory characterisation, provisional value assignment and uncertainty evaluation

Jakub Karasiński[1*], Andrzej Gawor[1], Anna Ruszczyńska[1], Marcin Wojciechowski[1], Andrii Tupys[1,4], Agnieszka Borowiec[1], Ryszard Dobrowolski[2], Joanna Reszko-Zygmunt[2], Joanna Dobrzyńska[2], Dariusz Banaś[3], Aldona Kubala-Kukuś[3] and Ewa Bulska[1]

[1]University of Warsaw, Faculty of Chemistry, Biological and Chemical Research Centre, Warsaw, Poland
[2]Maria Curie-Sklodowska University in Lublin, Faculty of Chemistry, Lublin, Poland
[3]Jan Kochanowski University of Kielce, Institute of Physics, Kielce, Poland
[4]Faculty of Environmental Sciences and Natural Resource Management, Norwegian University of Life Sciences (NMBU), Høgskoleveien 12, Ås 1433, Norway

Cite This: *J. Chem. Metrol.* 2026, 20(2), e26094022

Read Online

Abstract: Part I of the MultiBio CRM study described the production of paired strawberry leaf and fruit candidate reference materials, including controlled cultivation, matrix-specific dehydration, comminution, fortification with As, Cd, Pb and Hg, homogenisation, packaging and sterilisation. Part II presents the multilaboratory characterisation, provisional value assignment and uncertainty evaluation. Six laboratories contributed results obtained by ICP-MS, ICP-OES, AAS and TXRF. Because the participants' uncertainties were reported using non-equivalent conventions, independent laboratory results were equally weighted and their median was used as the provisional assigned value, with a characterisation uncertainty of $1.25s/\sqrt{p}$. This component was combined with conservative matrix-level homogeneity contributions of 12.5% (leaves) and 27% (fruit). The provisional mass fractions and interim expanded uncertainties U* (k = 2, stability excluded) were: leaves, As (0.55 ± 0.14) mg/kg, Cd (0.55 ± 0.16) mg/kg and Pb (1.25 ± 0.41) mg/kg; fruit, As (0.48 ± 0.29) mg/kg, Cd (0.37 ± 0.24) mg/kg and Pb (1.40 ± 0.87) mg/kg. Mercury, determined by two laboratories with one technique, remained indicative. At the current uncertainty level the materials support quality control, method development and detection of gross bias, whereas certification requires approved homogeneity, stability and traceability evidence.

Keywords: CRM, certified reference material, multilaboratory characterisation, metrological traceability

1 Introduction

A certified reference material (CRM) is not defined only by the physical material or by a consensus concentration. It is a reference material characterised by a metrologically valid procedure and accompanied by documentation stating one or more property values, the associated uncertainties and metrological traceability (ISO 17034, 2016; ISO 33401, 2024; ISO 33403, 2024; ISO 33405, 2024). For quantitative chemical properties, the uncertainty statement should be consistent with the principles of the Guide to the Expression of Uncertainty in Measurement (JCGM 100, 2008; JCGM GUM-1, 2023) and with their application to analytical measurement (Ellison & Williams, 2012). CRMs therefore provide a practical link between metrologically higher-order standards, validated measurement procedures and routine laboratory results.

In food, environmental and biological analysis, CRMs are used to evaluate trueness, validate digestion and measurement procedures, establish internal quality-control limits, support uncertainty estimation and demonstrate comparability between laboratories (Emons et al., 2006; Kiełbasa

*Corresponding Author: Jakub Karasiński.
Email: jkarasinski@cnbc.uw.edu.pl
Received: September 15, 2026
Revised: October 23, 2026
Accepted: October 24, 2026
Published: November 26, 2026

Copyright © 2026. Published by ACG Publications.

et al., 2016). Matrix-matched materials are especially valuable because an aqueous calibration solution cannot mimic all possible effects associated with a natural plant matrix, including incomplete decomposition, adsorption losses, carbon-related signal enhancement or suppression, polyatomic and spectral interferences, moisture correction and sub-sampling variability. These effects become proportionally more important when toxic elements are measured at trace levels.

The production and certification of a matrix CRM therefore involves a sequence of interrelated steps, including material preparation, homogeneity and stability assessment, characterisation, value assignment, and uncertainty evaluation. Material processing must generate a sufficiently homogeneous and stable batch; the minimum representative test portion must be defined; characterisation results must be technically valid and traceable; and the final uncertainty must combine contributions from characterisation, between-unit heterogeneity and stability (Ellison et al., 2001; ISO 33405, 2024; Linsinger et al., 2001b; van der Veen et al., 2001a, 2001b, 2001c). Homogeneity and stability cannot be treated as routine ancillary tests because either component may dominate the final uncertainty. Alternative study designs can be appropriate, but the producer must demonstrate that the chosen design is capable of detecting relevant heterogeneity at the intended sample intake (ISO 33405, 2024; Linsinger & Bacquart, 2019).

The production of new plant-based CRMs involves many challenges. Available plant and food CRMs include tea leaves, tobacco leaves, cabbage, lichen, rice flour, cucumber, milk powder, mushroom powder, tea powder and strawberry leaves (Analytika, 2020; Chew et al., 2016; Grimalt et al., 2015; Herzig et al., 2002; Miyashita et al., 2012; Quevauviller et al., 1996; Rastogi et al., 2024; Samczyński et al., 2012; Santos et al., 2020; Zhu et al., 2011, 2013). Their development shows recurring requirements: controlled particle size, explicit dry-mass correction, several independent techniques and laboratories, technically justified exclusion of outlying data, traceable calibration and an uncertainty budget that includes heterogeneity and stability. Fortification with aqueous element solutions has been used to obtain analytically relevant mass fractions, including in strawberry and cabbage candidate materials (Hietaniemi et al., 1995), and work on particulate matrices has shown that material processing is inseparable from the metrological quality of a CRM (Emteborg et al., 2024).

Part I of the MultiBio CRM study (Karasiński et al., 2026) established the production protocol of the paired strawberry materials, summarised in Section 2.1, but did not report the multilaboratory dataset or the property-value calculations. Part I also showed that the minimum test portion is part of the measurand specification and that relative standard deviations obtained during process screening cannot substitute for the between-unit uncertainty derived from a formal homogeneity study. In this paper, "provisional" refers to values and uncertainties established before the stability components have been approved and the final traceability review has been completed.

The present study addresses the transition from multilaboratory results to provisional values. Its objectives were to (i) link the dataset to the material-production history, (ii) identify independent laboratory-level estimates, (iii) evaluate the reported uncertainty information, (iv) assign provisional values for As, Cd and Pb with an interim uncertainty model, (v) evaluate the status of Hg, and (vi) define the applications for which the materials are fit at the current uncertainty level.

2 Materials and Methods

2.1 Candidate Materials, Production History and Measurands

2.1.1 Continuity with the Part I Production Workflow

Both candidate materials, powdered strawberry leaves (MultiBio-CRM-L) and powdered strawberry fruit (MultiBio-CRM-F), were produced at the University of Warsaw in accordance with ISO 17034 and the supporting ISO reference-material standards (ISO 17034, 2016; ISO 33401, 2024; ISO 33403, 2024; ISO 33405, 2024; Karasiński et al., 2026); the production evidence relevant to the present study is summarised in Table 1. Botanical identity, cultivation conditions, harvest information and bulk-lot identity were recorded; fruit was harvested successively at technological maturity and leaves as a defined batch, and a preliminary contamination survey excluded atypical local enrichment.

Leaves were air dried and fruit was lyophilised to a residual moisture below 3% (w/w). Moisture-uptake tests showed that, after 24 h of exposure to laboratory air, the fruit and leaf powders reached approximately 20% and 15% moisture, respectively, with the steepest uptake during the first hours (Karasiński et al., 2026); exposure of opened units must therefore be minimised.

Particle size was reduced by fragmentation in agate mortars, cutting mills with 40-mesh sieves, ball milling and, where needed, cryogenic grinding, to a practical criterion of 80 µm or less. Bulk mixing in rotating polyethylene drums lasted approximately 16 h and was extended whenever ICP-MS screening of replicate test portions indicated unacceptable homogeneity, i.e. when the RSD of the fortified elements clearly exceeded the analytical repeatability, for which the RSD of non-fortified elements (below 5% after cryogenic milling) served as a benchmark. This corresponds to the acceptance criterion predefined in Part I: between-sample variability not exceeding the analytical repeatability, without location-dependent trends (Dybczyński et al., 1998; Karasiński et al., 2026; López-Romero et al., 2005).

Aqueous single-element standard solutions of As, Cd, Pb and Hg (approximately 1 mL each) were added after preliminary cryogenic grinding in aliquots of 20 µL to 30 µL with intermediate mixing, and the material was re-mixed. The nominal enrichment targets of 0.5 mg/kg As, 0.4 mg/kg Cd, 1.0 mg/kg Pb and 0.2 mg/kg Hg (Karasiński et al., 2026) were design targets, not assigned values. The effect of mixing time was evaluated by ICP-MS screening of replicate test portions of 250 mg and 500 mg taken after the initial and the extended mixing (Table 4).

https://doi.org/10.25135/jcm.2609-4022
J. Chem. Metrol. 2026, 20(2):e26094022

Table 1. Production evidence established in Part I and its relevance to the present certification stage

Production stage	Evidence established in Part I	Direct relevance to Part II
Matrix and lot definition	Controlled cultivation; authenticated strawberry leaf and fruit matrices; separate bulk lots; contamination survey	Defines matrix identity, provenance, intended-use scope and the population to which assigned values apply
Dehydration and moisture control	Air-dried leaves; lyophilised fruit; residual moisture <3% (w/w); rapid moisture re-uptake demonstrated	Requires explicit reporting basis and strict handling instructions; affects mass fraction and uncertainty
Comminution	Agate fragmentation, 40-mesh cutting mill, ball milling and cryogenic grinding; practical target ≤80 µm	Supports representative sub-sampling and the declaration of a minimum test portion
Targeted enrichment	Incremental aqueous additions of As, Cd, Pb and Hg with intermediate mixing	Defines production targets but does not replace analytical value assignment
Corrective homogenisation	Extended mechanical mixing after elevated screening RSD values	Provides evidence of process control and explains the final minimum test portion
Packaging and sterilisation	25 mL glass containers, argon purge, sequential numbering, sealed units, 28 kGy electron-beam treatment	Supports chemical and microbiological stability and permits testing for filling-order effects
Homogeneity and stability framework	Replicate sub-sampling, ANOVA, regression for trends, short- and long-term stability studies	Defines the required uncertainty components, but numerical producer-approved components must be available for final certification

The final powders were filled into 25 mL amber borosilicate-glass containers under argon, hermetically closed and sequentially numbered (nominal unit mass at least 12 g for leaves and 10 g for fruit), and sterilised with an electron beam at 28 kGy. Pre- and post-irradiation moisture, appearance and elemental screening results did not differ significantly (Karasiński et al., 2026; Kawamoto et al., 2019; Wise et al., 2025).

2.1.2 Measurands and Minimum Test Portion

The measurands for the described certification exercise were the mass fractions of total arsenic, cadmium, lead and mercury in the final bottled materials, expressed on a dry-mass basis. The certificate must therefore specify the moisture-determination or conditioning procedure, because moisture correction can be a significant source of between-laboratory bias in powdered biological materials (Herzig et al., 2002; Rastogi et al., 2024; Zhu et al., 2013).

The participants applied their in-house dry-mass procedures, which were not harmonised with respect to drying temperature, drying endpoint and the use of a separate test portion. Because the residual moisture of the bottled materials was below 3% (w/w), the resulting differences should not exceed about 3% for properly handled units, which is small compared with the between-laboratory RSDs of 6% to 34% (Table 6); larger effects of moisture uptake during handling cannot be excluded and are implicitly included in u_{char}. For certification, a single procedure will be prescribed: drying at 105 °C to constant mass for leaves, as for the strawberry-leaf CRM METRANAL® AN-BM01 (Analytika, 2020), and a validated lower-temperature procedure (e.g. vacuum drying) for the sugar-rich fruit powder.

Based on the results described in Part I, the minimum test portion for elemental analysis was set at 500 mg (Karasiński et al., 2026). This is therefore a condition for the valid use of the assigned value and must be specified in the certificate. Users must mix the bottle content before sub-sampling and follow the storage, equilibration and re-closing instructions stated in the final certificate.

2.1.3 Link between the Part I and Part II Datasets

All results refer to the batches described in Part I, i.e. one bulk lot per matrix carried through fortification, homogenisation, bottling and sterilisation (Karasiński et al., 2026). The screening RSDs in Table 4 were obtained by the coordinating laboratory on the bulk powders during homogenisation, i.e. before bottling and sterilisation. The characterisation results (Tables 5–9) and the formal between-unit homogeneity study refer to bottled, sterilised units of the final batches, analysed with test portions of at least 500 mg.

2.2 Participating Laboratories and Analytical Coverage

Six laboratories experienced in trace-element analysis of biological materials contributed results. The characterisation protocol did not prescribe a common method: each laboratory used its own validated procedure for total element determination, and the diversity of measurement principles was intended to reduce the risk of an undetected common bias (ISO 33405, 2024). The technique underlying each laboratory-level result is given in Table 2. Evidence of competence (scope of accreditation or quality-management documentation, proficiency-testing results and results for matrix CRMs analysed with the samples) and procedural details (decomposition, calibration, blanks and recoveries)

Table 2. Participating laboratories and the analytical technique underlying each laboratory-level result used in the consensus

Laboratory	As	Cd	Pb	Hg	Additional elements
L1	ICP-OES and GF-AAS[a]	ICP-OES	ICP-OES and GF-AAS[a]	CV-AAS	ICP-OES
L2	HG-AAS	GF-AAS	GF-AAS	CV-AAS	F-AAS
L3	–	ICP-MS	ICP-MS	–	ICP-OES
L4	–	–	–	–	TXRF
L5	ICP-MS	ICP-MS	ICP-MS	–	ICP-MS
L6	ICP-MS	ICP-MS	ICP-MS	–	ICP-MS

Note: [a]Mean of the two procedures (Section 2.4). –, not determined. Each entry corresponds to one laboratory-level result per analyte and matrix (Table 5); results for additional elements are summarised in Table 9.

Table 3. Uncertainty-reporting status by participating laboratory

Laboratory	Reported form in workbook	Coverage information	Use in present consensus
L1	Uncertainty reported	Not stated whether standard or expanded	Not used as a statistical weight; type of uncertainty and coverage factor to be specified
L2	Expanded uncertainty	$k = 2$ stated	Retained for technical review; not used as a weight
L3	Expanded uncertainty	$k = 2$ stated	Retained for technical review; not used as a weight
L4	Expanded uncertainty	$k = 2$ stated	Additional elements only
L5	Expanded uncertainty	Coverage factor not stated	Retained for technical review; not used as a weight; coverage factor to be specified
L6	Standard deviation reported	Workbook applied factor 2	Retained for technical review; not used as a weight; GUM-compliant uncertainty budget to be provided

were not reported uniformly in the workbooks; they are being compiled in the certification file for the technical review that precedes certification, and no unreported details are inferred here.

2.3 Review of Uncertainty Statements

Uncertainty information was not reported using a harmonised convention (Table 3): L2, L3 and L4 reported expanded uncertainties with an explicitly stated $k = 2$; L5 did not state the coverage factor; L1 did not state whether its value was a standard or an expanded uncertainty; and L6 reported the standard deviation of replicates, which was multiplied by two in the workbook. A standard deviation of replicates describes only repeatability and is not a measurement uncertainty in the sense of the GUM and the Eurachem/CITAC guide (Ellison & Williams, 2012; JCGM 100, 2008); L6 will therefore be asked to provide an uncertainty budget evaluated according to these documents, and L1 and L5 to specify the type of uncertainty and the coverage factor, before the final technical review.

Equal weighting of independent laboratory-level estimates was therefore selected. Participant uncertainties nevertheless remain necessary for the technical review, for documenting traceability and competence, and for checking the compatibility of individual results with the assigned value, e.g. by $E_n = x_i - x_{ass}/\sqrt{(U_i^2 + U^{*2})}$, with $|E_n| \leq 1$ indicating compatibility (ISO 13528, 2022). E_n numbers were not applied here, because the participants' uncertainties are not yet available on a common basis; they are foreseen for the final technical review (Ellison et al., 2001; Rukhin, 1996; Yatkin et al., 2020).

2.4 Data Curation and Independence of Results

All values used in the calculations were checked against the raw-result worksheets rather than copied from the summary sheets; this revealed and corrected a misalignment of the leaf As data and ensured that each value was associated with the correct laboratory and procedure.

L1 reported two procedures for As and Pb (ICP-OES and GF-AAS). Because they share the same organisation, quality system and potentially sample preparation and calibration resources, they were not counted as independent laboratories; the two results were averaged, as neither procedure was technically preferred. Every median is therefore based on exactly one result per laboratory, obtained with the technique given in Table 2.

No result was excluded solely because it was numerically distant from the centre of the dataset. A result should be rejected from CRM characterisation only when a documented technical cause, transcription error, invalid procedure or statistically justified inconsistency is established and reviewed (Ellison et al., 2001; van der Veen et al., 2001c; Yatkin et al., 2020). With four or five independent estimates, formal outlier tests have low power and can produce unstable decisions. Robust estimation and transparent presentation of all values were therefore preferred.

https://doi.org/10.25135/jcm.2609-4022
J. Chem. Metrol. 2026, 20(2):e26094022

2.5 Provisional Value Assignment

For each matrix and analyte, the provisional assigned value x_{ass} (Eq. (1)) was calculated as the median of p independent laboratory-level estimates:

$$x_{ass} = \text{median}\left(x_1, x_2, \ldots, x_p\right) \quad (1)$$

The median was selected because the dataset combines different analytical principles and contains only four or five results for As, Cd and Pb; it limits the effect of a single high or low result without requiring an exclusion decision. It is a pragmatic choice for this small, heterogeneous dataset rather than a universally superior estimator (ISO 13528, 2022; Rukhin, 1996; Yatkin et al., 2020); the arithmetic mean was calculated as a sensitivity check.

The between-laboratory standard deviation, $s_{lab} = \sqrt{\left[\Sigma\left(x_i - \bar{x}\right)^2/(p-1)\right]}$, was calculated from the p independent estimates. Because the assigned value is a median rather than an arithmetic mean, the characterisation standard uncertainty was calculated as recommended in ISO 13528 (2022) for robust estimates of location (Eq. (2)):

$$u_{char} = 1.25 \times s_{lab}/\sqrt{p} \quad (2)$$

Because $s_{lab}/\sqrt{p}$ is the standard error of an arithmetic mean, it would underestimate the uncertainty of a median. For large normal samples the standard error of the median is $\sqrt{(\pi/2)} \approx 1.25$ times that of the mean, which is the basis of Eq. (2) (ISO 13528, 2022); for the present datasets, a Monte Carlo simulation of normally distributed data (2×10^6 trials) gave ratios of 1.09 ($p = 4$) and 1.20 ($p = 5$), so that the factor 1.25 is slightly conservative. The classical standard deviation was retained because s_{lab} exceeded the scaled median absolute deviation for every dataset. Compared with $s_{lab}/\sqrt{p}$, Eq. (2) increases u_{char} by 25% and the relative interim expanded uncertainties by 0.4 to 4.1 percentage points (Table 8).

2.6 Uncertainty Evaluation

The uncertainty evaluation followed the GUM (JCGM 100, 2008; JCGM GUM-1, 2023), the Eurachem/CITAC guide (Ellison & Williams, 2012) and the approach for CRMs characterised by interlaboratory study (Ellison et al., 2001; ISO 33405, 2024). Because the participants' uncertainty budgets were not reported on a common basis, the characterisation uncertainty was evaluated top-down from the dispersion of the laboratory-level results, which integrates all sources that differ between laboratories, including effects not recognised in individual budgets (Thompson & Ellison, 2011). The interim model combines u_{char} with a between-unit heterogeneity contribution (u_{bb}); the sources are summarised in Section 3.9, and stability contributions will be added once the stability studies have been approved.

Homogeneity screening was used in Part I as a production-control tool (Karasiński et al., 2026). As expected for a material fortified with small liquid aliquots, the RSDs of replicate 250 mg test portions were very high after fortification, although the mean mass fractions were close to the design levels; extended mixing and a test portion of 500 mg reduced them (Table 4).

The formal homogeneity design of Part I comprised sequentially numbered units, replicate portions, permuted measurement order, regression checks for filling-order effects and ANOVA separating between-unit and within-unit variability; for fruit, one unit from each of seven filling-sequence subsets (65 units in total) was analysed in four replicates (ISO 33405, 2024; ISO 5725-2, 2025; Karasiński et al., 2026; van der Veen & Pauwels, 2000; van der Veen et al., 2001a). Because the numerical evaluation of these studies had not been approved when the interim budget was prepared, matrix-level contributions were adopted.

The interim model therefore applied relative homogeneity uncertainty contributions (r_{bb}) of 12.5% for leaves and 27% for fruit. These are deliberately conservative assumptions, not ANOVA estimates. For leaves, r_{bb} equals the largest 500 mg screening RSD observed for As, Cd and Pb (12.4% for Pb, rounded). For fruit, r_{bb} was set at approximately twice this value, because the lyophilised fruit powder took up moisture faster (20% versus 15% after 24 h) and was more prone to agglomeration during fine grinding and mixing (Karasiński et al., 2026). Because screening RSDs also contain the analytical repeatability, both values are upper estimates. The same contribution was applied to As, Cd and Pb within each matrix (Eq. (3)):

$$u_{bb} = r_{bb} \times x_{ass} \quad (3)$$

where $r_{bb} = 0.125$ for leaves and 0.27 for fruit. Compliance with the acceptance criterion of Part I was not demonstrated at the screening stage, because the 500 mg RSDs of the fortified elements (7.7% to 12.4% for As, Cd and Pb; 20.0% for Hg) exceeded those of non-fortified elements (below 5%). The materials were therefore not declared homogeneous; a non-negligible heterogeneity contribution was included instead, as foreseen in ISO 33405 (2024). Because each laboratory analysed different units, part of the between-unit variability is already contained in s_{lab}, so that U^* is an upper estimate (Linsinger & Bacquart, 2019).

U^* is therefore not a certification-ready uncertainty: in the final certificate, analyte-specific u_{bb} values from the between-unit studies will replace the matrix-level contributions (ISO 33405, 2024; Linsinger et al., 2001b; van der Veen et al., 2001a).

The interim combined standard uncertainty u_c^* (Eq. (4)) and expanded uncertainty U^* (Eq. (5)) were calculated as:

$$u_c^* = \sqrt{\left(u_{char}^2 + u_{bb}^2\right)} \quad (4)$$

$$U^* = k \times u_c^*, k = 2 \quad (5)$$

The asterisk distinguishes the interim uncertainty, which includes the characterisation and homogeneity contributions, from the final CRM uncertainty, which will also include the stability contributions that depend on the defined transport conditions, storage temperature and validity period.

Table 4. Homogenisation-screening RSDs (%) of the fortified elements depending on the mixing procedure and test portion mass (bulk powder before bottling; data from Part I)

Element	As prepared *(see Section 2.1.1)*	After additional mixing *(see Section 2.1.1)*	
	250 mg	250 mg	500 mg
As	22.7	16.0	8.6
Cd	35.7	7.7	7.7
Pb	25.2	15.0	12.4
Hg	45.6	21.1	20.0

Note: RSDs of replicate test portions of the fortified bulk powder (coordinating laboratory, ICP-MS), as reported in Part I (Karasiński et al., 2026), where the screening data were not separated by matrix. They include the analytical repeatability, served only for process control and are not ANOVA-based between-unit standard deviations; non-fortified elements showed RSDs below 5%.

The final model will therefore be (Eq. (6)):

$$U_{CRM} = k \times \sqrt{(u_{char}^2 + u_{bb}^2 + u_{sts}^2 + u_{lts}^2)} \quad (6)$$

where u_{sts} and u_{lts} are standard uncertainty contributions from short-term transport and long-term storage stability, respectively. Isochronous designs reduce day-to-day analytical variation, but the uncertainty of a possible slope must still be propagated even when the estimated slope is not statistically significant (Lamberty et al., 1998; Linsinger et al., 2001a; van der Veen et al., 2001b).

2.7 Reporting and Rounding of Property Values

A value was proposed as a candidate certified value when at least four independent laboratory estimates obtained with more than one analytical principle were available, no unresolved technical invalidity was identified, and the remaining stability and traceability actions were defined. This is a project decision rule rather than a numerical requirement of ISO 17034 or ISO 33405, because the adequacy of a characterisation dataset depends on the quality and independence of the evidence (Ellison et al., 2001; ISO 17034, 2016; ISO 33405, 2024).

Values based on one or two laboratories, or lacking sufficient evidence, were classified as indicative. For the additional matrix elements, which are reported without uncertainty, a median and range were tabulated as an information value when at least three independent results were available and the highest result did not exceed the lowest by more than a factor of 1.5; datasets with a larger spread were classified as screening information only. Elements with fewer than three results, or with an individual result flagged as technically doubtful during the review, were not tabulated. Uncertainties were rounded to two significant digits and values to the same decimal place (ISO 33401, 2024; JCGM 100, 2008).

3 Results

3.1 Material Preparation and Readiness for Value Assignment

Before multilaboratory characterisation, the candidate materials underwent controlled preparation, homogenisation, packaging, sterilisation and evaluation of their suitability for certification (Figure 1). The mean mass fractions of the fortified elements were close to the intended ranges, whereas the initial dispersion between test portions was too high for reliable value assignment. Extended mixing and a minimum test portion of 500 mg reduced the screening RSDs of As, Cd and Pb to values between 7.7% and 12.4%, whereas Hg remained at 20.0% (Table 4), indicating greater residual microscale heterogeneity. The materials were therefore considered suitable for characterisation with respect to As, Cd and Pb, whereas Hg required a more conservative interpretation.

For Hg, the higher residual heterogeneity and the limited independence of the available results do not support certification, and an indicative value is reported instead (Section 3.7).

3.2 Data Completeness after Consolidation

After consolidation of the two L1 procedures, each matrix contained four independent estimates for As, five for Cd and Pb, and two for Hg (Table 5); L4 contributed only TXRF results for additional elements. The equality of the two leaf Hg results does not indicate negligible characterisation uncertainty, because $p = 2$ cannot provide a stable estimate of between-laboratory dispersion.

3.3 Descriptive Statistics and Estimator Sensitivity

The leaf material generally showed lower between-laboratory dispersion than the fruit material (Table 6). The best agreement was observed for As in leaves (0.535 mg/kg to 0.610 mg/kg; RSD 6.0%) and the largest dispersion for Cd in fruit (0.190 mg/kg to 0.530 mg/kg; RSD 34.3%). For all As, Cd and Pb datasets, the relative difference between the arithmetic mean and the median did not exceed 7.2%.

Moreover, the absolute mean–median differences were smaller than the characterisation standard uncertainty for

https://doi.org/10.25135/jcm.2609-4022
J. Chem. Metrol. 2026, 20(2):e26094022

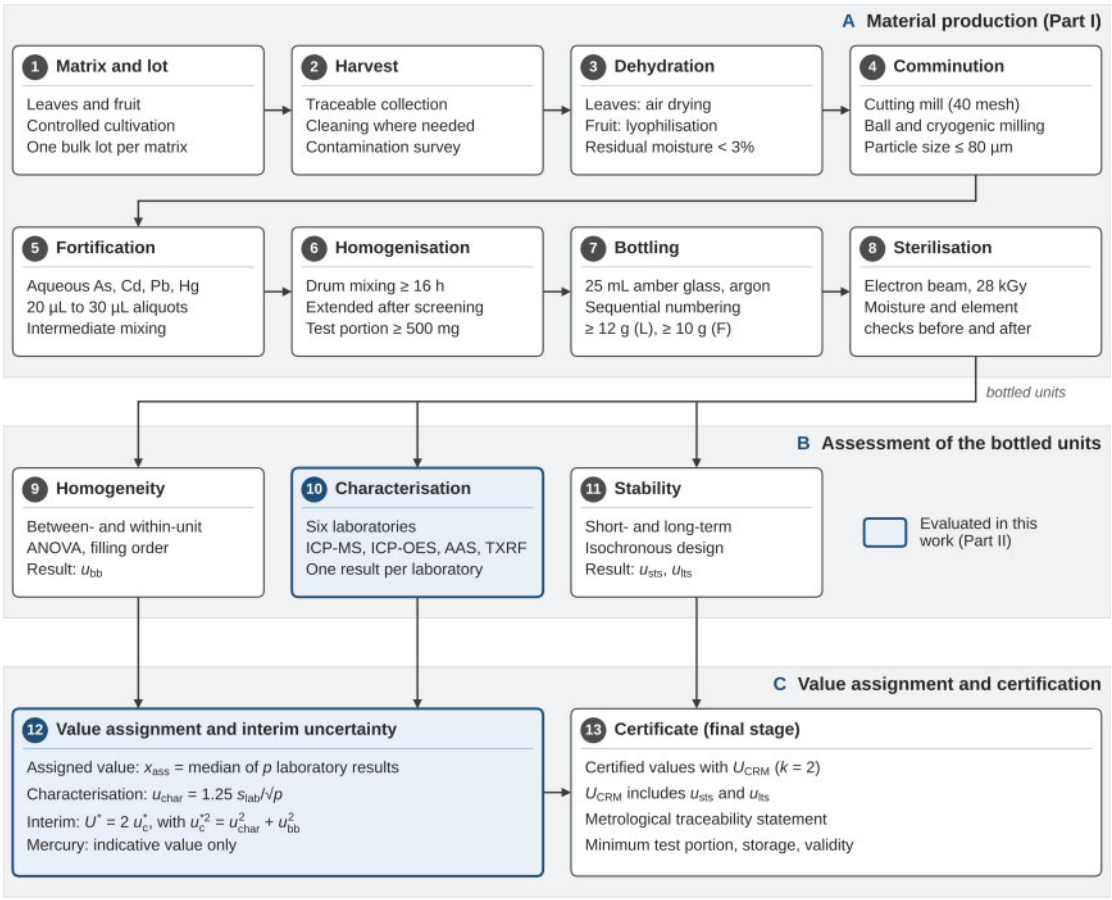

Figure 1. Production, metrological assessment and value-assignment workflow for the MultiBio CRM strawberry leaf and fruit materials. Steps 1–8 and the designs of steps 9 and 11 were described in Part I (Karasiński et al., 2026); steps 10 and 12 (shaded) are evaluated in the present work. L, leaves; F, fruit

Table 5. Independent laboratory-level target-element results obtained on the final bottled material, after consolidation of duplicate methods (mg/kg)

Matrix	Element	L1	L2	L3	L5	L6
Leaves	As	0.535	0.545	–	0.556	0.610
Leaves	Cd	0.550	0.582	0.410	0.570	0.460
Leaves	Pb	1.320	1.180	1.250	0.954	1.620
Leaves	Hg	0.130	0.130	–	–	–
Fruit	As	0.480	0.478	–	0.655	0.440
Fruit	Cd	0.370	0.422	0.190	0.530	0.320
Fruit	Pb	1.195	1.400	1.180	1.540	2.100
Fruit	Hg	0.198	0.166	–	–	–

Note: –, no result reported (the analyte was not determined by that laboratory); none of the entries represents a value below a limit of detection or quantification. L4 reported only additional matrix elements (TXRF). The technique underlying each entry is given in Table 2.

every analyte ($|\text{mean}-\text{median}|/u_{char}$ = 0.05 to 0.84) and therefore much smaller than U^*. Because U^* is dominated by the homogeneity contribution, the comparison with u_{char} is the more stringent test; it shows that the provisional assigned values are not strongly dependent on the estimator of location.

3.4 Arsenic

For leaves, the four As results (0.535 mg/kg to 0.610 mg/kg) were tightly grouped; the median of 0.5505 mg/kg was only 2.0% below the mean. The data combined consolidated ICP-OES/GF-AAS evidence from L1, HG-AAS from L2 and ICP-MS from L5 and L6, i.e. principles with different interference profiles. The interim expanded uncertainty was 0.1438 mg/kg, rounded to 0.14 mg/kg.

For fruit, the median As value was 0.4790 mg/kg. Three estimates clustered between 0.440 mg/kg and 0.480 mg/kg, while L5 reported 0.655 mg/kg; no documented technical reason justified its exclusion. The median limited its influence, while the dispersion remained represented in u_{char}. The interim expanded uncertainty (0.2853 mg/kg, 59.6%) is dominated by the 27% fruit homogeneity contribution.

3.5 Cadmium

Leaf Cd results ranged from 0.410 mg/kg to 0.582 mg/kg, with a median of 0.550 mg/kg and a lower mean (0.514 mg/kg) caused by two results below 0.50 mg/kg. The mean–median difference (−6.5%) was much smaller than the relative

Table 6. Descriptive statistics for independent laboratory-level results

Matrix	Element	*p*	Range (mg/kg)	Mean	Median	RSD (%)	Mean–median (%)
Leaves	As	4	0.535–0.610	0.5615	0.5505	6.0	+2.0
Leaves	Cd	5	0.410–0.582	0.5144	0.5500	14.7	−6.5
Leaves	Pb	5	0.954–1.620	1.2648	1.2500	19.1	+1.2
Fruit	As	4	0.440–0.655	0.5132	0.4790	18.8	+7.2
Fruit	Cd	5	0.190–0.530	0.3664	0.3700	34.3	−1.0
Fruit	Pb	5	1.180–2.100	1.4830	1.4000	25.4	+5.9

Note: Mean–median (%) = 100(mean–median)/median; a positive value means that the mean is higher than the median. RSD (%) = $100s_{lab}$/mean, where $s_{lab} = \sqrt{[\Sigma(x_i - \bar{x})^2/(p-1)]}$ is the standard deviation of the p independent laboratory-level estimates; the RSD therefore describes the between-laboratory dispersion and not the within-laboratory repeatability. Unit: mg/kg.

interim expanded uncertainty (29.3%). ICP-OES, GF-AAS and ICP-MS were represented, a method diversity comparable to that of established plant and food CRMs (Chew et al., 2016; Rastogi et al., 2024; Zhu et al., 2011).

Fruit Cd was the most dispersed dataset (0.190 mg/kg to 0.530 mg/kg; RSD 34.3%), with a closely agreeing median (0.370 mg/kg) and mean (0.366 mg/kg), indicating a broadly balanced spread rather than one extreme value. The interim expanded uncertainty was 0.2442 mg/kg, rounded to 0.24 mg/kg. The final review should examine digestion recovery, blank contribution, calibration level and dry-mass correction for each result before the Cd value is approved.

3.6 Lead

Leaf Pb results ranged from 0.954 mg/kg to 1.620 mg/kg (median 1.250 mg/kg, mean 1.265 mg/kg). The L1 result (1.320 mg/kg) was the mean of 1.22 mg/kg (ICP-OES) and 1.42 mg/kg (GF-AAS); replacing it by either individual result would change the median by at most 0.03 mg/kg (2.4%), i.e. by less than u_{char} (0.135 mg/kg). The interim expanded uncertainty was 0.413 mg/kg.

Fruit Pb results ranged from 1.180 mg/kg to 2.100 mg/kg (median 1.400 mg/kg, mean 1.483 mg/kg); the highest result increased the mean but did not determine the median. The interim expanded uncertainty was 0.865 mg/kg (61.8%). This does not invalidate the material, but shows that the adopted homogeneity contribution and the multilaboratory dispersion currently limit the precision of the Pb value in fruit.

3.7 Mercury

Mercury was reported only by L1 and L2, both using CV-AAS. The leaf results were identical (0.130 mg/kg) and the fruit results were 0.198 mg/kg and 0.166 mg/kg. Two results obtained by one analytical principle cannot establish an independent characterisation basis, because method-specific blank effects, reduction efficiency, memory effects or volatilisation losses could remain common or undetected. The recommended treatment is therefore an indicative value of 0.13 mg/kg for leaves and 0.18 mg/kg for fruit, without a certified uncertainty. Additional results should preferably include an independent principle such as isotope-dilution ICP-MS or direct mercury analysis with SI-traceable calibration.

3.8 Proposed Certified Values

The proposed values are summarised in Table 7. For As, Cd and Pb they are the medians of the p independent laboratory-level estimates, with $U^{\star}$ (k = 2) combining the characterisation and the adopted matrix-level homogeneity contributions. Because stability and the final traceability review have not yet been incorporated, they are candidate certified values rather than final certified values. Mercury is reported separately as an indicative value.

3.9 Components of the Uncertainty Budget

The homogeneity contribution was the largest component of the combined variance for every candidate certified value, accounting for 57% to 92% of u_c^{*2} for leaves and 67% to 82% for fruit (Table 8), a direct consequence of treating 12.5% and 27% as relative standard uncertainties. It influenced $U^{\star}$ more than the choice of location estimator or the factor applied to the uncertainty of the median. The structure of the budget is shown in Figure 2: the characterisation branch is evaluated top-down through s_{lab}, the homogeneity branch through r_{bb}, and the stability branch will be added in the final budget.

Additional characterisation laboratories would reduce u_{char} but would have little effect on $U^{\star}$ as long as u_{bb} dominates the budget (91.5% of u_c^{*2} for As in leaves; 76.4% for Pb in fruit). A substantial reduction of $U^{\star}$ therefore depends primarily on analyte-specific u_{bb} values from the formal between-unit study and, where achievable, on improved homogeneity at the declared minimum test portion. For fruit, however, the characterisation component alone corresponds to 25% to 38% ($2u_{char}/x_{ass}$), so that additional independent results and better harmonisation of digestion, dry-mass correction and calibration will also be required (Section 4.8).

3.10 Information Values for Additional Matrix Elements

The workbooks also contained results for major, minor and trace elements other than As, Cd, Hg and Pb. They are reported in Table 9 as preliminary information values, classified according to Section 2.7, to help users assess matrix matching; any future certification would require analyte-specific homogeneity, stability and characterisation review.

https://doi.org/10.25135/jcm.2609-4022
J. Chem. Metrol. 2026, 20(2):e26094022

Table 7. Proposed target-element values. $U^{\star}$ is an interim expanded uncertainty (k = 2) that excludes stability

Matrix	Element	Unit	p	Proposed value	U* (k = 2)	Recommended status
Leaves	As	mg/kg	4	0.55	0.14	Candidate certified value after stability and traceability approval
Leaves	Cd	mg/kg	5	0.55	0.16	Candidate certified value after stability and traceability approval
Leaves	Pb	mg/kg	5	1.25	0.41	Candidate certified value after stability and traceability approval
Leaves	Hg	mg/kg	2	0.13	Not assigned	Indicative value only; insufficient independent characterisation
Fruit	As	mg/kg	4	0.48	0.29	Candidate certified value after stability and traceability approval
Fruit	Cd	mg/kg	5	0.37	0.24	Candidate certified value after stability and traceability approval
Fruit	Pb	mg/kg	5	1.40	0.87	Candidate certified value after stability and traceability approval
Fruit	Hg	mg/kg	2	0.18	Not assigned	Indicative value only; insufficient independent characterisation

Note: Because the effective degrees of freedom could not be evaluated from the available participant data, $U^{\star} = 2u_c^{\star}$ is reported as an approximate expanded uncertainty, and no specific coverage probability is assigned.

Table 8. Components of the uncertainty budget and their relative contributions

Matrix	Element	p	x_{ass}	u_{char}	u_{bb}	Char. variance (%)	Homog. variance (%)	U*	$U^{\star}/x_{ass}$ (%)
Leaves	As	4	0.5505	0.0209	0.0688	8.5	91.5	0.1438	26.1
Leaves	Cd	5	0.5500	0.0422	0.0688	27.4	72.6	0.1613	29.3
Leaves	Pb	5	1.2500	0.1350	0.1562	42.7	57.3	0.4130	33.0
Fruit	As	4	0.4790	0.0602	0.1293	17.8	82.2	0.2853	59.6
Fruit	Cd	5	0.3700	0.0702	0.0999	33.1	66.9	0.2442	66.0
Fruit	Pb	5	1.4000	0.2102	0.3780	23.6	76.4	0.8650	61.8

Note: All absolute values are in mg/kg; u_{char} was calculated with Eq. (2). Variance percentages refer to u_{char}^2 and u_{bb}^2 as fractions of $u_c^{\star 2}$.

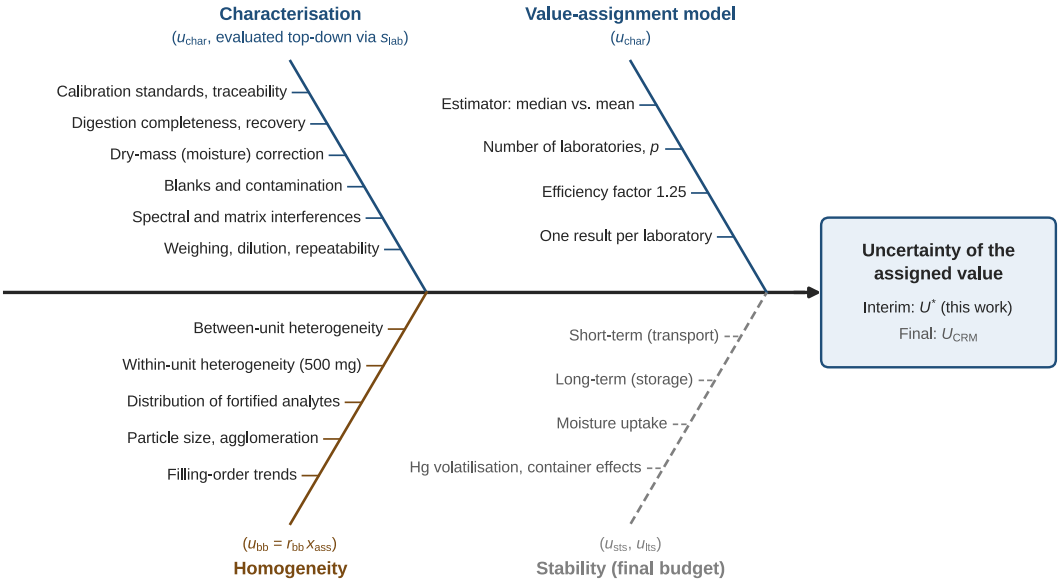

Figure 2. Cause-and-effect (fishbone) diagram of the uncertainty sources of the assigned value and their allocation to the components of the interim (u_{char}, u_{bb}) and final (u_{sts}, u_{lts}) uncertainty models. Sources in the characterisation branch were not quantified individually; their combined effect is captured by the between-laboratory dispersion (s_{lab}). The stability branch (dashed) will be quantified in the final certification budget

4 Discussion

4.1 Suitability of the Consensus Strategy

The dataset illustrates why statistical calculation cannot be separated from technical evaluation. A weighted mean would normally be preferable, but weighting non-equivalent uncertainty statements could give excessive influence to a laboratory that reported a standard rather than an expanded uncertainty, or whose coverage factor was not documented. Equal weighting avoids this distortion, the median reduces sensitivity to extreme results, and the mean–median comparison shows that the assigned values are not strongly estimator-dependent (Ellison et al., 2001; Rukhin, 1996; Yatkin et al., 2020). Once harmonised uncertainties are available, estimators that model between-laboratory effects

Table 9. Preliminary information values for additional matrix elements

Matrix	Element	Unit	*p*	Median	Reported range	Status
Leaves	Fe	mg/kg	6	128	113–141	
Leaves	Cu	mg/kg	5	8.26	7.42–10.9	
Leaves	Mn	mg/kg	5	226	216–270	
Leaves	Ca	g/kg	4	10.55	10.07–10.73	
Leaves	Cr	mg/kg	4	4.54	3.86–5.15	
Leaves	K	g/kg	4	16.8	14.21–18.84	Information value; full technical review required
Leaves	Mg	g/kg	4	2.48	2.36–2.7	
Leaves	Zn	mg/kg	4	19.82	17.6–22.3	
Fruit	Fe	mg/kg	6	45.9	42.77–50	
Fruit	Cu	mg/kg	4	2.495	2.43–2.57	
Fruit	Mn	mg/kg	5	47.5	45.8–54	
Fruit	Cr	mg/kg	5	1.92	0.498–2.85	High dispersion; screening information only
Fruit	K	g/kg	4	15.4	13.43–16.92	Information value; full technical review required
Fruit	Mg	g/kg	4	1.26	1.21–1.35	
Fruit	Zn	mg/kg	5	9.74	8.75–15.3	High dispersion; screening information only
Fruit	Ni	mg/kg	4	1.045	0.56–1.26	

Note: "Full technical review required": the median and range may be used as preliminary information values, but the underlying results, procedures, traceability and uncertainties must be reviewed before any value assignment. "Screening information only": highest/lowest result > 1.5; the range indicates only the approximate level. Not tabulated: Sr (in each matrix, one result was flagged as technically doubtful and is under investigation) and Co in fruit (only two results, differing by more than an order of magnitude).

explicitly will also be evaluated; agreement among valid models would strengthen the final value assignment, whereas disagreement would indicate laboratory bias requiring investigation (Koepke et al., 2017).

4.2 Method Independence and Pseudo-Replication

Counting procedures instead of laboratories would have inflated the number of As and Pb results and artificially reduced the standard error; the paired ICP-OES and GF-AAS results of L1 are nevertheless kept in the certification file as a within-laboratory method comparison. Method-principle diversity is particularly important for trace elements in biological matrices: ICP-MS is affected by polyatomic and carbon-related effects, ICP-OES by spectral interferences and GF-AAS by matrix modification and atomisation conditions, HG-AAS provides selective introduction of As, and CV-AAS is highly specific for Hg but provides no independence when all laboratories use it. Plant and food CRMs are therefore commonly characterised with several methods (Chew et al., 2016; Miyashita et al., 2012; Quevauviller et al., 1996; Rastogi et al., 2024; Zhu et al., 2011, 2013).

4.3 Matrix Dependence and the Role of Homogeneity

The larger relative uncertainty of the fruit material is driven mainly by its 27% homogeneity contribution, consistent with the Part I observations that fruit had a higher initial water content, absorbed moisture faster and agglomerated more readily during fine grinding (Karasiński et al., 2026).

The Part I RSD sequence (Table 4) shows that corrective mixing and the larger test portion improved representativeness for As, Cd and Pb, whereas the residual 20.0% for Hg identifies a persistent material challenge. Because such RSDs also reflect the analytical repeatability, u_{bb} must be linked to the declared test portion and derived from the approved between-unit variance model (ISO 33405, 2024; Linsinger & Bacquart, 2019; Polkowska-Motrenko et al., 2022; van der Veen et al., 2001a).

4.4 Fortified versus Native Analytes: Implications for Digestion-Recovery Studies

The target elements were added as aqueous inorganic solutions where native mass fractions were below analytically or legislatively relevant levels (Karasiński et al., 2026), so that much of each measurand originates from the added analyte. Added species are expected to be associated mainly with particle surfaces, whereas native elements are distributed within the tissue and partly bound in organic or refractory structures; native As may also include methylated species. Equivalent behaviour during digestion has therefore not been demonstrated; the only evidence is indirect, i.e. all laboratory-level results obtained with different decomposition and detection principles lie within $x_{ass} \pm U^*$.

After complete decomposition, e.g. closed-vessel microwave or pressure digestion (EN 13805, 2014), both native and added analytes should be released, so the materials are suitable for checking whole procedures for total element determination. Because surface-associated analytes are released more readily than matrix-bound ones, as known for spike recoveries (Thompson et al., 1999), recoveries obtained with these materials may overestimate those for native analytes when digestion is incomplete or the technique is species-sensitive (e.g. HG-AAS without complete mineralisation). The materials are therefore not proposed for demonstrating digestion efficiency for native analytes, and the certificate will state this limitation; a

https://doi.org/10.25135/jcm.2609-4022
J. Chem. Metrol. 2026, 20(2):e26094022

comparison with unfortified material is recommended for the certification phase.

4.5 Stability, Validity Period and Transport

Even when the target elements are not expected to degrade, changes in moisture, volatilisation (especially of Hg), container interaction or segregation can alter the measurand or the dry-mass correction. The uncertainty of a possible slope from the short- and long-term studies must be propagated over the transport and validity periods even when the slope is not statistically significant (Lamberty et al., 1998; Linsinger et al., 2001a; van der Veen et al., 2001b); the resulting u_{sts} and u_{lts} will be included in U_{CRM}, and the certificate will state the storage temperature, transport limits and period of validity (ISO 33401, 2024).

4.6 Metrological Traceability and Certificate Content

A numerical consensus does not by itself establish metrological traceability. The producer must document the measurement procedures, the calibration standards and their certificates, gravimetric and volumetric traceability and the route by which each laboratory result is linked to SI units (ISO 17034, 2016; JCGM 100, 2008). The certificate must identify the measurand as a mass fraction of total element on a stated basis, give the assigned value with its expanded uncertainty and coverage factor, the minimum test portion, handling and storage instructions, the intended use and its limitations (ISO 33401, 2024; ISO 33403, 2024).

For the present dataset, calibration was based on commercial single- or multi-element standard solutions supplied with certificates of analysis. The traceability of these calibrants to the SI, the matrix CRMs analysed for quality control and the corresponding uncertainties are verified for each laboratory in the technical review that precedes certification (Ellison & Williams, 2012). At the level of national metrology institutes, the comparability of such traceability chains is demonstrated in key comparisons, e.g. for elements in river water (Can et al., 2023) and polar contaminants in groundwater (Bilsel et al., 2022).

4.7 Comparison with Available Plant and Food CRMs

The MultiBio materials occupy a concentration range that is analytically relevant to routine toxic-element measurements. Available tea, tobacco, cabbage, lichen, rice, mushroom, tea-powder and strawberry-leaf CRMs show that certification can be achieved either with multiple independent methods within a national metrology institute or with a multilaboratory campaign involving several competent laboratories (Analytika, 2020; Chew et al., 2016; Herzig et al., 2002; Isleyen et al., 2024; Miyashita et al., 2012; Quevauviller et al., 1996; Rastogi et al., 2024; Samczyński et al., 2012; Zhu et al., 2011). Those materials typically report lower relative uncertainties because high-accuracy methods, extensive replication and analyte-specific homogeneity and stability data are available. The larger interim uncertainties for the MultiBio fruit material therefore reflect the current evidence rather than an intrinsic limitation of strawberry as a CRM matrix.

A distinctive strength of MultiBio is the paired leaf-fruit design from one crop system, which complements the existing strawberry-leaf CRM (Analytika, 2020) and allows methods to be evaluated across related matrices of different physical and chemical composition. The project also builds national capability for future plant CRMs, consistent with the broader role of complex-matrix reference materials in analytical metrology (Isleyen, 2017; Isleyen et al., 2024; Karasiński et al., 2026).

4.8 Fitness for Intended Use at the Current Uncertainty Level

A bias is detected with a CRM only if the absolute difference between the mean measured value and the assigned value, Δ_m, exceeds $U_\Delta = 2\sqrt{\left(u_m^2 + u_{CRM}^2\right)}$, where u_m and u_{CRM} are the standard uncertainties of the user's mean result and of the reference value (ISO 33403, 2024; Linsinger, 2010). With $u_{CRM} = u_c^*$ and a relative u_m of 5%, the smallest detectable relative bias is approximately 28% to 35% for leaves and 60% to 67% for fruit. At present, the leaf material can therefore reveal only gross errors (e.g. incomplete digestion or calibration errors of about one third of the value), the fruit material cannot be used for trueness assessment or method validation, and biases of 10% to 20% cannot be detected with either material.

Applications that do not rely on the accuracy of the assigned value are already supported: in-house quality control with control charts based on the laboratory's own results (test portions of at least 500 mg, preferably replicated), method development and within-laboratory comparison of procedures, assessment of matrix effects between the two matrices, and training.

With analyte-specific u_{bb} values of 3% to 5%, the relative expanded uncertainties would decrease to about 10% to 24% for leaves and 26% to 39% for fruit with the present characterisation data, so that the leaf material would support trueness assessment at bias levels of about 15% to 25%; for fruit, additional laboratories and harmonised procedures will also be required. For comparison, the strawberry-leaf CRM METRANAL® AN-BM01 is certified for As at (0.290 ± 0.021) mg/kg, Cd at (0.170 ± 0.011) mg/kg and Pb at (1.70 ± 0.14) mg/kg, i.e. with relative expanded uncertainties of 6% to 8% (Analytika, 2020).

5 Conclusion

For strawberry leaves, the proposed values are (0.55 ± 0.14) mg/kg As, (0.55 ± 0.16) mg/kg Cd and (1.25 ± 0.41) mg/kg Pb. For strawberry fruit, the proposed values are (0.48 ± 0.29) mg/kg As, (0.37 ± 0.24) mg/kg Cd and (1.40 ± 0.87) mg/kg Pb. These uncertainties are interim expanded uncertainties U^* with $k = 2$ and include characterisation and the currently adopted matrix-level homogeneity terms, but not numerical stability contributions. Mercury should remain indicative at 0.13 mg/kg for leaves and 0.18 mg/kg for fruit until additional independent characterisation and an approved Hg-specific homogeneity assessment are available.

The MultiBio materials were designed to support multi-technique elemental analysis in two related but compositionally distinct plant matrices, at enrichment levels measurable by ICP-MS, ICP-OES and atomic absorption spectrometry. At the current uncertainty level they are suitable for internal quality control, method development, inter-method comparison, assessment of matrix effects and detection of gross bias; their use for trueness assessment and method validation requires the certified values and uncertainties that will result from the completed homogeneity, stability and traceability studies.

These results support provisional value assignment for As, Cd and Pb, whereas Hg remains an indicative value because only two laboratories, both using CV-AAS, reported results. The principal limitations of the present evaluation are the incomplete harmonisation of the reporting basis and of the participants' uncertainty statements, the use of conservative matrix-wide rather than analyte-specific homogeneity contributions, and the lack of direct evidence that fortified and native analytes behave equivalently during digestion. Final certification therefore requires conversion of all results to a common dry-mass basis, completion of the technical and traceability review, incorporation of approved analyte-specific homogeneity contributions, and inclusion of short- and long-term stability contributions in the final uncertainty budget.

Acknowledgement

The authors acknowledge all laboratories that contributed analytical results to the MultiBio CRM characterisation exercise. The study was carried out at the Biological and Chemical Research Centre, University of Warsaw, established within a project co-financed by the European Union from the European Regional Development Fund under the Operational Programme Innovative Economy 2007–2013. Experimental work at the University of Warsaw was performed in laboratories operating under accredited quality-management systems in accordance with ISO/IEC 17025:2017 and ISO 15189:2022, with Polish Centre for Accreditation certificates AB 1525 and AM 025.

Funding Statement

This research was funded under the Polska Metrologia II programme of the Polish Ministry of Science and Higher Education within the project entitled "Multifunctional plant-based matrix reference material: strawberry leaves and fruits (MultiBio CRM) with certified content of metals and microplastics to ensure measurement traceability in Polish laboratories", contract No. PM-II/SP/0062/2024/02 of 14 February 2024. The research was also partially funded by the European Union under the Horizon Europe research and innovation programme (2021–2027), within the PATHFOOD project (Creating a Functional and Sustainable Food Chain: A Path to Climate Neutrality and Improved Health in the EU), grant agreement No. 101160011.

Author Contributions

Conceptualization: J.K., A.G., E.B.; Methodology: J.K., A.G., A.R., M.W., A.T., E.B.; Investigation: J.K., A.G., A.R., M.W., A.T., A.B., R.D., J.R.-Z., J.D., D.B., A.K.-K.; Data curation: J.K., A.R.; Formal analysis: J.K., A.T.; Validation: A.G., E.B.; Resources: J.K., E.B.; Writing – original draft: J.K., A.G.; Writing – review and editing: all authors; Supervision: E.B.; Project administration: J.K., A.G.; Funding acquisition: J.K., E.B. All authors have read and approved the final version of the manuscript.

Availability of Data and Materials

The data supporting the findings of this study are available from the corresponding author upon reasonable request. (No separate source-data file accompanies this article).

Artificial Intelligence (AI) Statement

The authors affirm that no Artificial Intelligence (AI) tools were employed in the preparation of this manuscript, which includes language editing, graphical abstract preparation, figure generation, or image editing.

Conflicts of Interest

The authors declare that they do not have any conflict of interest.

ORCID

Jakub Karasiński: 0000-0003-3143-9131
Andrzej Gawor: 0000-0001-9578-4968
Anna Ruszczyńska: 0000-0002-5199-9681
Marcin Wojciechowski: 0000-0001-9329-6135
Andrii Tupys: 0000-0002-8801-0961
Agnieszka Borowiec: 0000-0001-6483-3966
Ryszard Dobrowolski: 0000-0002-1625-5070
Joanna Reszko-Zygmunt: 0000-0002-5337-4458
Joanna Dobrzyńska: 0000-0002-6148-0699
Dariusz Banaś: 0000-0003-1566-5446
Aldona Kubala-Kukuś: 0000-0003-1547-3348
Ewa Bulska: 0000-0003-4872-5521

References

Analytika. (2020). *Certificate of Analysis: METRANAL® CRM ANBM01, Strawberry Leaves*. Prague: ANALYTIKA, spol. s r.o.

Bilsel, M., Gökçen, T., Binici, B., Isleyen, A., Piechotta, C., Karwai, A. C., Krylov, A., Miheeva, A., Beliakov, M., Palagina, M., Tkachenko, I, Perkola, N., Lewin, M. & Hua, T. (2022). High polarity analyte(s) in aqueous media: Determination of L-PFOA

and L-PFOS in ground water. *Metrologia*, *59*(1A), 08016. DOI: 10.1088/0026-1394/59/1A/08016.

Can, S. Z., Arı Engin, B., İşleyen, A., Jotanovic, A., Acosta, O., Prina, P., Schvartz, M., Savić, M., Stojanović, M., Forigua, D. A. A., Näykki, T., Sara-Aho, T., Vogl, J., Koenig, M., Rienitz, O., Noordmann, J., Pape, C., Towara, J., Kakoulides, E., Alexopoulos, C. et al. (2023). Determination of elements in river water. *Metrologia*, *60*(1A), 08001. DOI: 10.1088/0026-1394/60/1A/08001.

Chew, G., Sim, L. P., Ng, S. Y., Ding, Y., Shin, R. Y. C. & Lee, T. K. (2016). Development of a mushroom powder certified reference material for calcium, arsenic, cadmium and lead measurements. *Food Chemistry*, *190*(8), 293–299. DOI: 10.1016/j.foodchem.2015.05.071.

Dybczyński, R., Polkowska-Motrenko, H., Samczyński, Z. & Szopa, Z. (1998). Virginia tobacco leaves (CTA-VTL-2)–a new Polish CRM for inorganic trace analysis including microanalysis. *Fresenius' Journal of Analytical Chemistry*, *360*, 384–387. DOI: 10.1007/s002160050718.

Ellison, S. L. R., Burke, S., Walker, R. F., Heydorn, K., Månsson, M., Pauwels, J., Wegscheider, W. & te Nijenhuis, B. (2001). Uncertainty for reference materials certified by interlaboratory study: Recommendations of an international study group. *Accreditation and Quality Assurance*, *6*(6), 274–277. DOI: 10.1007/s007690100341.

Ellison, S. L. R. & Williams, A. (Eds.). (2012). *Eurachem/CITAC Guide: Quantifying Uncertainty in Analytical Measurement (3rd ed.)*, Teddington, UK: Eurachem/CITAC. ISBN 978-0-948926-30-3. https://www.eurachem.org.

Emons, H., Held, A. & Ulberth, F. (2006). Reference materials as crucial tools for quality assurance and control in food analysis. *Pure and Applied Chemistry*, *78*(1), 135–143. DOI: 10.1351/pac200678010135.

Emteborg, H., Seghers, J., García-Ruiz, S., Elordui-Zapatarietxe, S., Breidbach, A., Labibes, K., Charoud-Got, J. & Koeber, R. (2024). Paving the way for new and challenging matrix reference materials—particle suspensions at the core of material processing providing RMs for method development and method validation. *Analytical and Bioanalytical Chemistry*, *416*(9), 2079–2088. DOI: 10.1007/s00216-023-05046-2.

EN 13805:2014. (2014). *Foodstuffs—Determination of Trace Elements—Pressure Digestion*. Brussels: CEN.

Grimalt, S., Harbeck, S., Shegunova, P., Seghers, J., Sejerøe-Olsen, B., Emteborg, H. & Dabrio, M. (2015). Development of a new cucumber reference material for pesticide residue analysis: Feasibility study for material processing, homogeneity and stability assessment. *Analytical and Bioanalytical Chemistry*, *407*(11), 3083–3091. DOI: 10.1007/s00216-015-8476-x.

Herzig, R., Rehnert, A., Korhammer, S., Kumpulainen, J., Schramel, P., Muntau, H., Linsinger, T. P. J. & Quevauviller, P. (2002). Certification of a new cabbage reference material for the quality control of trace-element determinations with some considerations on moisture. *TrAC Trends in Analytical Chemistry*, *21*(11), 746–761. DOI: 10.1016/S0165-9936(02)01103-2.

Hietaniemi, V., Kostamo, A., Kumpulainen, J. & Pihlava, J.-M. (1995). Preparation and testing for homogeneity of spiked strawberry and cabbage candidate reference materials. *Fresenius' Journal of Analytical Chemistry*, *352*(1–2), 40–42. DOI: 10.1007/bf00322293.

Isleyen, A. (2017). Certified reference material production in Turkey. *Journal of Chemical Metrology*, *11*(2), 86–88. DOI: 10.25135/jcm.12.17.10.056.

Isleyen, A., Can, S. Z., Cankur, O., Ari Engin, B., Vogl, J., Koenig, M., Horvat, M., Jacimovic, R., Zuliani, T., Fajon, V., Jotanovic, A., Gaževic, L., Milosevic, M., Ochsenkuehn–Petropoulou, M., Tsopelas, F., Lymberopoulou, T., Tsakanika, L.-A., Serifi, O., Ochsenkuehn, K. M., Bulska, E. et al. (2024). Certification of the total element mass fractions in UME EnvCRM 03 soil sample via a joint research project. *Accreditation and Quality Assurance*, *29*, 293–301. DOI: 10.1007/s00769-024-01597-8.

ISO 13528:2022. (2022). *Statistical Methods for use in Proficiency Testing by Interlaboratory Comparison*. Geneva: ISO.

ISO 17034:2016. (2016). *General Requirements for the Competence of Reference Material Producers*. Geneva: ISO.

ISO 33401:2024. (2024). *Reference Materials—Contents of Certificates, Labels and Accompanying Documentation*. Geneva: ISO.

ISO 33403:2024. (2024). *Reference Materials—Requirements and Recommendations for Use*. Geneva: ISO.

ISO 33405:2024. (2024). *Reference Materials—Approaches for Characterization and Assessment of Homogeneity and Stability*. Geneva: ISO.

ISO 5725-2:2025. (2025). *Accuracy (Trueness and Precision) of Measurement Methods and Results—Part 2: Basic Method for the Determination of Repeatability and Reproducibility of a Standard Measurement Method*. Geneva: ISO.

JCGM 100:2008. (2008). *Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement*. Sèvres: Joint Committee for Guides in Metrology (BIPM).

JCGM GUM-1:2023. (2023). *Guide to the Expression of Uncertainty in Measurement—Part 1: Introduction*. Sèvres: Joint Committee for Guides in Metrology (BIPM). DOI: 10.59161/JCGMGUM-1-2023.

Karasiński, J., Gawor, A., Ruszczyńska, A., Wojciechowski, M., Tupys, A., Dobrowolski, R., Reszko-Zygmunt, J., Dobrzyńska, J., Banaś, D., Kubala-Kukuś, A. & Bulska, E. (2026). Metrological framework for the development of plant-based certified reference materials. Part I. *Journal of Chemical Metrology*, *20*(1), e26023803. DOI: 10.25135/jcm.2602.3803.

Kawamoto, M. S., de Souza, G. B. & de Araujo Nogueira, A. R. (2019). Preparation and evaluation of a new reference material for macro- and micronutrients in fish feed. *Microchemical Journal*, *149*, 104027. DOI: 10.1016/j.microc.2019.104027.

Kiełbasa, A., Gadzała-Kopciuch, R. & Buszewski, B. (2016). Reference materials: Significance, general requirements, and demand. *Critical Reviews in Analytical Chemistry*, *46*, 224–235. DOI: 10.1080/10408347.2015.1045120.

Koepke, A., Lafarge, T., Possolo, A. & Toman, B. (2017). Consensus building for interlaboratory studies, key comparisons, and meta-analysis. *Metrologia*, *54*(3), S34–S62. DOI: 10.1088/1681-7575/aa6c0e.

Lamberty, A., Schimmel, H. & Pauwels, J. (1998). The study of the stability of reference materials by isochronous measurements. *Fresenius' Journal of Analytical Chemistry*, *360*, 359–361. DOI: 10.1007/s002160050711.

Linsinger, T. (2010). *Comparison of a Measurement Result with the Certified Value (ERM Application Note 1, Revised 2010)*. Geel: European Commission, Joint Research Centre, Institute for Reference Materials and Measurements.

Linsinger, T. P. J. & Bacquart, T. (2019). Alternative designs for the assessment of homogeneity: Use of characterisation data. *Accreditation and Quality Assurance*, *24*(4), 281–287. DOI: 10.1007/s00769-019-01382-y.

Linsinger, T. P. J., Pauwels, J., Lamberty, A., Schimmel, H. G., van der Veen, A. M. H. & Siekmann, L. (2001a). Estimating the uncertainty of stability for matrix CRMs. *Fresenius' Journal of Analytical Chemistry*, *370*, 183–188. DOI: 10.1007/s002160100719.

Linsinger, T. P. J., Pauwels, J., van der Veen, A. M. H., Schimmel, H. & Lamberty, A. (2001b). Homogeneity and stability of reference materials. *Accreditation and Quality Assurance*, *6*(1), 20–25. DOI: 10.1007/s007690000261.

López-Romero, R. M., Etchevers, J. D. & Vaquera, H. H. (2005). How to prepare a good plant sample for internal reference material. *Communications in Soil Science and Plant Analysis*, *36*, 431–438. DOI: 10.1081/CSS-200043139.

Miyashita, S., Inagaki, K., Narukawa, T., Zhu, Y., Kuroiwa, T., Hioki, A. & Chiba, K. (2012). Development of a certified reference material (NMIJ CRM 7531-a) for the determination of trace cadmium and other elements in brown rice flour. *Analytical Sciences*, *28*, 1171–1177. DOI: 10.2116/analsci.28.1171.

Polkowska-Motrenko, H., Samczyński, Z., Dybczyński, R. S., Chajduk, E., Danko, B., Kalbarczyk, P., Krata, A. A., Pyszynska, M. & Zuba, I. (2022). Preparation of three new certified reference materials for food and environmental analysis and certification using laboratory intercomparison as well as primary reference measurement procedures. *Food Analytical Methods*, *15*, 377–390. DOI: 10.1007/s12161-021-02081-6.

Quevauviller, P., Herzig, R. & Muntau, H. (1996). Certified reference material of lichen (CRM 482) for the quality control of trace element biomonitoring. *Science of the Total Environment*, *187*(2), 143–152. DOI: 10.1016/0048-9697(96)05139-X.

Rastogi, L., Prasad, A. D., Krishna, D. S., Yadlapalli, S. & Dash, K. (2024). Certified reference material (CRM) of tea powder (BARC-D3201) for K, Ca, P, Mg, Mn, Al, Fe, Ba, Zn, Cu, Sr, Pb, As, Cd, and Hg: Method validation and its production. *Food Control*, *158*, 110241. DOI: 10.1016/j.foodcont.2023.110241.

Rukhin, A. L. (1996). Weighted estimates of multilaboratory consensus values. *Computational Statistics & Data Analysis*, *22*, 471–479. DOI: 10.1016/0167-9473(96)00005-9.

Samczyński, Z., Dybczyński, R. S., Polkowska-Motrenko, H., Chajduk, E., Pyszynska, M., Danko, B., Czerska, E., Kulisa, K., Doner, K. & Kalbarczyk, P. (2012). Two new reference materials based on tobacco leaves: Certification for over a dozen of toxic and essential elements. *The Scientific World Journal*, *2012*, 216380. DOI: 10.1100/2012/216380.

Santos, A. S., dos Santos, L. O., Pereira Junior, J. B., de Oliveira, F. M. & Ferreira, S. L. C. (2020). Development of reference material from powdered milk: Uncertainties and multilaboratory evaluation through confidence ellipses. *Microchemical Journal*, *159*, 105330. DOI: 10.1016/j.microc.2020.105330.

Thompson, M. & Ellison, S. L. R. (2011). Dark uncertainty. *Accreditation and Quality Assurance*, *16*(10), 483–487. DOI: 10.1007/s00769-011-0803-0.

Thompson, M., Ellison, S. L. R., Fajgelj, A., Willetts, P. & Wood, R. (1999). Harmonized guidelines for the use of recovery information in analytical measurement. *Pure and Applied Chemistry*, *71*(2), 337–348. DOI: 10.1351/pac199971020337.

van der Veen, A. M. H., Linsinger, T. P. J. & Pauwels, J. (2001a). Uncertainty calculations in the certification of reference materials. 2. Homogeneity study. *Accreditation and Quality Assurance*, *6*(1), 26–30. DOI: 10.1007/s007690000238.

van der Veen, A. M. H., Linsinger, T. P. J., Lamberty, A. & Pauwels, J. (2001b). Uncertainty calculations in the certification of reference materials. 3. Stability study. *Accreditation and Quality Assurance*, *6*(6), 257–263. DOI: 10.1007/s007690000292.

van der Veen, A. M. H., Linsinger, T. P. J., Schimmel, H., Lamberty, A. & Pauwels, J. (2001c). Uncertainty calculations in the certification of reference materials. 4. Characterisation and certification. *Accreditation and Quality Assurance*, *6*(7), 290–294. DOI: 10.1007/PL00010459.

van der Veen, A. M. H. & Pauwels, J. (2000). Uncertainty calculations in the certification of reference materials. 1. Principles of analysis of variance. *Accreditation and Quality Assurance*, *5*(12), 464–469. DOI: 10.1007/s007690000237.

Wise, S. A., Hosbas Coskun, S., Hayes, H. V., Wilson, W. B., Murray, J. A., Lippert, J. A., Burdette, C. Q., Schantz, M. M., Murphy, K. E., Christopher, S. J., Yu, L. L., Rimmer, C. A., Pasiakos, S. M. & Kuszak, A. J. (2025). Development of reference materials for dietary supplements—analytical challenges, use, limitations, and future needs. *Analytical and Bioanalytical Chemistry*, *417*(12), 2439–2471. DOI: 10.1007/s00216-025-05787-2.

Yatkin, S., Trzepla, K., Hyslop, N. P., White, W. H., Butler, O., Ancelet, T., Davy, P., Gerboles, M., Kohl, S. D., McWilliams, A., Saucedo, L., Van Der Haar, M. & Jonkers, A. (2020). Comparison of a priori and interlaboratory-measurement-consensus approaches for value assignment of multi-element reference materials on PTFE filters. *Microchemical Journal*, *158*(4), 105225. DOI: 10.1016/j.microc.2020.105225.

Zhu, Y., Narukawa, T., Inagaki, K., Kuroiwa, T. & Chiba, K. (2011). Development of a certified reference material (NMIJ CRM 7505-a) for the determination of trace elements in tea leaves. *Analytical Sciences*, *27*(11), 1149–1155. DOI: 10.2116/analsci.27.1149.

Zhu, Y., Narukawa, T., Miyashita, S., Kuroiwa, T., Inagaki, K., Chiba, K. & Hioki, A. (2013). Development of a certified reference material (NMIJ CRM 7512-a) for the determination of trace elements in milk powder. *Analytical Sciences*, *29*, 247–253. DOI: 10.2116/analsci.29.247.