A bilateral comparison of ^{227}Th activity standards between National Physical Laboratory and National Institute of Standards and Technology

To cite this article: Andrew J Fenwick *et al* 2025 *Metrologia* **62** 025002

OPEN ACCESS
IOP Publishing | Bureau International des Poids et Mesures Metrologia
Metrologia **62** (2025) 025002 (10pp) https://doi.org/10.1088/1681-7575/adac65

A bilateral comparison of ^{227}Th activity standards between National Physical Laboratory and National Institute of Standards and Technology

Andrew J Fenwick[1,*], Denis E Bergeron[2], Emma Bendall[1], Brittany A Broder[2], Jeffrey T Cessna[2], Seán M Collins[1,3], Leticia Pibida[2], Natasha Ramirez[1] and Elisa Napoli[4]

[1] National Physical Laboratory, Teddington, United Kingdom
[2] National Institute of Standards and Technology, Gaithersburg, United States of America
[3] School of Mathematics and Physics, University of Surrey, Guildford, United Kingdom
[4] Bayer AS, Oslo, Norway

E-mail: andrew.fenwick@npl.co.uk

Received 4 October 2024, revised 20 January 2025
Accepted for publication 21 January 2025
Published 3 February 2025

Abstract
The National Physical Laboratory (United Kingdom) and the National Institute of Standards and Technology (United States) each determined the massic activity of a common solution of ^{227}Th. Measurements at both laboratories were performed before radioactive equilibrium. The direct comparison showed good accord between the laboratories' activity standards with a *t*-test score of 0.15 based on ionization chamber measurements. Challenges associated with pre-equilibrium comparisons are addressed and multiple comparison approaches are presented. The use of a hybrid comparison methodology using a model to correct asynchronously acquired data to an intermediate timepoint mitigates major sources of uncertainty and potential bias in extreme cases.

Supplementary material for this article is available online

Keywords: Th-227, targeted alpha therapy, radionuclide calibrator, ionisation chamber, time-dependent activity calibration, international standards, comparison

* Author to whom any correspondence should be addressed.

1. Introduction

The therapeutic use of alpha-emitting radiopharmaceuticals has become of great interest due to the demonstrated efficacy of these therapies for cancer, such as the Food and Drug Administration (FDA) approved Xofigo^{TM} ($^{223}RaCl_2$)[5] for metastatic castration-resistant prostate cancer [1]. Due to its suitable physical and chemical characteristics thorium-227 (figure 1), the radioactive parent of ^{223}Ra, is currently being investigated for potential applications in the targeted treatment of breast cancer, prostate cancer, and other metastasised cancers [2].

To enable nuclear medicine departments and manufacturers to have confidence in the activity measured prior to patient administration, it is vital to establish primary activity standards and obtain international equivalence. During radiopharmaceutical drug approval, the FDA in the United States of America (USA) along with the Medicines and Healthcare Regulatory Authority in the United Kingdom (UK) and European Medicines Agency (EMA) in Europe, require traceability of activity measurement to be demonstrated. Furthermore, to improve the confidence in results obtained during clinical trials, for performing more accurate dosimetry calculations and for ongoing optimisation of the product, it is desirable to ensure harmonisation in activity measurement is achieved globally. To ensure equivalence of radioactivity measurements around the world, comparison exercises between participating countries are undertaken through intercomparisons, usually involving multiple laboratories and facilitated by the Consultative Committee on Ionising Radiation (CCRI), Regional Metrology Organizations, or by sending an ampoule to the International Bureau of Weights and Measures (BIPM) for comparison with the ionisation chamber-based international reference system (SIR) [3, 4].

For radionuclides which have decay progeny that do not achieve radioactive equilibrium after radiochemical preparation within a realistic operational time period (this will be radionuclide dependent), the observed response in an ionisation chamber (or a radionuclide calibrator) will not follow the expected exponential decay of the parent radionuclide, and therefore the usual route for comparison through the SIR is not possible until a robust methodology for such a comparison has been established. A key comparison exercise is also difficult to arrange due to the limited availability of these radionuclides on the open market, and the inherent complexity due to the ratio of decay progeny changing during preparation and shipment without an agreed method to calculate the time zero (t_0) of the parent-progeny 'clock' (i.e. the theoretical time when the parent exists with no progeny). The objective of this work was to show that such a comparison between National Metrology Institutes (NMIs) or Designated Institutes (DIs) is feasible and to show equivalence between the standards of the USA and UK.

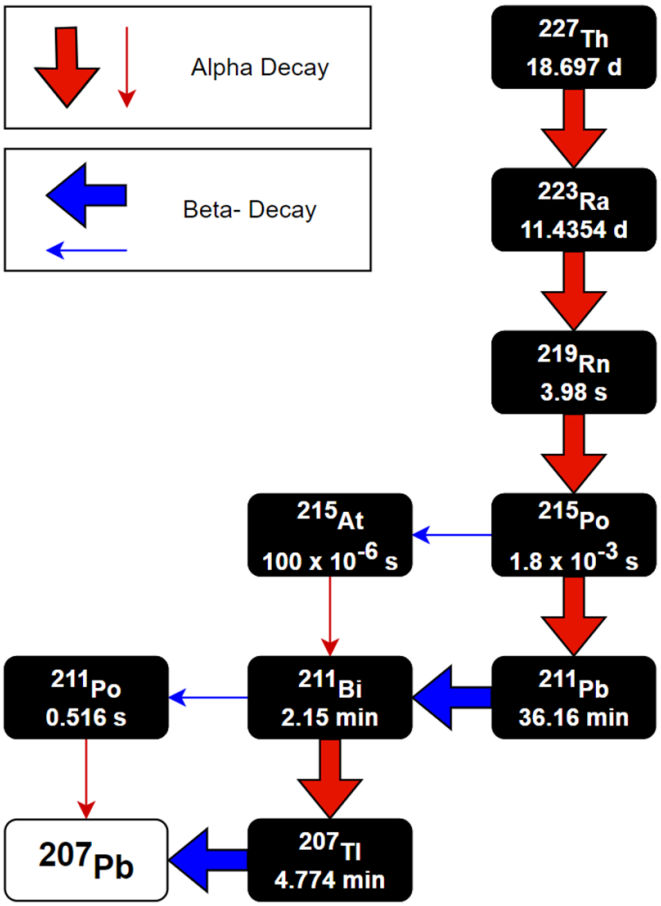

Figure 1. Decay chain for ^{227}Th including nominal half-life values taken from [11]. Large arrows indicate the principal decay pathway (branching ratio >99.7%).

The first reported primary standardisation of ^{227}Th activity was from the Physikalisch-Technische Bundesanstalt [5]. The National Physical Laboratory (NPL) and National Institute of Standards and Technology (NIST) have independently performed primary activity measurements of ^{227}Th and measurement of gamma emission probabilities [6–8]. Due to the complexities of performing a worldwide comparison at this time, a bi-lateral comparison was proposed between NPL and NIST to establish equivalence and test methods for performing comparisons of non-equilibrium radionuclides. Gravimetrically linked samples of ^{227}Th were prepared at NIST and sent to NPL for blind activity comparison. The comparison tested the ability of each laboratory to determine the activity and t_0 independently. Presented here is a detailed discussion of the advantages and disadvantages of several different comparison approaches, including the conventional 'direct' comparison with an exchange of material and alternative instrument-based approaches. This work follows a previous successful

[5] Certain commercial entities, equipment, or materials may be identified in this document in order to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

bi-lateral comparison between the two NMIs for ^{224}Ra in 2020 [9].

All uncertainties are stated as standard uncertainties or combined standard uncertainties as defined in the Guide to the Expression of Uncertainty in Measurement [10].

2. Method

2.1. Experimental method at NIST

2.1.1. Radiochemical purification and source preparation. A dry powder of ^{227}Th (provided by Bayer AS) was received by NIST in January 2023 for the comparison. The powder was dissolved in 3 mol l^{-1} HNO_3 at room temperature. The radiochemical isolation of the ^{227}Th from the decay progeny was performed by extraction chromatography utilising the TEVA resin [8, 12]. Approximately 50 MBq of ^{227}Th was loaded onto the column and rinsed using 3 mol l^{-1} HNO_3 to remove the decay progeny. The purified ^{227}Th was eluted using 10 ml of 3.1 mol l^{-1} HCl, with the time of the end of rinse (i.e. the end of the chemical separation) being recorded as a reference separation time (t_{sep}). This eluted solution was diluted to approximately 20 ml using the same HCl solution. The sample was then gravimetrically dispensed into Type I 5 ml flame-sealed ampoules [13] for ionisation chamber and gamma spectrometry measurements, and 10 ml ISO standard Schott© vials (VCDIN10R—hereinafter referred to as 10R vials) [14] for ionisation chamber measurement. A dilution was made using one of the prepared ampoules, bringing the massic activity to approximately 3 kBq g^{-1}. This was then gravimetrically dispensed into liquid scintillation (LS) sources for primary measurements which were prepared in standard 22 ml glass vials with 10 ml Hionic Flour. Two sources were prepared with approximately 0.1 g and 0.4 g of the ^{227}Th solution. To the 0.1 g source, 3 mol l^{-1} HCl was added to match the acid content of the 0.4 g source and distilled deionised water was added to each source to achieve a total aqueous fraction of approximately 0.09. A composition-matched blank was also prepared.

2.1.2. Primary measurement techniques. The sources shipped to NPL for the comparison were prepared with gravimetric links to sources measured by primary methods (figure S1). The ^{227}Th primary activity standardisation at NIST [6] was based on the triple-to-double coincidence ratio (TDCR) method of LS counting, with confirmation by CIEMAT/NIST efficiency tracing with ^{3}H [15] and live-timed $4\pi\alpha\beta$(LS)-γ(NaI) anticoincidence counting (LTAC) [16]. The ^{227}Th massic activity was determined by both TDCR and LTAC as part of this comparison with measurements being made between 1.8 d $< t_{sep} <$ 15.9 d. After ^{227}Th is separated from its progeny, LS efficiencies (defined as LS counts per ^{227}Th decay) change rapidly with the ingrowth of progeny. The efficiency model is thus time-dependent, requiring a good estimate for the LS efficiencies of the beta-emitting progeny, ^{211}Pb and ^{207}Tl, and for reliable decay data for the calculation of their relative activities using the Bateman equations [17].

Efficiency calculations for the β-decaying progeny of ^{227}Th were carried out with the MICELLE2 code [18]. A simplified decay scheme was adopted for ^{211}Pb with branch normalizations chosen to assign 'missing' decays to similar cascades. Conversion coefficients for the γ transitions were calculated using the internal conversion coefficient database known as 'BrIcc' [19]. For ^{207}Tl, the 0.271(10) % of β decays feeding excited levels of ^{207}Pb were ignored and LS efficiencies were calculated assuming 100% decay directly to the ground state. The LS efficiencies require further correction for the relatively short-lived progeny (^{219}Rn, ^{215}Po, ^{215}At, and ^{211}Po) to account for decays that occur during the deadtime triggered by the decay of their progenitor [5]. The single-figure-of-merit approach adopted for TDCR counting of ^{227}Th is presented in detail in a separate publication [6].

LTAC counting was performed with a three-photomultiplier tube detector with a NaI(Tl) detector positioned below the optical chamber. List mode data were acquired with a CAEN desktop digitizer (DT5724). Anticoincidence gates were set to cover peaks arising from both beta and alpha emitters, but the extrapolations used to determine activities reported here rely on a gate covering mostly the γ rays emitted by ^{211}Pb. This experimental configuration allows a good monitor of the β detection efficiency with reduced sensitivity to background.

2.1.3. Ionisation chamber measurements. All ampoules and vials were measured on the Vinten 671 (SN: 3–2) ionisation chamber (VIC), connected to a Keithley 6517B electrometer with a LabVIEW interface. Each VIC measurement included 1000 individual current readings with a 0.5 s integration time. One vial was measured repeatedly (for approximately 13 d) to monitor and test ingrowth correction data. Ionisation chamber measurements at NIST were performed between 0.08 d $< t_{sep} <$ 8 d.

The vial activity was determined from time-dependent calibration coefficients ($K_{VIC}(t)$, with units pA MBq^{1} to convert the instrument response to ^{227}Th and its progeny to ^{227}Th activity) determined in previous experiments [6]. For the bilateral comparison, the calibration coefficient for ^{227}Th in isolation, i.e. K_{VIC} at t_0, was determined by fitting experimental data see [6]. To generate the fits, the K_{VIC} for each nuclide was calculated based on gamma-ray emission intensity tabulations and a response curve generated from Monte Carlo simulations in a DOSRZnrc [20] model which is designed to score dose in a generalised cylindrical geometry [21]. The Bateman equations were solved using the appropriate half-lives and branching ratios taken from the Decay Data Evaluation Project (DDEP) [11] to obtain the time-evolving fractional activities relative to the parent nuclide. The total instrument response at a given time point was calculated by summing the products of the fractional activity and K_{VIC} for all nuclides. Th-227,

^{223}Ra, ^{219}Rn, ^{211}Pb, and ^{211}Bi all decay with gamma-ray emissions that contribute significantly to the ionisation chamber response. The validity of the model was established by comparing the K_{VIC} calculated for ^{223}Ra in equilibrium with its progeny to the experimental value [22]; the model gave $K_{\mathrm{VIC}} = 3.16(1)$ pA MBq^{-1} (where the stated uncertainty is from the TOPAS statistics only and does not include contributions from, e.g. input decay data), consistent with the experimental value of 3.17(4) pA MBq[1].

For ^{227}Th, K_{VIC} at t_{sep} was determined from a weighted non-linear least squares fit. Initially, the ^{227}Th and ^{223}Ra K_{VIC} were both considered free parameters and the K_{VIC}s for the other nuclides were estimated from the Monte Carlo model. The combined standard uncertainty on the ^{227}Th K_{VIC} was estimated from five components: uncertainty due to the background current, standard uncertainty on the ^{223}Ra K_{VIC}, an estimated uncertainty ($\pm$20 min) on the t_{sep}, the standard uncertainty on the ^{227}Th half-life, and the fit uncertainty. The uncertainty due to the ^{223}Ra K_{VIC} included the standard uncertainty from the experimental value and a component for the small difference between the K_{VIC} for ^{227}Th calculated with and without allowing the K_{VIC} for ^{223}Ra to vary in the fit.

2.1.4. Gamma spectrometry measurements. All gamma spectrometry measurements at NIST were performed using an n-type high-purity germanium (HPGe) detector located within a 10 cm thick lead shield with a cadmium and copper lining with reproducible source positioning jigs. Measurements were performed to determine the source impurities and to confirm t_0 is consistent with t_{sep}. The sources were measured for 2 h at different distances and were repeated a day later to confirm the initial measurement. Measurements were made at 25 cm and 40 cm source-to-detector measured from the bottom of the source to the detector end cap. The measurement distance depended on the source activity. The Genie-2000 software was used for peak fitting.

2.1.4.1. Effective time zero determination. The calculations for t_0 were performed following the method described in [23]. Gamma-ray emission data published by DDEP [11], ENSDF [24] and NPL [7, 25] were used in the determination of t_0.

2.2. Experimental method at NPL

2.2.1. Source preparation. Two 10R vials (table 1) were received at NPL on 25-01-2023. Following measurement by ionisation chamber, as much solution as possible was extracted from vial 22-0205.5(6) using an 11 gauge hypodermic needle piercing the septum with a plastic pycnometer inserted through. The solution was then gravimetrically dispensed from the pycnometer as a 1 g aliquot to a 2 ml ISO ampoule and the remainder (approx. 3 g) to a 5 ml ISO ampoule [26]. The vial was refilled with 4.0732(23) g inactive carrier solution (3 mol l^{-1} HCl) and the residual response was measured on the ionisation chamber.

Table 1. Details of vials received at NPL from NIST.

Vial ID	Mass (NIST)/g	Mass (NPL extracted)/g
22–0205.4(6)	3.9659(20)	—
22–0205.5(6)	3.9975(20)	3.9730 (42)

2.2.2. Ionisation chamber measurements. Calibration of the NPL Vinten 671 (SN: 3–5) ionisation chamber was performed in 2016 using sources traceable to primary ^{227}Th massic activity measurements carried out by NPL using LS counting techniques, with ingrowth correction curves measured for each geometry as described in [8].

The two 10R vials received by NPL (table 1) were measured in the Vinten 671 (SN: 3–5) ionisation chamber at NPL which is connected to an external feedback capacitor circuit via a Keithley 6514 electrometer and controlled using LabVIEW software. Following initial measurement, one vial (22-0205.5(6)) was opened and dispensed to other geometries as described previously. Further measurements of the 2 ml and 5 ml ISO ampoules prepared from vial number 22-0205.5(6) and vial 22-0205.4(6) were made on the Vinten 671 ionisation chamber. The 10R vial was measured on successive days between 7 d $< t_{\mathrm{sep}} <$ 30 d and the 5 ml ISO ampoule was measured on successive days between 7 d $< t_{\mathrm{sep}} <$ 16 d to test ingrowth correction and confirm results. Only the 5 ml ISO ampoule and 10R vial (22-0205.4(6)) were used in the final activity determinations.

2.2.3. Gamma spectrometry measurements. All gamma spectrometry measurements were performed using the 9.5% relative efficiency p-type semi-planar HPGe gamma-ray spectrometer 'THOR'. This detector had previously been calibrated for its full-energy peak (FEP) detection efficiency, for a source-to-detector geometry of 40.0 cm in a matched geometry and matrix for 1 g solution in a 2 ml ISO ampoule, using a suite of traceable primary and secondary gamma-ray emitting radionuclide standards covering the energy range 22 keV to 1836 keV. The source-to-detector distance provided a solid angle of 0.13%, which results in negligible true coincidence summing effects.

The detector was contained in a shield with 100 mm thick walls, with internal dimension of 1220 mm $\times$ 915.5 mm $\times$ 904.5 mm, with the internal walls covered with a 0.5 mm Cd and 0.7 mm Cu graded liner to reduce interferences from background radiation and Pb fluorescence x-rays in the spectrum. An aluminium optical breadboard was mounted in line with the detector along the horizontal plane with kinematic mounting plates holding a precision-engineered sample holder to provide a reproducible geometric source positioning.

Two spectra were collected on the 25-01-2023 (approximately 55 500 s live time) and 26-01-2023 (approximately 10 000 s live time) using a MIRION LYNX digital signal analyser operating with the loss-free counting option to correct for dead time and pulse-pile up losses during measurement.

Table 2. Gamma-ray emissions and nuclear decay data used to determine the t_0 at NPL.

Radionuclide	$T_{1/2}$ /d	Energy /keV	I_γ /100 decays
^{227}Th	18.697	210.6	1.1843(54)
		234.7 + 236.0	12.912(58)
		256.2	6.642(30)
		300.0	2.1522(80)
		329.9	2.696(10)
^{223}Ra	11.4354	144.3	3.481(12)
		154.2	6.032(24)
		269.5	13.312(57)
(^{219}Rn)		271.2	10.755(47)

The spectra were analysed using the MIRION GENIE 2000 v3.4.1 software package, utilising the Interactive Peak Fitting application to manually adjust the FEP fits where required.

2.2.3.1. Effective time zero determination. The t_0 and standard uncertainty of the solution were determined from the activity ratio of the parent and the decay progeny at the time of measurement using the equations in [23]. First, the activity of ^{227}Th and the combined activity of ^{223}Ra and ^{219}Rn progenies were determined using the half-lives and gamma-ray emission intensities determined at the NPL [7, 25, 27]. The gamma-ray emissions and nuclear decay data are provided in table 2. The FEPs of 234.7 keV and 236.0 keV were combined to remove possible errors in the deconvolution of the two FEPs. The 271.2 keV FEP from the decay of ^{219}Rn was used as this was considered to be in equilibrium with the ^{223}Ra throughout the measurement due to its 3.98(3) s half-life [11]. Multiple gamma-ray emissions were combined to derive the activity to reduce any bias to the t_0 imposed by any inaccuracies in any one gamma-ray emission intensity. The activities determined from each gamma-ray of ^{227}Th and the decay progenies in the initial measurement on 25-01-2023 are shown in figure 2.

From equation (11) in [23], the time from the start of the measurement to the t_0 was determined. The standard uncertainty of the t_0 was estimated using the extended propagation formula in section 3.2 of [23].

2.2.3.2. Activity measurement and contamination check. The use of HPGe gamma-ray spectrometry for determining the activity of ^{227}Th is relatively convenient as no corrections for the ingrowth of ^{223}Ra and decay progenies are required and the main characteristic gamma-ray emissions from ^{227}Th are distinct from its progeny. From the two measurements on THOR, the activity of the solution at t_0 was determined using the gamma-ray emissions and nuclear decay data in table 2. The use of the absolute gamma-ray emission intensities previously determined at NPL provide a link to the primary standard of ^{227}Th and thus provide another method to compare the status of the NPL and NIST standards [28].

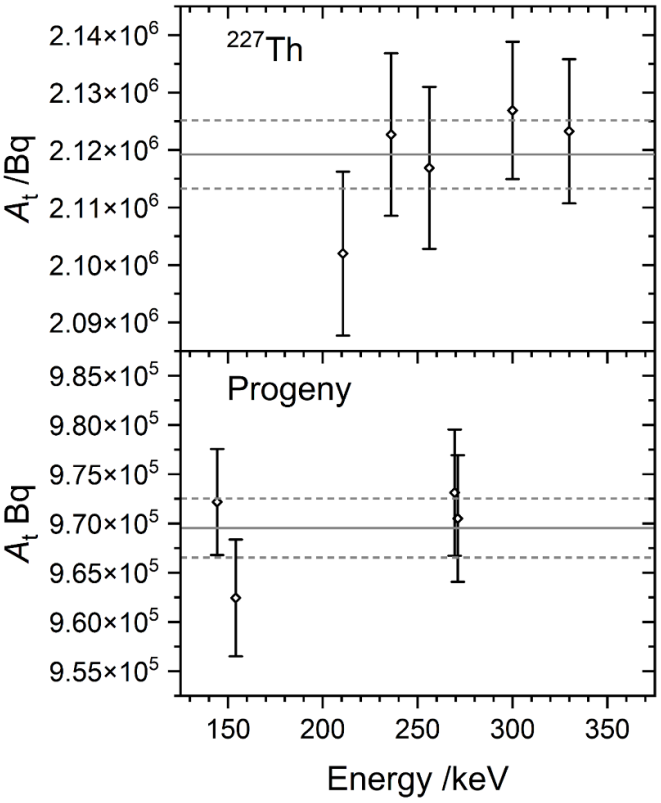

Figure 2. Activities determined for the individual gamma-ray emissions of ^{227}Th and decay progenies at 25-01-2023 17:45:25 UTC. The solid and dashed lines represent the weighted mean of the individual activities and the standard uncertainty of the weighted mean respectively.

No gamma-ray emitting contaminants, other than the decay progenies, were identified in the spectra. Due to the weak gamma-ray emissions of the ^{227}Ac precursor and the multitude of gamma rays from the decay of ^{227}Th and progeny, resulting in a large Compton continuum, any calculation of the detection limit is of limited use to provide a reasonable estimate of its presence. The comparison of the massic activity between the ionisation chamber, over multiple time points, and HPGe gamma-ray spectrometry provides evidence that any presence of ^{227}Ac was not significant.

3. Results

3.1. Primary measurements at NIST

The uncertainty budget for the TDCR-determined massic activity is shown in table 3. The counting statistics and efficiency model are the main contributors to the combined standard uncertainty and are estimated, in part, based on observations in two prior experiments which included more measurements for a more rigorous test of the efficiency model.

The LTAC results were considered confirmatory. Table 4 compares the LTAC results with the TDCR measurements

Table 3. Uncertainty budget for TDCR determined massic activity.

Uncertainty component	Type	Value/%
Decay data	B	0.01
Counting	A	0.22
Efficiency Model	B	0.19
Separation time (t_{sep})	B	0.08
Weighing	B	0.05
Combined standard uncertainty		0.31

Table 4. LTAC results relative to TDCR-determined activities. The stated uncertainty is estimated from the extrapolation uncertainty only.

$t - t_{sep}$/d	A_{LTAC}/A_{TDCR}
7.0	0.9952(13)
15.9	1.0000(9)

at two time points. Approximately one week after separation, the LTAC-determined massic activity was lower than the TDCR-determined massic activity, with the difference greater than the estimated uncertainty for either method. At 15.9 d after separation, however, the methods were in excellent accord. The TDCR and LTAC systems shared hardware in this study, and the priority was to obtain data with the proven (FPGA-based) TDCR system. This resulted in relatively few LTAC measurements so that there is not a good estimate for the repeatability or reproducibility. The stated uncertainties on the ratios given in table 4 are due to the extrapolation only.

3.2. Time zero (t_0)

For the NIST measurements, t_0 was calculated using the 154 keV and 235 keV lines as well as the 269 keV and 235 keV lines. The same procedure was used for measurements performed on the X-detector at 25 cm and 40 cm positions. A weighted mean t_0 of 17-01-2023 18:26 ± 00:27 was determined using the measurements in both positions.

From the NPL measurements, the t_0 determined from the two measurements were consistent, being (1) 17-01-2023 17:59:21 UTC ± 00:52:50 and (2) 17-01-2023 18:04:45 UTC ± 01:04:05. Due to the extensive correlations (specifically the FEP efficiency and gamma-ray emission intensities) between the two determinations the NPL t_0 has been taken from the initial measurement (1). The magnitude of the uncertainty is due to the time difference between the time of separation and the time of measurement.

Comparison of the time zero determined using the more precise published decay data by the NPL to those from the DDEP evaluations of ^{227}Th, ^{223}Ra and ^{219}Rn shows a vast improvement in the accuracy and precision. The time zero determined from measurement (1) using the DDEP evaluated data resulted in a value of 17-01-2023 15:34:29 UTC ± 07:31:59. This time zero is substantially different with a standard uncertainty that is over 8.5 times greater due to the less precise gamma-ray emission intensities.

All calculated values for t_0 were consistent with the 'end of rinse' time recorded by NIST during the separation (t_{sep} = 17-01-2023 18:40 UTC), suggesting that any breakthrough of progeny was insignificant. For the comparison, a singular common reference time was needed, so t_{sep} was adopted with a conservative uncertainty of 20 min.

3.3. Ionisation chamber results

At NIST, the mass ratios for all sources were compared to ionisation chamber responses to confirm solution homogeneity and quantitative transfers of activity. Agreement was consistently better than the statistical uncertainty on the measurements. For the vial that was measured for 13 d, the massic activity was estimated to be 2.912(9) MBq g^{-1} at the t_{sep}, based on the response observed during two previous primary standardisation campaigns at NIST. This massic activity agreed with the TDCR-determined massic activity to within ≈0.02%. Based on the TDCR-determined massic activity and the fitting procedure described in section 2.1.3, K_{VIC} for a 10R vial at t_0 was estimated as 1.1275(86) pA MBq^{-1} and the uncertainty budget for K_{VIC} is given in table 5.

At the NPL, the reference activity at the NPL t_0 was determined using the 5 ml ISO ampoule measured on the Vinten 671 system with a calibration factor (at t_0) of 1.1102(56) pA MBq^{-1} and a time-dependent correction factor derived from a weighted non-linear least squares fit of measured ingrowth data. The ampoule was measured seven times between 7 d and 16 d post-separation (figure 3) and the mean massic activity was calculated to be 2.916(19) MBq g^{-1} using the NPL determined t_0. The measurements indicated that corrections for ingrowth were consistent over this period. The unadulterated 10R 22-0205.4(6) was measured between 7 d and 30 d post-separation and massic activity was determined using the NIST-supplied mass and a calibration factor of 1.1083(55) pA MBq^{-1} combined with a time-dependent correction factor. The massic activity at the NPL determined t_0 was 2.915(19) MBq g^{-1}. A typical uncertainty budget for the massic activity determination for the 5 ml ISO ampoule is shown in table 6.

Table 5. Uncertainty budget for NIST K_{VIC} at t_0 in this bilateral comparison.

Uncertainty component	*Type*	*Value*	
Background	A	0.39	%
Ra-223 K_{VIC}	B	0.31	%
Time zero (t_0)	B	0.21	%
Half-life	B	0.013	%
Fit	A	0.54	%
Combined standard uncertainty		0.76	%

Table 6. Uncertainty budget for ^{227}Th massic activity determination at NPL (contributions less than 0.05% have been excluded).

Uncertainty component	Type	Value	
^{227}Th calibration factor	B	0.50	%
Correction factor fitting uncertainty	B	0.25	%
Time zero (t_0)	B	0.30	%
Volume correction	B	0.10	%
Statistical	A	0.02	%
Weighing	B	0.05	%
Capacitance	B	0.10	%
Combined standard uncertainty		**0.65**	%

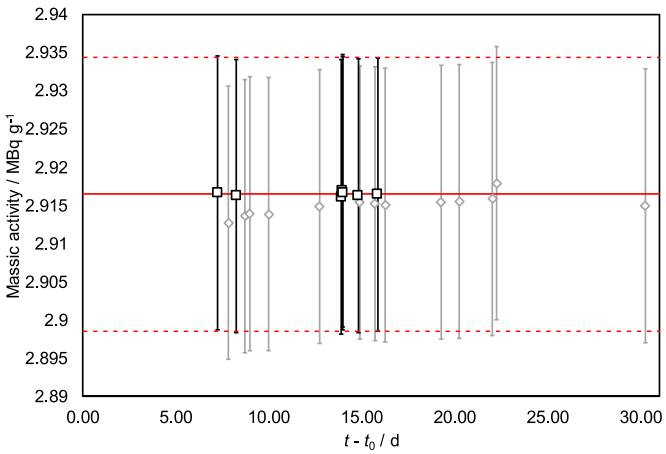

Figure 3. NPL massic activity determination by ionisation chamber (reported at t_0). Black squares are measurements of the 5 ml ISO ampoule, grey diamonds are measurements determined with the 10R vial as received from NIST using the NIST-supplied mass. The red lines indicate the mean value of the 5 ml ISO ampoule measurements with associated uncertainty.

3.4. Gamma spectrometry activity measurements

At NPL, a massic activity of 2.901(13) MBq g^{-1} was determined at the reference time of 17-01-2023 17:59:21 UTC, from a weighted mean of the gamma-ray emissions, with the standard uncertainty calculated to account for correlations between gamma-ray emissions. NIST did not calculate a massic activity using gamma spectrometry techniques.

4. Discussion

4.1. Comparison of activity values

The comparison of activities at t_{sep} is shown in figure 4 and demonstrates agreement between the values. The NPL ionisation chamber result gives a t-test score of 0.15 relative to the NIST result. The values were determined using independent time zero calculations and independent ionisation chamber response curves, demonstrating that this technique for performing a comparison is viable. The NPL and NIST calibration factors for a 10R vial on the Vinten ionisation chamber (determined for t_0) are significantly different (1.7%) and this

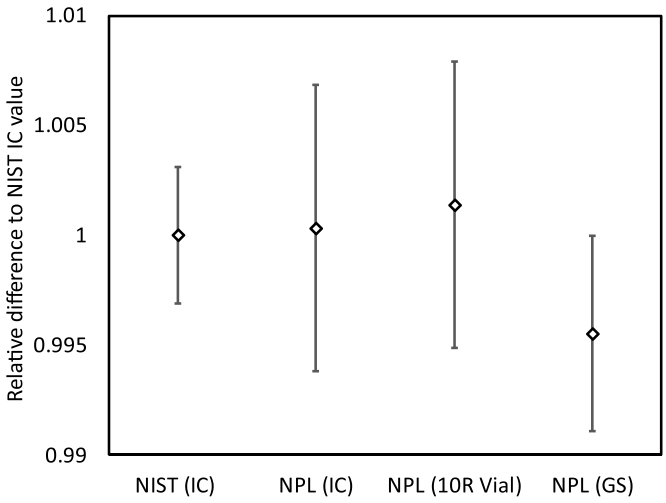

Figure 4. Results relative to the NIST value determined by ionisation chamber. (IC) indicates the NPL result using the 5 ml ISO ampoule in the ionisation chamber system and is the NPL reference value. (10R Vial) is the result determined using the unadulterated NIST vial and the NIST supplied vial mass. (GS) is the NPL result from gamma spectrometry measurements. Uncertainties are shown as standard uncertainty.

could be due to slight differences in the measurement setup affecting the response from the low energy x-ray emission from ^{227}Th.

This bilateral comparison was based on a direct exchange of material, which follows the approach taken for CCRI K2 Key Comparison exercises. However, other methods of comparison are possible, as exemplified by the familiar K1 comparisons facilitated by the SIR [3, 4]. The following two approaches consider methodologies for achieving such comparisons using a common instrument and tightly specified source geometry (container, volume, and chemical composition) for radionuclides that must be compared before radioactive equilibrium is reached. These are the extreme cases, illustrating the advantages of synchronous (with respect to $t - t_0$) (section 4.1.1) and model-assisted asynchronous (section 4.1.2) comparisons. We propose that the best comparison approach may ultimately represent a compromise between these two extremes.

Here we use data from our bilateral comparison of ^{227}Th to provide examples and borrow from the practices used in BIPM comparisons.

In all approaches, laboratories independently determine t_0 and report $A(t_0)$ and t_0.

4.1.1. Synchronous measurements. Synchronous measurements involve measurement on a common instrument (like the SIR or SIRTI) or on instruments with well-known response relationships. At NIST, experiment-to-experiment consistency was established by comparing values for $K_{\mathrm{VIC}}(t)$; analyses showed z-scores <1 in all cases, indicating accord. In the present NIST-NPL bilateral comparison, comparisons could be derived from data acquired on the Vinten 671 ionisation chambers at both laboratories which have previously been

Table 7. Ratio of the response of Vinten 671 ionisation chambers at NIST (3–2) and NPL (3–5) for 6 radionuclides.

	^{125}I	^{241}Am	^{57}Co	^{137}Cs	^{60}Co	^{226}Ra
$\frac{NIST\,Vinten\,(3-2)}{NPL\,Vinten\,(3-5)}$	0.9998	0.9998	1.0014	1.0014	1.0012	1.0017

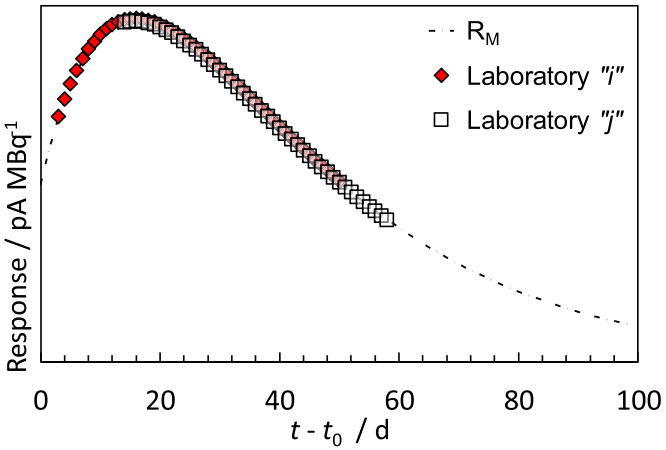

Figure 5. Figure showing the comparison of ionisation chamber measurements of ^{227}Th as a function of time between two laboratories and a theoretical model, $R_M(t)$.

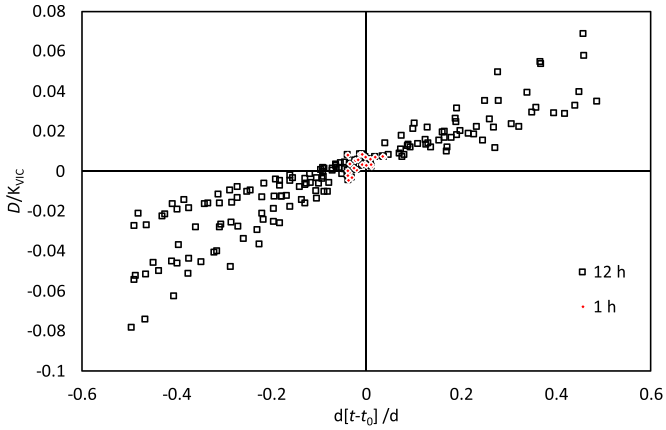

Figure 6. Figure demonstrating relative deviation ($D_{i,j}/K_{VIC}$) as a function of measurement time difference (d[t − t$_0$]). Black squares show data acquired with a time difference of 12 h and red diamonds depict a 1 h time difference.

compared in terms of current response for a range of radionuclides covering the typical operating energies of this chamber (table 7).

A source must be measured repeatedly over a period of time and the activity of ^{227}Th, $A_{Th}(t)$, is calculated for each measurement time, $(t-t_0)$, so that the IC response, $R(t)$, is calculated from the net measured current, $I(t)$, and the activity (decay-corrected to the measurement time using the ^{227}Th half-life (figure 5)).

$$R(t-t_0) = I(t-t_0)/A_{Th}(t-t_0). \tag{1}$$

For two laboratories, i and j, at any time, $(t-t_0)$, a difference between the comparing laboratories, $D_{ij}(t-t_0)$, can be calculated as

$$\begin{aligned} D_{ij}(t-t_0) &= R_i(t-t_{0,i}) - R_j(t-t_{0,j}) \\ &= \frac{I_i(t-t_{0,i})}{A_{Th,i}(t-t_{0,i})} - \frac{I_j(t-t_{0,j})}{A_{Th,j}(t-t_{0,j})}. \end{aligned} \tag{2}$$

The average difference is then taken as an average of $D_{ij}(t-t_0)$ for all available times. If there are N values of $(t-t_0)$ with measurements for both labs, then

$$D_{ij} = \frac{\sum D_{ij}(t-t_0)}{N}. \tag{3}$$

In practice, corrections for impurities may be required. In the case of a shared instrument, the equations could also be modified to include the use of a radium reference source to assure response constancy and facilitate normalization.

In this approach, the comparison is based on a time-dependent response, $\bar{R}(t-t_0)$, but need not incorporate any model if measurements with overlapping $(t-t_0)$ are available for any number (n) of comparing laboratories (figure 5)

$$\bar{R}(t-t_0) = \frac{\sum R_x(t-t_0)}{n} \tag{4}$$

where x refers to the laboratory (e.g. $x = i$ or j) and a D_i for laboratory i is calculated by replacing $R_j(t-t_0)$ in equation (2) with $\bar{R}(t-t_0)$.

In the present NIST-NPL bilateral comparison, data acquired with $(t-t_0) = 1.01$ d to 10.09 d were analysed to give $D_{ij} = 0.0091(94)$ pA MBq^{-1} where the stated uncertainty is only the standard deviation on $N = 23$ determinations of $D_{ij}(t-t_0)$. Expressed as a relative deviation, $\frac{D_{ij}(t-t_0)}{K_{VIC}(t-t_0)} = 0.0036(38)$. For this analysis, only data acquired with $(t-t_0)$ within 1 h of each other were included. A less strict cut on the data would include more points but allows increased bias as shown in figure 6. The bias increases with the absolute difference in measurement times (dt). In steeper regions of the ingrowth curve, the sensitivity of D_{ij} will be greater, so that there is substantial spread in the bias calculated at any given dt value (figure 6).

4.1.2. Model-assisted asynchronous measurements. When considering model-assisted asynchronous measurements, a modelled IC response curve, $R_M(t)$, is used to convert net measured currents, $I(t)$, into activities, $A_M(t)$. These

are decay-corrected with the ^{227}Th half-life to t_0 to give $A_M(t_0)$,

$$A_M(t_0) = \sum \frac{I_i(t-t_0)}{R_M(t-t_0)_i} C_{\text{decay}}(t)_i \tag{5}$$

where $C_{\text{decay}}(t) = e^{\lambda_{\text{Th}}(t-t_0)}$.

Normalized activities are approximated from the $A_x(t_0)$ reported by each laboratory as

$$A_i^0(t_0) = \frac{A_i(t_0)}{A_M(t_0)} \tag{6}$$

and the comparison between any number of laboratories is based on the weighted average of normalized activities, $\bar{A}(t_0)$. Agreement between the result of one laboratory and the comparison average, D_i, and differences between comparing laboratories, D_{ij}, are calculated according to

$$D_i = A_i^0(t_0) - \bar{A}(t_0) \tag{7}$$

$$D_{ij} = D_i - D_j = A_i^0(t_0) - A_j^0(t_0). \tag{8}$$

Here again, corrections for impurities and terms for normalization against a radium reference source could be added to the equations.

To a good approximation (see, e.g. [5]), after a few days, $R_M(t)$ can be calculated from the IC response for ^{227}Th and ^{223}Ra(+ progeny) and the decay constants (λ) for ^{227}Th and ^{223}Ra:

$$R_M(t) = Xe^{\left(-\lambda_{(\text{Th}-227)}t\right)} + Ye^{\left(-\lambda_{(\text{Ra}-223)}t\right)} \tag{9}$$

where X and Y are fit parameters (Kossert and Nähle used 'A' and 'B'; 'A' is avoided here to avoid confusion with terms for activity). Thus, in addition to uncertainty for $I(t-t_0)$, $A(t_0)$, $(t-t_0)$, and the half-life of ^{227}Th, components for the half-life of ^{223}Ra, the instrument response for ^{227}Th, the instrument response for ^{223}Ra(+ progeny), and the fit of the data to the response model would have to be evaluated.

In the present NIST-NPL bilateral comparison, applying the model-assisted asynchronous method, based on the $K_{\text{VIC}}(t_0)$ values described in sections 2.1.3 and 2.2.2, gives $D_{ij} = 0.019(10)$.

A concern in either comparison approach may be that an error by the submitting lab in the estimation of t_0 introduces difficult-to-detect bias. This problem is not unique to pre-equilibrium measurements, of course; an error in the reported reference time in any comparison introduces bias to a comparison result, but when the decay corrections involve only a single exponential, the situation is simpler. An error in t_0 estimation should be evident as a trend in $D_{ij}(t-t_0)$, but only if measurements are made at sufficiently spaced time points. The biggest advantage of the model-assisted asynchronous approach is that the comparison is based on a time-independent $\bar{A}(t_0)$, allowing submissions from different laboratories to be compared without the need for overlapping $(t-t_0)$ measurement windows. Superficially, this might also suggest that the asynchronous approach allows a comparison with fewer measurements. However, acquiring measurements over an extended period (e.g. one or two 10 min acquisitions per day for 10 d) allows any trends in $D_{ij}(t-t_0)$ or their variance to manifest regardless of the comparison approach.

A compromise between the two approaches could involve calculation of $D_{ij}(t-t_0)$ at multiple points but relaxing the constraint of well-matched values for $(t-t_0)$ by using a model to interpolate between measurement times for one or both laboratories. Effectively, this approach places the comparison reference time at a time other than t_0 (e.g. $t_{\text{ref}} = (t_0 + 10\text{ d})$). This offers the advantage of being able to acquire data without a strict 'schedule'. In the present comparison, application of this approach to a subset of VIC data acquired with $(t–t_0) = 5$ d to 20 d gave $D_{ij}(t_{10\text{ d}}) = 0.0029(96)$, where the stated uncertainty is statistical only. A full uncertainty would require treatment of the fit uncertainties, which will be less for interpolation than for extrapolation.

The goal of a comparison method is to introduce minimal uncertainty in the measurement process so that the combined standard uncertainty on the comparison values is almost entirely due to the uncertainty reported by the submitting laboratory. This may be possible with either approach presented here. In the present NIST-NPL bilateral comparison, the model and fit uncertainties introduced using the model-assisted asynchronous approach result in larger uncertainties and indicate a statistically significant discrepancy that was not observed with the synchronous approach or with the direct comparison. A hybrid approach that uses a model to correct asynchronously acquired data to a reference time later than t_0 offers an excellent compromise, mitigating the major sources of uncertainty and potential bias from extreme cases.

4.2. Comparison of effective time zero (t_0) determinations

The measurements and calculations performed at NPL indicate that it is feasible to determine a sufficiently accurate time zero using gamma spectrometry to allow ionisation chamber measurements to be corrected, even after more than seven days of progeny ingrowth. The NPL and NIST t_0 calculations are statistically in agreement with the NPL time being earlier than that of NIST but with a higher uncertainty. The NPL results utilised the more recently published data [7] and it was noted that when compared to using the DDEP data [11] the standard uncertainty of the t_0 was dramatically improved.

5. Conclusions

A comparison of the ^{227}Th activity between NPL and NIST using ionisation chamber systems demonstrated excellent statistical agreement using independently determined calibration factors and time zero measurements. These results were confirmed by the primary measurements performed at NIST and radionuclide calibrator measurements performed at both institutions. The results of the comparison, performed with multiple different approaches, indicate that there is no significant underlying bias between the NPL and NIST primary standards

Table 8. Summary of results for the NPL-NIST bilateral comparison of ^{227}Th activity.

Method	Comparison result
Direct	$t = 0.15$
Synchronous	$D_{ij}(t - t_0)/K_{\mathrm{VIC}}(t - t_0) = 0.0036(38)$
Model-assisted asynchronous	$D_{ij} = 0.019(10)$
Hybrid	$D_{ij}(t_{10\,\mathrm{d}}) = 0.0029(96)$

(table 8). The comparison methodology provides a blueprint for other NMIs or DIs to perform similar comparisons.

Acknowledgment

This work was funded by Bayer AS (Norway). We are grateful to Ryan Fitzgerald (NIST) for anticoincidence counting support and critical input and to Carine Michotte (BIPM) for critical input. The National Physical Laboratory is operated by NPL Management Ltd, a wholly owned company of the UK government Department for Science, Innovation & Technology.

ORCID iDs

Andrew J Fenwick https://orcid.org/0000-0002-4543-8933
Denis E Bergeron https://orcid.org/0000-0003-1150-7950

References

[1] Parker C, Nilsson S, Heinrich D, Helle S I, O'sullivan J M, Fosså S D, Chodacki A, Wiechno P, Logue J and Seke M 2013 Alpha emitter radium-223 and survival in metastatic prostate cancer *N. Engl. J. Med.* **369** 213–23
[2] Karlsson J, Schatz C A, Wengner A M, Hammer S, Scholz A, Cuthbertson A, Wagner V, Hennekes H, Jardine V and Hagemann U B 2023 Targeted thorium-227 conjugates as treatment options in oncology *Front. Med.* **9** 1071086
[3] Judge S M, Coulon R M, Cox M G, Karam L, Knoll P, Michotte C, Msimang Z and Zimmerman B E 2022 Traceability for nuclear medicine: the status of primary radioactivity standards *Metrologia* **60** 012001
[4] Ratel G 2007 The Systeme International de Reference and its application in key comparisons *Metrologia* **44** S7–S16
[5] Kossert K and Nähle O 2019 Determination of the activity and half-life of ^{227}Th *Appl. Radiat. Isot.* **145** 12–18
[6] Bergeron D E, Cessna J T, Broder B A, Pibida L, Fitzgerald R P, DiGiorgio M, Napoli E and Zimmerman B E 2024 Activity standard and calibrations for ^{227}Th with ingrowing progeny *Appl. Radiat. Isot.* **209** 111326
[7] Collins S M, Keightley J D, Ivanov P, Arinc A, Fenwick A J and Pearce A K 2019a The potential radio-immunotherapeutic α-emitter ^{227}Th—part II: absolute γ-ray emission intensities from the excited levels of ^{223}Ra *Appl. Radiat. Isot.* **145** 251–7
[8] Collins S M, Keightley J D, Ivanov P, Arinc A, Jerome S M, Fenwick A J and Pearce A K 2019b The potential radio-immunotherapeutic α-emitter ^{227}Th—part I: standardisation via primary liquid scintillation techniques and decay progeny ingrowth measurements *Appl. Radiat. Isot.* **145** 240–50
[9] Bergeron D E, Collins S M, Pibida L, Cessna J T, Fitzgerald R, Zimmerman B E, Ivanov P, Keightley J D and Napoli E 2021 Ra-224 activity, half-life, and 241 keV gamma ray absolute emission intensity: a NIST-NPL bilateral comparison *Appl. Radiat. Isot.* **170** 109572
[10] BIPM 2008 *Evaluation of Measurement data—Guide to the Expression of Uncertainty in Measurement BIPM* (BIPM)
[11] Bé M-M *et al* 2011 *Table of Radionuclides (Vol. 6—A = 22–242)* Monographie BIPM-5 6 pp 125–38
[12] Ivanov P I, Collins S M, van Es E M, García-Miranda M, Jerome S M and Russell B C 2017 Evaluation of the separation and purification of ^{227}Th from its decay progeny by anion exchange and extraction chromatography *Appl. Radiat. Isot.* **124** 100–5
[13] Collé R 2019 Ampoules for radioactivity standard reference materials (No. NISTIR 8254) (NIST)
[14] ISO 2018 ISO 8362–1:2018(E); Injection containers and accessories, Part 1: injection vials made of glass tubing
[15] Broda R, Cassette P and Kossert K 2007 Radionuclide metrology using liquid scintillation counting *Metrologia* **44** S36
[16] Fitzgerald R, Bailat C, Bobin C and Keightley J D 2015 Uncertainties in $4\pi\beta$–γ coincidence counting *Metrologia* **52** S86
[17] Bateman H, 1910. Solution of a system of differential equations occurrig in the theory of radioactive transformations *Proc. Cambridge Philosophical Society*
[18] Kossert K and Carles A G 2010 Improved method for the calculation of the counting efficiency of electron-capture nuclides in liquid scintillation samples *Appl. Radiat. Isot.* **68** 1482–8
[19] Kibédi T, Burrows T W, Trzhaskovskaya M B, Davidson P M and Nestor C W 2008 Evaluation of theoretical conversion coefficients using BrIcc *Nucl. Instrum. Methods Phys. Res.* A **589** 202–29
[20] Rogers D W O, Kawrakow I, Seuntjens J P, Walters B R B and Mainegra-Hing E 2023 *NRC User Codes for EGSnrc (No. NRCC Report PIRS-702(revC))* (NRC Canada)
[21] Broder B A, Bergeron D E, Fitzgerald R and Zimmerman B E 2023 Comparison of calibration coefficients for a vinten ionization chamber simulated using four Monte Carlo methods *Appl. Radiat. Isot.* **202** 111068
[22] Bergeron D E, Cessna J T and Zimmerman B E 2015 Secondary standards for ^{223}Ra revised *Appl. Radiat. Isot.* **101** 10–14
[23] Pommé S, Collins S M, Harms A and Jerome S M 2016 Fundamental uncertainty equations for nuclear dating applied to the ^{140}Ba-^{140}La and ^{227}Th-^{223}Ra chronometers *J. Environ. Radioact.* **162–163** 358–70
[24] Kondev F G, McCutchan E A, Singh B and Tuli J K 2016 Nuclear Data Sheets for A = 227 *Nucl. Data Sheets* **132** 257–354
[25] Collins S M, Pearce A K, Regan P H and Keightley J D 2015b Precise measurements of the absolute γ-ray emission probabilities of ^{223}Ra and decay progeny in equilibrium *Appl. Radiat. Isot.* **102** 15–28
[26] ISO 2010 Injection equipment for medical use—Part 2: one-point-cut (OPC) ampoules. ISO 9187–2:2010
[27] Collins S M, Pearce A K, Ferreira K M, Fenwick A J, Regan P H and Keightley J D 2015a Direct measurement of the half-life of ^{223}Ra *Appl. Radiat. Isot.* **99** 46–53
[28] Bergeron D E, Kossert K, Collins S M and Fenwick A J 2022 Realization and dissemination of activity standards for medically important alpha-emitting radionuclides *Appl. Radiat. Isot.* **184** 110161