EURADOS
European Radiation Dosimetry Group e. V.

EURADOS Report 2023-03
Oberschleißheim, October 2023

EIVIC-2020: European In-Vivo Intercomparison Exercise 2020 - Organisation of a European Interlaboratory Comparison on Whole-Body Counting

Didier Franck, Oliver Meisenberg, María Antonia López,
Tiffany Beaumont, Werner Buchholz,
Juan Francisco Navarro, Begoña Pérez,
Kerstin Hürkamp, Filip Vanhavere

ISSN 2226-8057
ISBN 978-3-943701-35-7
DOI 10.12768/hmma-d768

European Radiation Dosimetry Group e. V.

EURADOS Report 2023-03

Oberschleißheim, October 2023

EIVIC-2020: European In-Vivo Intercomparison Exercise 2020 - Organisation of a European Interlaboratory Comparison on Whole-Body Counting

Didier Franck[1], Oliver Meisenberg[2], María Antonia López[3], Tiffany Beaumont[1], Werner Buchholz[2], Juan Francisco Navarro[3], Begoña Pérez[3], Kerstin Hürkamp[4], Filip Vanhavere[5]

[1] Institut de Radioprotection et de Sûreté Nucléaire (IRSN), 92260 Fontenay-aux-Roses, France

[2] Federal Office for Radiation Protection (BfS), 85764 Oberschleißheim, Germany

[3] Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), 28040 Madrid, Spain

[4] European Radiation Dosimetry Group e.V. (EURADOS), 85764 Oberschleißheim, Germany

[5] Belgian Nuclear Research Centre (SCK CEN), 2400 Mol, Belgium

ISSN 2226-8057

ISBN 978-3-943701-35-7

DOI 10.12768/hmma-d768

Imprint

Issued by:
European Radiation Dosimetry Group e. V.
Ingolstädter Landstr. 1
85764 Oberschleißheim
Germany
office@eurados.org
www.eurados.org

The European Radiation Dosimetry Group e.V. is a non-profit organisation promoting research and development and European cooperation in the field of the dosimetry of ionising radiation. It is registered in the Register of Associations (Amtsgericht München, registry number VR 207982) and certified to be of non-profit character (Finanzamt München, notification from 2023-03-14).

Liability Disclaimer

Content

Abstract

The objective of the EIVIC-2020 project was to assess the implementation of the individual monitoring requirements of the Basic Safety Standards (BSS) Directive in EU Member States based on in-vivo measurements and receive an overview of the capabilities and performance of whole-body counters in Europe.

This exercise was supported by the European Commission (Directorate-General for Energy) in Luxembourg under the contract ENER/2019/NUCL/SI2.811157 and organised by EURADOS (European Radiation Dosimetry Group e.V., Germany), the Institut de Radioprotection et de Sûreté Nucléaire (IRSN, France) and the Federal Office for Radiation Protection (BfS, Germany) in collaboration with the Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT, Spain) and the Karlsruhe Institute of Technology (KIT, Germany).

It was organised between October 2019 and June 2022 and dedicated to whole-body measurement of gamma emitters in several tasks selected that cover the range of such possible measurements associated to different intake scenarios. In total, 43 installations from 21 countries took part in the intercomparison exercise.

The measured data were compared with reference activity values to evaluate the corresponding bias according to the standards ISO 28218 and ISO 13528. This report gives a summary of the results of the different tasks. In general, the results are good, and most facilities are in conformity with the criteria for the bias and z-scores in the ISO standards. Additionally, it was tested if the results can be attributed to several organisational and metrological properties of the participating laboratories.

1. Introduction and context

In 2018, the European Commission published in its Radiation Protection Series the document Technical Recommendations for Monitoring Individuals for Occupational Intakes of Radionuclides as RP 188 [RP 2018]. This guidance document emphasises that for quality assurance of the measurement results, it is essential that the laboratories performing whole-body counting regularly participate in suitable interlaboratory comparisons.

In this frame, the European Commission decided in 2019 to launch a call for tender (ENER/D3/2019-158), with the objective to assess the implementation of the individual monitoring requirements of the BSS Directive in EU Member States based on in-vivo measurements and receive an overview of the performance of in-vivo measurements using whole-body counters.

In accordance with the tender specifications the project "European In-vivo Intercomparison Exercise 2020" (EIVIC-2020) was initiated jointly by the European Radiation Dosimetry Group (EURADOS), the Federal Office for Radiation Protection (BfS, Germany) and the Institut de Radioprotection et de Sûreté Nucléaire (IRSN, France).

The objective of the present report is to present the results of the exercises put in place and the conclusion of the European intercomparison of in-vivo monitoring in Europe.

2. General organisation

The principle of intercomparison for in-vivo measurement is to circulate the object subjected to the test from one laboratory to another. The duration of the intercomparison is therefore dependent on the number of participants and the availability of facilities. The schedule followed for the organisation of the whole-body intercomparison is detailed in Table 1.

Table 1: Schedule followed for the organisation of the whole-body intercomparison.

Description	Date of accomplishment
Inception Report including description of methodology of the intercomparison exercise	10/2019
Kick-off meeting in Luxembourg	10/2019
List of participants and further contacts of whole-body counting laboratories in Europe	02/2021
Phantoms prepared and ready for transport/shipment	04/2021
Participants' meeting (online)	06/2021
Progress Report including reports on source production, their quality assurance and the final schedule of the measurement campaign	07/2021
End of measurement campaign	12/2021
Analysis of the results provided by participants from the questionnaires and communication of their results to participants	02/2022
Participants' final workshop	06/2022
Draft final report including compilation of results, their statistical evaluation and a summary of the participants' workshop	06/2022
Report: Proceedings of the Participants' workshop including lessons learned from the intercomparison exercise	07/2022
Presentation of the results to the Article 31 group of experts in Luxembourg report)	11/2022
Final reports (publishable scientific report and administrative report)	03/2023

Thirty-seven laboratories were contacted from 21 countries (+ European Commission and IAEA) and were officially registered (see Table 2). Several laboratories conducted measurements with more than one whole-body counter, so that more than 37 results were received.

The circuit was organised by shipment and by attended transport. Principal point for the decision between attended transport or shipment were the distance, prospective travelling time, the

accessibility (possibly also of nearby facilities), temporal availability and ideally an in-line connection of the institutions in one of the tours, as well as the possible requirements of customs clearance. The geographical distribution of participating laboratories in Europe can be seen in Figure 1.

The measurement campaign started in the first week of May 2021 and finished at the end of November 2021.

Table 2: List of participants of the EIVIC exercise.

Institution	City	Country	A/S*
Seibersdorf Labor GmbH	Seibersdorf	Austria	A
LKH University Hospital Graz	Graz	Austria	A
SCK CEN Belgian Nuclear Research Centre	Mol	Belgium	A
Kozloduy Nuclear Power Plant	Kozloduy	Bulgaria	S
University of Zagreb School of Medicine	Zagreb	Croatia	A
SÚRO National Radiation Protection Institute	Prague	Czech Republic	A
SIS National Institute of Radiation Protection	Herlev	Denmark	A
EC Joint Research Centre	Ispra	European Commission (Italy)	A
STUK Finnish Radiation and Nuclear Safety Authority	Helsinki	Finland	S
SPRA French Defense Radiation Protection Service	Clamart	France	A
ORANO	La Hague	France	cancelled
CEA French Alternative Energies and Atomic Energy Commission	Saclay	France	S
FZJ Research Centre Jülich	Jülich	Germany	A
LIA Institute for Work Design	Düsseldorf	Germany	A
University Hospital Leipzig	Leipzig	Germany	A
MTA Hungarian Academy of Sciences	Budapest	Hungary	S
National Public Health Centre	Budapest	Hungary	A

Institution	City	Country	A/S*
IAEA	Vienna	IAEA (Austria)	A
SOGIN	Caorso	Italy	A
ENEA	Rome	Italy	S
Radiation Protection Centre	Vilnius	Lithuania	S
NRG Nuclear Research and Consultancy Group	Petten	Netherlands	A
IFE Institute of Energy and Technology	Kjeller	Norway	A
National Centre for Nuclear Reseach	Otwock	Poland	S
IFIN-HH National Institute of Physics and Nuclear Engineering	Magurele	Romania	cancelled
Mochovce Nuclear Power Plant	Mochovce	Slovakia	S
JAVYS Nuclear and Decommissioning Company	Bratislava	Slovakia	A
Tecnatom	Madrid	Spain	S
FOI Swedish Defence Research Agency	Umeå	Sweden	S
University of Lund	Malmö	Sweden	A
Ringhals Nuclear Power Plant	Väröbacka	Sweden	A
Forsmark Nuclear Power Plant	Forsmark	Sweden	S
Barsebäck Nuclear Power Plant	Löddeköpinge	Sweden	A
Bundesamt für Bevölkerungsschutz	Spiez	Switzerland	A
PSI Paul Scherrer Institute	Villigen	Switzerland	A
AWE Atomic Weapons Establishment	Reading	United Kingdom	A
UKHSA UK Health Security Agency	Chilton	United Kingdom	A

*A: personal attendance, S: shipment.

When registering participants, a specific programme for this intercomparison was sent. It specified in particular the assembly procedure of the phantom under test (Annex I to III).

Figure 1: Distribution of participating laboratories, in red circuit organised by attended transport, in blue circuit organised by shipment. Map data: © 2022 Google, GeoBasis-DE/BKG, Inst. Geogr. Nacional.

3. Phantoms and sources

3.1 Phantoms

The intercomparison has been carried out using anthropomorphic phantoms equipped with sealed radioactive sources. The use of such phantoms is a common method for the calibration of whole-body counters and the examination of their proficiency. These phantoms can be set up in various heights and masses to simulate a variety of different human shapes and sizes and can be equipped with various radionuclides. The type of phantom selected is the Saint-Petersburg brick phantom [Kovtun 2000]. This phantom consists of rectangular bricks made from polyethylene, which can be set up in six shapes (P1–P6) resembling persons of weight 12 kg to 110 kg. The bricks contain holes, which can be filled with rod sources of known activities. These phantoms are an unofficial de-facto standard that is used worldwide by many laboratories for their calibrations and that is considered an appropriate method for calibration also by ICRU [ICRU 1992, ICRU 2003]. In contrast to a bottle phantom like the Bottle Mannequin Absorber (BOMAB) phantom with a liquid content of radionuclide solution, the brick phantom comprises solid radioactive sources and therefore offers better protection against the risk of spills. With its larger number of components, it is also more suitable to be set up in different postures such as lying in a stretcher geometry, sitting in chair and inclined chair geometries or standing.

For the EIVIC intercomparison two sizes of phantoms have been chosen: P4 and P5 corresponding to 70 kg and 90 kg persons respectively in order to correspond to the weight ranges generally observed for the workers measured in the whole-body laboratories. Photographs and diagrams illustrate the construction of the phantom in the stretcher (and likewise in the standing) geometry in Figure 2 and Figure 3. Each participant received a booklet detailing instructions about phantoms assembly (see Annex III).

For each of the two circuits (shipment and attended transport), one phantom was used. The two phantoms featured identical properties and the radioactive sources for each of the phantoms were selected in such way that the differences of their activities were negligible.

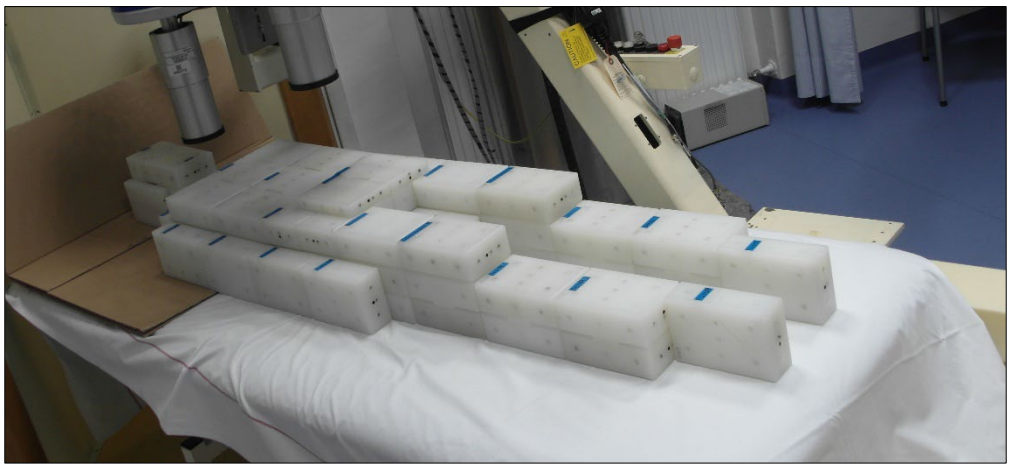


Figure 2: Photo of the phantom in the P4/70 kg configuration, erected in a whole-body counter with stretcher geometry.

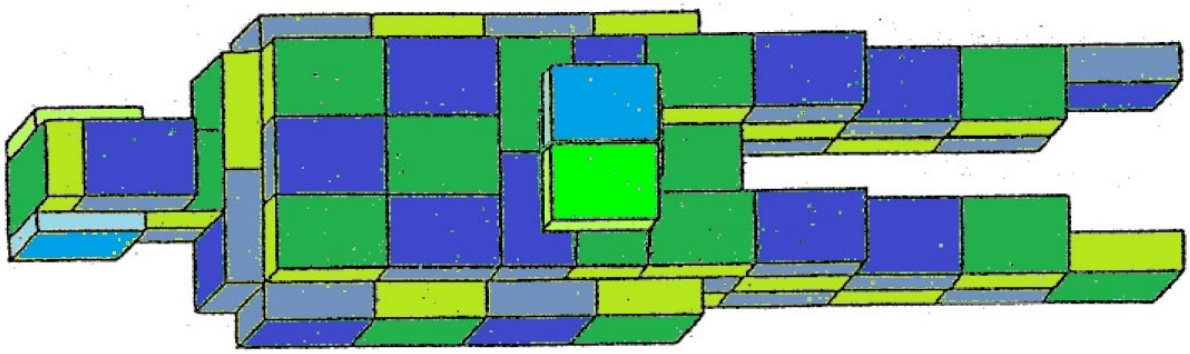


Figure 3: Sketch of the phantom in the P4/70 kg configuration (adapted from Kovtun, 1995).

3.2 Radionuclides, sources and tasks

3.2.1 Description of the measurement tasks

One of the objectives of the intercomparison was to simulate measurements that are relevant for the occupational monitoring programmes of individuals exposed to intakes of gamma emitters at the workplace. Several tasks have been selected that cover the range of possible measurements associated to different intake scenarios. In order to simulate occupational internal exposures, different phantoms regarding their size (limited to phantoms resembling adults) and different radionuclides have been chosen. The selected radionuclides are realistic for the internal monitoring of individuals (Tasks 2 and 3) and/or feature characteristics that are advantageous for the assessment of the proficiency of whole-body counters (all tasks).

For this intercomparison, four tasks were defined concerning measurements of phantoms equipped with radionuclide sources. For each phantom measurement task, one specific set of radionuclide sources has been used. Each set contains a mixture of those radionuclides that are to be measured in the respective measurement task. The measurement tasks comprise the following nuclides:

- Task 1: ^{60}Co, ^{133}Ba and ^{137}Cs. Phantom size P4/70 kg (called Victor as these are the usual sources for the IRSN proficiency test phantom called Victor),
- Task 2: ^{134}Cs and ^{137}Cs. P5/90 kg (called Emergency),
- Task 3: ^{68}Ge (with its daughter nuclide ^{68}Ga in secular equilibrium) and ^{88}Y. P4/70 kg (called Medicine),
- Task 4: ^{133}Ba and ^{152}Eu. P4 and P5 configurations (called Calibration).

Task 1 comprises radionuclides that are frequently used in proficiency tests of whole-body counters because their gamma-ray emissions cover a wide range of energy (80 to 1332 keV) and do not interfere with each other, even when measured with low-resolution sodium iodide (NaI(Tl)) detectors.

Task 2 comprises radionuclides that are relevant for the monitoring of members of the population after nuclear accidents. A metrological challenge is the correction of the peak area of ^{137}Cs at 662 keV taking into account the overlapping contribution of the 605 keV emission of ^{134}Cs when measured with NaI(Tl) detectors.

Task 3 featured ^{68}Ge and its decay product ^{68}Ga in secular equilibrium, which are common radionuclides in nuclear medical diagnostics. These radionuclides emit positron radiation, resulting in the production of annihilation radiation of 511 keV at an abundance of 178%. They also emit gamma radiation at an energy of 1077 keV with an abundance of 3.2%. A definite identification of the radionuclides is only possible from the 1077 keV emission, but this emission was not observed by all participants because of the small abundance. At 511 keV, background counts must be subtracted from the gross peak area. Additionally, the peak shape differs from the shape of usual gamma radiation peaks.

This task also featured ^{88}Y, which emits gamma radiation at energies of 898 keV (abundance 93.7%) and 1836 keV (99.3%). Whereas 898 keV might be well within the energy range that is taken into account in the efficiency calibration of most whole-body counters, 1836 keV is beyond the upper limit of the energy range of the efficiency calibration if ^{133}Ba and ^{152}Eu are used for the calibration (highest energy 1408 keV). This task was suitable both for NaI(Tl) and for HPGe detectors.

The sets of sources were replaced in the middle of the campaign by new sources in the attended tour as well as for the shipment because of the short half-lives and small abundances of these radionuclides. Thus, activities in a similar order of magnitude could be offered to all participants. However, inevitably those participants at the end of the application period of each set had to measure sources with significantly smaller activities than those at the beginning. This allowed to assess the influence of the activities at the time of the measurement on the measurement results.

The 70 kg phantom (P4) was used for this task.

Task 4 comprises radionuclides that feature a large variety of gamma-ray emissions over a wide range of energies (80 to 1408 keV). Therefore, measuring the phantoms of this task enables the participating laboratories to establish an efficiency calibration based on the reported reference activities.

During all measurement tasks, the phantoms were also prepared with ^{40}K rods in order to simulate the natural radiation background of human bodies. Tasks 1 and 3 (as well as Task 4a on a voluntary basis) were conducted with the P4/70 kg phantom, Tasks 2 and 4b were conducted with the P5/90 kg phantom.

Table 3 reports the activities of the different phantoms used for this work, based on gamma-spectroscopy measurements. As described in Section 4.1, these activities were not used as the reference activities for the calculation of the z-scores. For the measurement Tasks 1, 2, 3 (set 1) and 4, the reference date was 01/05/2021, and for the Task 3 (set 2) it was 10/08/2021.

Table 3: Nuclides and reference activities (based on gamma-spectroscopy measurements) for the different tasks.

	^{60}Co	^{68}Ge	^{88}Y	^{133}Ba	^{134}Cs	^{137}Cs	^{152}Eu	^{40}K		
Task 1: Victor P4	**1100**			**2720**		**3850**		**4210**	activity	in Bq
reference date:	20			60		90		210	1σ uncertainty	in Bq
01.05.2021	1.9			2.2		2.3		5.0	1σ uncertainty	in %
Task 2: Emergency P5					**3490**	**3220**		**5460**		
01.05.2021					40	40		270		
					1.1	1.2		5.0		
Task 3: Medicine P4		**4190**	**4720**					**4210**		
set 1		410	120					210		
01.05.2021		9.8	2.5					5.0		
set 2		**4630**	**4400**					**4210**		
10.08.2021		450	110					210		
		9.7	2.5					5.0		
Task 4a: Calibration P4				**21750**			**25720**	**4210**		
01.05.2021				290			600	210		
				1.3			2.3	5.0		
Task 4b: Calibration P5				**27930**			**33050**	**5460**		
01.05.2021				390			750	270		
				1.4			2.3	5.0		
Task 5: Person (only visitation)						**≈ 95**		**≈ 4300**	activity	in Bq
						≈ 30		≈ 95	1σ uncertainty	in Bq
						≈ 1.2		≈ 53.7	specific activity	in Bq/kg
						≈ 0.4		≈ 1.2	1σ uncertainty	in Bq/kg

3.2.2 Production of the sources

Rod sources for Task 1 were taken from the stock of IRSN, which uses them frequently for proficiency tests with their brick phantom "Victor". Rod sources for Tasks 2, 3 and 4 were produced specifically for EIVIC-2020 in the laboratories of BfS according to the method described by Woidy and Meisenberg [Woidy 2022]. The sources produced according to this method consist of tubes of rigid polyvinyl chloride that are filled with hardened epoxy resin blended with radionuclide solution. Both ends of the tubes are plugged with inactive epoxy resin. The sources featured a length of 16.5 cm, of which about 14 cm were radioactive filling, an outer diameter of 6 mm and an inner diameter of 5.4 mm. For the production of each single source, radionuclide solution (with volumes between 10 and 100 µl) was pipetted into epoxy resin, blended with hardener and filled into the tubes where the resin was left for hardening. This method allowed the production of about 10 sources per day, with a complete set for one phantom consisting of up to 108 sources. The radionuclide solutions were purchased from Eckert & Ziegler Nuclitec, Germany. Their activities were traceable to national standards of the United States National Institute of Standards and Technology (NIST).

An additional exemplary set of sources was tested for certification as sealed radionuclide sources according to ISO 2919 [ISO 2012]. The following tests were conducted:

- Impact test with a weight of 57 g dropped from a height of 1 m,
- Puncture test with a weight of 57 g dropped from a height of 1 m,
- Temperature tests at -66 °C and +80 °C,
- External pressure test at 14 kPa,
- Bending test with a mass of 12 kg acting on the middle of the sources.

All tests were passed so that the sources qualify as sealed sources according to Class 2 as defined in ISO 2919 [ISO 2012]. Beside the practical advantages of using solid sources rather than liquid ones, this certification facilitates the application of the sources regarding legal aspects since e.g. some participating laboratories might be authorised to handle only sealed radionuclide sources and no contamination tests of worktops are required.

3.2.3 Quality assurance

During the production and quality-assurance process, the activities of all produced rod sources were determined by several independent methods:

- Calculation of the activities based on the net weight of the radionuclide standard:
- After each step of the preparation of the blend of radionuclide solution and epoxy resin, the blend was weighed with a precision balance (CP124S with draught shield, resolution 0.1 mg, Sartorius, Germany), which was subject to annual quality-assurance checks by an accredited service. Additionally, the amount of blend that was left in the mixing vessel after pipetting the blend into the tubes was weighed and taken into account in the calculation of the amount of radionuclide solution in each single rod source. Since the activity of the radionuclide solution was traceable, this yields a traceable activity of each single source.
- Measurement of each single source on a gamma-spectroscopy detector:
- Each single produced rod source was measured with an HPGe gamma-spectroscopy detector (GMX series, n-type, Ortec, USA, calibrated with ^{60}Co, ^{133}Ba and ^{137}Cs with traceability). For this purpose, two geometries were used: a high-efficiency geometry where the source was placed in close contact horizontally on the detector, with its ends protruding over the edge of the detector (diameter of detector crystal: 7.4 cm, length of the sources: 16.5 cm); a low-efficiency geometry with approximately 13 cm between the horizontal source and the detector. The high-efficiency geometry was possible because additionally the spatial homogeneity of the activity along the length of the sources was checked and confirmed with an exemplary selection of sources of each set by measuring single section of the sources using a collimator. This geometry featured smaller counting uncertainties than the low-efficiency geometry, but the measurement was subject to the effect of coincidence summing, which affects the measurement of radionuclides that emit gamma radiation in several steps one after the other (in the selected tasks all nuclides except of ^{137}Cs).
- Measurement of phantoms equipped with a whole set of sources in whole-body counters: All sets of sources were measured inside the phantoms that were assembled according to the respective measurement tasks. The measurements were conducted with the whole-body counters of the organisers of the project.
- BfS and IRSN performed measurements with germanium detectors (HPGe) and CIEMAT and KIT performed measurements with HPGe detectors (broad-energy/BEGe model) and with NaI(Tl) detectors. Measurement times and calibration curves (P4 or P5) were those that are usually used at the respective laboratory. The results of the quality assurance (QA) measurements show acceptable agreement for the radionuclides of study when comparing with reference values.

Although all of these methods were found suitable for the quality assurance of the produced sources (e.g. in order to identify sources that needed to be discarded because of deviating activity), they did not prove sufficiently precise in order to determine the reference activities. The calculation based on the net weight was affected by the uncertainty of the weighing of small masses in the order of 10 mg. The measurement of single sources on gamma-spectroscopy detectors was subject to the

mentioned effect of coincidence summing and also of a possible slightly inhomogeneous spatial distribution of the activity in each single source.

Besides, measurements of the phantoms with the whole-body counters of the organisers were subject to the usual influences of slightly different geometries between measurement and calibration. Therefore, the results of the single different methods differed by up to 10%. This was deemed sufficiently small to confirm the quality of the sources and in particular the homogeneity of the activities throughout the different sources of each set. However, it led to the conclusion not to use the activities determined by one of the described quality-assurance methods as the reference activity for the certification of the proficiency of the participants. As shown in the next chapter the use of robust mean was preferred to be used as target value.

As a result, as presented in the next chapter the use of robust mean was preferred to be used as target value.

4. Methodology of data evaluation

The data provided by the participating laboratories have been treated statistically by IRSN using the software ProLab™. For the whole-body intercomparison, the statistical processing was the following:

- Distribution of the results (chart of the population),
- Search for aberrant values (Grubbs method),
- Relative bias: assessment of the laboratory performances according to standard ISO 28218 [ISO 2010],
- Z-score: assessment of the laboratory performances according to standard ISO 13528 [ISO 2022].

The relative bias is a measure of how close the assessed activity is to the measured activity. According to the standard ISO 28218, in the service laboratory internal quality control, the bias shall be between -25% and +50%. When the bias is outside the range of -25% to +50% in internal quality control checks of service laboratories, the service laboratory shall make appropriate corrections in phantom calibration or measurement protocols to reduce or eliminate the bias.

Another estimator, called zeta-score, could be used as recommended in the standard ISO 13528. This estimator is based on the uncertainties of measurement and relevant only for measurement done in the same condition (measurement and calibration). In the case of this intercomparison, it was decided not to use this estimator because of the large diversity of the whole-body facilities in terms of detection system (NaI(Tl) or HPGe), kind of phantom and protocols used.

The performance criteria used are detailed below. The conformity of the results with these various criteria helps to qualify the proficiency of the laboratories and to identify ways to improve the methods used.

4.1 Assigned value (ISO 13528)

According to the ISO 13528 [ISO 2022], four methods to determine the assigned value can be considered for this kind of intercomparison:

- The reference value of the certificate,
- The value from one laboratory,
- The consensus value from expert laboratories,
- The consensus value from participants, using a robust statistical method.

The choice between these methods is crucial and required for a realistic analysis of results.

As explained in the section 3, excepted for the "Victor" sources, the rod sources have been produced specifically for the EIVIC-2020 by the laboratory of BfS. This laboratory is not accredited to produce sealed sources. Regarding the rod sources of Task 1 and the ^{40}K rods, it will be possible to use the reference value of the certificate. Nevertheless, to guarantee a homogeneous analysis between the tasks, this method was not adopted. Consequently, the method using the reference value with a certificate was not used.

Because of the discrepancies up to 10% between the activities of the sources calculated from the weight of the radionuclide solutions, measured on gamma spectrometry systems and measured in complete phantoms during the QA process, it was decided that the reference values given by BfS were not used as assigned value.

Regarding the consensus value method from expert laboratories, the measurement of Tasks 1, 2 and 4 were performed by 3 experts (KIT, CIEMAT and IRSN). Unfortunately, for the Task 3, due to the short half-life of the radionuclide, only one expert (CIEMAT) performed the measurement. As the result, this method cannot guarantee the homogeneous analysis between the tasks and was not used.

Because of the high statistic (43 facilities) and because no additional measurements are required to obtain the assigned value, it was decided to use the consensus value from participants, determined with a robust method.

Robust mean refers to the arithmetic mean of the reported values without outliers and was calculated using the Q/Hampel method[1]. The method known as Q/Hampel uses the Q method for the calculation of the robust standard deviation s* together with the Hampel estimator for the calculation of the robust location parameter x*. This method is applied for the statistical analysis of interlaboratory studies. It was used to guarantee a homogeneous and robust analysis between the tasks.

According to ISO 13528 [ISO 2022], the assigned value (robust mean) was compared with an independent reference value for each task and each radionuclide. The difference (%) was calculated using the BfS value (from gamma spectroscopy measurements) as the reference for the Task 2, 3 and 4 and using the IRSN value (from the certificate of the sources) as the reference for the Task 1. For each task, the differences between the robust mean and the reference value were noted in Section 5.

As recommended in the ISO 13528, an additional action was carried out to check the validity of the method used to determine the assigned value. For example, for the ^{137}Cs (Task 2), the arithmetic mean of the expert laboratories (3040 ± 128 Bq) is very close to the value of robust mean (2996 ± 60 Bq). This observation is verified in the other tasks.

A graphical example of the representation of the raw data and the assigned value are presented in Figure 4.

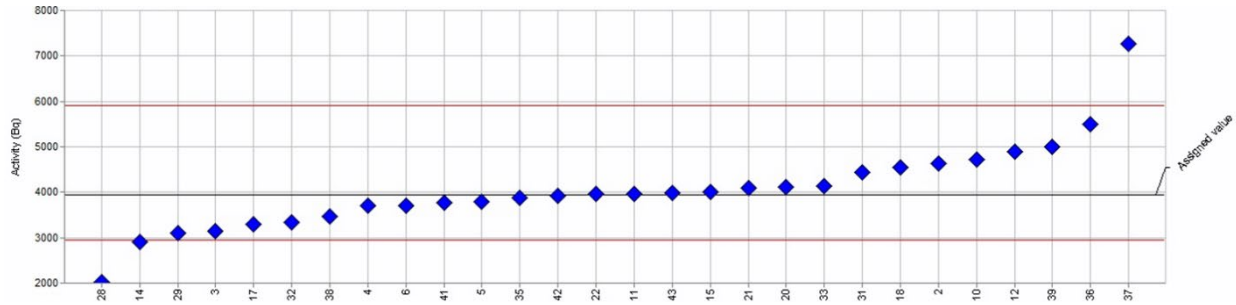


Figure 4: Graphical example of the representation of the raw data. Red lines: -25% and +50% performance criteria.

[1] Q/Hampel method according to the ISO 13528: https://quodata.de/en/web%C2%ADservices/QHampel.html#0

4.2 Estimated bias (ISO 28218)

The laboratory bias estimate is defined as a percentage. This performance test is calculated as follows:

$$Bias(\%) = \frac{x - X}{X} \times 100$$

- x: Result of the participating facility
- X: Activity of the target value (assigned value)

According to the recommendations of ISO 28218 "Performance criteria for radiobioassay" [ISO 2010], the relative bias error must be within a range of -25% to +50% relative to the target value.

A graphical example of the representation of the laboratory bias estimate is presented in Figure 5.

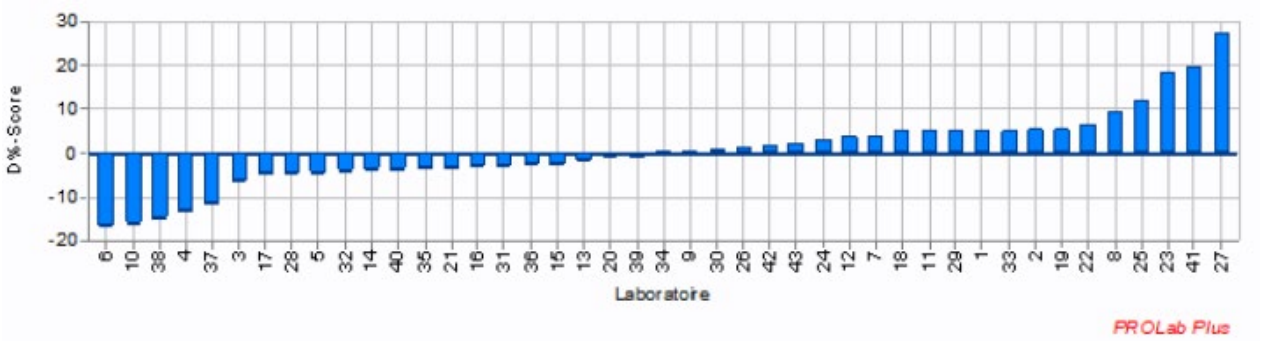


Figure 5: Graphical example of the representation of laboratory bias estimate.

4.3 Outliers (Grubbs Test)

Each data set was subjected to the Grubbs test in order to detect possible outliers at the ends of the distribution.

The test consists in calculating, for n values classified in ascending order of x_1, x_2, ..., x_n, the test statistic Gp:

to test x_1 $\quad G_p = \dfrac{\bar{x} - x_1}{s}$ $\qquad$ to test x_n $\quad G_p = \dfrac{x_n - \bar{x}}{s}$

with

- S: inter-laboratory standard deviation
- $x1$: Lowest population value
- xn: Highest population value
- $\bar{x}$: Mean of the n values of the population

The value of Gp is compared with a critical value that depends on the number n of values. If one of the extreme values is identified as an outlier, this value is discarded and the test is repeated with the remaining set of values until no value is identified as an outlier anymore.

4.4 Z-Score estimation (ISO 13528)

The z-score is an indicator of the laboratory proficiency compared to that of the other laboratories because it is correlated with the robust standard deviation. Thus, it depends directly on the dispersion of the results from the laboratories. The z-score is calculated by means of the following formula:

$$z = \frac{x - X}{\hat{\sigma}}$$

with:

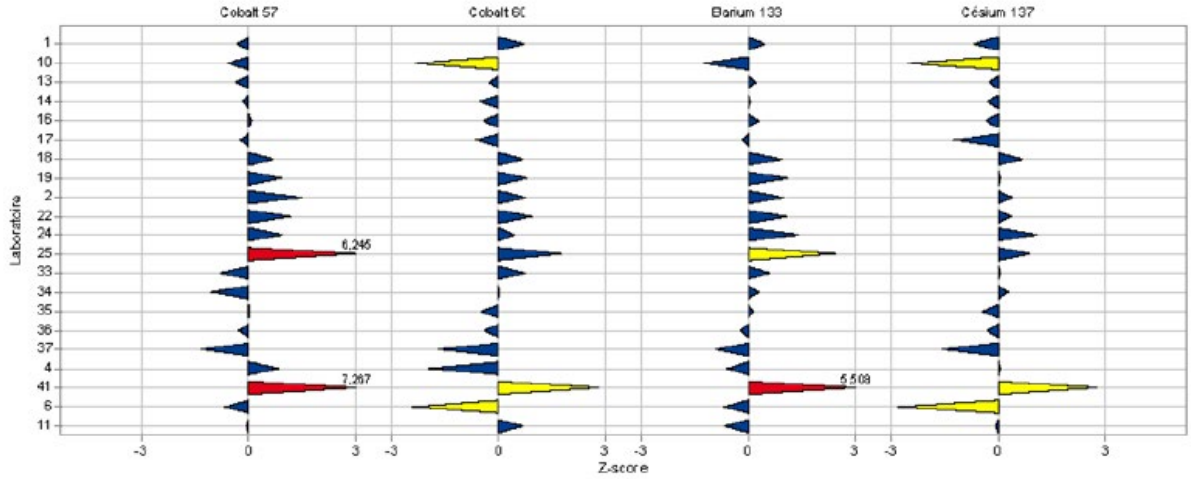
- x: Result of the participating facility,
- X: Activity of the target value,
- $\hat{\sigma}$: Robust standard deviation for proficiency evaluation.

According to the recommendations of ISO 13528 "Statistical methods for use in proficiency testing by intercomparison" [ISO 2022], the current z-score criteria are:

- $|z - score| \leq 2$: the result is satisfactory,
- $2 < |z - score| < 3$: the result is considered to give a warning signal,
- $|z - score| \geq 3$: the result is considered unacceptable (action signal).

It should be noted that the presence of a single action signal or several warning signals in two successive cycles must be regarded as evidence of an abnormality that requires remedial action. The z-score (Figure 6) depends directly on the dispersion of the results from the facilities.

Figure 6: Graphical example of the representation of the z-score. Yellow: warning signal, red: action signal.



4.5 Expression of the results

The laboratories had to identify and quantify the radionuclides present in the phantom. The activity and the uncertainty related to each result had to be expressed in becquerel (Bq). The latter is given as the expanded uncertainty at 2σ indicating a coverage factor *k* equivalent to 2.

The laboratories reported results that were valid at the date of the measurement.

To compare all the results, decay correction to a reference date was conducted by the EIVIC-2020 team. This was done because decay correction is not a task that is required in in-vivo internal monitoring. Decay correction by the EIVIC-2020 team based on an identical half-life for all laboratories and each radionuclide ensured that no additional source of possible errors was introduced in this step.

For Task 1, 2 and 4 (a and b) the results of activity and their associated uncertainties were calculated at the reference date set at 01/05/2021.

For the Task 3, given the short half-life, it was impossible to conduct the whole circuit with unique set of radionuclides. The first part of the EIVIC campaign was carried out with the set 1, as follows:

- Shipment tour: 01/05/21 to 30/07/2021
- Attended tour: 01/07/2021 to 06/08/2021

The reference date for the radionuclides of set 1 (^{68}Ge/^{68}Ga, ^{88}Y and ^{40}K) is 01/05/2021.

The second part of the EIVIC campaign was carried out with the set 2, as follows:

- Shipment tour: 02/08/2021 to 20/11/2021
- Attended tour: 10/08/2021 to 31/10/2021

The reference date for the short half-life radionuclides (^{68}Ge/^{68}Ga and ^{88}Y) is 10/08/2021. For the ^{40}K, the reference date is 01/05/2021, as in the previous tasks.

It has to be noted for the facility '15', the measurements of Task 3 were carried out at the end of the EIVIC campaign with both sets together. The results of Task 3 are not included in the analysis for this facility.

5. Results of the EIVIC exercise

5.1 Participation in the EIVIC intercomparison for the different tasks

The EIVIC intercomparison exercise was performed by 35 facilities from 21 countries. During this intercomparison, each participant laboratory had to choose the tasks they wanted to participate in function of their technical possibilities and skills. The participation of the facilities for the different tasks proposed is given in Table 4.

Each facility is represented by an anonymous code which has been assigned for the presentation of the results called 'ID Lab' in this report and sent to each participant.

Table 4: Summary of participation at the EIVIC intercomparison for the different tasks (*)

ID Lab.	Task 1 - Victor	Task 2 - Emergency	Task 3.1 – Medicine	Task 3.2 - Medicine	Task 4a – Calib. 70 kg	Task 4b – Calib. 90 kg
1	X	X	X			X
2	X	X		X	X	X
3	X	X		X	X	X
4	X	X	X		X	X
5	X	X	X		X	X
6	X	X		X	X	X
7	X	X		X		X
8	X	X	X		X	X
9	X	X		X	X	X
10	X	X		X		X
11	X	X		X		
12	X	X		X	X	X
13	X	X	X			X
14	X	X		X	X	X
15	X	X				
16		X				X
17	X	X		X	X	X
18	X	X		X		X
19	X	X		X		X

ID Lab.	Task 1 - Victor	Task 2 - Emergency	Task 3.1 – Medicine	Task 3.2 - Medicine	Task 4a – Calib. 70 kg	Task 4b – Calib. 90 kg
20	X	X		X	X	X
21	X	X		X	X	X
22	X	X	X			
23	X					
24	X	X		X	X	X
25	X	X	X		X	X
*26**						
27	X	X		X		
28	X	X		X		
29	X	X		X	X	X
30	X	X	X		X	X
31	X	X		X	X	X
32	X	X	X		X	X
33	X	X		X	X	X
34	X	X	X			
35	X	X		X		X
36	X	X		X	X	X
37	X	X	X			
38	X	X	X			X
39	X	X		X		X
*40**						
41	X	X	X		X	X
42	X	X	X			
43	X	X	X			

(*) no data submitted by labs 26 and 40 after conducting the measurements.

5.2 Task 1 – Victor: ^{60}Co, ^{133}Ba, ^{137}Cs and ^{40}K

The sources of the Task 1 contained three radionuclides: ^{60}Co, ^{133}Ba and ^{137}Cs. Most facilities have also submitted results of ^{40}K.

The activity has been reported by the participant at measurement date. Nevertheless, the EIVIC team used the date of the start of measurement campaign as 'reference date'. The activities have been corrected by the radioactive decay. The radionuclide parameters [BIPM, 2020] and the sources reference dates are given in Table 5.

Table 5: List of radionuclide parameters and sources reference dates of the Task 1.

Radionuclide	^{60}Co	^{133}Ba	^{137}Cs	^{40}K
Half-life (y)	5.27	10.54	30.05	1.25×10^9
Energy of γ emissions (keV)	1173, 1332	81, 276, 302, 356, 384	662	1461
Reference date (EIVIC)	01/05/2021	01/05/2021	01/05/2021	01/05/2021

5.2.1 Assigned value and statistic parameters

The reference value (Bq) of sources, given in section 3.2, was compared with the robust mean (Bq) of participants for each radionuclide, as recommended in ISO 13528 [ISO 2022] (cf. Table 6). The robust mean was determined with the Q/Hampel robust method implemented in the 'ProLab™' intercomparison software. The statistic parameters used are summarised in the Table 7.

Table 6: Comparison between the reference value and the robust mean of participants for Task 1 "Victor".

Task 1 – Victor	^{60}Co	^{133}Ba	^{137}Cs	^{40}K
Reference value (IRSN) (Bq)	1100 ± 20	2720 ± 60	3850 ± 90	4210 ± 210
Robust mean (Bq)	1183 ± 25	2836 ± 62	3787 ± 65	3941 ± 137
Difference (%)	7.12	4.19	-1.52	-6.41

The difference between the reference value and the robust mean does not exceed 7%.

Even if these differences do not reflect a systematic bias, the robust mean of participants was used as the assigned value, for each radionuclide to ensure a uniform treatment of results between the various tasks.

Table 7: ProLab™ statistic parameters for Task 1 "Victor".

ProLab™ statistic parameters – P4 Victor	^{60}Co	^{133}Ba	^{137}Cs	^{40}K
Statistic method	Q / Hampel	Q / Hampel	Q / Hampel	Q / Hampel
Number of facilities reporting results	39	39	40	28
Number of participants (following plan)	43	43	43	43
Assigned value (Bq)	1183	2836	3787	3941
Robust mean (Bq)	1183	2836	3787	3941
Standard Deviation (SD) (Bq)	155	388	412	725
Relative SD (%)	13.12	13.68	10.87	18.38
Uncertainty of assigned value (Bq)	25	62	65	137

The raw data of participants are represented for ^{60}Co, ^{133}Ba, ^{137}Cs and ^{40}K in Figure 7.

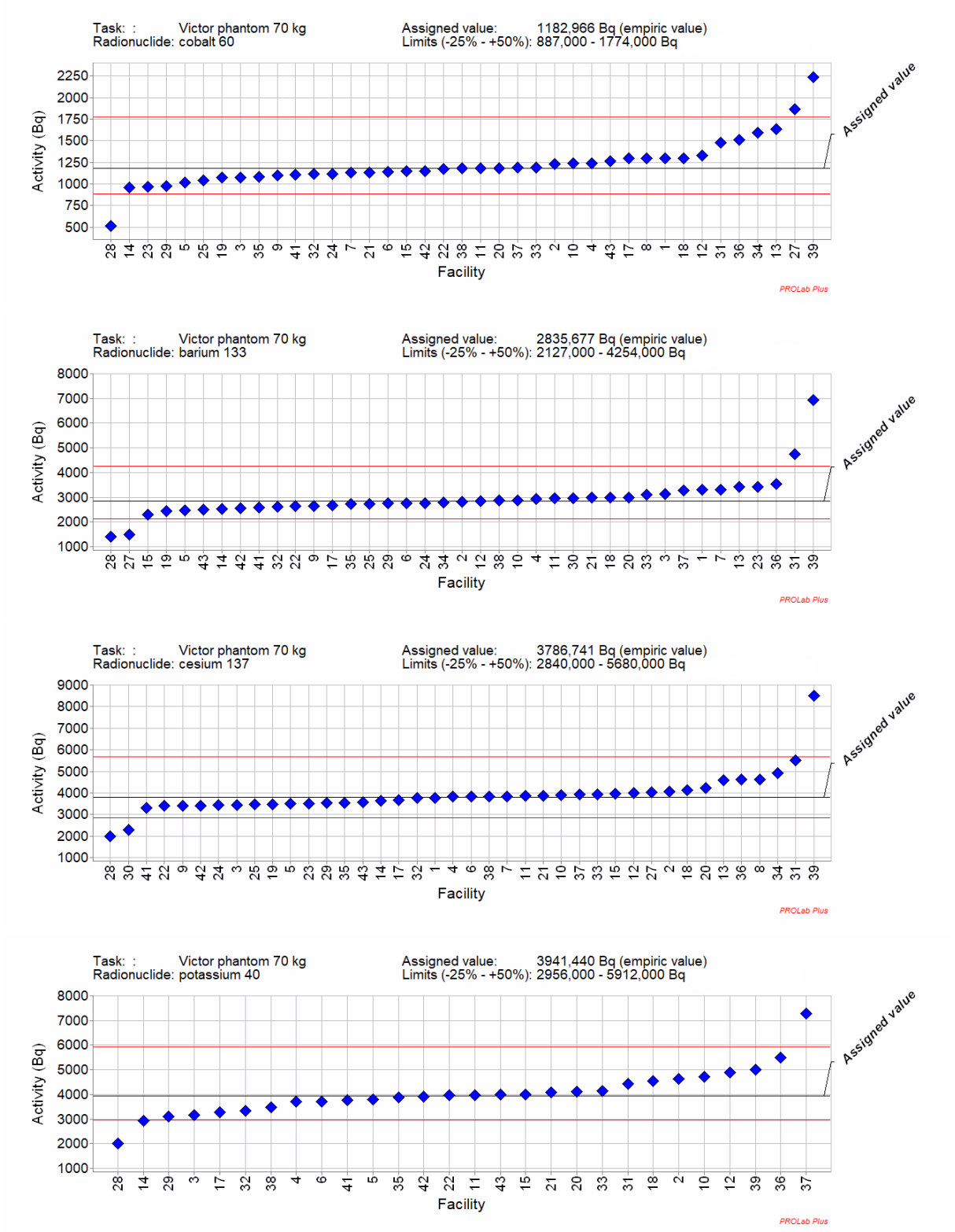


Figure 7: Representation of the raw data of participants and the assigned value: ^{60}Co, ^{133}Ba, ^{137}Cs and ^{40}K for Task 1 "Victor". Red lines: -25% and +50% performance criteria.

5.2.2 Bias and outliers (ISO 28218)

The results of the participants, treated with the standard ISO 28218 [ISO 2010], are compared to the robust mean corresponding to the assigned value as described in section 5.2.1.

Table 8 summarises the bias obtained for the measurement of ^{60}Co, ^{133}Ba, ^{137}Cs and ^{40}K. The bias of participants is shown in Figure 8.

Table 8: Bias of participants for Task 1 "Victor"

ID Lab.	^{60}Co	^{133}Ba	^{137}Cs	^{40}K
1	9.8%	16.6%	-0.3%	
2	4.2%	-1.1%	7.0%	17.6%
3	-9.0%	10.0%	-9.2%	-20.2%
4	4.5%	3.8%	0.9%	-6.1%
5	-13.9%	-13.2%	-7.9%	-3.6%
6	-3.5%	-3.0%	1.0%	-6.1%
7	-4.3%	16.9%	1.6%	
8	9.6%		21.9%	
9	-7.3%	-7.0%	-10.0%	
10	4.5%	1.4%	2.5%	20.0%
11	0.0%	3.9%	2.1%	0.6%
12	12.7%	0.0%	5.7%	23.9%
13	38.1%	20.5%	21.4%	
14	-18.6%	-11.0%	-3.9%	-26.0%
15	-3.0%	-18.5%	4.7%	1.6%
17	9.6%	-6.3%	-3.2%	-16.6%
18	10.0%	5.2%	9.0%	15.5%
19	-9.4%	-14.4%	-8.3%	
20	0.0%	5.2%	11.8%	4.3%
21	-4.0%	5.1%	2.3%	3.8%
22	-0.6%	-7.1%	-10.4%	0.4%
23	-17.8%	20.7%	-7.4%	
24	-5.8%	-3.0%	-9.4%	
25	-11.9%	-3.4%	-8.5%	
27	57.9% Suspicious (Grubbs)	-47.2%	6.2%	
28	-56.5% Suspicious (Grubbs)	-50.2% Suspicious (Grubbs)	-47.8%	-48.7%

ID Lab.	^{60}Co	^{133}Ba	^{137}Cs	^{40}K
29	-17.6%	-3.0%	-7.0%	-21.3%
30		4.2%	-39.4%	
31	25.1%	67.4% Outlier (Grubbs)	45.2%	12.6%
32	-6.0%	-7.4%	-0.8%	-15.1%
33	0.8%	9.3%	3.6%	5.2%
34	34.8%	-1.5%	29.7%	
35	-8.5%	-3.7%	-6.3%	-1.8%
36	27.4%	25.0%	21.8%	39.4%
37	0.4%	15.1%	3.4%	84.3% Outlier (Grubbs)
38	-0.4%	1.3%	1.1%	-11.8%
39	88.8% Outlier (Grubbs)	144.1% Outlier (Grubbs)	124.7% Outlier (Grubbs)	26.9%
41	-6.1%	-9.0%	-12.8%	-4.3%
42	-2.7%	-9.7%	-9.9%	-0.7%
43	7.2%	-11.6%	-5.8%	1.4%

The ^{60}Co results show that three facilities [27, 28 and 39] are in non-conformity according to the [-25%; +50%] criteria of ISO 28218. The other facilities are in conformity. The facilities '27' and '28' have been detected as a suspicious value and the facility '39' has been detected as an outlier (Grubbs test).

The ^{133}Ba results show that four facilities [27, 28, 31 and 39] are in non-conformity according to the [-25%; +50%] criteria ISO 28218 [ISO 2010]. The other facilities are in conformity. Facility '28' has been detected as a suspicious value and the facilities '31' and '39' have been detected as outliers (Grubbs test).

The ^{137}Cs results show that three facilities [28, 30 and 39] are in non-conformity according to the [-25%; +50%] criteria ISO 28218 [ISO 2010]. The other facilities are in conformity. The facility '39' has been detected as an outlier (Grubbs test).

The ^{40}K results show that three facilities [14, 28 and 37] are in non-conformity according to the [-25%; +50%] criteria ISO 28218 [ISO 2010]. The other facilities are in conformity. The facility '37' has been detected as an outlier (Grubbs test).

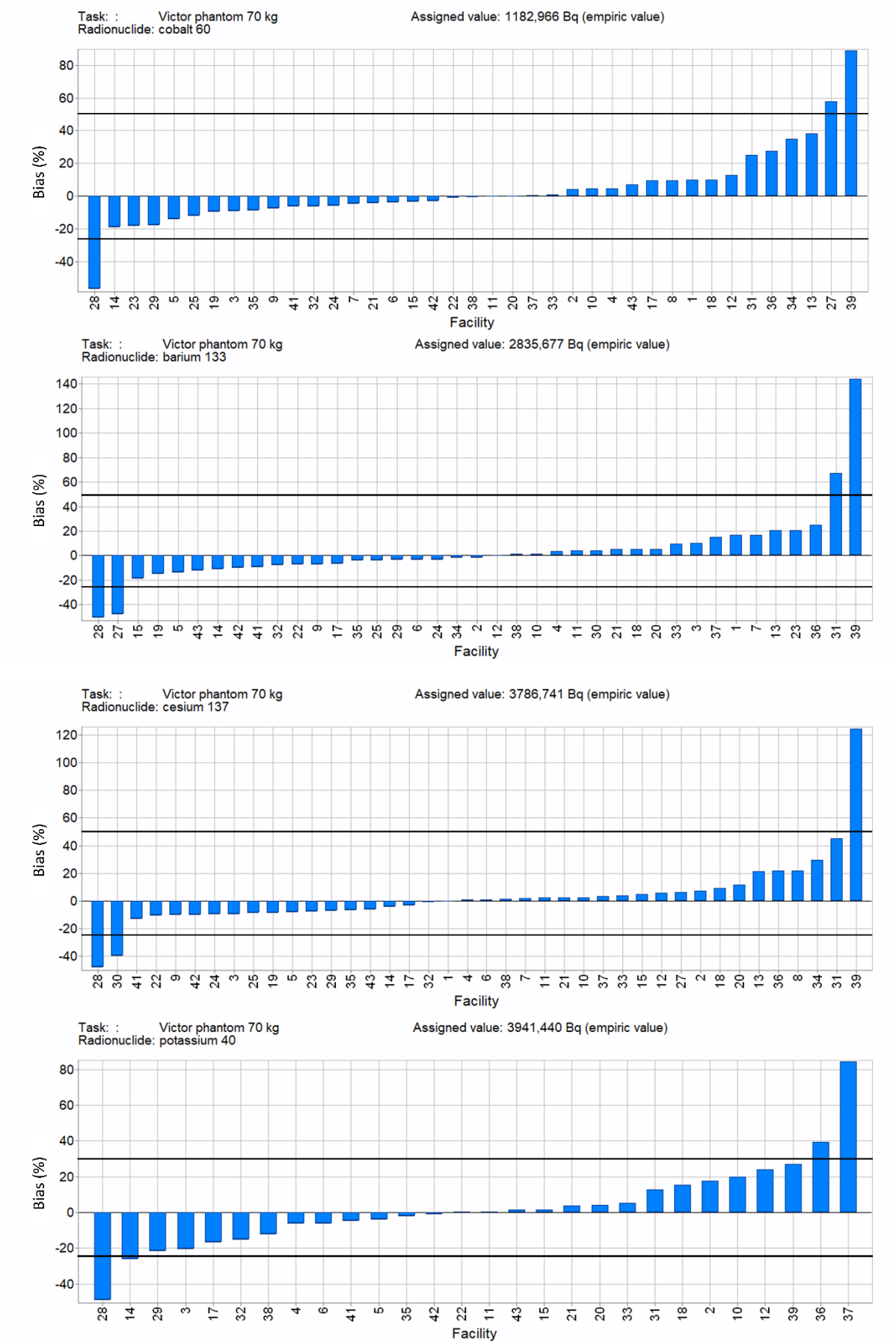


Figure 8: Representation of the bias (%) for the radionuclides of Task 1 "Victor". Black lines: -25% and +50% performance criteria.

5.2.3 Z-score

Table 9 summarises the z-scores obtained for the ^{60}Co, ^{133}Ba, ^{137}Cs and ^{40}K measurements. The z-scores indicated in yellow and red, respectively represent the results to be considered as giving a warning signal and an action signal. The z-scores are shown in Figure 9.

The results for ^{60}Co show that the z-score of the facilities '27', '28' and '39' are larger than 3 ($|z|\geq 3$), give an action signal and the facilities '13', '34' and '36' give a warning signal ($2<|z|<3$) according to ISO 13528 [ISO 2022]. All other facilities are considered satisfactory ($|z|\leq 2$).

The results for ^{133}Ba show that the z-score of the facilities '27', '28', '31' and '39' over 3 ($|z|\geq 3$) give an action signal according to ISO 13528 [ISO 2022]. All other facilities are considered satisfactory ($|z|\leq 2$).

The results for ^{137}Cs show that the z-score of the facilities '28','30', '31' and '39' are larger than 3 ($|z|\geq 3$) and give an action signal and the facilities '8', '34' and '36' give a warning signal ($2<|z|<3$) according to ISO 13528 [ISO 2022]. All other facilities are considered satisfactory ($|z|\leq 2$).

The results for ^{40}K show that the z-score of the facility '37' is over 3 ($|z|\geq 3$), giving an action signal and the facilities '28' and '36' give a warning signal ($2<|z|<3$) according to ISO 13528 [ISO 2022]. All other facilities are considered satisfactory ($|z|\leq 2$).

Table 9: Z-score of participants for Task 1 "Victor".

ID Lab.	^{60}Co	^{133}Ba	^{137}Cs	^{40}K
1	0.75	1.22	-0.03	
2	0.32	-0.08	0.64	0.96
3	-0.68	0.73	-0.84	-1.10
4	0.34	0.28	0.08	-0.33
5	-1.06	-0.97	-0.73	-0.20
6	-0.27	-0.22	0.09	-0.33
7	-0.33	1.23	0.15	
8	0.73		2.01	
9	-0.55	-0.51	-0.92	
10	0.34	0.10	0.23	1.09
11	0.00	0.28	0.19	0.03
12	0.97	0.00	0.52	1.30
13	2.91	1.50	1.97	
14	-1.42	-0.80	-0.36	-1.42
15	-0.23	-1.36	0.43	0.09
17	0.73	-0.46	-0.29	-0.90
18	0.76	0.38	0.83	0.84

ID Lab.	^{60}Co	^{133}Ba	^{137}Cs	^{40}K
19	-0.72	-1.05	-0.76	
20	0.00	0.38	1.08	0.23
21	-0.30	0.37	0.21	0.21
22	-0.04	-0.52	-0.95	0.02
23	-1.36	1.51	-0.68	
24	-0.44	-0.22	-0.86	
25	-0.91	-0.25	-0.78	
27	4.41	-3.45	0.57	
28	-4.30	-3.67	-4.40	-2.65
29	-1.34	-0.22	-0.64	-1.16
30		0.31	-3.62	
31	1.91	4.93	4.16	0.69
32	-0.46	-0.54	-0.07	-0.82
33	0.06	0.68	0.33	0.28
34	2.66	-0.11	2.74	
35	-0.65	-0.27	-0.58	-0.10
36	2.09	1.83	2.00	2.14
37	0.03	1.10	0.31	4.59
38	-0.03	0.10	0.10	-0.64
39	6.77	10.54	11.47	1.46
41	-0.46	-0.66	-1.18	-0.24
42	-0.21	-0.71	-0.91	-0.04
43	0.55	-0.84	-0.53	0.07

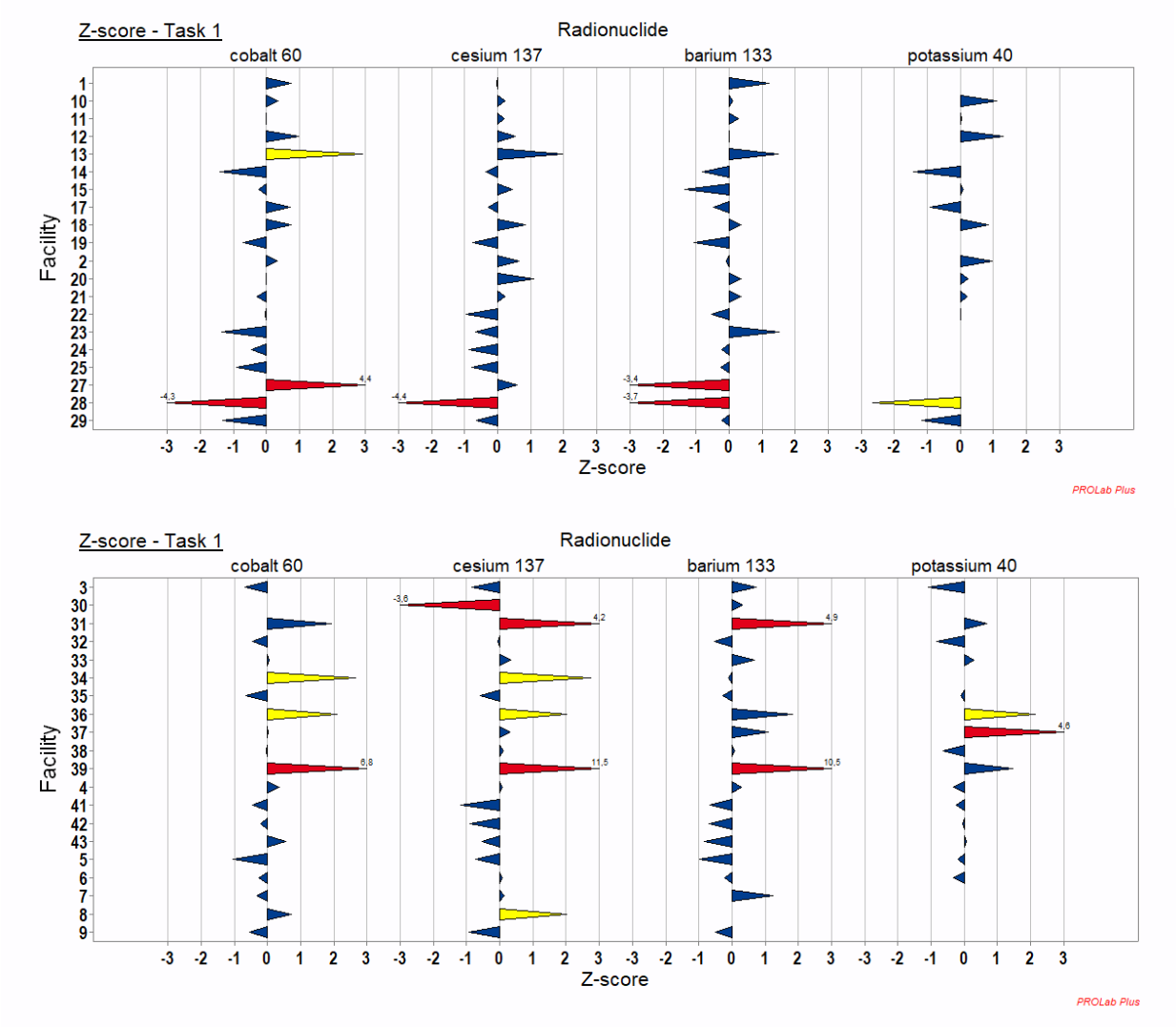


Figure 9: Representation of the z-score for the radionuclides of Task 1 "Victor". Yellow: warning signal, red: action signal.

5.3 Task 2 – Emergency (90 kg): ^{134}Cs, ^{137}Cs and ^{40}K

The sources of the Task 2 contained two radionuclides: ^{134}Cs and ^{137}Cs. Most facilities have also submitted results of ^{40}K. As for Task 1, the activities have been reported by the participant at measurement date and corrected for the radioactive decay. The radionuclide parameters [BIPM, 2020] and sources reference dates are given in Table 10.

Table 10: List of radionuclide parameters and sources reference dates of Task 2 "Emergency".

Radionuclide	^{134}Cs	^{137}Cs	^{40}K
Half-life (y)	2.06	30.05	1.25×10^9
Energy of γ emissions (keV)	605, 796	662	1461
Reference date (EIVIC)	01/05/2021	01/05/2021	01/05/2021

5.3.1 Assigned value and statistic parameters

The reference value of sources, given in section 3.2, was compared with the robust mean of participants for each radionuclide, as recommended in ISO 13528 [ISO 2022] (cf. Table 11). The robust mean was determined with the Q/Hampel robust method implemented in the ProLab™ intercomparison software. The statistic parameters used are summarised in the Table 12.

Table 11: Comparison between the reference value and the robust mean of participants for Task 2 "Emergency".

Task 2 – Emergency	^{134}Cs	^{137}Cs	^{40}K
Reference value (BfS) (Bq)	3490 ± 40	3220 ± 40	5460 ± 270
Robust mean (Bq)	3455 ± 50	2996 ± 60	4981 ± 193
Difference (%)	-1.04	-7.09	-8.68

The difference between the reference value and the robust mean does not exceed 9%.

Even if there is a small systematic underestimation, these differences do not reflect a systematic bias, the robust mean of participants was used as the assigned value for each radionuclide, to ensure a uniform treatment of results between the various tasks.

The raw data of participants are represented for ^{134}Cs, ^{137}Cs and ^{40}K in Figure 10.

Table 12: ProLab™ statistic parameters for Task 2 "Emergency".

ProLab™ statistic parameters – P5 Emergency	^{134}Cs	^{137}Cs	^{40}K
Statistic method	Q / Hampel	Q / Hampel	Q / Hampel
Number of facilities reporting results	40	40	28
Number of participants (following plan)	43	43	43
Assigned value (Bq)	3455	2996	4981
Robust mean (Bq)	3455	2996	4981
Standard Deviation (SD) (Bq)	313	382	1021
Relative SD (%)	9.07	12.75	20.49
Uncertainty of assigned value (Bq)	50	60	193

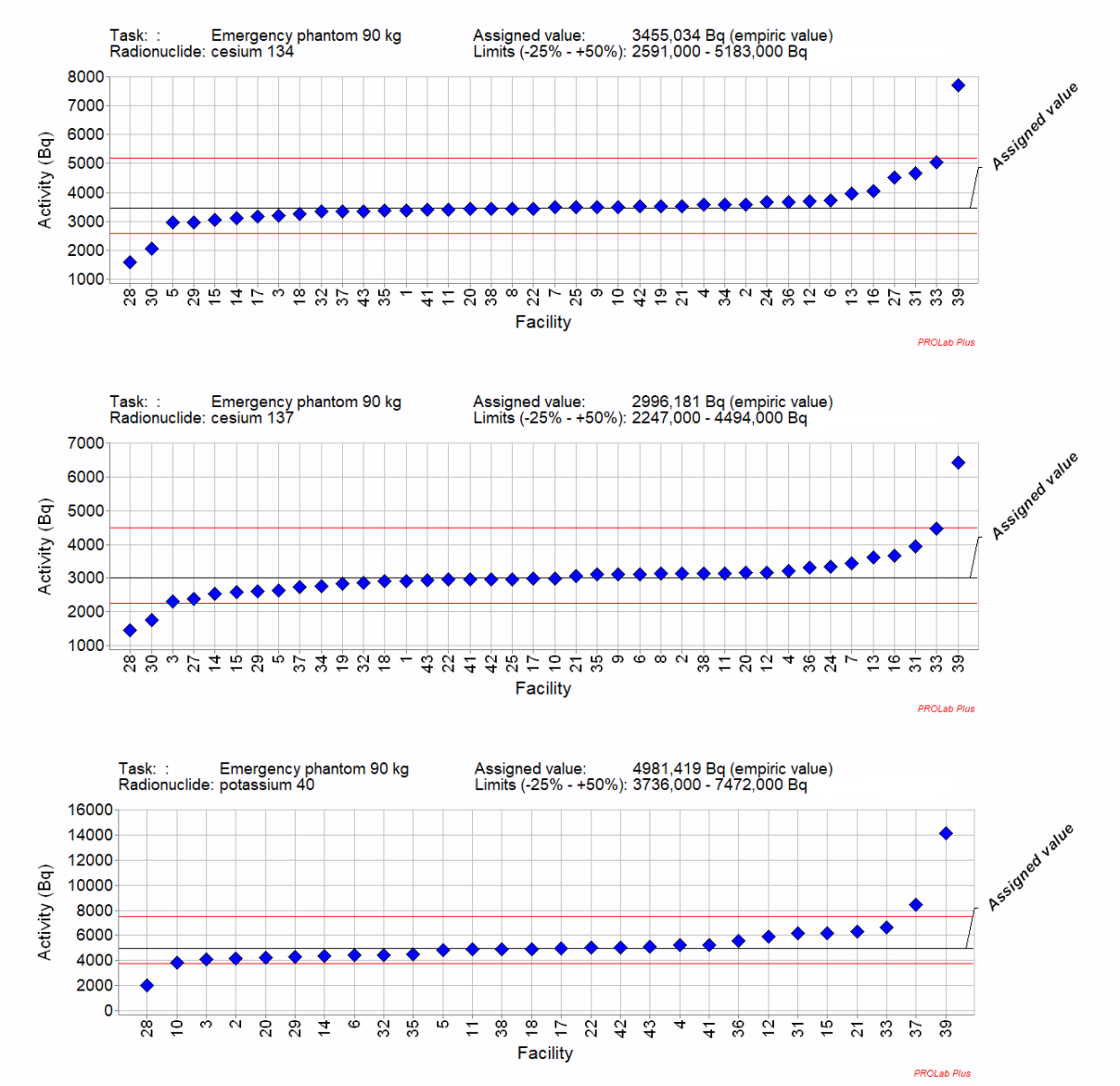


Figure 10: Representation of the raw data of participants and the assigned value: ^{134}Cs, ^{137}Cs and ^{40}K for Task 2 "Emergency". Red lines: -25% and +50% performance criteria.

5.3.2 Bias and outliers (ISO 28218)

The results of the participants, treated with the standard ISO 28218 [ISO 2010], are compared to the robust mean corresponding to the assigned value as described in section 5.3.1.

Table 13 summarises the bias obtained for the measurement of ^{134}Cs, ^{137}Cs and ^{40}K. The bias of participants is shown in Figure 11.

Table 13: Bias of participants for Task 2 "Emergency".

ID Lab.	^{134}Cs	^{137}Cs	^{40}K
1	-2.2%	-2.7%	
2	3.9%	4.6%	-16.7%
3	-7.4%	-23.3%	-18.0%
4	3.1%	7.4%	4.4%
5	-14.5%	-12.1%	-3.6%
6	8.0%	4.1%	-11.7%
7	1.2%	15.0%	
8	-0.5%	4.4%	
9	1.2%	3.9%	
10	1.3%	-0.1%	-22.9%
11	-1.1%	4.8%	-1.8%
12	6.7%	5.3%	18.5%
13	14.5%	20.5%	
14	-10.3%	-15.2%	-12.3%
15	-11.7%	-13.8%	24.0%
16	16.7%	22.0%	
17	-8.2%	-0.3%	-1.0%
18	-6.1%	-3.2%	-1.3%
19	1.8%	-5.5%	
20	-0.9%	5.2%	-15.2%
21	1.8%	2.0%	25.7%
22	-0.4%	-1.4%	1.1%

ID Lab.	^{134}Cs	^{137}Cs	^{40}K
24	5.7%	11.5%	
25	1.2%	-1.0%	
27	30.7%	-20.6%	
28	-54.3% Suspicious (Grubbs)	-51.9%	-59.0%
29	-14.1%	-12.7%	-13.7%
30	-40.6%	-41.2%	
31	35.2%	31.2%	23.4%
32	-3.4%	-4.2%	-11.2%
33	45.8%	48.8%	33.3%
34	3.6%	-8.2%	
35	-2.5%	3.9%	-9.7%
36	6.4%	10.1%	12.1%
37	-3.0%	-8.5%	69.5%
38	-0.8%	4.6%	-1.3%
39	123.0% Outlier (Grubbs)	114.0%	183.1% Outlier (Grubbs)
41	-1.7%	-1.2%	4.7%
42	1.5%	-1.0%	1.4%
43	-2.9%	-1.7%	2.2%

The ^{134}Cs results show that three facilities [28, 30 and 39] are in non-conformity according to the [-25%; +50%] criteria of the standard ISO 28218 [ISO 2010]. The other facilities are in conformity. For the facility '28' a suspicious value has been detected and the facility '39' has been considered as an outlier (Grubbs test).

The ^{137}Cs results show that three facilities [28, 30 and 39] are in non-conformity according to the [-25%; +50%] criteria of the standard ISO 28218 [ISO 2010]. The other facilities are in conformity.

The ^{40}K results show that three facilities [28, 37 and 39] are in non-conformity according to the [-25%; +50%] criteria of the standard ISO 28218 [ISO 2010]. The other facilities are in conformity. The facility '39' has been detected as an outlier (Grubbs test).

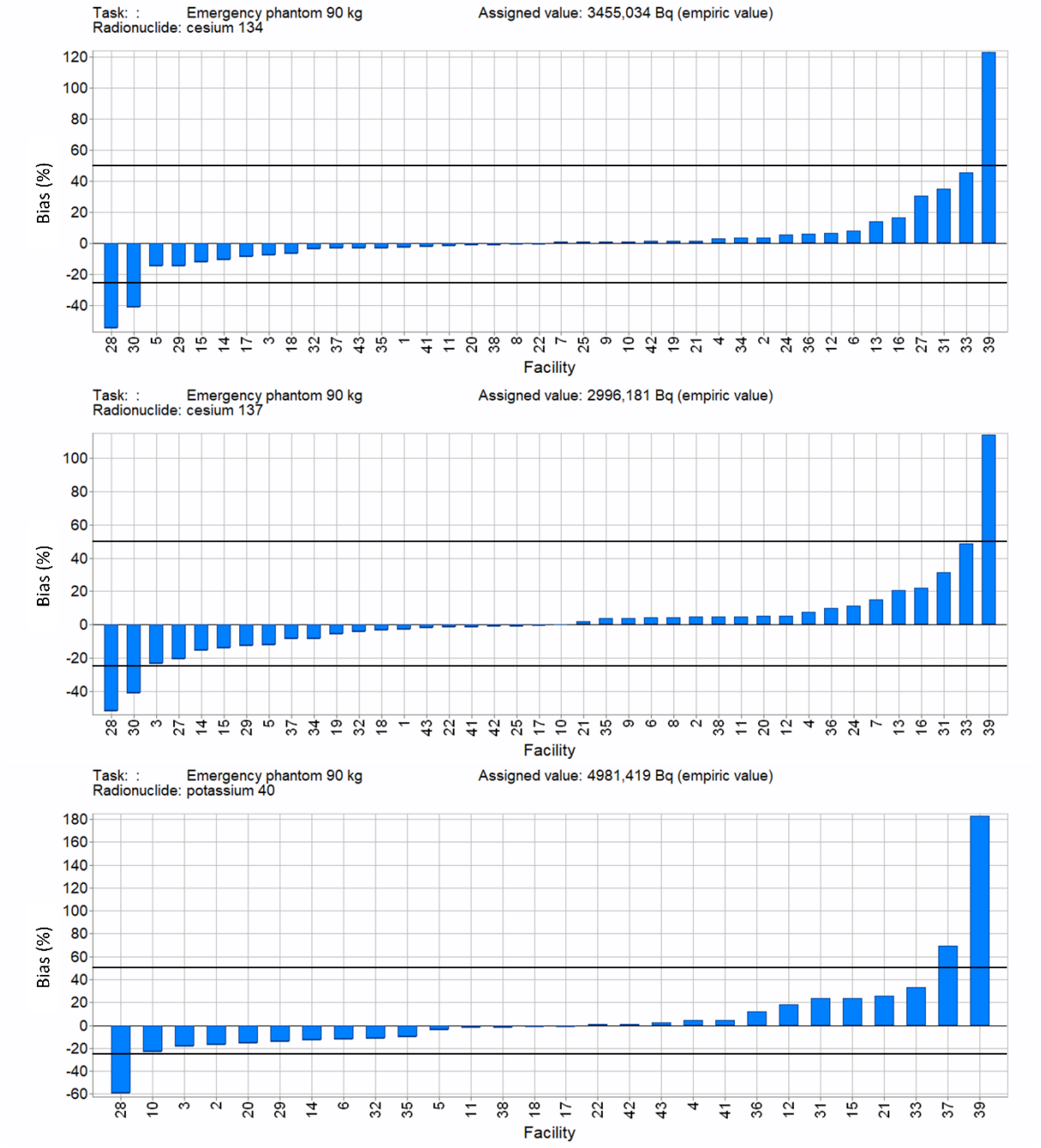


Figure 11: Representation of the bias (%) for the radionuclides of Task 2 "Emergency". Black lines: -25% and +50% performance criteria.

5.3.3 Z-score (ISO 13528)

Table 14 summarises the z-scores obtained for the ^{134}Cs, ^{137}Cs and ^{40}K measurements. The z-scores indicated in yellow and red, respectively, represent the results to be considered as giving a warning signal and an action signal. The z-scores are shown in Figure 12.

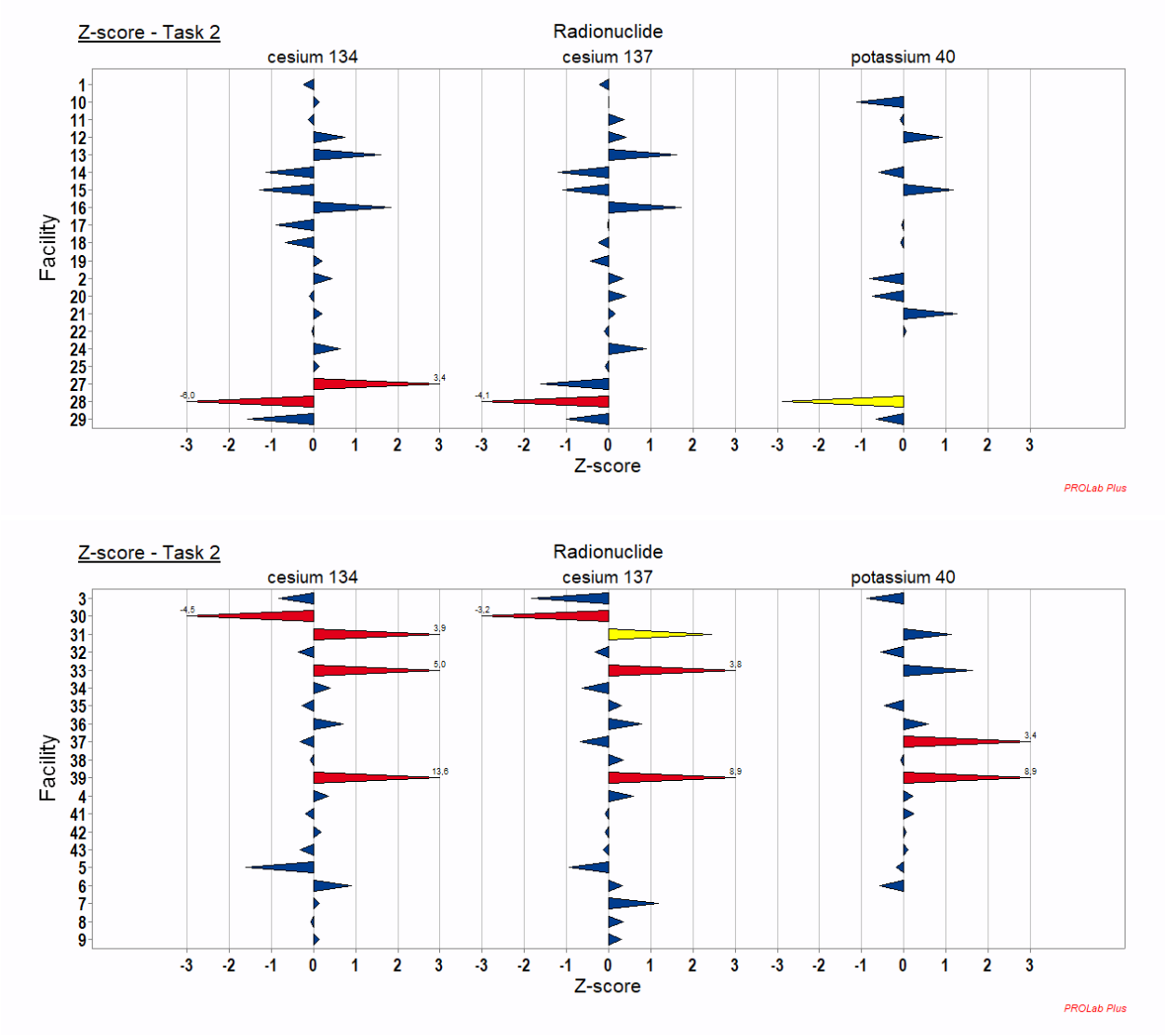


Figure 12: Representation of the z-score for the radionuclides of Task 2 "Emergency". Yellow: warning signal, red: action signal.

The results for ^{134}Cs show that the z-scores of the facilities '27', '28', '30', '31', '33' and '39' are larger than 3 ($|z| \geq 3$) and give an action signal according to ISO 13528 [ISO 2022]. All other facilities are considered satisfactory ($|Z| \leq 2$).

The results for ^{137}Cs show that the z-scores of the facilities '28','30', '33' and '39' are larger than 3 ($|z| \geq 3$), give an action signal and the facility '31' gives a warning signal ($2 < |z| < 3$) according to ISO 13528 [ISO 2022]. All other facilities are considered satisfactory ($|z| \leq 2$).

The results for ^{40}K show that the z-scores of the facilities '37' and '39' are over 3 ($|z| \geq 3$) and give an action signal and the facility '28' gives a warning signal ($2 < |z| < 3$) according to ISO 13528 [ISO 2022]. All other facilities are considered satisfactory ($|z| \leq 2$).

Table 14: Z-scores of participants for Task 2 "Emergency".

ID Lab.	^{134}Cs	^{137}Cs	^{40}K
1	-0.24	-0.22	-
2	0.43	0.36	-0.81
3	-0.82	-1.83	-0.88
4	0.34	0.58	0.21
5	-1.60	-0.95	-0.18
6	0.89	0.32	-0.57
7	0.13	1.18	-
8	-0.05	0.34	-
9	0.14	0.31	-
10	0.14	-0.01	-1.12
11	-0.12	0.38	-0.09
12	0.74	0.42	0.90
13	1.60	1.61	-
14	-1.13	-1.19	-0.60
15	-1.29	-1.08	1.17
16	1.84	1.72	-
17	-0.90	-0.02	-0.05
18	-0.67	-0.25	-0.06
19	0.20	-0.44	-
20	-0.10	0.41	-0.74
21	0.20	0.16	1.26
22	-0.04	-0.11	0.05
24	0.63	0.90	-
25	0.13	-0.08	-
27	3.39	-1.62	-
28	-5.98	-4.07	-2.88
29	-1.56	-0.99	-0.67

ID Lab.	^{134}Cs	^{137}Cs	^{40}K
30	-4.47	-3.23	-
31	3.88	2.45	1.14
32	-0.38	-0.33	-0.55
33	5.05	3.82	1.62
34	0.40	-0.64	-
35	-0.28	0.31	-0.47
36	0.71	0.79	0.59
37	-0.33	-0.67	3.39
38	-0.09	0.36	-0.06
39	13.56	8.95	8.93
41	-0.19	-0.09	0.23
42	0.17	-0.08	0.07
43	-0.32	-0.13	0.11

5.4 Task 3 – Medicine #1: ^{68}Ge/^{68}Ga, ^{88}Y and ^{40}K

The sources of Task 2 contained two radionuclides: ^{68}Ge/^{68}Ga and ^{88}Y. Most facilities have also received results of ^{40}K. As for Task 1, the activities have been reported by the participant at measurement date and have been corrected for the radioactive decay. The radionuclide parameters [BIPM, 2020] and sources reference dates are given in Table 15.

Table 15: List of radionuclide parameters and sources reference dates of Task 3 "Medicine".

Radionuclide	^{68}Ge/^{68}Ga	^{88}Y	^{40}K
Half-life	270.95 d	106.63 d	1.25x10^{9} y
Energy of γ emissions (keV)	511, 1077	898, 1836	1461
Reference date (EIVIC)	01/05/2021	01/05/2021	01/05/2021

5.4.1 Assigned value and statistic parameters

The reference value (Bq) of sources, given in section 3.2, was compared with the robust mean (Bq) of participants, for each radionuclide, as recommended in ISO 13528 [ISO 2022] (cf. Table 16). The robust mean was determined with the Q/Hampel robust method implemented in the 'ProLab™' intercomparison software. The statistic parameters used are summarised in Table 17.

Table 16: Comparison between the reference value and the robust mean of participants for Task 3 "Medicine" #1.

Task 3 – Medicine #1	^{68}Ge/^{68}Ga	^{88}Y	^{40}K
Reference value (BfS) (Bq)	4192 ± 410	4722 ± 119	4210 ± 210
Robust mean (Bq)	3741 ± 114	4707 ± 163	3862 ± 127
Difference (%)	-10.76	-0.32	-8.27

The difference between the reference value and the robust mean does not exceed 11%.

Even if these differences do not reflect a systematic bias, the robust mean of participants was used as the assigned value, for each radionuclide, to ensure a uniform treatment of results between the various tasks.

Table 17: ProLab™ statistic parameters for Task 3 "Medicine" #1.

ProLab™ statistic parameters – P4 Medicine #1	^{68}Ge/^{68}Ga	^{88}Y	^{40}K
Statistic method	Q / Hampel	Q / Hampel	Q / Hampel
Number of facilities reporting results	9	15	11
Number of participants (following plan)	15	15	15
Assigned value (Bq)	3741	4707	3862
Robust mean (Bq)	3741	4707	3862
Standard Deviation (SD) (Bq)	341	630	422
Relative SD (%)	9.11	13.38	10.38
Uncertainty of assigned value (Bq)	114	163	127

The raw data of participants are represented for ^{68}Ge, ^{88}Y and ^{40}K in Figure 13.

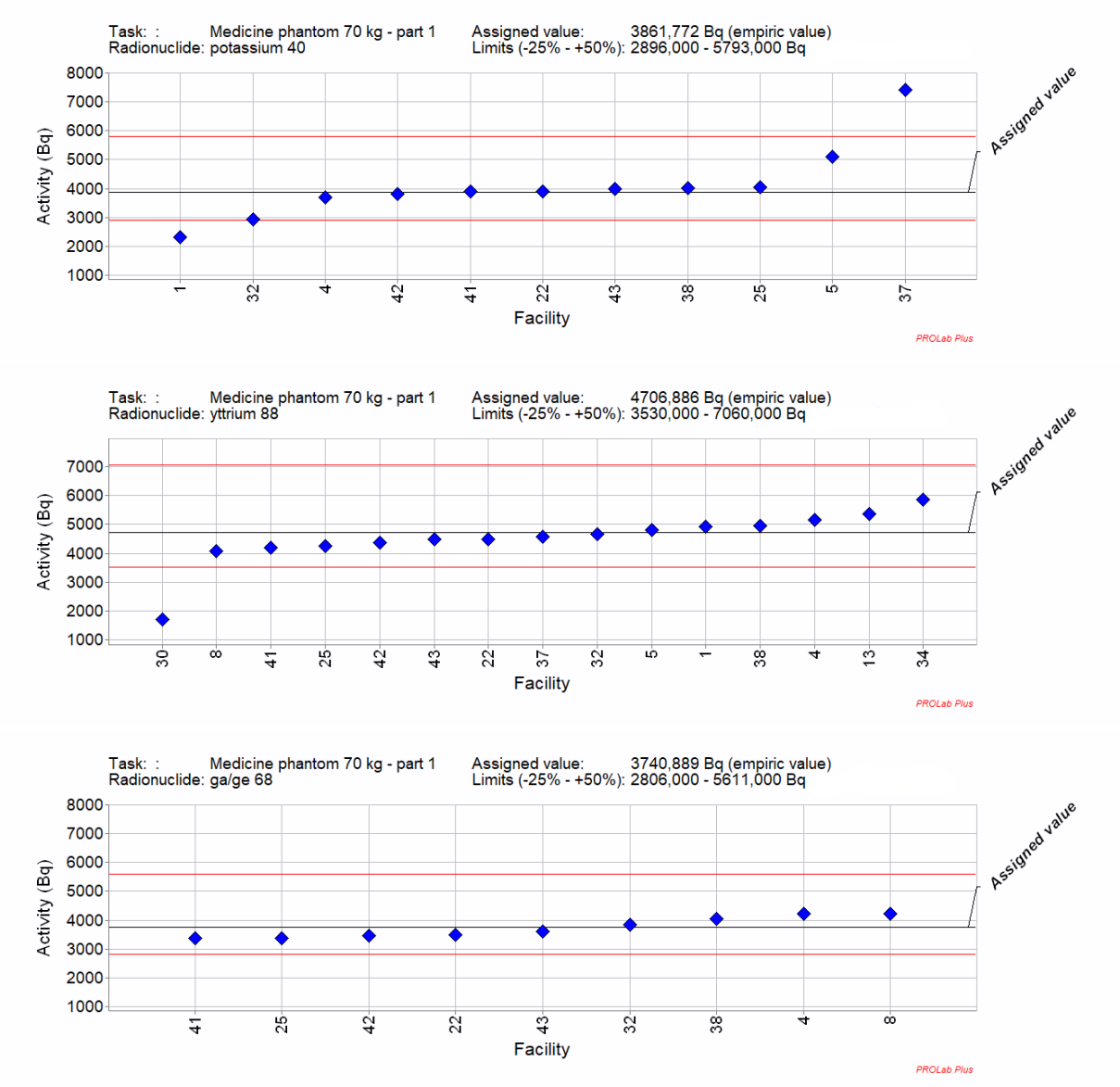


Figure 13: Representation of the raw data of participants and the assigned value: $^{68}Ge/^{68}Ga$ and ^{88}Y and ^{40}K for Task 3 "Medicine" #1. Red lines: -25% and +50% performance criteria.

5.4.2 Bias and outliers (ISO 28218)

The results of the participants, treated with the standard ISO 28218 [ISO 2010], are compared to the robust mean corresponding to the assigned value as described in section 5.4.1.

Table 18 summarises the bias obtained for the measurement of ^{68}Ge/^{68}Ga, ^{88}Y and ^{40}K. The bias of participants is shown in Figure 14.

The ^{68}Ge results show that all facilities are in conformity according to the [-25%; +50%] criteria of the standard ISO 28218 [ISO 2010].

The ^{88}Y results show that one facility [30] is in non-conformity according to the [-25%; +50%] criteria of the standard ISO 28218 [ISO 2010]. The other facilities are in conformity.

The ^{40}K results show that two facilities [1 and 37] are in non-conformity according to the [-25%; +50%] criteria of the standard ISO 28218 [ISO 2010]. Other facilities are in conformity. The result of the facility '37' has been detected as an outlier (Grubbs test).

Table 18: Bias of participants for Task 3 "Medicine" #1.

ID Lab.	^{68}Ge/^{68}Ga	^{88}Y	^{40}K
1		4.7%	-40.3%
4	12.5%	9.7%	-4.2%
5		1.9%	32.1%
8	12.9%	-13.5%	
13		14.2%	
22	-6.4%	-4.6%	1.2%
25	-9.7%	-9.9%	5.1%
30		-63.6%	
32	2.9%	-1.2%	-24.0%
34		24.5%	
37		-3.0%	92.2% Outlier (Grubbs)
38	8.3%	5.0%	4.1%
41	-9.9%	-10.9%	0.9%
42	-7.2%	-7.5%	-1.5%
43	-3.3%	-4.9%	3.3%

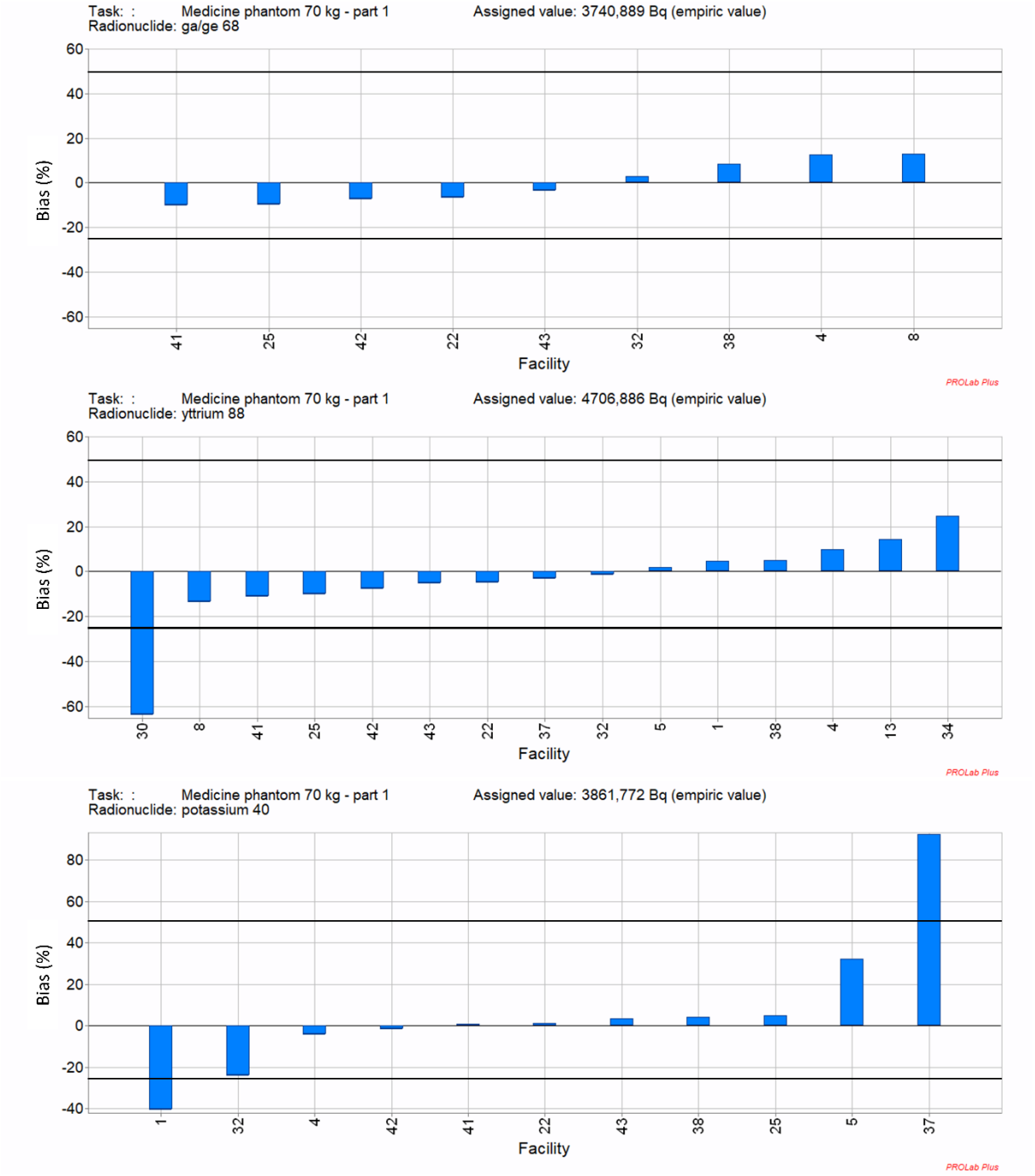


Figure 14: Representation of the bias (%) for the radionuclides of Task 3 "Medicine". Black lines: -25% and +50% performance criteria.

5.4.3 Z-score (ISO 13528)

Table 19 summarises the z-scores obtained for the ^{68}Ge/^{68}Ga, ^{88}Y and ^{40}K measurements. The z-scores, in yellow and red, respectively represent the results to be considered as giving a warning signal and an action signal. The z-scores are shown in Figure 15 .

Table 19: Z-score of participants for Task 3 "Medicine" #1.

ID Lab.	^{68}Ge/Ga	^{88}Y	^{40}K
1		0.4	-3.7
4	1.4	0.7	-0.4
5		0.1	2.9
8	1.4	-1.0	
13		1.1	
22	-0.7	-0.3	0.1
25	-1.1	-0.7	0.5
30		-4.8	
32	0.3	-0.1	-2.2
34		1.8	
37		-0.2	8.4
38	0.9	0.4	0.4
41	-1.1	-0.8	0.1
42	-0.8	-0.6	-0.1
43	-0.4	-0.4	0.3

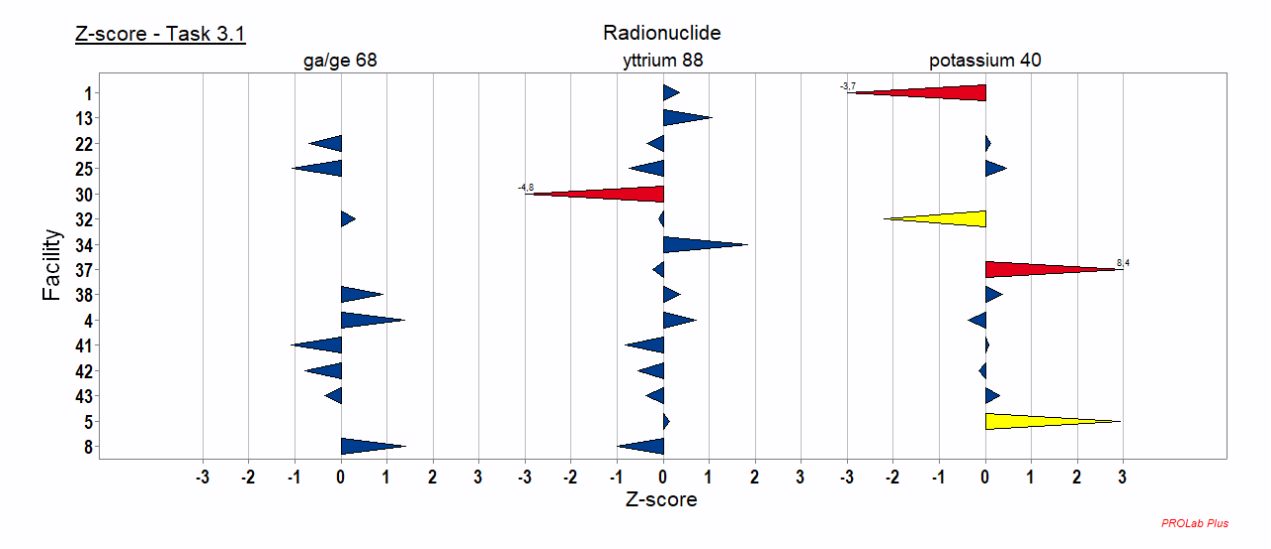


Figure 15: Representation of the z-score for the radionuclides of Task 3 "Medicine" #1. Yellow: warning signal, red: action signal.

The results for ^{68}Ge/^{68}Ga show, that the z-score of all facilities is considered satisfactory (|z|≤2).

The results for ^{88}Y show that the z-score of the facility '30' is over 3 (|z|≥3) and give an action signal according to ISO 13528. All other facilities are considered satisfactory (|z|≤2).

The results for ^{40}K show that the z-score of the facilities '1' and '37' are over 3 (|z|≥3), give an action signal and the facility '5' and '32' give a warning signal (2<|z|<3) according to ISO 13528. All other facilities are considered satisfactory (|z|≤2).

It should be noted that the z-score evaluation, for the ^{68}Ge/^{68}Ga, must be interpreted carefully, because the number of facilities reporting results is less than 10.

5.5 Task 3 – Medicine #2: ^{68}Ge/^{68}Ga, ^{88}Y and ^{40}K

The sources of Task 3 #2 contained two radionuclides: ^{68}Ge/^{68}Ga and ^{88}Y. Most facilities have also received results of ^{40}K. As for previous tasks, the activities have been reported by the participants at measurement date and corrected by the radioactive decay. The radionuclide parameters [BIPM, 2020] and sources reference dates are given in Table 20.

Table 20: List of radionuclide parameters and sources reference dates of Task 3 "Medicine" #2.

Radionuclide	^{68}Ge/^{68}Ga	^{88}Y	^{40}K
Half-life	270.95 d	106.63 d	1.25x10^{9} y
Energy of γ emissions (keV)	511, 1077	898, 1836	1461
Reference date (EIVIC)	10/08/2021	10/08/2021	01/05/2021

5.5.1 Assigned value and statistic parameters

The reference value (Bq) of sources, given in section 3.2, was compared with the robust mean (Bq) of participants for each radionuclide, as recommended in ISO 13528 [ISO 2022] (cf. Table 21). The robust mean was determined with the Q/Hampel robust method implemented in the 'ProLab™' intercomparison software. The statistic parameters used are summarised in the Table 22.

Table 21: Comparison between the reference value and the robust mean of participants for Task 3 "Medicine" #2.

Task 3 – Medicine #2	^{68}Ge	^{88}Y	^{40}K
Reference value (BfS) (Bq)	4625 ± 452	4396 ± 110	4210 ± 210
Robust mean (Bq)	4219 ± 15	4263 ± 127	4163 ± 201
Difference (%)	-8.79	-3.02	-1.13

The difference between the reference value and the robust mean does not exceed 9%.

Even if these differences do not reflect a systematic bias, the robust mean of participants was used as the assigned value for each radionuclide, to ensure a uniform treatment of results between the various tasks.

Table 22: ProLab™ statistic parameters for Task 3 "Medicine" #2.

ProLab™ statistic parameters – P4 Medicine #2	^{68}Ge	^{88}Y	^{40}K
Statistic method	Q / Hampel	Q / Hampel	Q / Hampel
Number of facilities reporting results	8	22	18
Number of participants (following plan)	23	23	23
Assigned value (Bq)	4219	4263	4163
Robust mean (Bq)	4219	4263	4163
Standard Deviation (SD) (Bq)	428	595	853
Relative SD (%)	10.15	13.96	20.49
Uncertainty of assigned value (Bq)	151	127	201

The raw data of participants are represented for ^{68}Ge/^{68}Ga, ^{88}Y and ^{40}K in Figure 16.

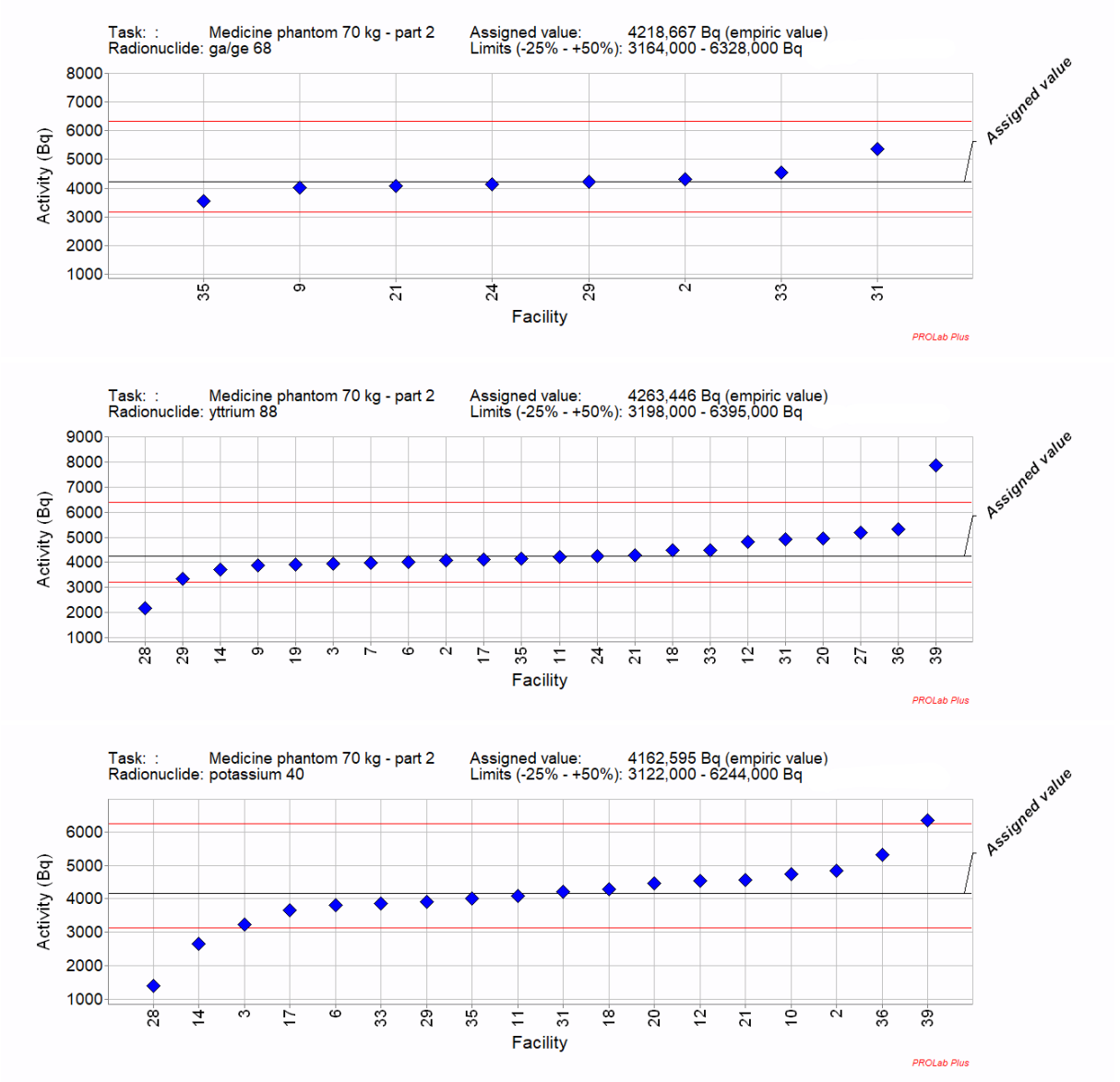


Figure 16: Representation of the raw data of participants and the assigned value: ^{68}Ge/^{68}Ga, ^{88}Y and ^{40}K for Task 3 "Medicine" #2. Red lines: -25% and +50% performance criteria.

5.5.2 Bias and outliers (ISO 28218)

The results of the participants, treated with the standard ISO 28218 [ISO 2010], are compared to the robust mean corresponding to the assigned value as described in section 5.5.1.

Table 23 summarises the bias obtained for the measurement of ^{68}Ge/^{68}Ga, ^{88}Y and ^{40}K. The bias of participants is shown in Figure 17.

Table 23: Bias of participants for Task 3 "Medicine" #2.

ID Lab.	^{68}Ge/^{68}Ga	^{88}Y	^{40}K
2	2.3%	-4.2%	16.1%
3		-7.6%	-22.0%
6		-5.8%	-8.7%
7		-6.8%	
9	-4.8%	-9.1%	
10			13.6%
11		-1.2%	-2.0%
12		12.8%	9.4%
14		-13.3%	-36.2%
17		-3.5%	11.8%
18		4.7%	3.0%
19		-8.0%	
20		16.2%	7.1%
21	-3.4%	0.7%	9.7%
24	-2.1%	0.0%	
27		21.6%	
28		-49.1% suspicious (Grubbs)	-66.4%
29	0.04%	-21.8%	-6.3%
31	27.2%	15.4%	1.2%
33	7.9%	5.1%	-7.4%
35	-16.1%	-2.5%	-3.9%
36		24.5%	27.9%
39		84.5% outlier (Grubbs)	52.5%

The ^{68}Ge results show that all facilities are in conformity according to the [-25%; +50%] criteria of the standard ISO 28218 [ISO 2010].

The ^{88}Y results show that two facilities [28, 39] are in non-conformity according to the criteria [-25%; +50%] of the standard ISO 28218 [ISO 2010]. The other facilities are in conformity. The facility '39' has been detected as an outlier and '28' as a suspicious value (Grubbs test).

The ^{40}K results show that three facilities [14, 28 and 39] are in non-conformity according to the [-25%; +50%] criteria of the standard ISO 28218 [ISO 2010]. Other facilities are in conformity.

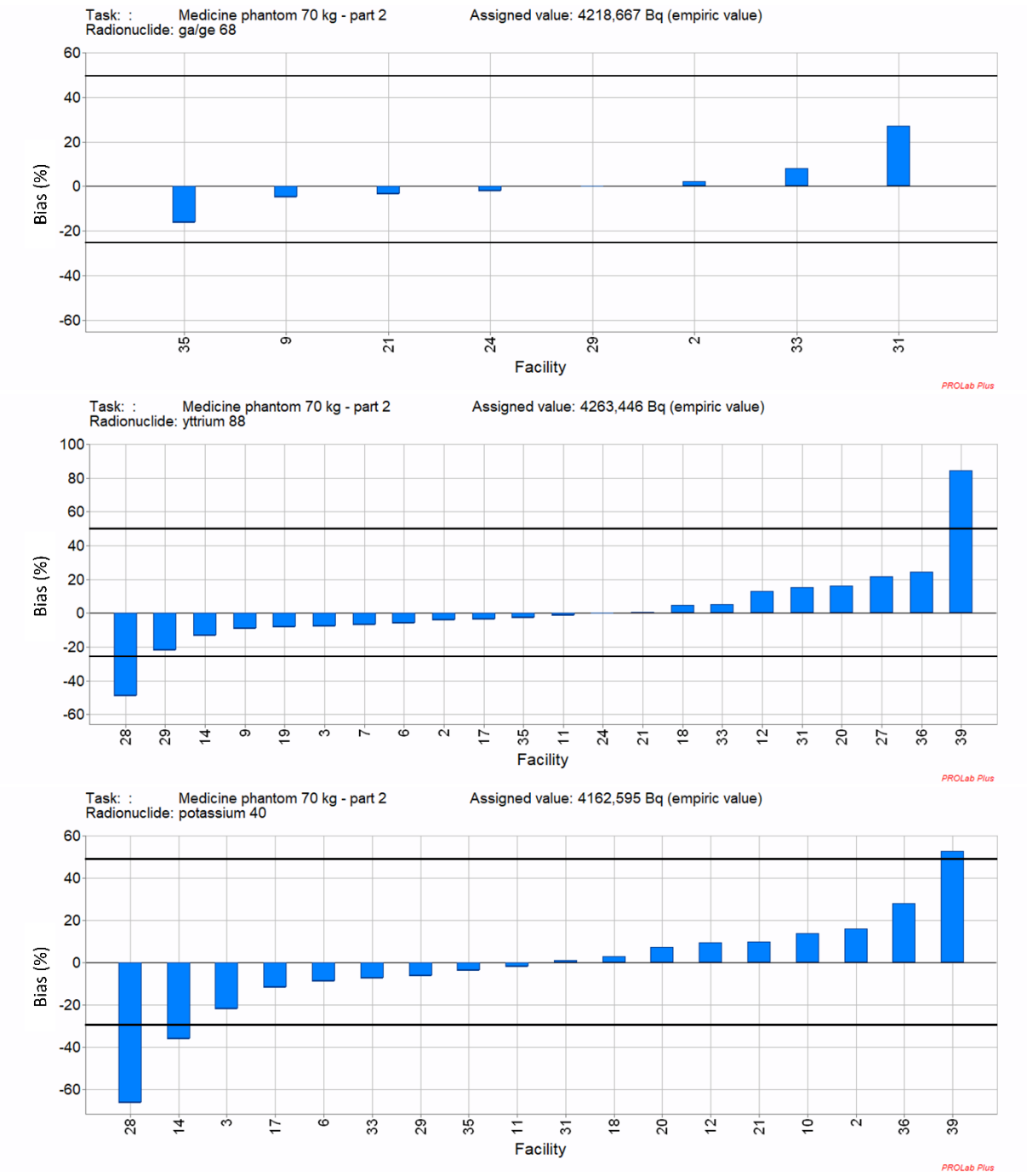


Figure 17: Representation of the bias (%) for the radionuclides of the Task 3 "Medicine" #2. Black lines: -25% and +50% performance criteria.

5.5.3 Z-score (ISO 13528)

Table 24 summarises the z-scores obtained for the $^{68}Ge/^{68}Ga$, ^{88}Y and ^{40}K measurements. The z-scores indicated in yellow and red, respectively, represent the results to be considered as giving a warning signal and an action signal. The z-scores are shown in Figure 18.

Table 24: Z-score of participants for Task 3 "Medicine" #2.

ID Lab.	$^{68}Ge/^{68}Ga$	^{88}Y	^{40}K
2	0.2	-0.3	0.8
3		-0.5	-1.1
6		-0.4	-0.4
7		-0.5	
9	-0.5	-0.7	
10			0.7
11		-0.1	-0.1
12		0.9	0.5
14		-0.9	-1.8
17		-0.3	-0.6
18		0.3	0.1
19		-0.6	
20		1.2	0.3
21	-0.3	0.0	0.5
24	-0.2	0.0	
27		1.5	
28		-3.5	-3.2
29	0.0	-1.6	-0.3
31	2.7	1.1	0.1
33	0.8	0.4	-0.4
35	-1.6	-0.2	-0.2
36		1.8	1.4
39		6.1	2.6

The results for $^{68}Ge/^{68}Ga$ show that the z-score of facility '31' gives a warning signal (2<|z|<3). All other facilities are considered satisfactory (|z|<2).

The results for ^{88}Y show that the z-score of facility '28 and '39' are over 3 (|z|≥3), give an action signal according to ISO 13528 [ISO 2022]. All other facilities are considered satisfactory (|z|≤ 2).

The results for ^{40}K show that the z-score of the facility '28' is larger than 3 (|z|≥3), gives an action signal and the facilities '39' gives a warning signal (2<|z|<3) according to ISO 13528. All other facilities are considered satisfactory (|z|≤2).

It should be noted that the z-score evaluation, for the $^{68}Ge/^{68}Ga$, must be interpreted carefully because the number of facilities reporting results is less than 10.

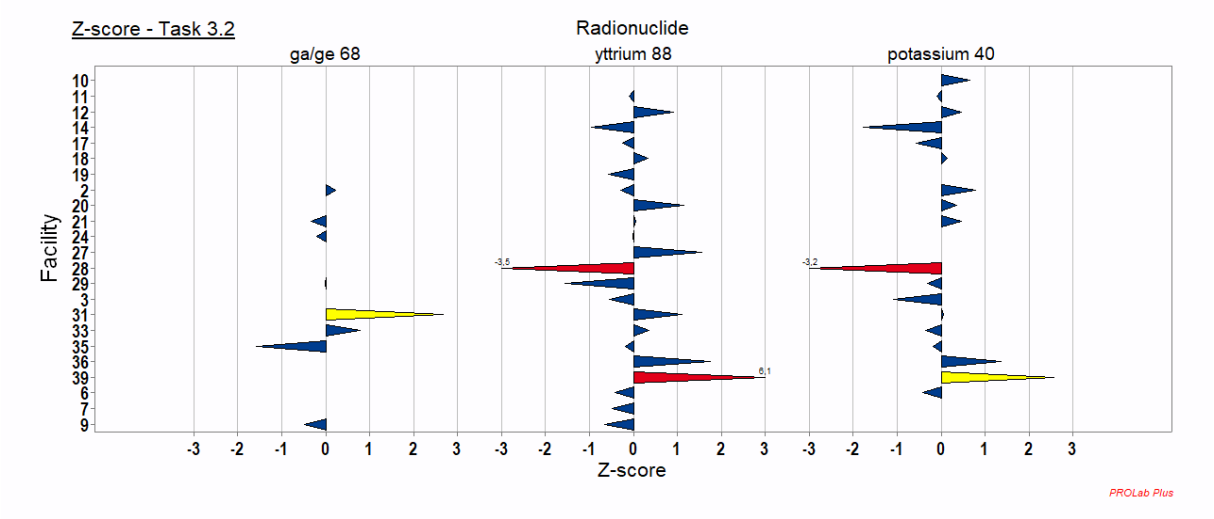


Figure 18: Representation of the z-score for the radionuclides of Task 3 "Medicine" #2. Yellow: warning signal, red: action signal.

5.6 Task 4a – Calibration 70 kg: ^{133}Ba, ^{152}Eu and ^{40}K

The sources of Task 4a contained two radionuclides: ^{133}Ba and ^{152}Eu. Most facilities have also received results of ^{40}K. As for previous tasks, the activities have been reported by the participant at measurement date and have been corrected for the radioactive decay. The radionuclide parameters [BIPM, 2020] and sources reference dates are given in the Table 25.

Table 25: List of radionuclide parameters and sources reference dates of Task 4a "Calibration".

Radionuclide	^{133}Ba	^{152}Eu	^{40}K
Half-life (y)	10.54	13.52	$1.25x10^9$
Main energy of γ emissions (keV)	81, 276, 302, 356, 384	122, 245, 344, 779, 867, 964, 1086, 1112, 1408	1461
Reference date (EIVIC)	01/05/2021	01/05/2021	01/05/2021

5.6.1 Assigned value and statistic parameters

The reference value (Bq) of sources, given in section 3.2, was compared with the robust mean (Bq) of participants for each radionuclide, as recommended in ISO 13528 [ISO 2022] (cf. Table 26). The robust mean was determined with the Q/Hampel robust method implemented in the ProLab™ intercomparison software. The statistic parameters used are summarised in the Table 27.

Table 26: Comparison between the reference value and the robust mean of participants for Task 4a "Calibration".

Task 4.a - Calibration	^{133}Ba	^{152}Eu	^{40}K
Reference value (BfS) (Bq)	21750 ± 290	25720 ± 600	4210 ± 210
Robust mean (Bq)	20535 ± 460	25730 ± 728	3770 ± 297
Difference (%)	-5.57	-0.06	-10.47

The difference between the reference value and the robust mean does not exceed 10%. Even if there is a small systematic underestimation, these differences do not reflect a systematic bias, the robust mean of participants was used as the assigned value, for each radionuclide, to ensure a uniform treatment of results between the various tasks.

Table 27: ProLab™ statistic parameters for Task 4a "Calibration".

ProLab statistic parameters – P4 Calibration A	^{133}Ba	^{152}Eu	^{40}K
Statistic method	Q / Hampel	Q / Hampel	Q / Hampel
Number of facilities reporting results	21	20	17
Number of participants (following plan)	25	25	25
Assigned value (Bq)	20535	25730	3770
Robust mean (Bq)	20535	25730	3770
Standard Deviation (SD) (Bq)	2106	3258	1225
Relative SD (%)	10.26	12.66	32.49
Uncertainty of assigned value (Bq)	460	728	297

The raw data of participants are represented for ^{133}Ba, ^{152}Eu and ^{40}K in the Figure 19.

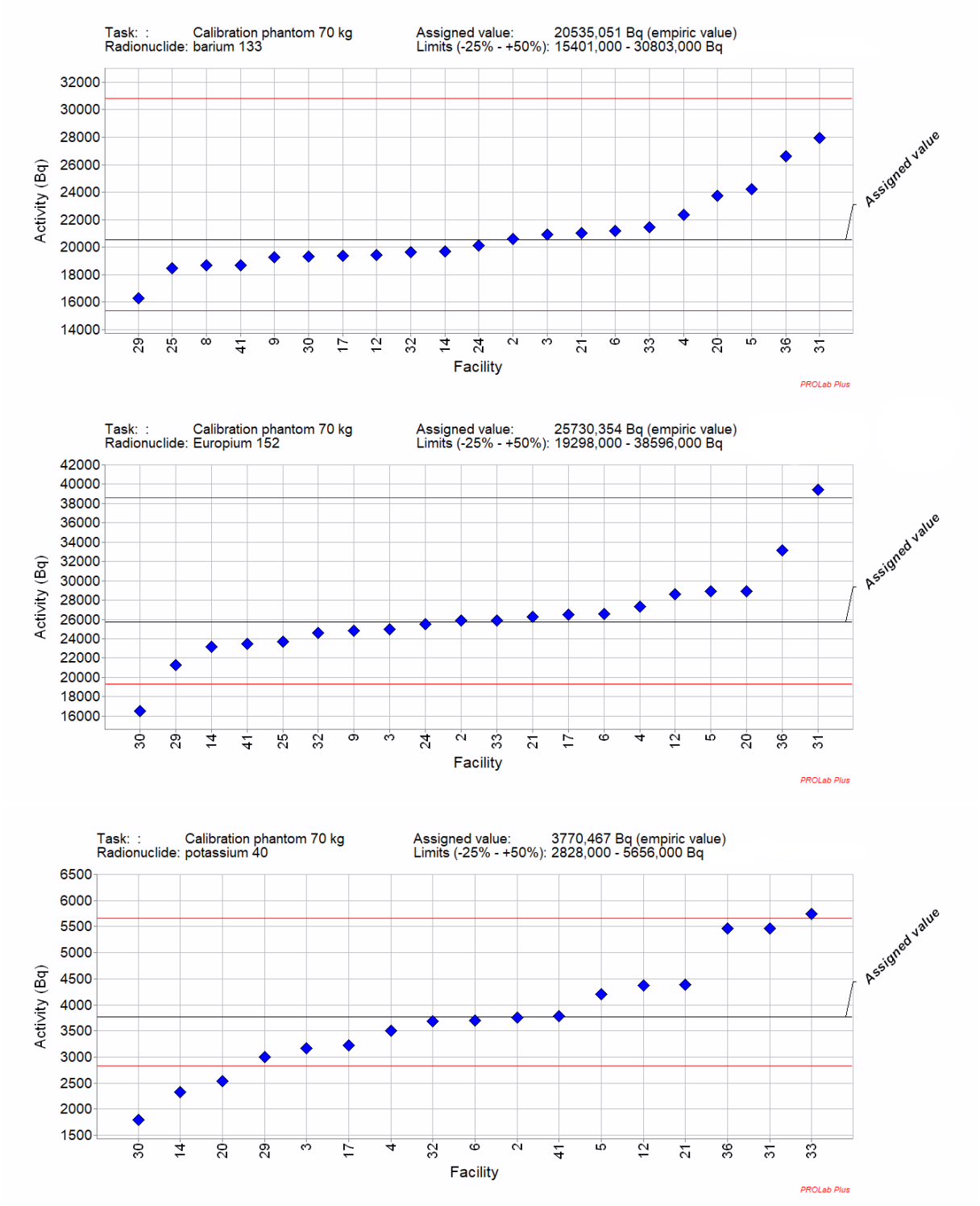


Figure 19: Representation of the raw data of participants and the assigned value: ^{133}Ba, ^{152}Eu and ^{40}K for Task 4a "Calibration". Red lines: -25% and +50% performance criteria.

5.6.2 Bias and outliers (ISO 28218)

The results of the participants, treated with the standard ISO 28218 [ISO 2010], are compared to the robust mean corresponding to the assigned value as described in section 5.6.1.

Table 28 summarises the bias obtained for the measurement of ^{133}Ba, ^{152}Eu and ^{40}K. The bias of participants is shown in the Figure 20.

The ^{133}Ba results show that all facilities are in conformity according to the [-25%; +50%] criteria of the standard ISO 28218 [ISO 2010].

The ^{152}Eu results show that two facilities [30, 31] are in non-conformity according to the [-25%; +50%] criteria of the standard ISO 28218 [ISO 2010]. The other facilities are in conformity. The facility '31' has been detected as an outlier (Grubbs test).

The ^{40}K results show that four facilities [14, 20, 30 and 33] are in non-conformity according to the [-25%; +50%] criteria of the standard ISO 28218 [ISO 2010]. Other facilities are in conformity.

Table 28: Bias of participants for Task 4a "Calibration".

ID Lab.	^{133}Ba	^{152}Eu	^{40}K
2	0.3%	0.7%	-0.3%
3	1.8%	-2.8%	-16.2%
4	8.7%	6.2%	-7.2%
5	17.9%	12.3%	11.4%
6	3.2%	3.3%	-1.9%
8	-9.1%		
9	-6.2%	-3.6%	
12	-5.3%	11.2%	16.0%
14	-4.1%	-9.9%	-38.1%
17	-5.7%	2.9%	-14.5%
20	15.6%	12.3%	-32.7%
21	2.3%	2.0%	16.3%
24	-2.1%	-0.9%	
25	-9.9%	-7.9%	
29	-20.6%	-17.2%	-20.4%
30	-5.8%	-35.7%	-52.4%
31	36.1%	53.3% Suspicious (Grubbs)	44.9%
32	-4.3%	-4.4%	-2.3%
33	4.5%	0.7%	52.2%
36	29.7%	28.8%	44.8%
41	-9,0%	-8,7%	0,30%

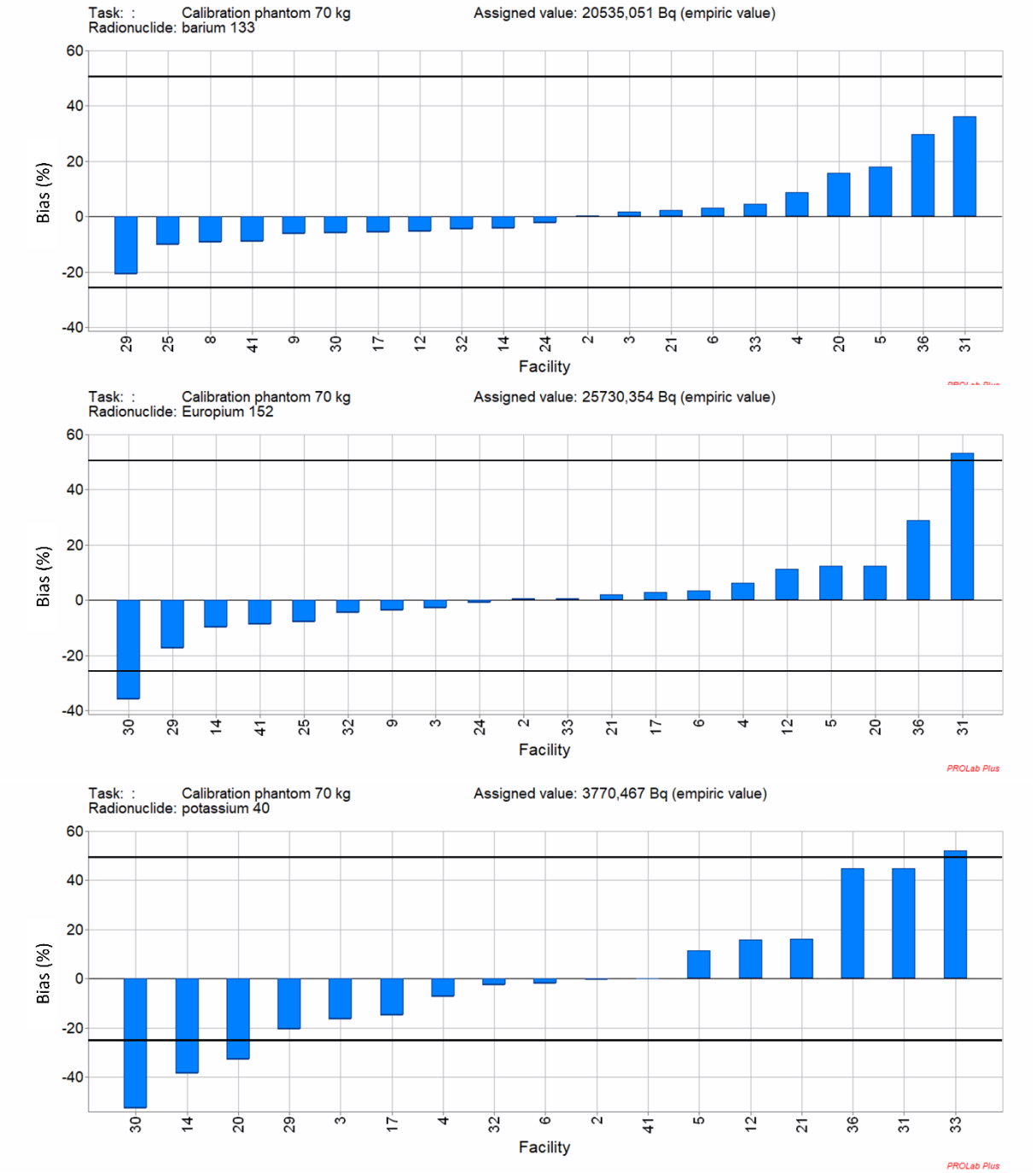


Figure 20: Representation of the bias (%) for the radionuclides of Task 4a "Calibration". Black lines: -25% and +50% performance criteria.

5.6.3 Z-score (ISO 13528)

Table 29 summarises the z-scores obtained for ^{133}Ba, ^{152}Eu and ^{40}K measurements. The z-scores indicated in yellow and red, respectively, represent the results to be considered as giving a warning signal and an action signal. The z-scores are shown in Figure 21.

The results for ^{133}Ba show that the z-score of facility '31' is larger than 3 ($|z|\geq 3$), gives an action signal and the facilities '29' and '36' give a warning signal ($2<|z|<3$) according to ISO 13528 [ISO 2022]. All other facilities are considered satisfactory ($|z|\leq 2$).

The results for ^{152}Eu show that the z-score of facility '31' is over 3 (|z|≥3) gives an action signal and the facilities '30' and '36' give a warning signal (2<|z|<3) according to ISO 13528 [ISO 2022]. All other facilities are considered satisfactory (|z|≤2).

The results for ^{40}K show that all facilities are considered satisfactory (|z|≤2).

Table 29: Z-score of participants for Task 4a "Calibration".

ID Lab.	^{152}Eu	^{133}Ba	^{40}K
2	0.05	0.03	-0.01
3	-0.22	0.18	-0.50
4	0.49	0.85	-0.22
5	0.97	1.75	0.35
6	0.26	0.31	-0.06
8		-0.89	
9	-0.28	-0.61	
12	0.89	-0.52	0.49
14	-0.78	-0.40	-1.17
17	0.23	-0.55	-0.45
20	0.98	1.52	-1.01
21	0.16	0.23	0.50
24	-0.07	-0.21	
25	-0.62	-0.97	
29	-1.36	-2.01	-0.63
30	-2.82	-0.57	-1.61
31	4.21	3.52	1.38
32	-0.34	-0.42	-0.07
33	0.05	0.44	1.61
36	2.27	2.90	1.38
41	-0.69	-0.87	0.01

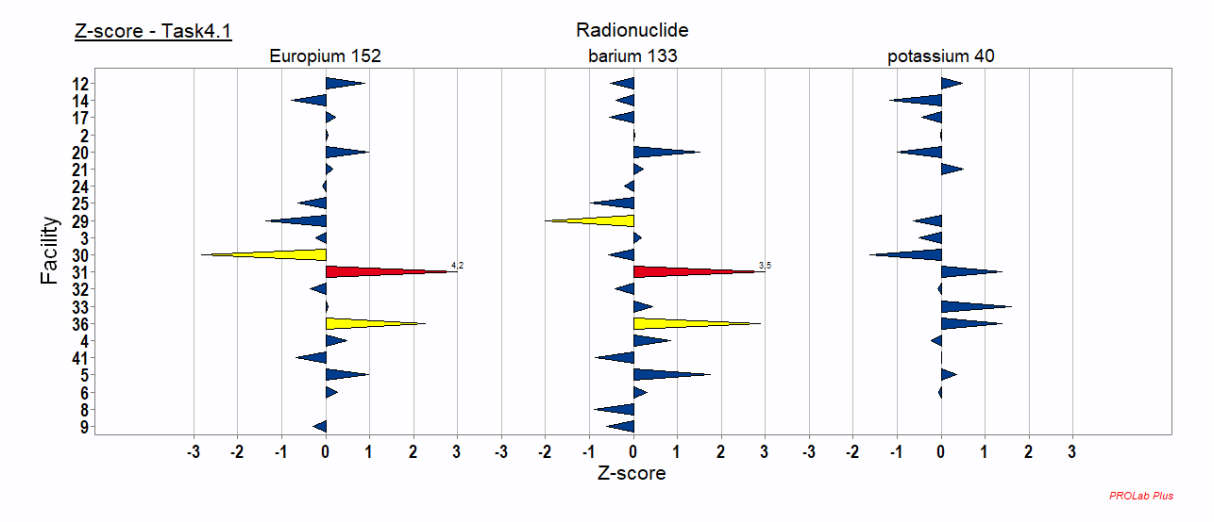


Figure 21: Representation of the z-score for the radionuclides of Task 4a "Calibration". Yellow: warning signal, red: action signal.

5.7 Task 4b – Calibration 90 kg: ^{133}Ba, ^{152}Eu and ^{40}K

The sources of Task 4b contained two radionuclides: ^{133}Ba and ^{152}Eu. Most facilities have also received rods of ^{40}K. As for previous tasks, the activities have been reported by the participant at measurement date and have been corrected for the radioactive decay. The radionuclide parameters [BIPM, 2020] and sources reference dates are given in the Table 30.

Table 30: List of radionuclide parameters and sources reference dates of Task 4b "Calibration".

Radionuclide	^{133}Ba	^{152}Eu	^{40}K
Half-life (y)	10.54	13.52	1.25×10^9
Main energy of γ emissions (keV)	81, 276, 302, 356, 384	122, 245, 344, 779, 867, 964, 1086, 1112, 1408	1461
Reference date (EIVIC)	01/05/2021	01/05/2021	01/05/2021

5.7.1 Assigned value and statistic parameters

The reference value (Bq) of sources, given in section 3.2, was compared with the robust mean (Bq) of participants for each radionuclide, as recommended in ISO 13528 [ISO 2022] (Table 31). The robust mean was determined with the Q/Hampel robust method implemented in the ProLab™ intercomparison software. The statistic parameters used are summarised in the Table 32.

Table 31: Comparison between the reference value and the robust mean of participants for Task 4b "Calibration".

Task 4.b - Calibration	^{133}Ba	^{152}Eu	^{40}K
Reference value (BfS) (Bq)	27930 ± 390	33050 ± 750	5460 ± 270
Robust mean (Bq)	25782 ± 496	32668 ± 685	4692 ± 240
Difference (%)	-7.68	-1.14	-14.00

The difference between the reference value and the robust mean does not exceed 14%. Even if there is a small systematic underestimation, these differences do not reflect a systematic bias, the robust mean of participants was used as the assigned value for each radionuclide, to ensure a uniform treatment of results between the various tasks.

It must be noted that the assigned values of the Task 4b were determined using all data of participants. Even if several calibration curves were applied by the participants to quantify the activity, due to the low statistic and the chosen method (Q/Hampel), it was not possible to determine the robust mean excluding the data with 70 kg calibration curve. Nevertheless, the arithmetic means of expert laboratories are very close to the robust mean: 26050 Bq vs. 25782 Bq for ^{133}Ba and 31807 Bq vs. 32668 Bq for ^{152}Eu. It was then considered that the robust mean with all data is a good estimator for this analysis.

Table 32: ProLab™ statistic parameters for Task 4b "Calibration".

ProLab statistic parameters – P5 Calibration	^{133}Ba	^{152}Eu	^{40}K
Statistic method	Q / Hampel	Q / Hampel	Q / Hampel
Number of facilities reporting results	31	30	21
Number of participants (following plan)	33	33	33
Assigned value (Bq)	25782	32668	4692
Robust mean (Bq)	25782	32668	4692
Standard Deviation (SD) (Bq)	2760	3750	1101
Relative SD (%)	10.71	11.48	23.47
Uncertainty of assigned value (Bq)	496	685	240

The raw data of participants are represented for ^{133}Ba, ^{152}Eu and ^{40}K in the Figure 22.

5.7.2 Bias and outliers (ISO 28218)

The results of the participants, treated with the standard ISO 28218 [ISO 2010], are compared to the robust mean corresponding to the assigned value as described in section 5.7.1.

Table 33 summarises the bias obtained for the measurement of ^{133}Ba, ^{152}Eu and ^{40}K. The bias of participants is shown in the Figure 22.

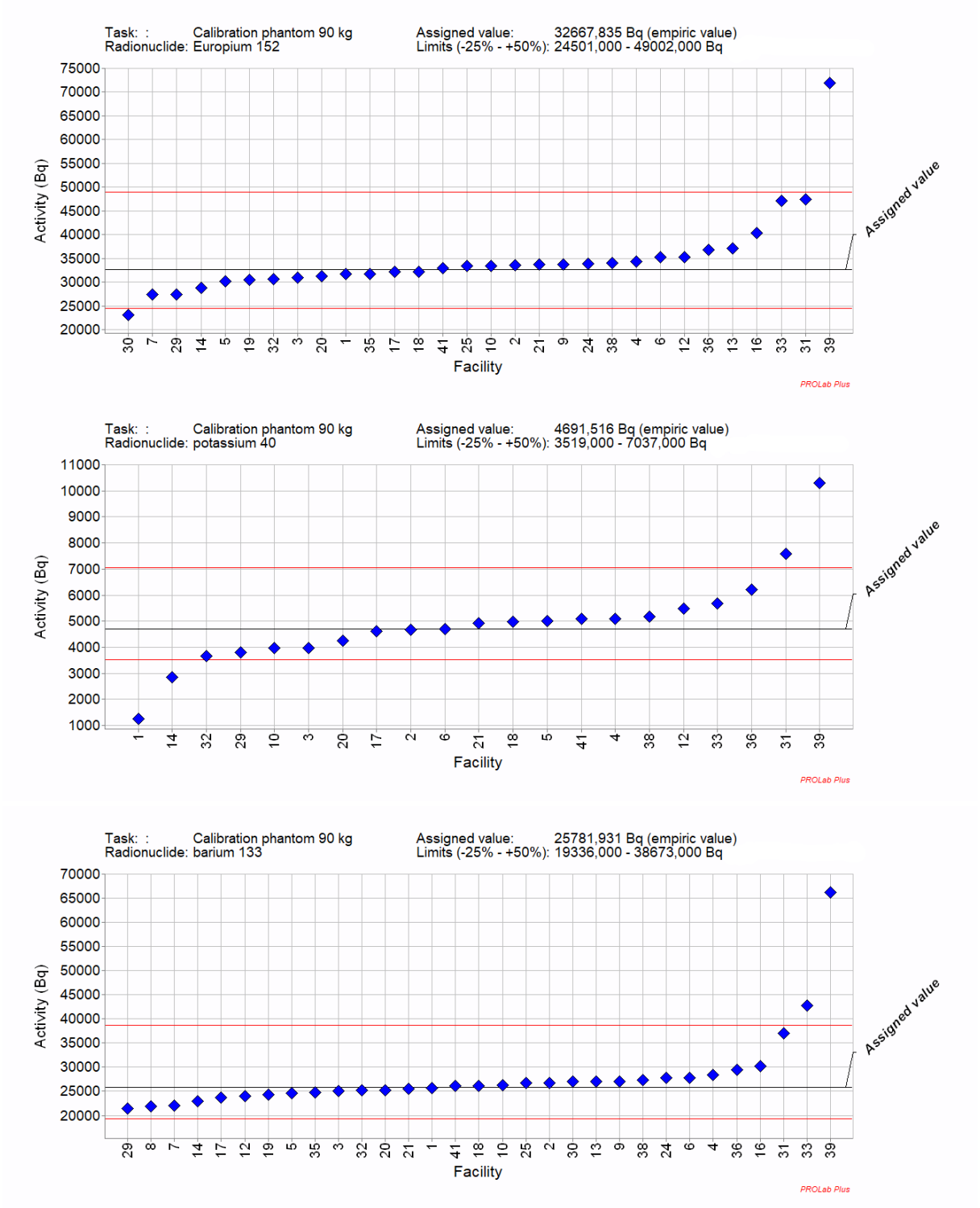


Figure 22: Representation of the raw data of participants and the assigned value: ^{133}Ba, ^{152}Eu and ^{40}K for Task 4b "'Calibration". Red lines: -25% and +50% performance criteria.

Table 33: Bias of participants for Task 4b "Calibration".

ID Lab.	^{152}Eu	^{133}Ba	^{40}K
1	-3.0%	-0.4%	-73.5%
2	2.8%	3.9%	-0.6%
3	-5.3%	-3.0%	-15.2%
4	5.3%	10.2%	8.7%
5	-7.8%	-4.5%	6.6%
6	7.8%	7.9%	0.2%
7	-16.2%	-14.4%	
8		-15.0%	
9	3.2%	5.1%	
10	2.2%	2.0%	-15.2%
12	7.9%	-6.9%	16.6%
13	13.5%	4.9%	
14	-11.9%	-10.7%	-39.3%
16	23.6%	17.0%	
17	-1.5%	-7.9%	-1.5%
18	-1.4%	1.6%	6.2%
19	-6.7%	-5.9%	
20	-4.2%	-2.1%	-9.5%
21	3.0%	-0.9%	5.0%
24	3.7%	7.7%	
25	2.1%	3.6%	
29	-16.2%	-17.0%	-19.0%
30	-29.2%	4.7%	
31	45.4%	43.8% Outlier (Grubbs)	61.3%
32	-6.0%	-2.4%	-21.9%
33	44.2%	66.1% Outlier (Grubbs)	20.8%
35	-2.9%	-3.7%	
36	12.8%	14.2%	32.5%
38	4.2%	6.3%	10.0%

ID Lab.	^{152}Eu	^{133}Ba	^{40}K
39	120.0% Outlier (Grubbs)	156.6% Outlier (Grubbs)	119.5% Outlier (Grubbs)
41	0.9%	1.1%	8.3%

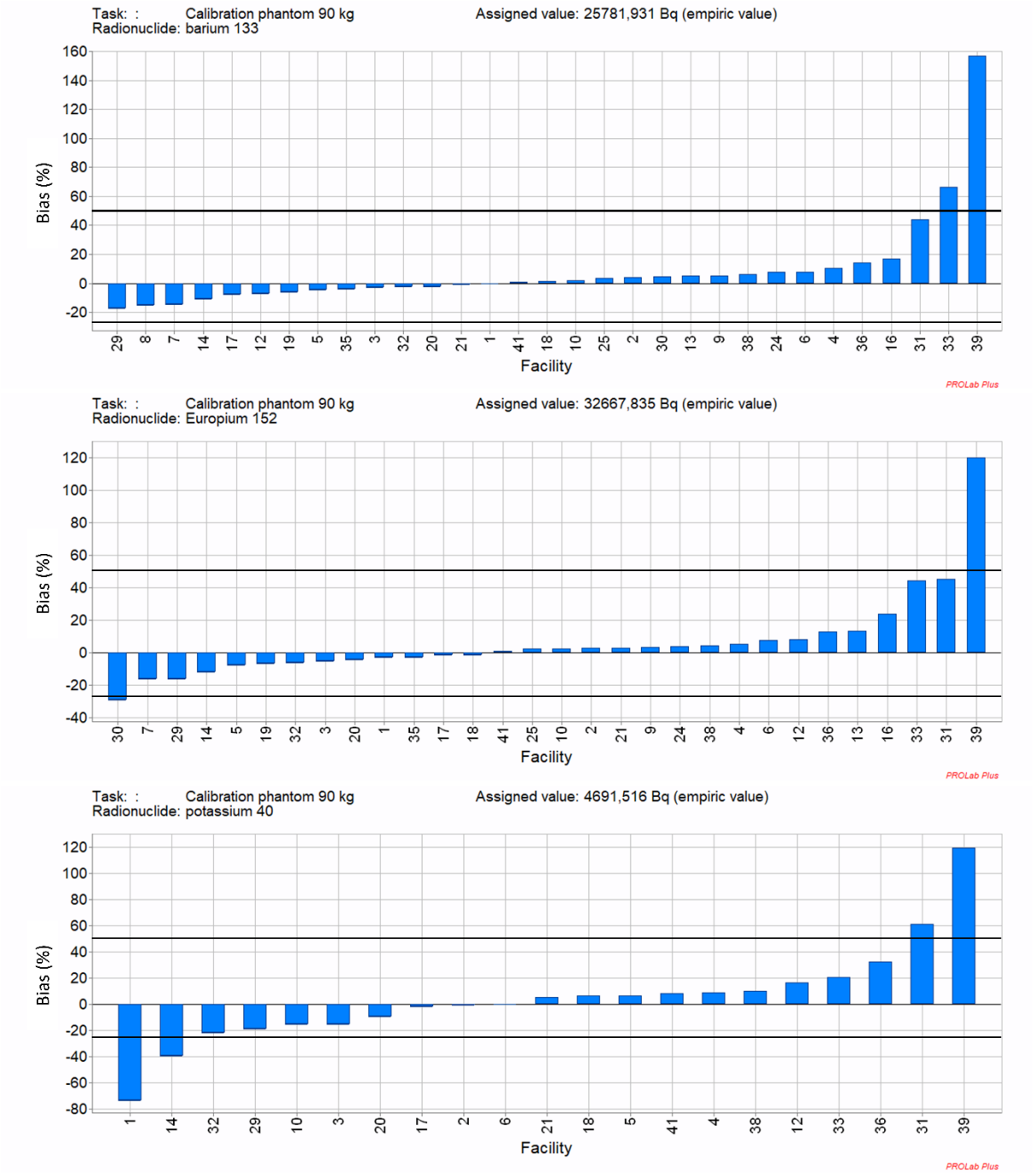


Figure 23: Representation of the bias (%) for the radionuclides of the Task 4b "Calibration". Black lines: -25% and +50% performance criteria.

The ^{133}Ba results show that two facilities [33 and 39] are in non-conformity according to the [-25%; +50%] criteria of the standard ISO 28218 [ISO 2010]. The other facilities are in conformity. The values of three facilities [31, 33 and 39] have been detected as outliers (Grubbs test).

The ^{152}Eu results show that two facilities [30, 39] are in non-conformity according to the [-25%; +50%] criteria of the standard ISO 28218 [ISO 2010]. The other facilities are in conformity. The result of facility '39' has been detected as an outlier (Grubbs test).

The ^{40}K results show that four facilities [1, 14, 31 and 39] are in non-conformity according to the [-25%; +50%] criteria of the standard ISO 28218 [ISO 2010]. Other facilities are in conformity. The value of facility '39' has been detected as an outlier (Grubbs test).

5.7.3 Z-score (ISO 13528)

Table 34 summarises the z-scores obtained for the ^{133}Ba, ^{152}Eu and ^{40}K measurements. The z-scores indicated in yellow and red, respectively, represent the results to be considered as giving a warning signal and an action signal. The z-scores are shown in the Figure 24.

Table 34: Z-score of participants for Task 4b "Calibration".

ID Lab.	^{152}Eu	^{133}Ba	^{40}K
1	-0.26	-0.04	-3.13
2	0.24	0.37	-0.03
3	-0.46	-0.28	-0.65
4	0.46	0.96	0.37
5	-0.68	-0.42	0.28
6	0.68	0.73	0.01
7	-1.41	-1.34	
8		-1.40	
9	0.28	0.47	
10	0.19	0.19	-0.65
12	0.69	-0.64	0.71
13	1.18	0.46	
14	-1.04	-1.00	-1.68
16	2.06	1.59	
17	-0.13	-0.73	-0.06
18	-0.12	0.15	0.27
19	-0.59	-0.55	
20	-0.37	-0.20	-0.40
21	0.26	-0.08	0.21

ID Lab.	^{152}Eu	^{133}Ba	^{40}K
24	0.32	0.72	
25	0.19	0.34	
29	-1.41	-1.59	-0.81
30	-2.54	0.44	
31	3.95	4.09	2.61
32	-0.52	-0.22	-0.94
33	3.85	6.18	0.89
35	-0.25	-0.35	
36	1.11	1.32	1.38
38	0.37	0.59	0.43
39	10.45	14.63	5.09
41	0.08	0.10	0.35

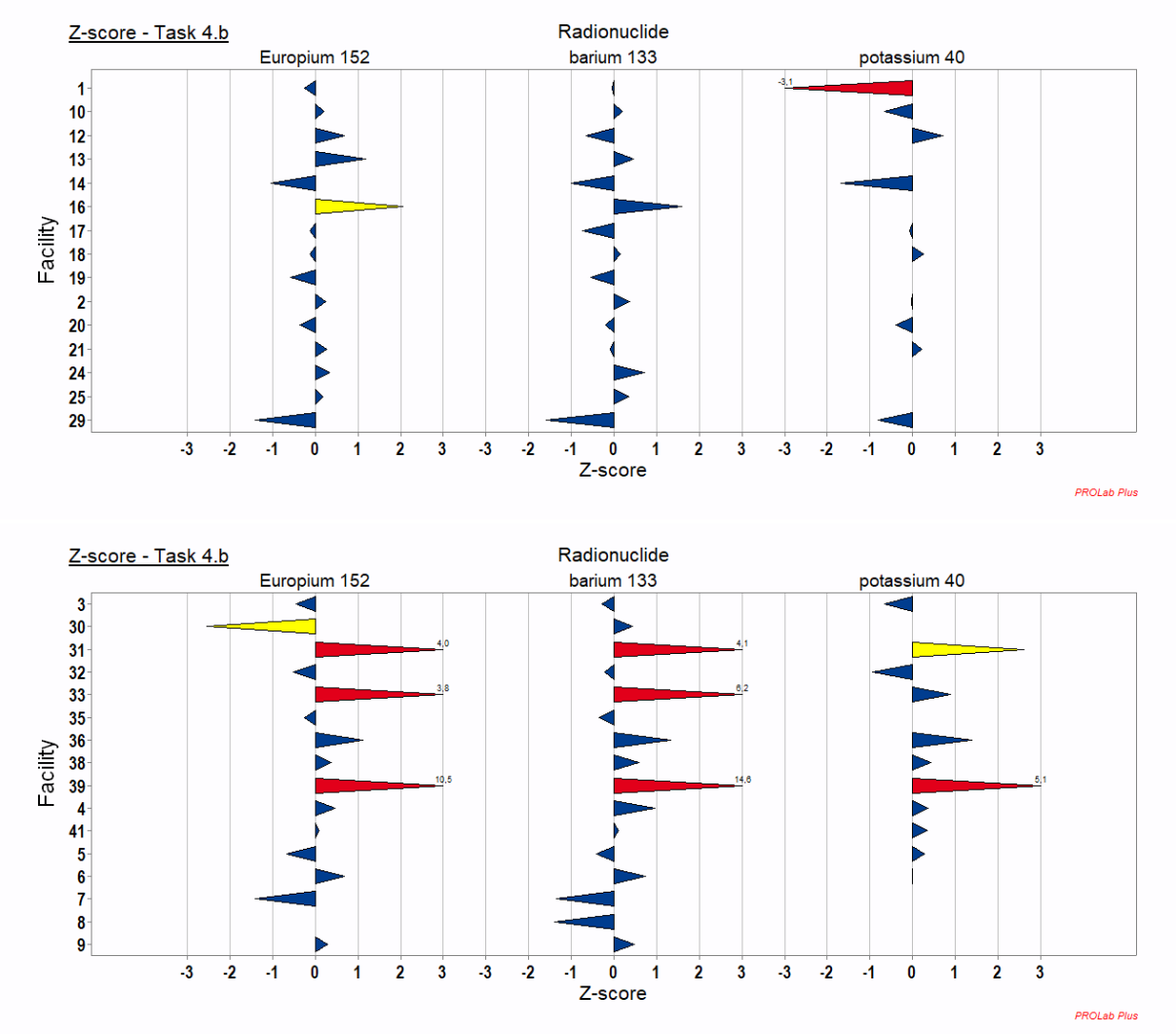


Figure 24: Representation of the z-score for the radionuclides of Task 4b "Calibration". Yellow: warning signal, red: action signal.

The results for ^{133}Ba show that the z-score of the facilities '31', '33' and '39' are over 3 ($|z|\geq 3$), give an action signal according to ISO 13528 [ISO 2022]. All other facilities are considered satisfactory ($|z|\leq 2$).

The results for ^{152}Eu show that the z-score of the facilities '31', '33' and '39' are over 3 ($|z|\geq 3$) give an action signal and the facilities '16' and '30' give a warning signal ($2<|z|<3$) according to ISO 13528 [ISO 2022]. All other facilities are considered satisfactory ($|z|\leq 2$).

The results for ^{40}K show that the z-score of facilities '1' and '39' are larger than 3 ($|z|\geq 3$) and gives an action signal and the facility '31' gives a warning signal ($2<|z|<3$) according to ISO 13528 [ISO 2022]. All other facilities are considered satisfactory ($|z|\leq 2$).

5.8 Conclusion on the results for the measurement Tasks 1, 2, 3.1, 3.2, 4a and 4b

The conclusion for this intercomparison can be summarised by the compliance reports for the different geometries for the participants of each task, with regard to the ISO 28218 [ISO 2010] and ISO 13528 [ISO 2022] standards. They are summarised in Table 35 to Table 40.

Depending on the normative reference applied, it should be noted that a difference in conformity exists for several installations. Several reasons can generally explain this difference:

- The tolerance intervals are more restrictive according to the standard ISO 13528 [ISO 2022] than to the standard ISO 28218 [ISO 2010].
- The bias (ISO 28218) is a criterion which allows to assess the performance of an installation in relation to the "target" value, and therefore independently of the other participants. The z-score is a performance estimator which depends on the dispersion of the results of the participants. It therefore allows to evaluate a facility compared to all the participating facilities (use of the robust standard deviation for capability evaluation).
- The intercomparison is representative of the variability of the materials and methods used. In fact, the measurements are not carried out under the same conditions and with different installations, particularly in terms of detection system (NaI(Tl) or germanium detectors of different sizes), calibration curves used (70 kg systematically or adapted to the configuration), the duration of the measurement, the detector-patient distances and the use of more or less realistic anthropomorphic phantoms.

For the two laboratories with the most extreme biases over all tasks, i.e. facilities '28' and '39', a discussion of the results with the laboratory staff yielded information that they applied a calibration for lung measurements and tried to adapt the results to the whole-body geometry of the intercomparison.

Non-conforming results are therefore to be interpreted with care and must be considered as complementary elements allowing the laboratory to evaluate itself compared to other participants.

A further analysis would therefore be to categorise the participants for the calculation of the z-score in order to compare similar methods and measurement systems. A first study has been done and is presented in the Chapter 6. Nevertheless, to date, the great variability of the materials and methods used does not allow us to have sufficiently large populations to carry out a statistical analysis for all the parameters.

Table 35: Compliance report for Task 1 "Victor".

ID	ISO 28218	ISO 13528
1	Conform (^{60}Co/^{133}Ba/^{137}Cs)	Acceptable (^{60}Co/^{133}Ba/^{137}Cs)
2	Conform	Acceptable
3	Conform	Acceptable
4	Conform	Acceptable
5	Conform	Acceptable
6	Conform	Acceptable
7	Conform (^{60}Co/^{133}Ba/^{137}Cs)	Acceptable (^{60}Co/^{133}Ba/^{137}Cs)
8	Conform (^{60}Co/^{137}Cs)	Acceptable (^{60}Co) | Warning signal (^{137}Cs)
9	Conform (^{60}Co/^{133}Ba/^{137}Cs)	Acceptable (^{60}Co/^{133}Ba/^{137}Cs)
10	Conform	Acceptable
11	Conform	Acceptable
12	Conform	Acceptable
13	Conform (^{60}Co/^{133}Ba/^{137}Cs)	Acceptable (^{133}Ba/^{137}Cs) | Warning signal (^{60}Co)
14	Conform (^{60}Co/^{133}Ba/^{137}Cs) | Not conform (^{40}K)	Acceptable
15	Conform	Acceptable
16	-	-
17	Conform	Acceptable
18	Conform	Acceptable
19	Conform (^{60}Co/^{133}Ba/^{137}Cs)	Acceptable (^{60}Co/^{133}Ba/^{137}Cs)
20	Conform	Acceptable
21	Conform	Acceptable
22	Conform	Acceptable
23	Conform (^{60}Co/^{133}Ba/^{137}Cs)	Acceptable (^{60}Co/^{133}Ba/^{137}Cs)
24	Conform (^{60}Co/^{133}Ba/^{137}Cs)	Acceptable (^{60}Co/^{133}Ba/^{137}Cs)
25	Conform (^{60}Co/^{133}Ba/^{137}Cs)	Acceptable (^{60}Co/^{133}Ba/^{137}Cs)
26	-	-
27	Conform (^{137}Cs) | Not conform (^{60}Co/^{133}Ba)	Accept. (^{137}Cs) | Action signal (^{60}Co/^{133}Ba)
28	Not conform (^{60}Co/^{133}Ba/^{137}Cs/^{40}K)	Warning signal (^{40}K) | Action signal (^{60}Co/^{133}Ba/^{137}Cs)
29	Conform	Acceptable

ID	ISO 28218		ISO 13528	
30	Conform (^{133}Ba)	Not conform (^{137}Cs)	Acceptable (^{133}Ba)	Action signal (^{137}Cs)
31	Conform (^{60}Co/^{137}Cs/^{40}K)	Not conform (^{133}Ba)	Acceptable (^{60}Co/^{40}K)	Action signal (^{133}Ba/^{137}Cs)
32	Conform		Acceptable	
33	Conform		Acceptable	
34	Conform (^{60}Co/^{133}Ba/^{137}Cs)		Acceptable (^{133}Ba)	Warning signal (^{60}Co/^{137}Cs)
35	Conform		Acceptable	
36	Conform		Acceptable (^{133}Ba)	Warning signal (^{60}Co/^{137}Cs/^{40}K)
37	Conform (^{60}Co/^{133}Ba/^{137}Cs) / Not conform (^{40}K)		Acceptable (^{60}Co/^{133}Ba/^{137}Cs) / Action signal (^{40}K)	
38	Conform		Acceptable	
39	Conform (^{40}K)	Not conform (^{60}Co/^{133}Ba/^{137}Cs)	Acceptable (^{40}K)	Action signal (^{60}Co/^{133}Ba/^{137}Cs)
40	-		-	
41	Conform		Acceptable	
42	Conform		Acceptable	
43	Conform		Acceptable	

Table 36: Compliance report for Task 2 "Emergency".

ID	ISO 28218	ISO 13528
1	Conform (^{134}Cs/^{137}Cs)	Acceptable (^{134}Cs/^{137}Cs)
2	Conform	Acceptable
3	Conform	Acceptable
4	Conform	Acceptable
5	Conform	Acceptable
6	Conform	Acceptable
7	Conform (^{134}Cs/^{137}Cs)	Acceptable (^{134}Cs/^{137}Cs)
8	Conform (^{134}Cs/^{137}Cs)	Acceptable (^{134}Cs/^{137}Cs)
9	Conform (^{134}Cs/^{137}Cs)	Acceptable (^{134}Cs/^{137}Cs)
10	Conform	Acceptable

ID	ISO 28218	ISO 13528
11	Conform	Acceptable
12	Conform	Acceptable
13	Conform (^{134}Cs/^{137}Cs)	Acceptable (^{134}Cs/^{137}Cs)
14	Conform	Acceptable
15	Conform	Acceptable
16	Conform (^{134}Cs/^{137}Cs)	Acceptable (^{134}Cs/^{137}Cs)
17	Conform	Acceptable
18	Conform	Acceptable
19	Conform (^{134}Cs/^{137}Cs)	Acceptable (^{134}Cs/^{137}Cs)
20	Conform	Acceptable
21	Conform	Acceptable
22	Conform	Acceptable
23	-	-
24	Conform (^{134}Cs/^{137}Cs)	Acceptable (^{134}Cs/^{137}Cs)
25	Conform (^{134}Cs/^{137}Cs)	Acceptable (^{134}Cs/^{137}Cs)
26	-	-
27	Conform (^{134}Cs/^{137}Cs)	Acceptable (^{137}Cs) / Action signal (^{134}Cs)
28	Not conform (^{134}Cs/^{137}Cs/^{40}K)	Warning signal (^{40}K) / Action signal (^{134}Cs/^{137}Cs)
29	Conform	Acceptable
30	Not conform (^{134}Cs/^{137}Cs)	Action signal (^{134}Cs /^{137}Cs)
31	Conform	Acceptable (^{40}K) / Warning signal (^{137}Cs) / Action signal (^{134}Cs)
32	Conform	Acceptable
33	Conform	Acceptable (^{40}K) / Action signal (^{134}Cs/^{137}Cs)
34	Conform (^{134}Cs/^{137}Cs)	Acceptable (^{134}Cs/^{137}Cs)
35	Conform	Acceptable
36	Conform	Acceptable
37	Conform (^{134}Cs/^{137}Cs) / Not conform (^{40}K)	Acceptable (^{134}Cs/^{137}Cs) / Action signal (^{40}K)
38	Conform	Acceptable
39	Not conform (^{134}Cs/^{137}Cs /^{40}K)	Action signal (^{134}Cs/^{137}Cs /^{40}K)

ID	ISO 28218	ISO 13528
40	-	-
41	Conform	Acceptable
42	Conform	Acceptable
43	Conform	Acceptable

Table 37: Compliance report for Task 3.1 "Medicine".

ID	ISO 28218	ISO 13528
1	Conform (^{88}Y) / Not conform (^{40}K)	Acceptable (^{88}Y) / Action signal (^{40}K)
4	Conform	Acceptable
5	Conform (^{88}Y /^{40}K)	Acceptable (^{88}Y) / Warning signal (^{40}K)
8	Conform (^{68}Ge /^{88}Y)	Acceptable (^{68}Ge /^{88}Y)
13	Conform (^{88}Y)	Acceptable (^{88}Y)
22	Conform	Acceptable
25	Conform	Acceptable
30	Not conform (^{88}Y)	Action signal (^{88}Y)
32	Conform	Acceptable (^{88}Y) / Warning signal (^{40}K)
34	Conform (^{88}Y)	Acceptable (^{88}Y)
37	Conform (^{88}Y) / Not conform (^{40}K)	Acceptable (^{88}Y) / Action signal (^{40}K)
38	Conform	Acceptable
41	Conform	Acceptable
42	Conform	Acceptable
43	Conform	Acceptable

Table 38: Compliance report for Task 3.2 "Medicine".

ID	ISO 28218	ISO 13528
2	Conform	Acceptable
3	Conform (^{88}Y/^{40}K)	Acceptable (^{88}Y/^{40}K)
6	Conform (^{88}Y/^{40}K)	Acceptable (^{88}Y/^{40}K)
7	Conform (^{88}Y)	Acceptable (^{88}Y)

ID	ISO 28218	ISO 13528
9	Conform (^{68}Ge / ^{88}Y)	Acceptable (^{68}Ge/^{88}Y)
10	Conform (^{40}K)	Acceptable (^{40}K)
11	Conform (^{88}Y/^{40}K)	Acceptable (^{88}Y/^{40}K)
12	Conform (^{88}Y/^{40}K)	Acceptable (^{88}Y/^{40}K)
14	Conform (^{88}Y) / Not conform (^{40}K)	Acceptable (^{88}Y/^{40}K)
17	Conform (^{88}Y/^{40}K)	Acceptable (^{88}Y/^{40}K)
18	Conform (^{88}Y/^{40}K)	Acceptable (^{88}Y/^{40}K)
19	Conform (^{88}Y)	Acceptable (^{88}Y)
20	Conform (^{88}Y/^{40}K)	Acceptable (^{88}Y/^{40}K)
21	Conform	Acceptable
24	Conform (^{68}Ge/^{88}Y)	Acceptable (^{68}Ge/^{88}Y)
27	Conform (^{88}Y)	Acceptable (^{88}Y)
28	Not conform (^{88}Y/^{40}K)	Action signal (^{88}Y /^{40}K)
29	Conform	Acceptable
31	Conform	Acceptable (^{88}Y/^{40}K) / Warning signal (^{68}Ge)
33	Conform	Acceptable
35	Conform	Acceptable
36	Conform (^{88}Y/^{40}K)	Acceptable (^{88}Y/^{40}K)
39	Not conform (^{88}Y/^{40}K)	Warning signal (^{40}K) / Action signal (^{88}Y)

Table 39: Compliance report for Task 4a “Calibration” (P4).

ID	ISO 28218	ISO 13528
2	Conform	Acceptable
3	Conform	Acceptable
4	Conform	Acceptable
5	Conform	Acceptable
6	Conform	Acceptable
8	Conform (^{133}Ba)	Acceptable (^{133}Ba)
9	Conform (^{133}Ba/^{152}Eu)	Acceptable (^{133}Ba/^{152}Eu)

ID	ISO 28218	ISO 13528
12	Conform	Acceptable
14	Conform (^{133}Ba/^{152}Eu) / Not conform (^{40}K)	Acceptable
17	Conform	Acceptable
20	Conform (^{133}Ba/^{152}Eu) / Not conform (^{40}K)	Acceptable
21	Conform	Acceptable
24	Conform (^{133}Ba/^{152}Eu)	Acceptable (^{133}Ba/^{152}Eu)
25	Conform (^{133}Ba/^{152}Eu)	Acceptable (^{133}Ba/^{152}Eu)
29	Conform	Acceptable (^{152}Eu/^{40}K) / Warning signal (^{133}Ba)
30	Conform (^{133}Ba) / Not conform (^{152}Eu/^{40}K)	Acceptable (^{133}Ba/^{40}K) / Warning signal (^{152}Eu)
31	Conform (^{133}Ba/^{40}K) / Not conform (^{152}Eu)	Acceptable (^{40}K) / Action signal (^{133}Ba/^{152}Eu)
32	Conform	Acceptable
33	Conform (^{133}Ba/^{152}Eu) / Not conform (^{40}K)	Acceptable
36	Conform	Acceptable (^{40}K) / Warning signal (^{133}Ba/^{152}Eu)
41	Conform	Acceptable

Table 40: Compliance report for Task 4b "Calibration" (P5).

ID	ISO 28218	ISO 13528
1	Conform (^{133}Ba/^{152}Eu) / Not conform (^{40}K)	Acceptable (^{133}Ba/^{152}Eu) / Action signal (^{40}K)
2	Conform	Acceptable
3	Conform	Acceptable
4	Conform	Acceptable
5	Conform	Acceptable
6	Conform	Acceptable
7	Conform (^{133}Ba/^{152}Eu)	Acceptable (^{133}Ba/^{152}Eu)
8	Conform (^{133}Ba)	Acceptable (^{133}Ba)
9	Conform (^{133}Ba/^{152}Eu)	Acceptable (^{133}Ba/^{152}Eu)

ID	ISO 28218	ISO 13528
10	Conform	Acceptable
11	-	-
12	Conform	Acceptable
13	Conform (^{133}Ba/^{152}Eu)	Acceptable (^{133}Ba/^{152}Eu)
14	Conform (^{133}Ba/^{152}Eu) / Not conform (^{40}K)	Acceptable
15	-	-
16	Conform (^{133}Ba/^{152}Eu)	Acceptable (^{133}Ba) / Warning signal (^{152}Eu)
17	Conform	Acceptable
18	Conform	Acceptable
19	Conform (^{133}Ba/^{152}Eu)	Acceptable (^{133}Ba/^{152}Eu)
20	Conform	Acceptable
21	Conform	Acceptable
22	-	-
23	-	-
24	Conform (^{133}Ba/^{152}Eu)	Acceptable (^{133}Ba/^{152}Eu)
25	Conform (^{133}Ba/^{152}Eu)	Acceptable (^{133}Ba/^{152}Eu)
26	-	-
27	-	-
28	-	-
29	Conform	Acceptable
30	Conform (^{133}Ba) / Not conform (^{152}Eu)	Acceptable (^{133}Ba) / Warning signal (^{152}Eu)
31	Conform (^{133}Ba/^{152}Eu) / Not conform (^{40}K)	Warning signal (^{40}K) / Action signal (^{133}Ba/^{152}Eu)
32	Conform	Acceptable
33	Conform (^{152}Eu/^{40}K) / Not conform (^{133}Ba)	Acceptable (^{40}K) / Action signal (^{133}Ba/^{152}Eu)
34	-	-
35	Conform (^{133}Ba/^{152}Eu)	Acceptable (^{133}Ba/^{152}Eu)
36	Conform	Acceptable
37	-	-
38	Conform	Acceptable

ID	ISO 28218	ISO 13528
39	Not Conform (^{133}Ba/^{152}Eu/^{40}K)	Action signal (^{133}Ba/^{152}Eu/^{40}K)
40	-	-
41	Conform	Acceptable
42	-	-
43	-	-

6. Review of the main metrological and organisational characteristics of the facilities

Additional to the results, information about technical and organisational characteristics of the participating laboratories were collected. Several of these characteristics were used to test if they had a significant attribution with the quality of the reported results. For this purpose, certain statistical tests were applied. The tests were applied on the z-scores except those that were identified as outliers.

Before conducting the statistical tests, the complete set of reported z-scores (except outliers) was tested if it features a normal distribution. For this purpose, a Shapiro-Wilk normality test was conducted. The test failed (p-value $\ll 0.001$), leading to the assumption that the set of data does not feature a normal distribution. The quantile-quantile plot (Q-Q plot) (Figure 25), in which data from a normal distribution would exhibit a straight line, shows that extreme values occur more often in the set of z-scores than it is expected for a normal distribution. This suggests that the most extremely biased results are compromised by characteristics of the participating laboratories as explained in Section 5.8 rather than being an outcome of random deviations.

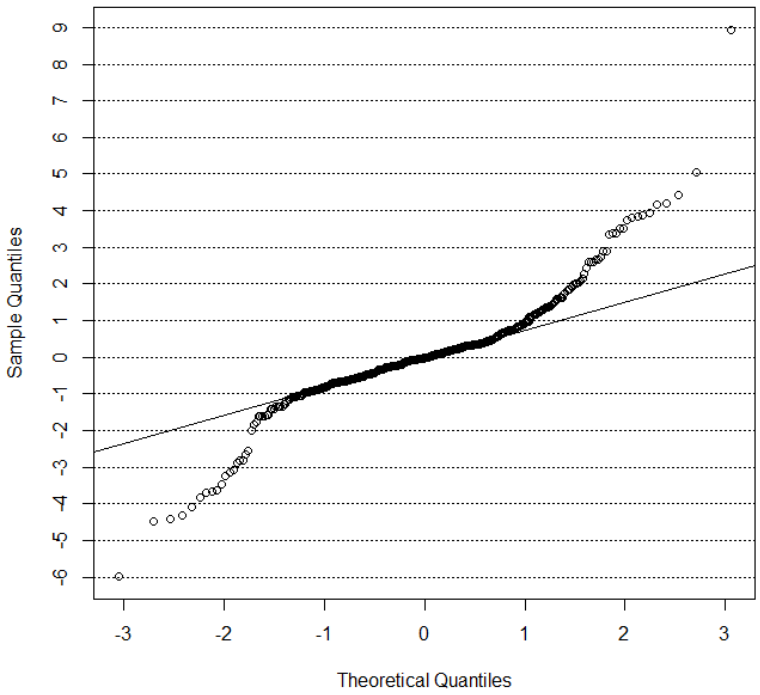


Figure 25: Quantile-quantile plot of all z-scores except outliers. The straight line marks the plot that is expected for a normal distribution. The data with sample quantiles smaller than about -1.5 and greater than about +1 deviate from the straight line.

Because of the deviation from a normal distribution, non-parametric tests were applied. These tests do not require the set of data to follow a certain distribution. The tests were performed using the R software package, version 4.0.2 [R Core Team 2020]. If not otherwise stated below, all reported z-scores (except outliers) from all four measurement tasks that involved phantoms (Tasks 1–4) were used for the tests. The following tests were conducted:

Mann-Whitney U test to compare the central tendency of the values: This test predicates if data from the one subset are significantly greater or smaller than data from the other subset (alternate hypothesis) or not (null hypothesis). This test is comparable to a *t*-test for normally distributed values. A two-sided test with correction for tied values was conducted.

Siegel-Tukey test to compare the dispersion of the values: This test predicates if data from the one subset are significantly more or less dispersed than data from the other subset (alternate hypothesis) or not (null hypothesis). This test is comparable to an *F*-test for normally distributed values. A two-sided test with correction for tied values and with adjustment of the medians was applied.

Reported results are the *p*-values, which are the maximum probabilities of obtaining the actual samples under the assumption of the samples originating from the same population. Small *p*-values indicate a significant difference between the sets of data, with the threshold set at 0.05 (i.e. confidence level of 95%). If the laboratories could be divided into more than two subsets (e.g. in the case of the measurement geometry, which was stretcher, inclined chair, chair and standing), the first characteristic serves as the reference and data for all other characteristics are compared with the data for the reference characteristic.

Beside the *p*-values, the number n_{lab} of laboratories with the respective characteristic, the number n_{result} of results from these laboratories that were applied in the statistical tests, the arithmetic mean, the median and the standard deviation are tabulated. In cases where the sum of n_{lab} is less than the number of whole-body counters that reported results (41), the respective characteristic is unknown for one or several whole-body counters. Additional to these values, box plots are presented: The box indicates the first quartile, median and third quartile, the whiskers denote the minimum and maximum value.

6.1 Type of participation in the intercomparison (personal attendance or shipment)

Twenty-five whole-body counters took part in the attended tour whereas 16 whole-body counters received the phantom by shipment. The activities of the sources were as similar as possible in both phantoms for all tasks ($\Delta < 0.05\%$). Therefore, possible differences could have been caused in particular by the assistance during the setup of the phantoms in the attended tour, but also by the longer time that was available for the measurements in the shipment tour.

The difference between the central tendencies between attended tour and shipment is mostly caused by some very small z-scores (i.e. strong underestimation) in the attended tour (Figure 26, Table 41). This might be only a coincidence between those few labs that tended to strongly underestimate the results and the participation of these labs in the attended tour but not a causal attribution between underestimation and attended tour. However, the greater number of reported results per laboratory in the shipment tour (12.6) as compared to the attended tour (10.5) can be explained: Because of the short time that was available in the attended tour, not all laboratories conducted both measurements of Task 4 “Calibration” (with a phantom of 70 kg and with one of 90 kg).

Table 41: Statistical parameters of the attribution of z-scores with the type of participation.

	personal attendance	shipment
n_{lab}	25	16
n_{result}	262	201
n_{result}/n_{lab}	10.5	12.6
mean	0.04	0.05
median	0.09	-0.13
standard deviation	1.53	1.12
p-values		
Mann-Whitney		0.017 sign. diff.
Siegel-Tukey		0.055 not sign. diff.

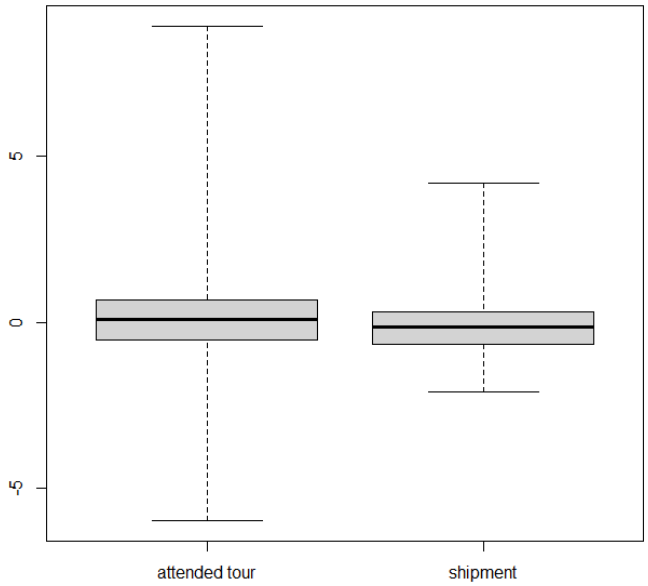


Figure 26: Box plot of the z-scores discriminated according to the type of participation.

6.2 Type of detector

Thirty whole-body counters conducted measurements with high-purity germanium (HPGe) detectors, nine whole-body counters conducted measurements with sodium iodide (NaI(Tl)) detectors. NaI(Tl) detectors feature a reduced energy resolution compared to HPGe detectors, impeding the discrimination and identification of radionuclides when the source features many gamma-radiation emissions or emissions with similar energies. Two whole-body counters conducted the identification of the radionuclides with HPGe detectors and measured the activity of the identified radionuclides with additional NaI(Tl) detectors.

Table 42: Statistical parameters of the attribution of z-scores with the type of participation.

	HPGe	NaI(Tl)	HPGe + NaI(Tl)
n_{lab}	30	9	2
n_{result}	367	69	27
n_{result}/n_{lab}	12.2	7.7	13.5
mean	-0.08	0.18	1.33
median	-0.06	-0.04	0.69
standard deviation	1.33	1.22	1.53
p-values *Mann-Whitney*		0.39 not sign. diff.	
Siegel-Tukey		0.082 not sign. diff.	

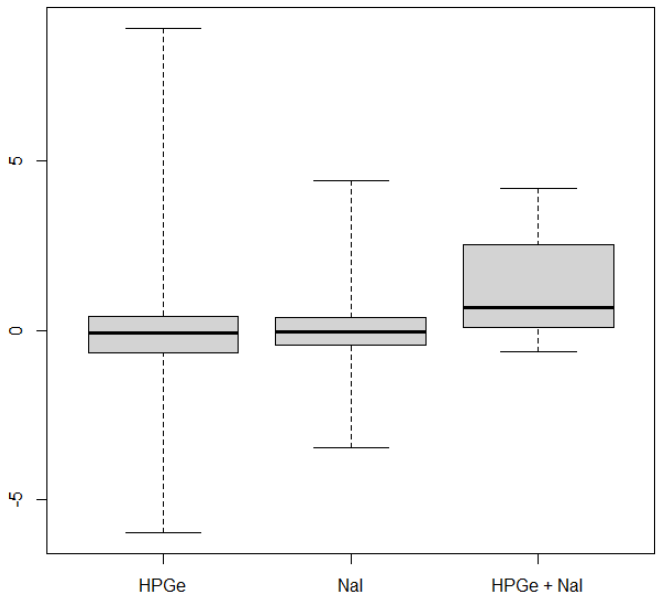


Figure 27: Box plot of the z-scores discriminated according to the type of detector.

It can be seen that the performance of HPGe and of NaI(Tl) detectors is similar (Figure 27, Table 42). The differing results for the combination of HPGe and NaI(Tl) detectors are not significant because of the small number of results.

It was also tested if different results between HPGe and NaI(Tl) detectors can be identified for ^{134}Cs (Table 43, Figure 28) and ^{137}Cs (Table 44, Figure 29) for Emergency (Task 2) and for Victor (Task 1) (Table 45, Figure 30).

The relevance of these nuclides in emergency response makes it important that they can be measured with NaI(Tl) detectors, which are often held available for emergency measurements, with good accuracy. Each single of these nuclides should not pose a difficulty for NaI(Tl) detectors because

of the small number of gamma emissions. However, in the Emergency task both nuclides were present together as it is usual after releases from nuclear reactors so that the peak of ^{137}Cs is overlapped by a peak of ^{134}Cs in NaI(Tl) spectra. The activity of ^{134}Cs can be calculated from undisturbed peaks whereas that is not possible for ^{137}Cs since this nuclide features only one gamma emission. In contrast, ^{137}Cs in the Victor task is not disturbed by the other two nuclides ^{60}Co and ^{133}Ba, which are present in this task, so that it is conceivable that the activity of ^{137}Cs in this task could have been calculated with better precision using NaI(Tl) detectors.

Table 43: Statistical parameters of the attribution of z-scores with the type of detector for ^{134}Cs in the Emergency task.

	HPGe	NaI(Tl)
n_{lab}	30	9
n_{result}	28	9
n_{result}/n_{lab}	0.9	1.0
mean	-0.25	0.41
median	0.04	-0.04
standard deviation	1.86	1.39
p-values		
Mann-Whitney		0.61 not sign. diff.
Siegel-Tukey		0.82 not sign. diff.

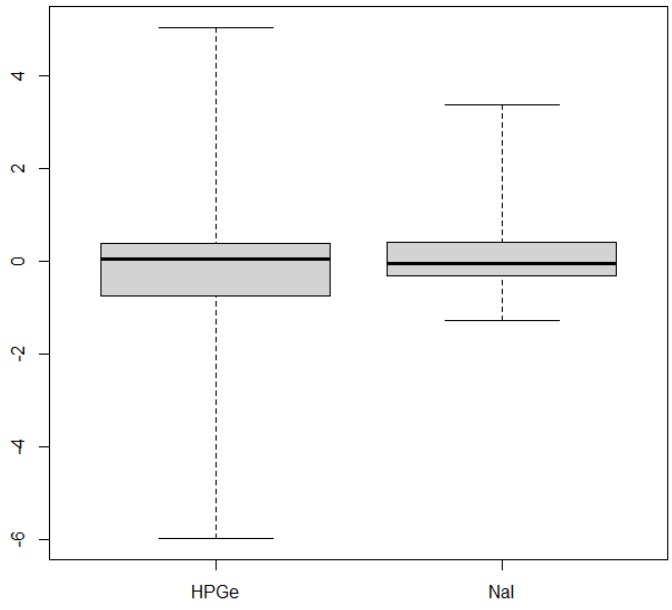


Figure 28: Box plot of the z-scores discriminated according to the type of detector for ^{134}Cs in the Emergency task.

Table 44: Statistical parameters of the attribution of z-scores with the type of detector for ^{137}Cs in the Emergency task, in which also ^{134}Cs was present.

	HPGe	NaI(Tl)
n_{lab}	30	9
n_{result}	29	9
n_{result}/n_{lab}	1.0	1.0
mean	0.23	-0.25
median	0.16	-0.13
standard deviation	2.19	0.95
p-values		
Mann-Whitney		0.26 not sign. diff.
Siegel-Tukey		0.84 not sign. diff.

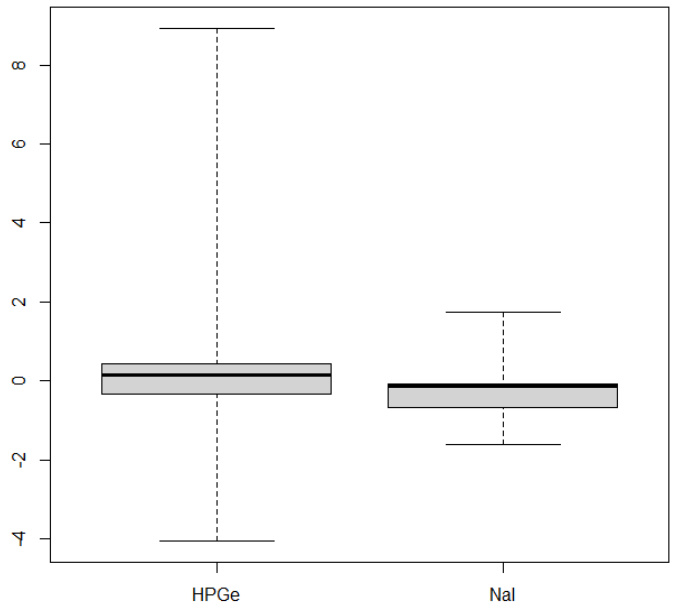


Figure 29: Box plot of the z-scores discriminated according to the type of detector for ^{137}Cs in the Emergency task, in which also ^{134}Cs was present.

Table 45: Statistical parameters of the attribution of z-scores with the type of detector for ^{137}Cs in the Victor task.

	HPGe	NaI(Tl)
n_{lab}	30	9
n_{result}	29	8
n_{result}/n_{lab}	1.0	0.9
mean	-0.23	0.23
median	-0.07	0.25
standard deviation	1.37	1.18
p-values *Mann-Whitney*		0.55 not sign. diff.
Siegel-Tukey		0.63 not sign. diff.

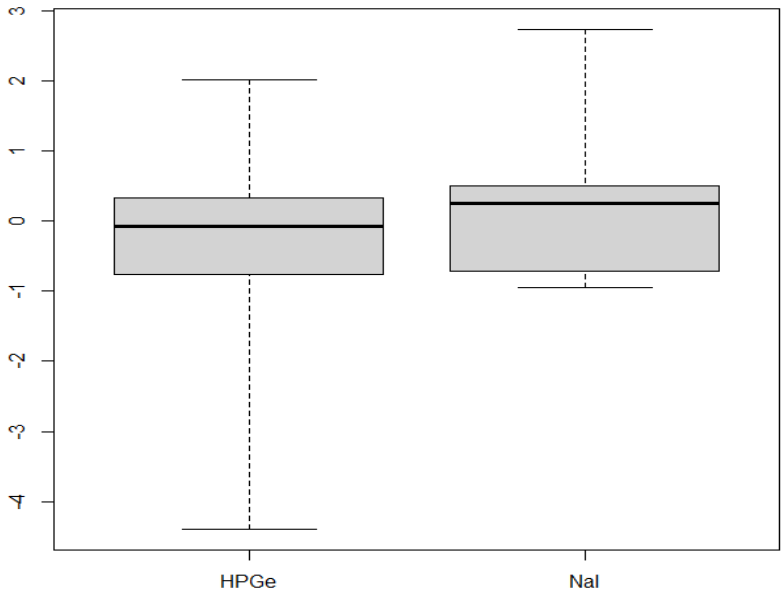


Figure 30: Box plot of the z-scores discriminated according to the type of detector for ^{137}Cs in the Victor task.

It can be seen that ^{137}Cs in the Emergency task was measured with NaI(Tl) detectors with a slight, yet not significant underestimation as compared to HPGe detectors. This is particularly interesting since in the Victor task, which can serve as a reference because of the undisturbed measurability of ^{137}Cs, facilities with NaI(Tl) detectors tended to report slightly bigger results than facilities with HPGe detectors (although not significantly different either). It must be noted that the reference value was calculated as the robust mean of all reported results (excluding outliers) for this nuclide so that the smaller values of NaI(Tl) detectors decreased the reference value. The underestimation in the

Emergency task could have been caused by an excessive subtraction of the count rate of the combined ^{134}Cs and ^{137}Cs peak to calculate the activity of ^{137}Cs only, which was necessary with NaI(Tl) detectors.

6.3 Measurement of a 90 kg phantom with a 70 kg calibration

A phantom of 90 kg was used for the Emergency task. However, several laboratories calibrate their whole-body counters only with a phantom of 70 kg and apply that calibration for all masses of people and phantoms to be measured; others conduct calibrations with different phantoms of up to 70 kg. On the other hand, several laboratories conduct calibrations also with bigger phantoms of 90 kg. For the results of the Emergency task, the performance of laboratories with a calibration for up to 70 kg or 70 kg only and of those with a calibration also for 90 kg was compared.

As shown in the Table 46 and Figure 31, the z-scores of those laboratories that calibrate their whole-body counters only at a mass of 70 kg are significantly smaller than the z-scores of those labs that applied a 90 kg calibration for the 90 kg measurement. It must be noted again that the reference value is the robust mean of all results so that no information can be derived which of the two groups reported better results. It is sensible that smaller results are gained if a 90 kg phantom is measured with a 70 kg calibration because of the stronger attenuation of the emitted gamma radiation by the bigger phantom.

Table 46: Statistical parameters of the attribution of z-scores with the phantom masses for calibration measurements.

	also 90 kg	only 70 kg
n_{lab}	14 (9x HPGe, 3x NaI(Tl), 2x both)	18 (13x HPGe, 5x NaI(Tl))
n_{result}	36	49
n_{result}/n_{lab}	2.6	2.7
mean	0.84	-0.14
median	0.35	-0.08
standard deviation	1.46	0.63
p-values *Mann-Whitney*		2.52×10^{-4} sign. diff.
Siegel-Tukey		3.39×10^{-2} sign. diff.

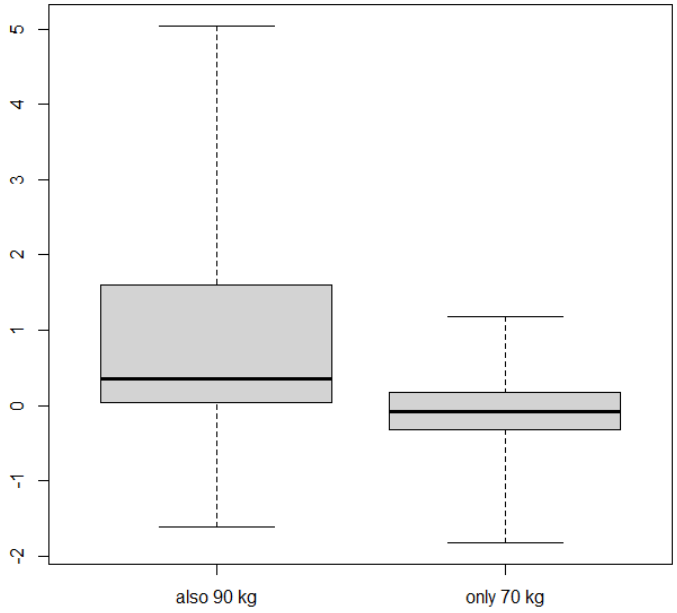


Figure 30: Box plot of the z scores discriminated according to the phantom masses for calibration measurements.

6.4 Measurement geometry

The whole-body counters that participated in the intercomparison conducted measurements in different geometries: stretcher/lying, chair/sitting, inclined chair, standing. Because of the lack of statistical power, the one laboratory with standing geometry was excluded from the statistical analysis of the association between geometry and results. The statistical parameters are given in Table 47 and the comparison in Figure 32. As shown no difference was found between the different type geometries used.

Table 47: Statistical parameters of the attribution of z-scores with the type of geometry.

	stretcher	inclined	chair	standing
n_{lab}	28	4	7	1
n_{result}	317	53	72	
n_{result}/n_{lab}	11.3	13.3	10.3	
mean	0.00	0.15	0.11	
median	-0.05	0.09	0.04	
standard deviation	1.40	0.80	1.57	
p-values *Mann-Whitney*		0.16 not sign. diff.	0.99 not sign. diff.	
Siegel-Tukey		0.13 not sign. diff.	0.53 not sign. diff.	

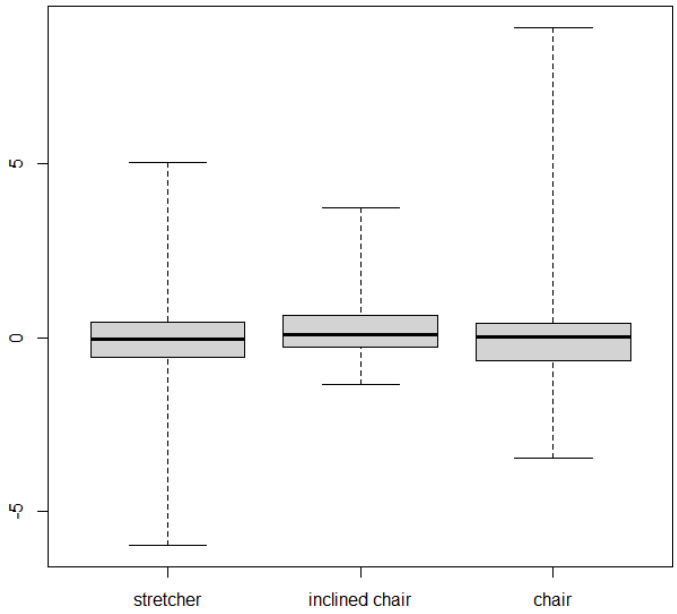


Figure 31: Box plot of the z-scores discriminated according to the type of geometry.

6.5 Type of calibration phantom

The participating laboratories used different types of phantoms for the calibration of their whole-body counters: brick phantom (equal or similar to the one that was used in this intercomparison), bottle mannequin absorber phantom BOMAB, other types of bottle phantoms, Canberra Transfer Phantom, computational phantoms for Monte-Carlo simulation, the Lawrence-Livermore Lung Phantom (LLNL) and self-made phantoms. In order to increase the statistical power of the comparison between bottle and brick phantoms, all types of bottle phantoms were summarised. Because of the small number of laboratories, LLNL lung phantoms and computational phantoms were excluded from the statistical analysis.

It can be seen in Table 48 and Figure 33 the bottle and brick phantoms showed similar results regarding the central tendency but different results regarding the dispersion (despite the fact that the standard deviation is quite similar). As can be seen in the box plot, the difference in the dispersion was influenced by some rather big under- and overestimations from laboratories with a brick phantom. With the Canberra phantom, results tended to be underestimated and with own phantoms results tended to be overestimated (yet with small dispersion despite the different makeups of these phantoms).

Table 48: Statistical parameters of the attribution of z-scores with the type of calibration phantom.

	brick	BOMAB	other bottle	Bottle (BOMAB and other)	Canberra	own	LLNL lung	Computa -tional
n_{lab}	12	5	9	14	4	4	2	2
n_{result}	135			165	42	40		
n_{result}/n_{lab}	11.3			11.8	10.5	10.0		
mean	0.19			0.35	-0.68	0.60		
median	0.08			0.07	-0.41	0.46		
standard deviation	1.19			1.25	1.66	1.31		
p-values Mann-Whitney				0.59 not sign. diff.	0.00035 sign. diff.	0.0063 sign. diff.		
Siegel-Tukey				0.00093 sign. diff.	0.0071 sign. diff.	0.022 sign. diff.		

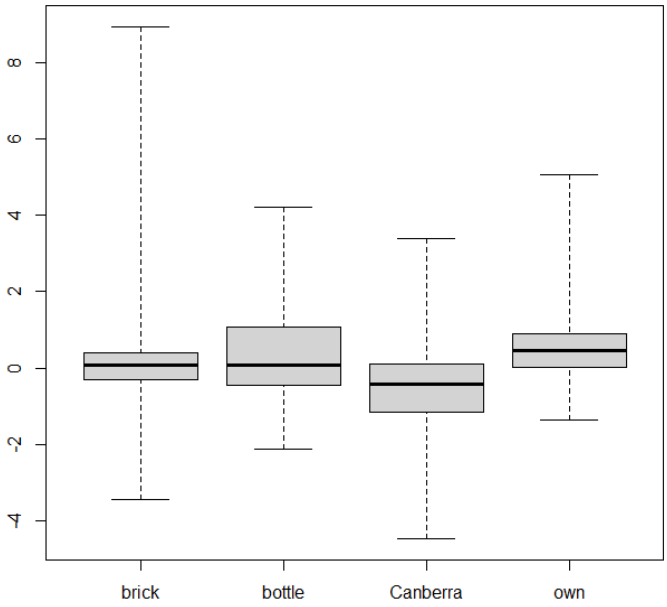


Figure 32: Box plot of the z-scores discriminated according to the type of calibration phantom.

6.6 Identification of the 1077 keV peak of ^{68}Ge

^{68}Ge (via its daughter nuclide ^{68}Ga) emits gamma radiation of an energy of 1077 keV with an abundance of 3.2%. Additionally, it produces secondary annihilation radiation of an energy of 511 keV with an abundance of 178%. Because of the great differences in the abundances, several laboratories were able to detect only the peak at 511 keV. However, this peak is influenced by background counts, which need to be subtracted to calculate the net count rate. Because of the rather small number of laboratories that were able to identify ^{68}Ge, statistical tests were not conducted.

It can be seen in Table 49 and Figure *34* that the dispersion of the results that were calculated with taking into account the peak at 1077 keV is smaller than that of the results that were calculated from 511 keV only.

Table 49: Statistical parameters of the attribution of z-scores for ^{68}Ge with those peaks that were applied for the calculation of the activity.

	511 keV only	511 keV and 1077 keV	1077 keV only
n_{lab}	4	5	2
n_{result}	4	5	2
n_{result}/n_{lab}	1.0	1.0	1.0
mean	0.51	0.57	-0.01
median	0.48	0.78	-0.01
standard deviation	1.90	0.66	0.46

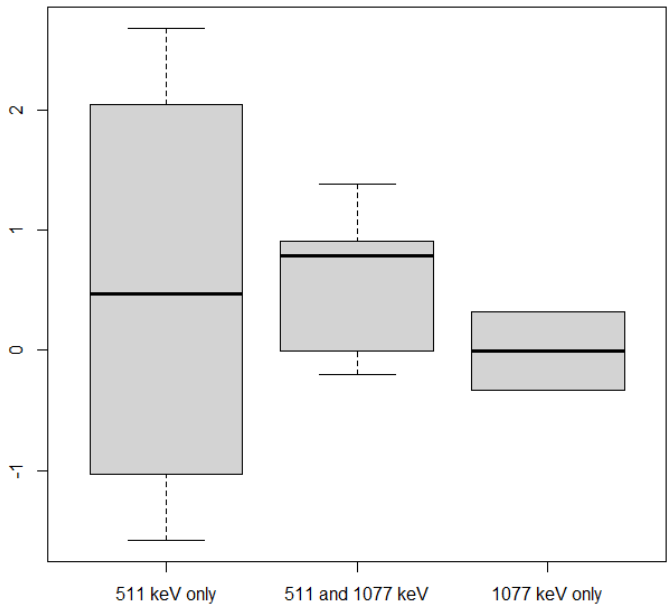


Figure 33: Box plot of the z-scores for ^{68}Ge discriminated according to those peaks that were applied for the calculation of the activity.

6.7 Accreditation of the laboratory

Accreditation according to ISO 17 025 [ISO 2017] is a certification that specific requirements regarding the management system are fulfilled. These requirements comprise organisational and technical aspects, such as traceability of the calibration, regular internal and external audits and regular participation in proficiency tests or inter-laboratory comparisons. In several countries, accreditation is mandatory for internal monitoring laboratories.

For the statistical evaluation, all laboratories that were not accredited at the time of the measurements were counted as not accredited, no matter if accreditation was prepared or if a quality-management system without accreditation was in force at the laboratory.

Although laboratories that were not accredited reported similar results on average, the dispersion of their results is significantly bigger (Table 50 and Figure 35). The set of results was divided into three groups of approximately equal size: laboratories with less than 100 measurements, with 100 to 500 measurements and with more than 500 measurements (reference for this test) in a year.

Table 50: Statistical parameters of the attribution of z-scores with the accreditation of laboratories.

	accredited	Not accredited
n_{lab}	22	16
n_{result}	247	179
n_{result}/n_{lab}	11.2	11.2
mean	0.09	-0.04
median	-0.03	-0.04
standard deviation	0.92	1.82
p-values		
Mann-Whitney		0.49 not sign. diff.
Siegel-Tukey		1.13×10^{-12} sign. diff.

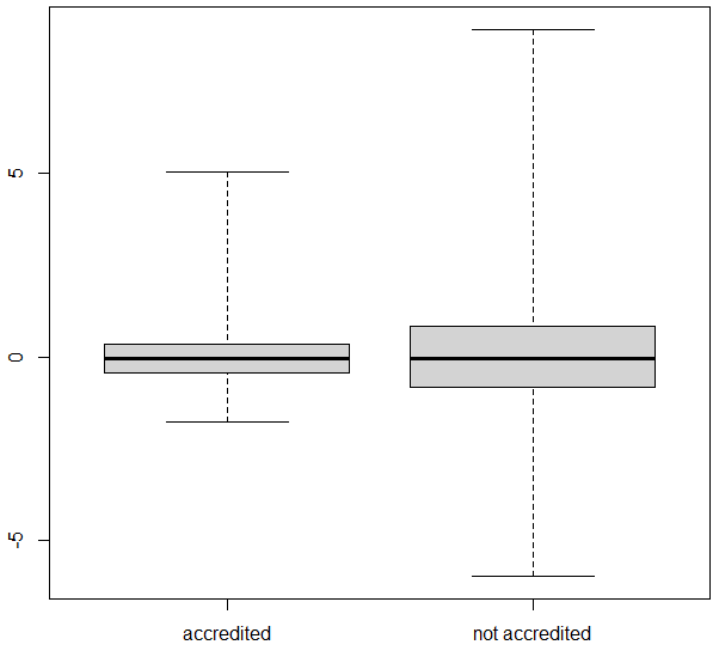


Figure 34: Box plot of the z-scores discriminated according to the fact if the laboratories were accredited or not.

Table 51 and Figure 36 indicate that laboratories with a greater number of routine measurements yielded z-scores closer to zero on average and with a smaller dispersion. However, the results of smaller laboratories were also generally acceptable. The pursuit of a dense network of whole-body counters, which is promoted in many countries in order to be prepared for measurements of a large number of people in an emergency, is justified by the results that were yielded by the smaller laboratories.

Table 51: Statistical parameters of the attribution of z-scores with the number of annual routine measurements.

	>500	100-500	<100
n_{lab}	13	14	13
n_{result}	152	156	139
n_{result}/n_{lab}	11.7	11.1	10.7
mean	-0.02	0.60	-0.53
median	-0.03	0.24	-0.47
standard deviation	0.71	1.56	1.51
p-values			
Mann-Whitney		1.41×10^{-4} sign. diff.	6.73×10^{-5} sign. diff.
Siegel-Tukey		1.42×10^{-5} sign. diff.	9.72×10^{-3} sign. diff.

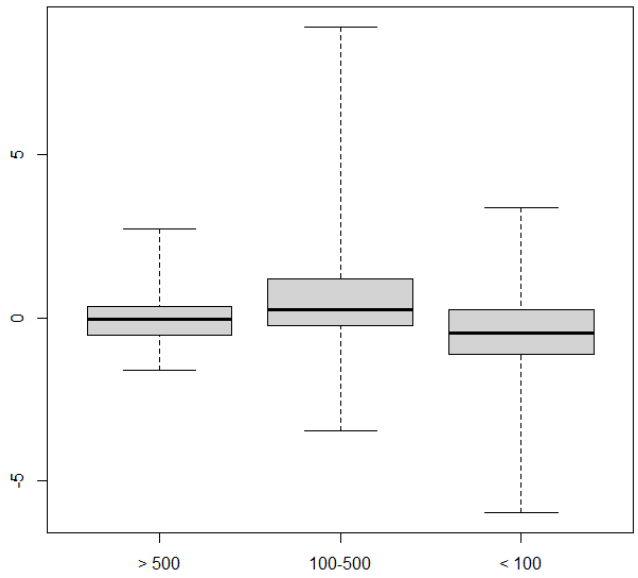


Figure 35: Box plot of the z-scores discriminated according to the number of annual routine measurements.

7. General conclusion of the intercomparison

The objective of the EIVIC project was to assess the implementation of the individual monitoring requirements in EU Member States based on in-vivo measurements and receive an overview of the capabilities and performance of whole-body counters in Europe. It was organised between October 2019 and June 2022 and dedicated to whole-body measurement of gamma emitters in several tasks selected that cover the range of such possible measurements associated to different intake scenarios. In total, 43 installations from 21 countries took part in the proposed measurements.

For this intercomparison, four tasks were defined with measurements of phantoms equipped with radionuclide sources. For each phantom measurement task, one specific set of radionuclide sources has been used. Each set contains a mixture of those radionuclides that are to be measured in the respective measurement task. The measurement tasks comprise the following nuclides:

- Task 1: ^{60}Co, ^{133}Ba and ^{137}Cs. P4/70 kg (called Victor as these are the usual sources for the IRSN proficiency test phantom called Victor),
- Task 2: ^{134}Cs and ^{137}Cs. P5/90 kg (called Emergency),
- Task 3: ^{68}Ge and ^{88}Y. P4/70 kg (called Medicine),
- Task 4: ^{133}Ba and ^{152}Eu. P4 & P5 configurations (called Calibration).

The analysis of the results was carried out with regard to the criteria of standard ISO 28218 and standard ISO 13528. The summary of the results is given below following these two standards.

7.1 Standard ISO 28218

According to standard ISO 28218, the conformity of the installation requires that the bias between the assigned value and the value of the participant must be between [-25%; +50%].

7.1.1 Task 1: ^{60}Co, ^{133}Ba and ^{137}Cs. P4/70 kg

- ^{60}Co: three facilities [27, 28 and 39] are in non-conformity. The other facilities are in conformity. The facilities '27' and '28' have been detected as a suspicious value and the facility '39' has been detected as an outlier (Grubbs test).
- ^{133}Ba: four facilities [27, 28, 31 and 39] are in non-conformity. The other facilities are in conformity. The facility '28' has been detected as a suspicious value and the facilities '31' and '39' have been detected as outliers (Grubbs test).
- ^{137}Cs: three facilities [28, 30 and 39] are in non-conformity. The other facilities are in conformity. Facility '39' has been detected as an outlier (Grubbs test).
- ^{40}K: three facilities [14, 28 and 37] are in non-conformity. The other facilities are in conformity. The facility '37 has been detected as an outlier (Grubbs test).

7.1.2 Task 2: ^{134}Cs and ^{137}Cs. P5/90 kg

- ^{134}Cs: three facilities [28, 30 and 39] are in non-conformity. The other facilities are in conformity. For the facility '28' a suspicious value has been detected and the facility '39' has been considered as an outlier (Grubbs test).
- ^{137}Cs: three facilities [28, 30 and 39] are in non-conformity. The other facilities are in conformity.
- ^{40}K: three facilities [28, 37 and 39] are in non-conformity. The other facilities are in conformity. The facility '39' has been detected as an outlier (Grubbs test).

7.1.3 Task 3: ^{68}Ge and ^{88}Y. P4/70 kg #1

- ^{68}Ge: all facilities are in conformity.
- ^{88}Y: one facility [30] is in non-conformity. The other facilities are in conformity.
- ^{40}K: two facilities [1 and 37] are in non-conformity. Other facilities are in conformity. The result of the facility '37' has been detected as an outlier (Grubbs test).

7.1.4 Task 3: ^{68}Ge and ^{88}Y. P4/70 kg #2

- ^{68}Ge: all facilities are in conformity.
- ^{88}Y: two facilities [28, 39] are in non-conformity. The other facilities are in conformity. The facility '39' has been detected as an outlier and '28' as a suspicious value (Grubbs test).
- ^{40}K: three facilities [14, 28 and 39] are in non-conformity. Other facilities are in conformity.

7.1.5 Task 4a: ^{133}Ba and ^{152}Eu. P4/70 kg

- ^{133}Ba: all facilities are in conformity.
- ^{152}Eu: two facilities [30, 31] are in non-conformity. The other facilities are in conformity. The facility '31' has been detected as an outlier (Grubbs test).
- ^{40}K: four facilities [14, 20, 30 and 33] are in non-conformity. Other facilities are in conformity.

7.1.6 Task 4b: ^{133}Ba and ^{152}Eu. P5/90 kg

- ^{133}Ba: two facilities [33 and 39] are in non-conformity. The other facilities are in conformity. The values of three facilities [31, 33 and 39] have been detected as outliers (Grubbs test).
- ^{152}Eu: two facilities [30, 39] are in non-conformity. The other facilities are in conformity. The result of the facility '39' has been detected as an outlier (Grubbs test).
- ^{40}K: four facilities [1, 14, 31 and 39] are in non-conformity. Other facilities are in conformity. The value of facility '39' has been detected as an outlier (Grubbs test).

7.2 Standard ISO 13528

Concerning the results analysed with standard ISO 13528, they should be interpreted with caution because the measurements were not carried out under the same conditions. The z-scores should be considered as additional elements allowing the laboratory to assess itself in relation to other participants.

According to this criterion:

7.2.1 Task 1: ^{60}Co, ^{133}Ba and ^{137}Cs. P4/70 kg

- ^{60}Co: the z-score of the facilities '27', '28' and '39' is larger than 3 ($|z|>3$), gives an action signal and for the facilities '13', '34' and '36' it gives a warning signal ($2<|z|$ 3). All other facilities are considered satisfactory ($|z|<2$).
- ^{133}Ba: the z-score of the facilities '27', '28', '31' and '39' over 3 ($|z|>3$) gives an action signal. All other facilities are considered satisfactory ($|z|< 2$).
- ^{137}Cs: the z-score of the facilities '28', '30', '31' and '39' is larger than 3 ($|z|>3$) and gives an action signal and for the facilities '8', '34' and '36' it gives a warning signal ($2<|Z|$ 3). All other facilities are considered satisfactory ($|z|<2$).

- ^{40}K: the z-score of facility '37' is over 3 (|z|>3), giving an action signal and for the facilities '28' and '36' it gives a warning signal (2<|z| 3). All other facilities are considered satisfactory (|z|<2).

7.2.2 Task 2: ^{134}Cs and ^{137}Cs. P5/90 kg

- ^{134}Cs: the z-score of the facilities '27', '28', '30', '31', '33' and '39' is larger than 3 (|z|>3) and gives an action signal. All other facilities are considered satisfactory (|z|<2).
- ^{137}Cs: the z-score of the facilities '28','30', '33' and '39' is larger than 3 (|z|>3), gives an action signal and the for the facility '31' it gives a warning signal (2<|z| 3). All other facilities are considered satisfactory (|z|<2).
- ^{40}K: the z-score of the facilities '37' and '39' is over 3 (|z|>3) and gives an action signal and for the facility '28' it gives a warning signal (2<|z| 3). All other facilities are considered satisfactory (|z|<2).

7.2.3 Task 3: ^{68}Ge and ^{88}Y. P4/70 kg #1

- ^{68}Ge: the z-scores of all facilities are considered satisfactory (|z|<2).
- ^{88}Y: the z-score of the facility '30' exceed 3 (|z|>3) and gives an action signal. All other facilities are considered satisfactory (|z|<2).
- ^{40}K: the z-score of the facilities '1' and '37' are over 3 (|z|>3), give an action signal and the facility '5' and '32' give a warning signal (2<|z| 3). All other facilities are considered satisfactory (|z|<2).

7.2.4 Task 3: ^{68}Ge and ^{88}Y. P4/70 kg #2

- ^{68}Ge: the z-score of the facility '31' gives a warning signal (2<|z|<3). All other facilities are considered satisfactory (|z|<2).
- ^{88}Y: the z-score of facilities '28' and '39' is larger than 3 (|z|>3), gives an action signal. All other facilities are considered satisfactory (|z|<2).
- ^{40}K: the z-score of facility '28' is over 3 (|z|>3), gives an action signal and the one for the facility '39' gives a warning signal (2<|z|<3). All other facilities are considered satisfactory (|z|<2).

7.2.5 Task 4a: ^{133}Ba and ^{152}Eu. P4/70 kg

- ^{133}Ba: the z-score of facility '31' is higher than 3 (|z|>3), gives an action signal and for the facilities '30' and '36' it gives a warning signal (2<|z|<3). All other facilities are considered satisfactory (|z|<2).
- ^{152}Eu: that the z-score of facility '31 is over 3 (|Z|>3) and gives an action signal and for the facilities '29' and '36' it gives a warning signal (2<|z|<3). All other facilities are considered satisfactory (|z|<2).
- ^{40}K: all facilities are considered satisfactory (|z|<2).

7.2.6 Task 4b: ^{133}Ba and ^{152}Eu. P5/90 kg

- ^{133}Ba: the z-score of the facilities '31', '33' and '39' is over 3 (|z|>3), gives an action signal. All other facilities are considered satisfactory (|z|<2).
- ^{152}Eu: the z-score of the facilities '31', '33' and '39' exceeds 3 (|z|>3), gives an action signal and for the facilities '16' and '30' it gives a warning signal (2<|z|<3). All other facilities are considered satisfactory (|z|<2).

- ^{40}K: the z-score of facility '1' is larger than 3 (|z|>3) and gives an action signal and for the facility '31' it gives a warning signal (2<|z|<3). All other facilities are considered satisfactory (|z|<2).

7.3 Statistical evaluation

The quality of most of the participating laboratories was rather independent from the metrological and organisational characteristics. The dispersion of the results within each investigated property was stronger than the difference between different properties. Therefore, attributable differences of these properties are small (no matter if significant or not).

Laboratories with NaI(Tl) detectors reported results that did not differ significantly from those with HPGe detectors, showing that the use of NaI(Tl) detectors is still justified despite the trend of the last years to HPGe detectors. Likewise, the type of geometry (stretcher, chair etc.) did not show a significant attributable influence on the results. Regarding the type of phantom, the statistical evaluation leads to the recommendation to use the most prevalent types of phantoms, namely brick phantoms or bottle phantoms instead of less common or even self-made phantoms.

Laboratories with a small number of measurements showed significantly different results than laboratories with more frequent measurements, however the interest in a dense network of internal service laboratories that exists in many countries might be a strong reason for maintaining also small laboratories. Accreditation also had a beneficial effect on the quality of the reported results. Of the two participants with the most extreme biases over all tasks, '28' falls in the group of small laboratories and '39' falls in the group of laboratories with a medium number of measurements; both were not accredited. Furthermore, it is observed that for Task 3, the most difficult exercise, the number of participants who reported results was small and almost exclusively limited to frequent participants in previous intercomparisons organised by IRSN or BfS.

8. References

[BIPM 2020] Bureau International des Poids et Mesures (editor), Table of Radionuclides, Volumes 1-9, Monographie BIPM-5. Sèvres, France, 2004-2020. http://www.lnhb.fr/nuclear-data/nuclear-data-table/.

[ICRU 1992] International Commission on Radiation Units and Measurements. Phantoms and Computational Models in Therapy, Diagnosis and Protection. ICRU Report 48 (ICRU: Bethesda) (1992).

[ICRU 2003] International Commission of Radiation Units and Measurements. Direct Determination of the Body Content of Radionuclides. ICRU Report 69, Journal of the ICRU 3 (1), (2003).

[ISO 2010] International Organization for Standardization. Radiation Protection — Performance criteria for radiobioassay. ISO 28218:2010 (ISO: Geneva) (2010).

[ISO 2012] International Organization for Standardization. Sealed radioactive sources — General requirements and classification. ISO 2919:2012 (ISO: Geneva) (2012).

[ISO 2022] International Organization for Standardization. Statistical methods for use in proficiency testing by interlaboratory comparisons. ISO 13528:2022 (ISO: Geneva) (2022).

[Kovtun 2000] N. Kovtun, V. B. Firsanov, V. I. Fominykh and G. A. Isaakyan. Metrological Parameters of the Unified Calibration Whole Body Phantom with Gamma-emitting Radionuclides. Radiat. Prot. Dosim., vol 89, pp. 239-242 (2000).

[R Core Team, 2020] R: A language and environment for statistical computing. Version 4.0.2. Vienna, Austria: R Foundation for Statistical Computing (2020).

[RP 2018] Radiation Protection N° 188, Technical Recommendations for Monitoring Individuals for Occupational Intakes of Radionuclides, Luxembourg, Publications Office of the European Union, ISBN 978-92- 79-86304-2 (2018).

[Woidy 2022] P. Woidy and O. Meisenberg. Production of sealed rod sources made from epoxy resin for the Saint-Petersburg brick phantom for the calibration of whole-body counters. Radiat. Environ. Biophys., vol. 61, pp. 391-398 (2022).

Annexes

Annex I: Invitation letter for potential participants

Annex II: Letter accompanying shipment of phantoms and sources

Annex III: Instructions about EIVIC phantoms assembly

Annex IV: Template for the Certificate of Attendance

Annex I: Invitation letter for potential participants

European Radiation Dosimetry Group

EURADOS

European Radiation Dosimetry Group e.V. • Postfach 11 29 • D-85758 Neuherberg

Your reference:
Your letter dated:
Our letter dated:

Our reference: EURADOS/2020/02/DF
Phone number: +33 (0)1 58 35 78 74
Fax number:
E-Mail: didier.franck@irsn.fr
Date: 17.01.2020

Invitation Letter to participate in EIVIC2020 - "Organisation of a European Interlaboratory Comparison on Whole Body Counting"

Dear Sir/Madam,

EIVIC2020 is a European Commission Project (EC, DG-ENER) to be developed on 2020-2021 (ref ENER/2019/NUCL/SI2.811157). Its aim is to organize the 2020 European Intercomparison of Whole Body Counters for the measurement of gamma emitters in total body.

The exercise is being coordinated by EURADOS (European Radiation Dosimetry Group, www.eurados.org, Chair Werner Rühm; Kerstin Hürkamp) together with IRSN-France (Chair Didier Franck; Tiffany Beaumont), BfS-Germay (Oliver Meisenberg, Werner Buchholz), counting with the collaboration of CIEMAT-Spain (M.A. López, J.F. Navarro, B. Pérez) and KIT-Germany (B. Breustedt).

Two brick phantom units simulating the internal contamination of an adult male will be transported or sent to the different in vivo laboratories all around Europe between autumn 2020 and summer 2021.

Please send your answer confirming/refusing participation by e-mail to EIVIC2020@eurados.org before 15 February 2020.

Further information regarding the protocol of measurements, information to be provided by your facility and the dates proposed (roadmap of the exercise) will be distributed later.

Sincerely,

Didier Franck Ph.D.
EIVIC Chair
IRSN
Deputy Head of the Research Dosimetry Department
PSE-SANTE/SDOS
B.P. 17 - 92262 Fontenay-aux-Roses Cedex
e-mail didier.franck@irsn.fr

Maria Antonia Lopez Ph.D.
EIVIC WP1 Leader
CIEMAT
Head of Internal Dosimetry Group
Avda. Complutense 40
28040 Madrid, Spain
e-mail ma.lopez@ciemat.es

Annex II: Letter accompanying shipment of phantoms and sources

EIVIC_letter.docx [Lecture seule] - Word

EURADOS

IRSN
INSTITUT DE RADIOPROTECTION ET DE SÛRETÉ NUCLÉAIRE

Bundesamt für Strahlenschutz

City, 05 May 2021

Laboratory Name
Laboratory Adresse

EIVIC Staff
General organization
didier.frank@irsn.fr
ma.lopez@ciemat.es

Transport/measurements
wbuchholz@bfs.de
omeisenberg@bfs.de

Results/analysis
tiffany.beaumont@irsn.fr

Subject EIVIC organization

Dear participant,

Your lab participated in the EIVIC project. You will find below any information on practical organization.

1 - Organization

This intercomparison will begin in second quarter of 2021.

- The circuit, with the running order, will be sent as soon as possible;
- Between two participants the transport of the phantom sources will be of the BFS's responsibility;
- Each laboratory has available 4 days to carry out measurements.

2 -Phantom and sources

The phantom « whole-body » and the measurements should be made with the usual methods (counting time, positioning ...).

Four tasks are proposed:

- **Task 1 "Victor"**, suitable for NaI(Tl) and germanium detectors: P5 phantom
- **Task 2 "accident"**, suitable for NaI(Tl) and germanium detectors: P4 phantom
- **Task 3 "medicine"**, suitable for NaI(Tl) and germanium detectors: P5 phantom
- **Task 4 "calibration"**, suitable only with germanium detectors: P4 and P5 phantom

A more detailed program will be sent as soon as possible.

3 - Results

EIVIC provides to each laboratory a template to submit the results "EIVIC_Results_*LaboratoryName*.xls". For the laboratory with many installations, a number equivalent of template to the number of installation will be provide.

The results must be signed by the person in charge of the measurements and send within a maximum of fifteen day after the departure of the phantom.

The template must be returned by e-mail, in ".xls" format, at tiffany.beaumont@irsn.fr and ma.lopez@ciemat.es.

Yours faithfully.

EIVIC Staff

Annex III: Instructions about EIVIC phantoms assembly

CIEMAT/DRD/DPI/07/21

EIVIC Project 2020-2021 European in-Vivo Monitoring Intercomparison Exercise

ENER/2019/NUCL/SI2.811157

"Organization of a European Interlaboratory Comparison on Whole Body Counting"

IGOR-OLGA EIVIC Phantoms Instructions about brick phantom assembly

B. Pérez López, J.F. Navarro Amaro, M.A. López Ponte, W. Buchholz, O. Meisenberg

Whole Body Counter Laboratory- Internal Dosimetry Service (CIEMAT)

The following materials (texts, tables, videos) were provided to the participants on the BfS project data cloud (EIVIC 2020 resources → document manuals folders) in order to properly set up the phantoms in their labs and contain instructions on packaging and shipment of the phantoms:

Table 52: Materials provided to the participants on the BfS project data cloud.

File name	Description	
Pictures folder	How to pack the cases New_2021: photos of the phantom Old 2020: photos of the phantom	22 ipg-files 62 jpg-files 32 jpg-files
Lying 70 kg phantom.mp4 Lying 90 kg phantom.mp4 Sitting 70 kg phantom.mp4 Sitting 90 kg phantom.mp4	Video on the lying 70 kg phantom Video on the lying 90 kg phantom Video on the sitting 70 kg phantom Video on the sitting 90 kg phantom	1:26 min 1:14 min 1:28 min 2:32 min
p4_p5-lying_manual.xlsx p4_p5-sitting-upright_manual.xlsx p4_p5-standing_manual.xlsx	Manuals on the setup of the standing, sitting and lying phantom	2 Excel sheets 2 Excel sheets 3 Excel sheets
phantoms_EIVIC2020-shipping_wb0925.xlsx	Manual on the contents of the boxes and phantom constructions	8 Excel sheets
phantomsetting_EIVIC2021_english.docx	Description Phantom Setup EU In-vivo Intercomparison Exercise	4 p.
PHANTOM_REPORT_EIVICV1.docx	IGOR-OLGA EIVIC Phantoms - Instructions about brick phantom assembly	19 p.

Annex IV: Template for the Certificate of Attendance

City, 12 September 2022

Laboratory Name
Laboratory Adresse

EIVIC Staff
General organization
didier.frank@irsn.fr
ma.lopez@ciemat.es

Transport/measurements
wbuchholz@bfs.de
omeisenberg@bfs.de

Results/analysis
tiffany.beaumont@irsn.fr

Subject EIVIC certificate of participation

Dear participant,

Your laboratory participated in the European In-Vivo Intercomparison Exercise organized within the framework of the EIVIC project, funding by the European Commission[1].

Your anonymized participation code, which will be used to find your results in the final report, is XX. For each task, the reference values and their standard (k=2) uncertainties are given (Bq) as well as the bias and the z-score of your results.

The analysis of results was carried out and the conformity report of the facilities are given, for each configuration, according to the criteria of ISO 28 218 and ISO 13 528.

According to the recommendations of ISO 28 218 "Performance criteria for radiobioassay", the relative bias shall be between [-25% to +50%] relative to the target value (reference value).

According to the recommendations of ISO 13 528 "Statistical methods for use in proficiency testing by intercomparison", the current z-score criteria are:

- $|z - score| \leq 2$: the result is acceptable;
- $2 < |z - score| \leq 3$: The result is considered to give a warning signal;
- $|z - score| > 3$: The result is considered to be unacceptable (action signal)

Please find below the results of your facility.

1 - Results Task 1 'Victor' (phantom size P4/70kg)

Task 1	^{60}Co	^{133}Ba	^{137}Cs	^{40}K
*Reference Value (Bq)**	1183 ± 25	2836 ± 62	3787 ± 65	3941 ± 137
Bias (%)				
Z-score				

**determined with the robust mean at the reference date (May 1st, 2021 - Beginning of the IC campaign)*

[1] All results have been produced under contract ENER/2019/NUCL/SI2.811157 with the European Atomic Energy Community.

2 - Results Task 2 'Accident' (P5/90kg)

Task 2	^{134}Cs	^{137}Cs	^{40}K
*Reference Value (Bq)**	3455 ± 50	2996 ± 60	4981 ± 193
Bias (%)			
Z-score			

**determined with the robust mean at the reference date (May 1st, 2021 - Beginning of the IC campaign)*

3 - Results Task 3 'Medicine' #1#2 (P4/70kg)

Task 3	^{68}Ge/Ga	^{88}Y	^{40}K
*Reference Value (Bq)**			
Bias (%)			
Z-score			

[1]*determined with the robust mean at the reference date (May 1st, 2021 - Beginning of the IC campaign)*

[2]*determined with the robust mean at the reference date (May 1st, 2021 for ^{40}K and August 10th 2021 for other radionuclides)*

4 - Results Task 4a 'Calibration' (P4/70kg)

Task 4a	^{133}Ba	^{152}Eu	^{40}K
*Reference Value (Bq)**	20535 ± 460	25730 ± 728	3770 ± 297
Bias (%)			
Z-score			

**determined with the robust mean at the reference date (May 1st, 2021 - Beginning of the IC campaign)*

5 - Results Task 4b 'Calibration' (P5/90kg)

Task 4b	^{133}Ba	^{152}Eu	^{40}K
*Reference Value (Bq)**	25782 ± 496	32668 ± 685	4692 ± 240
Bias (%)			
Z-score			

**determined with the robust mean at the reference date (May 1st, 2021 - Beginning of the IC campaign)*

2/3

6 - Conclusion/Sum up of the results

TASK	ISO 28 218	ISO 13 528
1 'Victor'	Conform / Not Conform	
2 'Emergency'		
3 ' Medicine'		
4a 'Calibration'		
4b 'Calibration'		

Yours faithfully.

EIVIC Staff

3/3