NKS-258
ISBN 978-87-7893-330-0

Nordic proficiency test for whole body counting facilities

Lilián del Risco Norrlid (1)
Oskar Halldorsson (2)
Soren Holm (3)
Jussi Huikari (4)
Mats Isaksson (5)
Bjorn Lind (6)

1. Swedish Radiation Safety Authority, SSM
2. Icelandic Radiation Safety Authority, GR
3. NM and PET at Copenhagen's University Hospital
4. Radiation and Nuclear Safety Authority, STUK
5. Dep. Radiation Physics Sahlgren Academy at Göteborg University
6. Norwegian Radiation Protection Authority, NRPA

January 2012

Abstract

The PIANOLIB activity aims to harmonize the calibrations of the measurement equipment in the region and to evaluate the quality of this kind of measurement by means of a proficiency test exercise. In this report the results of the proficiency test are presented. The exercise consisted in determining the activity of a phantom filled with two sets of certified radioactive materials, K-40 and Cs-137, in radioactive rods, uniformly spaced inside of the phantom. Most of the participants were able to quantify correctly the activities of K-40 and Cs-137 and by comparison with the results of a previous exercise of this kind, the overall performance is equally good. The problems experienced by laboratories which submitted non-acceptable results could generally be attributed to the calibration of their systems. It is important to keep the practice of intercomparison and NKS continues to be the best framework for supporting this kind of activity.

Key words

Whole body counting, proficiency test

NKS-258
ISBN 978-87-7893-330-0

Electronic report, January 2012

NKS Secretariat
P.O. Box 49
DK - 4000 Roskilde, Denmark

Phone +45 4677 4041
www.nks.org
e-mail nks@nks.org

Nordic proficiency test for whole body counting facilities

Report from the NKS-B PIANOLIB in 2011 (Contract: AFT/B (1) 6)

Lilián del Risco Norrlid[1]
Óskar Halldórsson[2]
Soren Holm[3]
Jussi Huikari[4]
Mats Isaksson[5]
Bjorn Lindh[6]

[1]Swedish Radiation Safety Authority - SSM
[2] Icelandic Radiation Safety Authority - GR
[3] NM and PET at Copenhagen's University Hospital
[4] Radiation and Nuclear Safety Authority - STUK
[5] Dep. Radiation Physics, Sahlgren academy at Göteborg University
[6] Norwegian Radiation Protection Authority – NRPA

Table of contents

1. Introduction

Whole body counting remains one of the most important tools for internal dosimetry, offering the possibility to quantify the internally deposited radionuclides directly in a speedy way and to detect insoluble materials with long retention times.

A whole body counting system can detect levels of most gamma emitters (>200 keV) at levels far below that which would cause adverse health effects in man. A typical detection limit for radioactive caesium is around 40 Bq. The Annual Limit on Intake (based on the worker dose limit that is 20 mSv) is about 2 MBq. The amount of naturally occurring radioactive potassium, present in humans is also easily detectable.

The measurement process is, however, complex and includes many sources of uncertainty, which makes an accurate determination of the internal radioactivity quite challenging. Quality assurance in whole body measurement is therefore of particular importance.

The activity "Phantom-based Intercomparison among Nordic whole body counting facilities and the development of a Nordic Phantom Library Website" (PIANOLIB) is providing the insight for the situation regarding resources, capabilities and calibration at *in vivo* laboratories in the region. In this report, the results of a proficiency test conducted in the framework of the PIANOLIB activity are presented.

2. Objectives and scope

The exercise consisted in determining the activity of a phantom filled with two sets of certified radioactive materials in radioactive rods, uniformly spaced inside of the phantom. This mimics a homogenous activity distribution in man. One set was caesium-137 and the other potassium-40. The objective was to evaluate the measurement quality in the region.

The program included 19 facilities with their counting systems, in the five Nordic countries, that is, *in-vivo* laboratories at nuclear facilities, laboratories with responsibilities in the national emergency response and other *in-vivo* laboratories at universities and hospitals, see the report for PIANOLIB in 2010 (del Risco Norrlid L. et al., 2010).

3. Materials and methods

3.1 Carrying out the proficiency test exercise

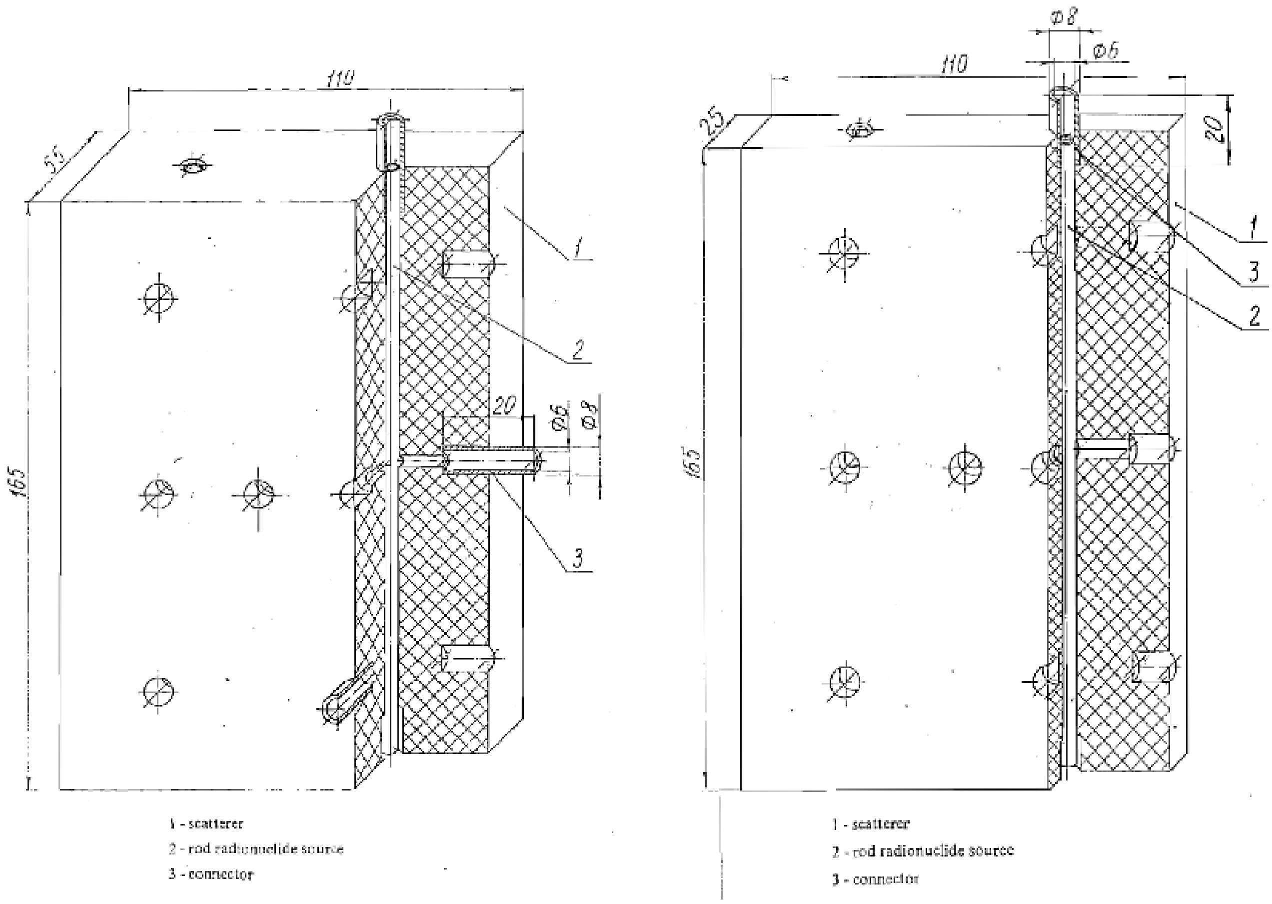
The shipping of the phantom for the PIANOLIB activity started in 2010. Five days were allowed for performing the measurements and other five days for the transport to the next laboratory. As average the phantom remained at the facilities 6.2 days and the average shipping time was 5.3 days. The transportation costs were covered by NKS in the framework of the PIANOLIB activity. The circulation of the phantom finished in June 2011 when the phantom arrived back in the Swedish Radiation Safety Authority.

3.2 The whole body phantom

The phantom IRINA (RIISH/STC, 1995) comprises tissue equivalent blocks of polyethylene and uses only solid source components, which significantly facilitates the transportation and reduces the risk of contamination during transport. The component scattering blocks of the phantom with rod radionuclide sources inserted in the through holes form a module of one-piece or half- piece radioactivity, Figures 1 a) and b).

The radioactive material in the rods is an active powder (^{40}K) or sand coated with active resin (^{137}Cs), encapsulated in plastic rods 165 mm long and 6 mm in diameter. The activities of the sets of rods have been certified by Mendeleyev Institute for Metrology (VNIIM, www.vniim.ru), which holds the primary standard for the unit of activity of radionuclide sources in the Russian Federation.

This proficiency test had as target activities those corresponding to the configuration P5 of IRINA. This size of the phantom represents a standard man of a weight of 77.8 kg. The number of blocks for configuration P5 of the phantom is given in Table 1. The activity is considered homogenously distributed when the mounted phantom is completely loaded with the radioactive rods.



a) b)

Figure 1. Module of scatter block with, a). One piece of radioactivity; b) Half piece of radioactivity.

One of the participants requested to mount the geometry P4 which is smaller than P5, because of space constrains. The weight of P4 is 61.5 kg and the number of blocks for this configuration is also given in Table 1. The certified activities for both P4 and P5 sizes of the phantom have as reference date 1996-08-01 and are given in Table 2. These values are given with a combined standard uncertainty of 5 % at the 95 % confidence level. The half-life of K-40 and Cs-137 are 1.28 Gyr and 30 yr respectively.

3.1 Data analysis and evaluation of performance

The participants were asked to perform stability measurements and activity determinations together with the corresponding measurement uncertainty, for the phantom loaded with the set of rods of Cs-137 and the set of rods of K-40. They were also asked to report the minimum detectable activity.

For stability, 5 repeated measurements had to be done without moving the phantom. A report on the gross and net counts in the ROI for K-40 was requested for each replicate measurement. The observed precision in percent was estimated as:

$$OP = \frac{\sqrt{\sum_{i=1}^{N} \frac{(C_i - M)^2}{N-1}} \cdot 100}{M}, \qquad (1)$$

where, *OP* is the observed precision; *N* is the number of repeated measurements, in this case is five; C_i is the counts observed at measurement *i*; *M* is the mean of the set of five measurements.

Table 1: Number of scatter blocks to mount the configuration P5 of the phantom IRINA

	P4 (61.5 kg) Big blocks	P4 (61.5 kg) Small blocks	P5 (77.8 kg) Big blocks	P5 (77.8 kg) Small blocks
Head & Neck	4	2	4	2
Chest	20	-	20	12
Arms	8	-	8	-
Abdomen	13	-	14	8
Thighs	14	-	14	10
Legs	10	-	12	4
Total	69	2	72	36

Table 2: Certified activities for configuration P4 and P5 of Irina phantom

Radionuclide	Activity (Bq) P4	Activity (Bq) P5
Cs-137	33900	43600
K-40	5660	7270

The *z*-score and *u*-score tests were run as biasing and significance tests. The reports are also evaluated regarding the values of maximum acceptable bias (MAB) and the limit of acceptable precision (LAP). MAB and LAP in this proficiency test exercise have been set taking into account the complexity of the measurement and the many sources of uncertainty in the process. The evaluation parameters MAB and LAP are shown in Table 3.

For the *z*-score testing each reported result is converted to the corresponding value of *z* in standard normal distribution. The standard normal distribution has a mean of 0 and a standard deviation of 1. Thus, the *z*-score is a measurement of the deviation, in standard deviation units, of the result from the true value. The *z*-score is determined as:

$$z = \frac{(x - X)}{s} , \qquad (3)$$

where, *x* is an individual reported result decay corrected to the date of the reference activities, that is, 1996-08-01, *X* is the true value, that is the certified activity and *s* is the estimate for the variation of the true value, which is its standard deviation. The z-score is interpreted as: if $|z| \leq 2$ the result is considered satisfactory, if $2 < |z| < 3$ the result is considered questionable and if $|z| \geq 3$ the result is considered unsatisfactory. Also, $2 < |z| < 3$ can be considered as a warning value, i.e. a revision of the result is desirable (Thompson M. et al., 2006).

The *u*-score is about whether the reported value is significantly different from the target at a given level of probability. The quantity *u* is compared with critical values of the t-statistics tables. The choice of significance level is 95 % and it was known that at least three measurements have been done for reporting a value as an average. That gives a table value equal to 3.18 (Taylor J.K., 1990). The *u*-score is determined as:

$$u = \frac{|x - X|}{\sqrt{\sigma_x^2 + s^2}} , \qquad (4)$$

where, *x* is an individual reported result, *X* is the true value, that is the certified activity, σ_x is the standard deviation of the result based on its reported combined standard uncertainty and *s* is the standard deviation of the true value.

Two other parameters for evaluation of the accuracy of the reported values are defined as:

$$A1 = |x - X| \text{ and}$$
$$A2 = 2.58\sqrt{unc_x^2 + unc_X^2} \qquad (5)$$

Table 3: Values for the maximum acceptable bias (MAB) and the limit of acceptable precision (LAP) applied in this proficiency test.

Radionuclide	MAB (%)	LAP (%)

Cs-137	20	20
K-40	20	20

where *unc* are the combined uncertainties at the 68 % confidence level for the report from the participant laboratory and the certified value, respectively.

Precision is evaluated according to:

$$P = \sqrt{\left(\frac{unc_x}{x}\right)^2 + \left(\frac{unc_X}{X}\right)^2} \cdot 100 \ (\%) \qquad (6)$$

The result is acceptable for trueness if $A1 \leq A2$ and acceptable for precision if $P \leq$ LAP, see Table 3. A combined evaluation of trueness and precision is done by setting the mark **A** (acceptable) if both criteria for trueness and precision are met, **N** (not acceptable) in the opposite case, when neither trueness and precision criteria are fulfilled and **W** (warning) in the case when only one of the criteria is met, given that the relative bias is below MAB. The relative bias in percent is $100 \cdot A1/X$.

The instruction for providing the minimum detectable activity (MDA) at the 95 % confidence level was to report the activity of K-40 and Cs-137 from measurements of the phantom not containing radioactive rods. The MDAs are then $MDA_{K40} \cong 3\sqrt{x0_{K40}}$ and $MDA_{Cs137} \cong 3\sqrt{x0_{Cs137}}$, where $x0$ are the "blank" activities for K-40 and Cs-137, respectively.

4. Results and discussion

4.1 Stability and precision

The repeatability tests showed good stability for all the participant's measurement systems. The observed precision (OP) has been compared to the Poisson precision (PP). The Poisson precision, in percent, can be estimated as:

$$PP = \frac{\sqrt{M} \cdot 100}{M} \qquad (7)$$

In figure 2, the ratio between the OP and PP is plotted based on the gross counts data for the K-40 photo peak. All values were normalized to a measurement time equal to 1800 seconds. The line at 1.34 is the average of the set of all the gross count reports. Unity would be the ideal value since it is well know that radioactive decay is governed by Poisson statistics and so the ideal counter should theoretically have precision close to Poisson precision. The dashed lines are the two standard deviations from the average of the set.

Most the participants submitted gross count data with precision behavior close to Poisson and within two standard deviation of the mean of the reported data.

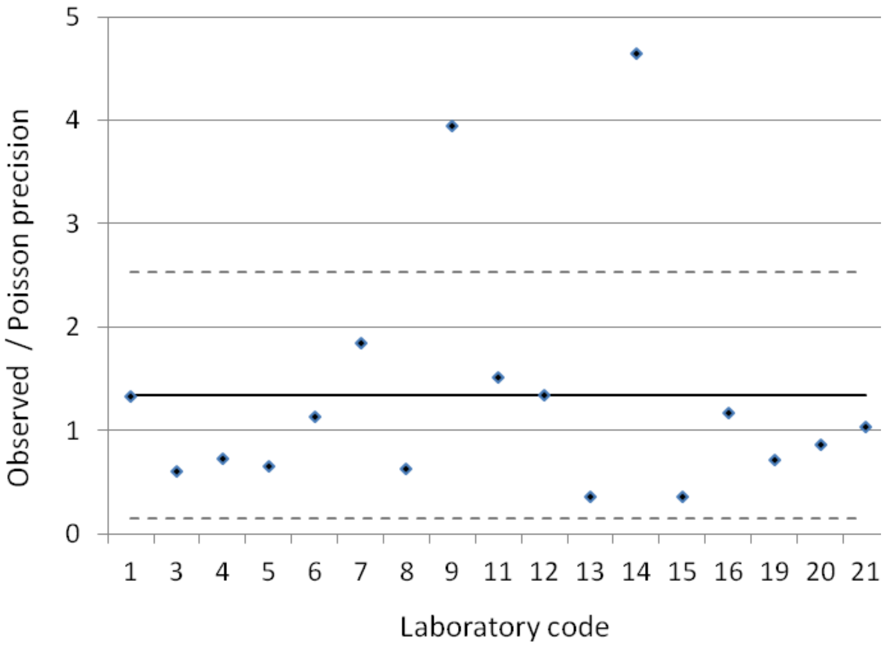


Figure 2. Ratio observed to Poisson precision

4.2 Bias and significance testing

Laboratories 7 and 8 didn't submit results for the activity of K-40. Results for the *z*- and *u*-scores are shown in figures 3 – 6. For K-40, the *z*-score testing showed three laboratories, 1, 5 and 6, with non-acceptable reports, that is, $z > |3|$ and two with questionable reports with $|2| < z < |3|$, figure 3. The *u*-testing showed, with 95 % confidence, that laboratories 5 and 6 reported activity values significantly different from the certified reference activity for K-40, figure 4.

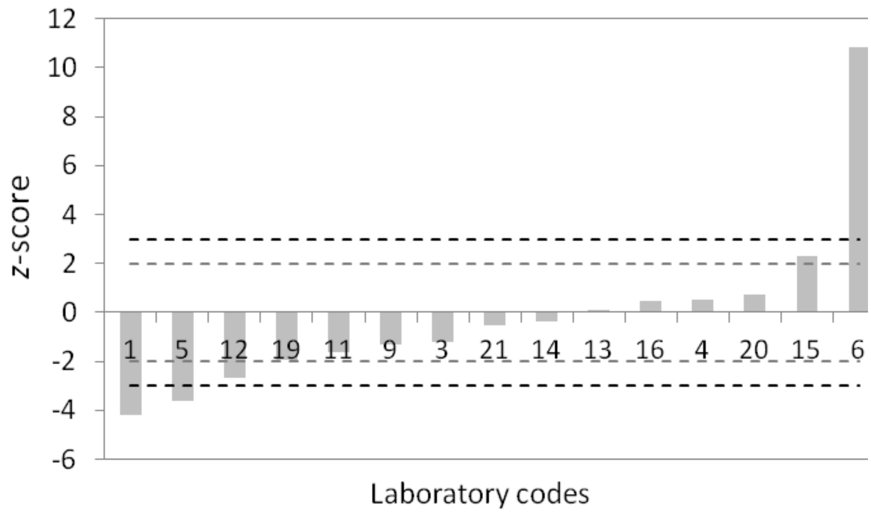


Figure 3. Biasing *z-score* testing for the reported activities of K-40.

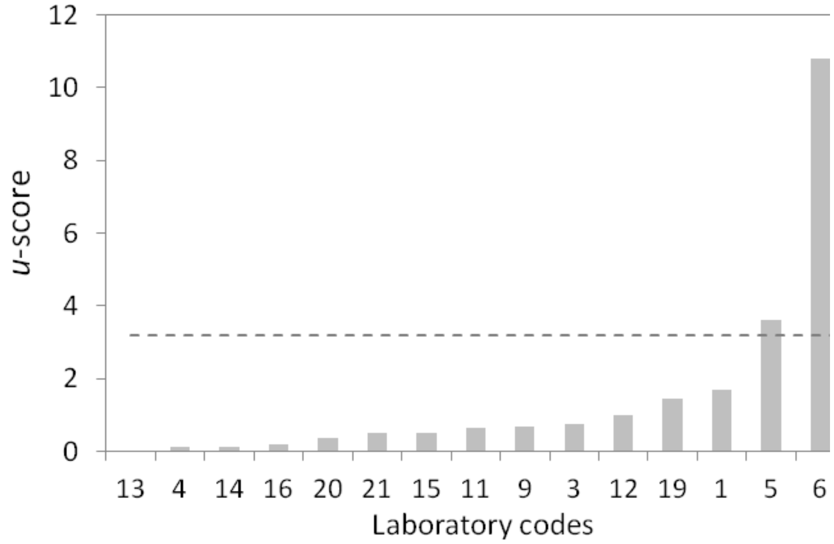


Figure 4. Significance *u*-score testing for the reported activities of K-40.

For Cs-137, the score testing showed two laboratories, 6 and 15, with $z > |3|$ and two with questionable reports with $|2| < z < |3|$, figure 5. The *u*-score testing showed also that laboratories 6 and 15 reported activity values significantly different from the certified reference value, figure 6.

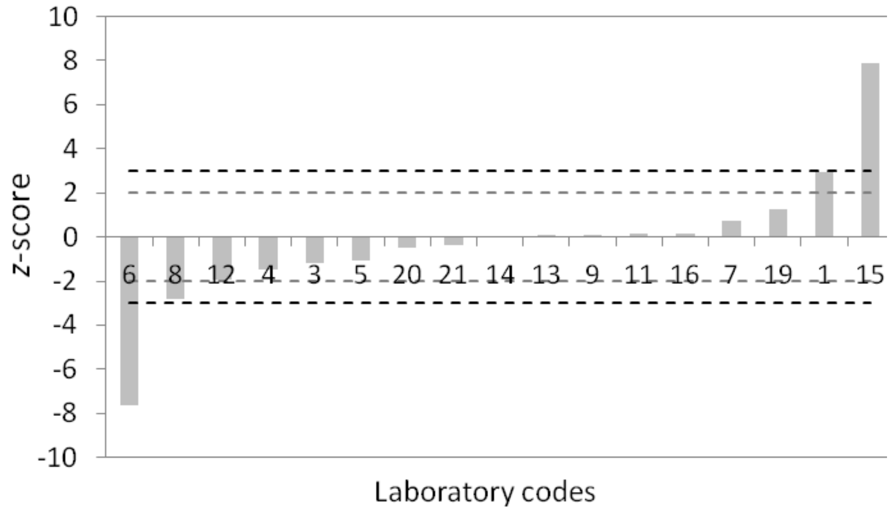


Figure 5. Biasing *z*-score testing for the reported activities of Cs-137.

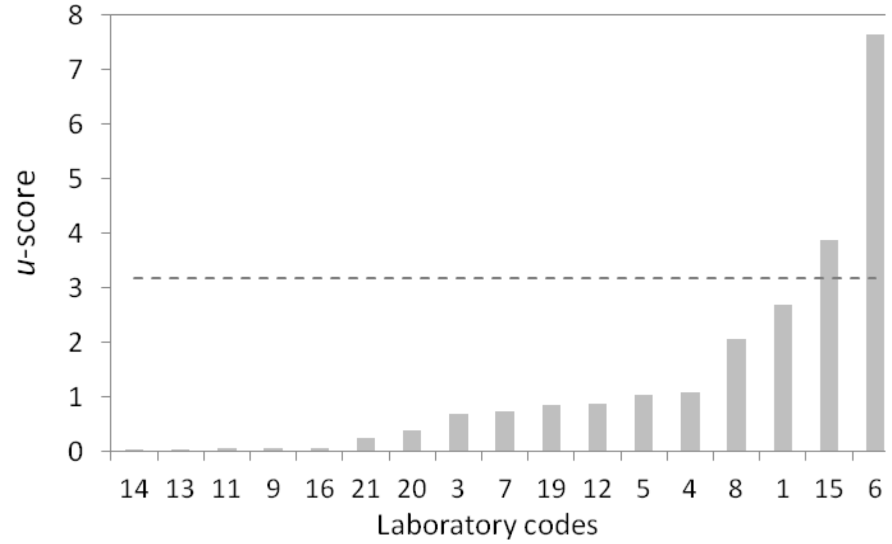


Figure 6. Significance *u*-score testing for the reported activities of Cs-137.

4.3 Combined evaluation based on accuracy and precision criteria

The plot ranking the accuracy of the reported values shows the results in percentage of the reference value in figure 7, for K-40 and in figure 8, for Cs-137. The 100 % line represents the true value. The reason for using the relative bias is that the reference activities were not the same for all participants; one laboratory used the configuration P4.

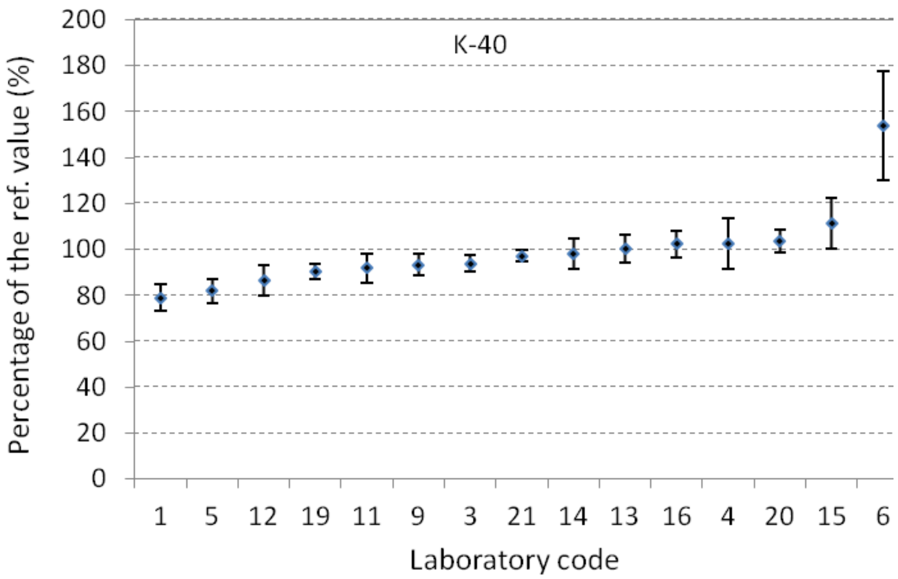


Figure 7. Relative bias for the reported values for the activity of K-40.

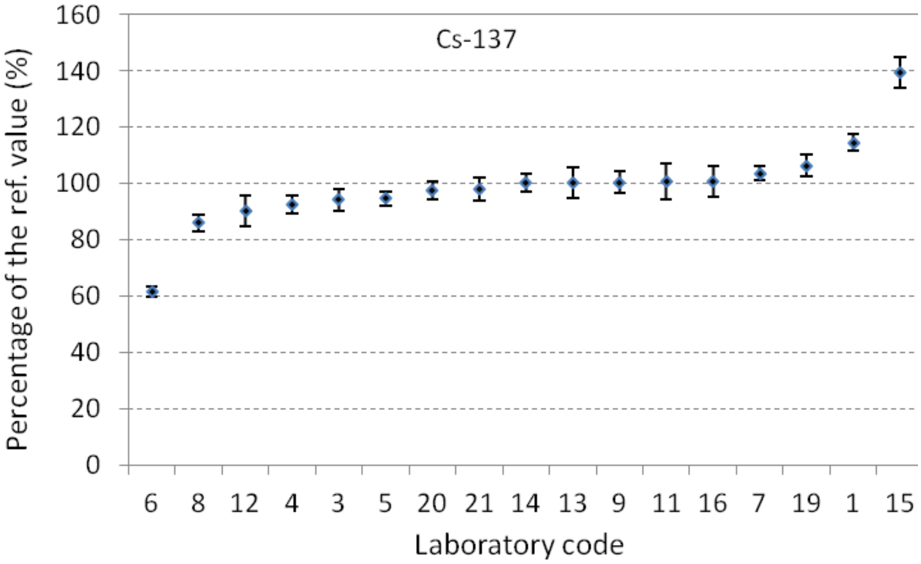


Figure 8. Relative bias for the reported values for the activity of Cs-137.

Most of the participants reported activity values within the acceptable bias. A combined evaluation is presented in Table 4, for K-40 and Table 5, for Cs-137. In these tables the reported uncertainty corresponds to the total combined uncertainty reported at one sigma, that is, with 68 % confidence. The comparison is has been performed between the reported values and the certified values, corrected for radioactive decay to the date of the measurement, for each participant laboratory.

In Table 5 that is for Cs-137, the decay corrected activity to the date of the measurement is specified. For the case of K-40, decay can be neglected because of the very long half-life of K-40.

Laboratories 5 and 6 didn't report the values for the combine standard uncertainties and therefore the evaluation was done only based in the relative bias of the reported activities with respect to the certified reference values.

The combined evaluation is consistent with the results of the tests run for significance and biasing scores (*u*- and *z*-score). Two laboratories didn't meet the reference values. Laboratories 5 and 6 reported non acceptable results for K-40 and laboratories 6 and 15 for Cs-137. Non acceptable performance are basically connected to inadequate calibration of the measurement system. That's why an obvious outcome from this proficiency test has been the recommendations to these laboratories to re-calibrate their system.

The result accepted with a warning reflects two situations. The first is a result with small measurement uncertainty or without any report about the uncertainty that has a relative bias still within the accepted interval. The second situation appears when the result is within the accepted interval for biasing and even closer to the target than in the first situation but the reported uncertainty is large. The laboratories giving acceptable reports but with a warning, were labs 1 and 4 for K-40, and labs 5 and 12 for Cs-137. These laboratories should review the uncertainty budget and eventually engage in re-calibration.

Table 4: Combined performance evaluation for the determination of K-40.

Laboratory code	Reported value Act. (Bq)	unc (Bq)	Performance score for K-40
1	5744	821	AW
3	6830	440	A
4	7464	1573	AW
5*	4640	-	NA
6	11200	-	NA
7**	-	-	-
8**	-	-	-
9	6790	600	A
11	6670	860	A
12	6300	900	A
13	7300	800	A
14	7140	910	A
15	8100	1563	A
16	7444	772	A
19	6560	330	A
20	7530	600	A
21	7070	150	A

* Laboratory 5 used Irina geometry P4

** Laboratories 7 and 8 didn't participate in the determination of K-40

Table 5: Combined performance evaluation for the determination of Cs-137

Laboratory code	Reported value		Target values decay corrected		Performance score for Cs-137
	Act. (Bq)	unc (Bq)	Act. (Bq)	unc (Bq)	
1	36452	696	31796	1590	A
3	29900	2100	31738	1587	A
4	29368	1482	31710	1586	A
5*	23300	-	24579	1229	AW
6	19500	-	31558	1578	NA
7	32700	1000	31530	1576	A
8	27100	1450	31502	1575	A
9	31600	1900	31446	1572	A
11	31600	3600	31385	1569	A
12	28200	3100	31268	1563	AW
13	31400	3100	31268	1563	A
14	31300	1300	31235	1562	A
15	43520	2770	31205	1560	NA
16	31426	3143	31178	1559	A
19	33070	1700	31083	1554	A
20	30220	1200	30978	1549	A
21	30800	1900	31417	1571	A

* Laboratory 5 used Irina geometry P4

4.4 Detection limits

The detection limits for the determination of K-40 ranges from 40–10000 Bq and for Cs-137 from 20–900 Bq. That a larger counting time gives lower MDA is the trend verified for laboratories with similar measurement geometry and detectors. In general, it was observed that the laboratories equipped with NaI detectors or combined measurement systems with NaI together with plastic scintillators and/or HPGe detectors showed the lowest MDAs. Also, the systems with more than one detector showed, in general, smaller MDA values than the systems with only one detector. In figure 9 and 10, the MDA reports are summarized for K-40 and Cs-137 respectively.

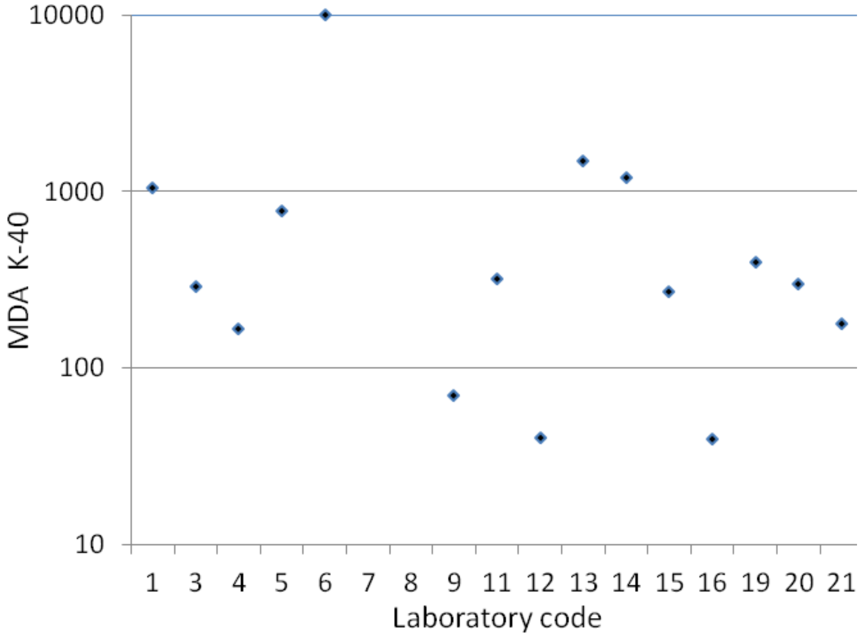


Figure 9. Minimum detectable activity reported for K-40 from all participants (labs 7 and 8 didn’t participate in K-40 determination).

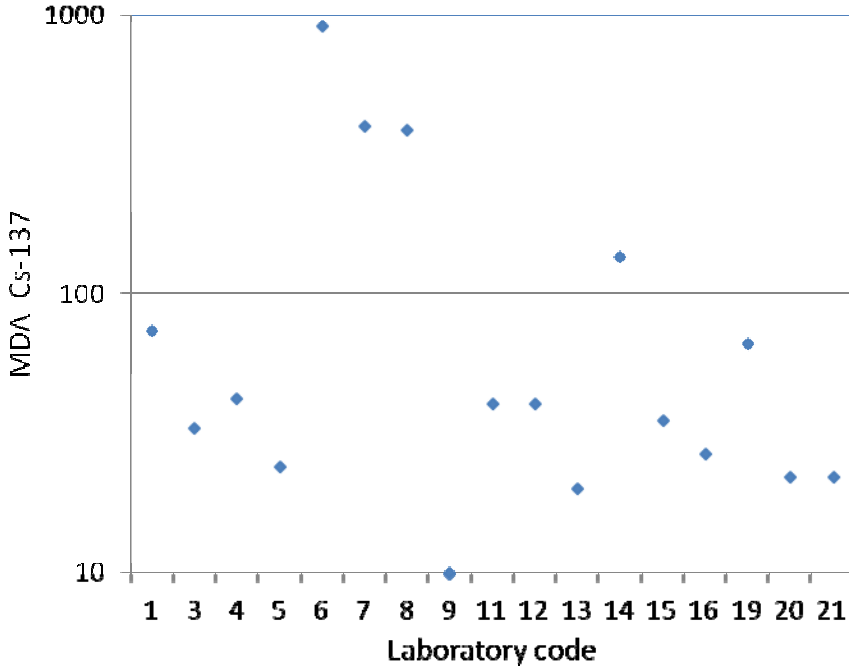


Figure 10. Minimum detectable activity reported for Cs-137 from all participants.

5. Conclusions

The proficiency test in the framework of the PIANOLIB activity was successfully completed with a satisfactory level of participation. From a total of 21 measurements systems registered from nineteen laboratories, we received reports from 17 measurement systems in fifteen laboratories, that is, 81 %. All participating laboratories were able to quantify the activity of Cs-137 and 15 quantified the activity of K-40. It should be pointed out that the target activity for K-40 was as low as encountered naturally in the human body, which is challenging.

Most of the participants were able to quantify correctly the activities of K-40 and Cs-137 and by comparison with the results of a previous exercise of this kind performed in 2006 (Rahola T. et al., 2006) the overall performance is equally good. The problems experienced by laboratories which submitted non-acceptable results could generally be attributed to the calibration of their systems.

The results of the proficiency test were presented and discussed in a Workshop organized in the framework of the activity with the consequent benefit for all participants of PIANOLIB of getting insights and exchange of experiences. The workshop was held in Gothenburg University, September 15th and 16th 2011. All the presentations of the PIANOLIB Workshop are available at the NKS website.

Acknowledgements

We like to thank all the participant laboratories for the cooperation during the activity. Our acknowledgement to Mats Isaksson at Gothenburg University for hosting the final workshop of this activity and to the workshop's speakers.

NKS conveys its gratitude to all organizations and persons who by means of financial support or contributions in kind have made the work presented in this report possible.

References

Del Risco Norrlid L. et al., 2010, “In-vivo whole body measurement of internal radioactivity in the Nordic countries”, NKS-238, ISBN 978-87-7893-310-2, Roskilde, Denmark

Rahola T., Falk R. and Tillander M., 1994, “Intercalibration of whole-body counting systems”, In: Ed. H. Dahlgaared, Nordic Radioecology, *The transfer of radionuclides through Nordic ecosystems to man*, 407 – 424.

Rahola T. et al, 2006, “Intercomparison exercise for whole body measurements in the Nordic countries”, In: Ed NKS, *Assessment of Internal Doses in Emergency Situations*, NKS-128, ISBN 87-7893-190-8 Roskilde, Denmark.

RIISH/STC., 1995, Technical documents for human whole body phantom with reference samples of radionuclides potassium-40, cobalt-60, cesium-137; Set UPH-07T, Saint Petersburg, Russia.

Thompson M., Ellison S.R. and Wood R., 2006, “The international harmonizes protocol for the proficiency testing of analytical chemistry laboratories”, *Pure Appl. Chem.* (78), 145-196.

Taylor J.K., 1990, *Statistical techniques for data analysis*, Lewis Publishers Inc., Chelsea.

Title	Nordic proficiency test for whole body counting facilities
Author(s)	Lilián del Risco Norrlid (1), Oskar Halldorsson (2), Soren Holm (3), Jussi Huikari (4), Mats Isaksson (5), Bjorn Lind (6)
Affiliation(s)	1. Swedish Radiation Safety Authority, SSM 2. Icelandic Radiation Safety Authority, GR 3. NM and PET at Copenhagen's University Hospital 4. Radiation and Nuclear Safety Authority, STUK 5. Dep. Radiation Physics Sahlgren Academy at Göteborg University 6. Norwegian Radiation Protection Authority, NRPA
ISBN	978-87-7893-330-0
Date	January 2012
Project	NKS-B / PIANOLIB
No. of pages	17
No. of tables	5
No. of illustrations	10
No. of references	6
Abstract	The PIANOLIB activity aims to harmonize the calibrations of the measurement equipment in the region and to evaluate the quality of this kind of measurement by means of a proficiency test exercise. In this report the results of the proficiency test are presented. The exercise consisted in determining the activity of a phantom filled with two sets of certified radioactive materials, K-40 and Cs-137, in radioactive rods, uniformly spaced inside of the phantom. Most of the participants were able to quantify correctly the activities of K-40 and Cs-137 and by comparison with the results of a previous exercise of this kind, the overall performance is equally good. The problems experienced by laboratories which submitted non-acceptable results could generally be attributed to the calibration of their systems. It is important to keep the practice of intercomparison and NKS continues to be the best framework for supporting this kind of activity.
Key words	Whole body counting, proficiency test

Available on request from the NKS Secretariat, P.O.Box 49, DK-4000 Roskilde, Denmark.
Phone (+45) 4677 4041, e-mail nks@nks.org, www.nks.org