
This is the **accepted version** of the journal article:

González Mesa, Jorge Ernesto; Holladay, Bret; Higuera Hernández, Manuel; [et al.]. «Assessment methods for inter-laboratory comparisons of the dicentric assay». *International Journal of Radiation Biology*, Published online: 20 Jul 2022. 8 pàg. DOI 10.1080/09553002.2022.2094021

This version is available at <https://ddd.uab.cat/record/265281>

under the terms of the  ^{IN}
COPYRIGHT license

Title: Assessment methods for interlaboratory comparisons of the dicentric assay

Authors: Jorge E. Gonzalez^a, Bert A. Holladay^b, Manuel Higuera^{c,d}, Marina Di Giorgio^e, Joan F. Barquinero^f

Affiliations:

^a Centro de Protección e Higiene de las Radiaciones, CP 11300, Playa, La Habana, Cuba

^b Statistics Department, California Polytechnic State University, San Luis Obispo, USA

^c Scientific Computation & Technological Innovation Center (SCoTIC), Universidad de La Rioja, Logroño, Spain

^d Departamento de Matemáticas y Computación, Universidad de La Rioja, Logroño, Spain

^e Autoridad Regulatoria Nuclear, Buenos Aires, Argentina

^f Departamento de Biología Animal, Biología Vegetal y Ecología, Universitat Autònoma de Barcelona, Barcelona, Spain

Corresponding author: Jorge E Gonzalez. cphrjorge@ceniai.inf.cu

Keywords: Dicentric assay, laboratory comparisons, robust algorithms, Shiny application.

Biographical notes

Jorge E. Gonzalez, pharmacist, is the research director at the Center for Radiation Protection and Hygiene (CPHR) of Cuba.

Bret A. Holladay, PhD, is an assistant professor in the Statistics Department at California Polytechnic State University, San Luis Obispo, USA.

Marina Di Giorgio, biologist, is the head and technical director of the biological dosimetry, radiopathology and internal dosimetry laboratories at the Nuclear Regulatory Authority (ARN) of Argentina.

Manuel Higuera, PhD, is an associate professor within the Mathematics and Computation Department at the La Rioja University

Joan F. Barquinero, PhD, is a Professor (Docent) of Human Genetics and Senior Researcher at the Department of Animal Biology, Plant Biology and Ecology, Faculty of Biosciences, Autonomous University of Barcelona, Cerdanyola del Valles, Spain.

ABSTRACT

Purpose:

To test the performance of different algorithms that can be used in interlaboratory comparisons based on dicentric chromosome analysis, and to evaluate the impact of considering *a priori* values different to calculate individual laboratory performance based on the ionizing radiation dose estimation.

Methods:

Mean and standard deviation estimations in inter-laboratory comparisons are tested on simulated data and data from previously published inter-laboratory comparisons using three robust algorithms, algorithm A, Algorithm B and Q/Hampel, all programmed in R-project language and implemented in a Shiny application. The simulated data were generated assuming three different probabilities to contaminate inter-laboratory comparisons samples with atypical dose values. Comparison between different algorithms was also done using published exercises where blood samples were irradiated at 0 and 0.7 Gy that represent a challenge for the assessment of an inter-laboratory comparison.

Results:

The best performance was obtained with the Q/Hampel algorithm for the estimation of the dose mean and with the algorithm B for the estimation of the dose standard deviation under the conditions tested in the simulations. The Q/Hampel algorithm showed the best performance when non-irradiated samples were evaluated and there was a high proportion of identical values. The presence identical values causes the Algorithm B to fail. Real examples illustrating the need to consider standard deviation priors, and the need to use algorithms resistant to a high proportion of identical values are presented.

Conclusions:

Q/Hampel algorithm is a serious candidate to estimate the dose mean in the inter-laboratory comparisons, and to estimate both parameters when the proportion of identical values equals or higher than the half of the results. When the proportion of identical values is less than the half of the

results, the Algorithm B should be considered as a candidate to estimate the standard deviation in the inter-laboratory comparisons with small number of laboratories. We remark that special attention is needed to establish prior definitions of standard deviation in the assessment of inter-laboratory dicentric assay comparisons.

1. Introduction

The main goals of an inter-comparison among laboratories is to give an independent view of laboratory performance and laboratory training. The successful participation of a laboratory in inter-laboratory comparisons (ILCs) is a traceable manner to demonstrate competence to a customer or a regulatory body (Koch 2004). Biological dosimetry laboratories have conducted dicentric assay inter-laboratory comparison (ILC) exercises since the 80s (Lloyd *et al.* 1987). During the last decade, there has been a noticeable increase in the number of ILCs (Di Giorgio *et al.* 2011; Wilkins *et al.* 2015; Kulka *et al.* 2017; Gregorie *et al.* 2021). The comparisons may involve the distribution of blood samples (Lloyd *et al.* 1987; Wilkins *et al.* 2015; Gregorie *et al.* 2021), slides (Garcia *et al.* 1995; Di Giorgio *et al.* 2011), or metaphase images (Livingston *et al.* 2011; Garcia *et al.* 2013; Romm *et al.* 2014). Depending on what is distributed among laboratories, different putative sources of uncertainties can be considered (IAEA, 2011). The threat of the malevolent use of a radiation device or other catastrophic event with a large number of potentially exposed people, and the low number of people that can be simultaneously analyzed in a single laboratory, have stimulated the creation of laboratory networks that can mutually assist. However, before any joint work, participant laboratories need to achieve a similar accuracy and precision in dose-assessment.

In biodosimetry ILCs to determine the performance of participants requires to have a value against which to compare the results of the participants. Most of the ILCs compare the participant's deviation from this assigned value with a numerical criterion which is used to decide whether the deviation is a cause of concern or not. For biodosimetry proficiency tests, the assigned value can be a consensus value or a reference value:

- a) **Consensus value from participant results:** the assigned value for the proficiency test is the location estimate (*e.g.*, mean from a robust algorithm, median, or arithmetic mean) from the results reported by participants.

b) **Reference value:** the assigned value is the delivered dose determined by a Primary or Secondary Standards Dosimetry Laboratory, where the individual dose-assessment reported by each participant laboratory is compared to the delivered dose, which is considered as the true location parameter.

In general, choosing assigned values independently of participant results offer advantages. An ILC also needs to define a dispersion or standard deviation value, and several criteria have been used. The dispersion can be a value defined by experts or can be issued from the ongoing or from previous ILCs. In the latter case, arithmetic or robust algorithms to calculate the standard deviation can be used.

<< table 1 >>

Currently, to reduce the influence of outliers, robust methods described in ISO 5725-5: 1998, ISO 13528: 2015 and ISO 17043: 2010 are selected for the purpose of inter-laboratory comparison under quality systems. Robust Algorithm A is applied to obtain the mean (x^*) of the results reported by the participants (x_i) and the standard deviation (s^*). The parameters obtained from robust algorithms are notated with an asterisks (*) symbol.

In most of the dicentric chromosome based ILCs, the algorithm A has been used (Di Giorgio *et al.* 2011, Romm *et al.* 2014; Gregoire *et al.* 2021). However, this algorithm faces convergence problems when identical values are reported. To determine the performance of the laboratories, the z-score test is applied to the data reported.

The assigned value and the dispersion parameter(s) are then used to evaluate the performance of each participant laboratory using a z-score test, which normalizes each laboratory estimated dose. Depending on the z-score value obtained, the estimated doses by each participant laboratory are then classified as “satisfactory”, “questionable” or “unsatisfactory”. Note-worthily, the z-score obtained not only depends on the reported doses, but also on how it is calculated, in this case the standard deviation greatly influences the result, and a simplistic interpretation of the z-score value can be misleading (Gregoire *et al.* 2021).

In this article, we present a simulation study of the performance of several robust algorithms presented in the ISO 13258 and DIN 38402-45 standards for ILCs based on dicentric chromosome assay. Particularly, we present the Q/Hampel algorithm, treated in depth in the ISO 13258 and DIN 38402-45 standards, that uses a three-parameter weighing function. The Q/Hampel algorithm has the characteristic to be resistant to a higher proportion of atypical and identical values. We also include an algorithm denominated as algorithm B that uses a logistic weighing function which is recommended in the ISO 13258 for comparisons with a small number of laboratories. We also suggest some solutions that can be implemented to solve the convergence problems. Additionally, we tested the performance of these algorithms with previous published ILCs based on dicentric assay.

2. Methods

All calculations of the present study were performed using the R-project programming language software, version 4 (R Core Team 2021).

2.1. Algorithms for the mean and standard deviation estimation.

For the arithmetic algorithm the *mean* and *sd* functions from R-project were used. Moreover, three robust algorithms were evaluated: **a)** the Algorithm A that applies the *hubers* function from the *MASS* (Venables and Ripley, 2002) R-project package (for cases with a majority presence of identical values a non-iterative solution was included to calculate it); **b)** the Algorithm B (Rousseeuw and Verboven 2002; Szewczak and Bondarzewski 2016), also named as the logistic M-estimator in section D.1.4.2 of the ISO 13528; and **c)** the Q/Hampel algorithm, described in detail in the section C5 of the ISO/IEC 13528 and in the German DIN 38402-45. The Shiny (Chang *et al.* 2021) application *InterLabComparison* which obtains the results of these algorithms from any ILC data is available at <http://shinur.unirioja.es/apps/InterLabComparison/> and <https://manu2h.shinyapps.io/InterLabComparison/>

2.2. Simulation study of arithmetic and robust algorithms performance.

Simulated ILC datasets were generated with the R function *rnorm.contam* (implemented in supplementary file **s1**). This function generates n values, that represent the doses reported by n laboratories, where each dose has the same defined probability to be a contaminant or outlier value. As example, table 2 shows three simulated datasets that were generated mimicking a real ILC that involved 10 laboratories, and where the blood was irradiated at 2.34 Gy (Lloyd *et al.* 1987).

<< insert table 2 >>

To evaluate the performance of the different algorithms indicated above, 1000 synthetic ILCs were generated. Then, the values obtained in each synthetic ILC were used to calculate the location (x^* or *mean*) and standard deviation (s^* or *sd*) parameters by the different algorithms. The distance of the obtained results with respect to the expected dose *mean* and *sd* were calculated for each simulated ILC and tabulated to select the closest result to the expected dose *mean* and *sd*.

2.3. Real example datasets and the z score calculation.

The real examples of ILCs selected are the 0.7 Gy data published by Lloyd *et al.* 1987, and the 0 Gy data published by Gregoire *et al.* 2021. These examples represent a challenge for the assessment of an inter-laboratory comparison. The 0 Gy example is representative of comparison with a high proportion of identical values (e.g., 0 values), and the 0.7 Gy example present an inter-laboratory comparison where all participant laboratories overestimate the delivered dose. The statistic to compare the different algorithms was the z score statistic. The z score is a standardized measure of performance, calculated using the participant result x , the assigned value (x_{pt}) and the standard deviation (σ_{pt}) (ISO 13528). The general expression to obtain a z score is:

$$z = (x - x_{pt})/\sigma_{pt}$$

When robust algorithms are applied the notation x_{pt} and σ_{pt} are substituted for x^* and s^* . When prior fitness for purpose values is defined by the experts, the notations x_{ffp} and s_{ffp} are used.

Three different expert criteria have been evaluated:

- a) Defining a maximum permissible error of 30% of the physical dose. The s_{ffp} value can be calculated dividing the permissible deviation by 3.0, which corresponds to three standard deviations from the physical dose. The 3.0 factor is consequent with the criteria that consider as unsatisfactory result a z value higher than 3. Consequently, if we define a prior permissible deviation of 30% from the physical dose then a s_{ffp} value of 10% from the physical dose is obtained. For example, if the physical dose is 2 Gy, the s_{ffp} will be 0.2 Gy.
- b) Defining a maximum permissible error of 0.5 Gy for delivered doses lower than 3 Gy, and 1 Gy from doses equal or higher than 3 Gy. Considering the deviation as equivalent to three standard deviations from the physical dose, this can be calculated dividing the permissible deviation 0.5 Gy or 1 Gy by a 3.0 factor. Consequently, if we define a permissible deviation of 1 Gy from the physical dose, then the s_{ffp} value is 0.333 Gy. Again this, 3.0 factor is consequent with the criteria in considering z values higher than 3 as unsatisfactory.
- c) Defining as the maximum permissible error based on the Poisson distribution (ISO/IEC 13528; Gregoire *et al.* 2021). This can be estimated by using a selected calibration curve (*e.g.*, the one indicated in the IAEA, 2011) to obtain a prior yield of dicentric expected from the delivered dose proposed to be tested in the ILC. With this yield and a concrete number of cells from the culture, the *rpois* function generates a Poisson sample, the sample is splitted in *n.samp* subsamples simulating that all possible subsamples are scored. Then, the yield of each subsample is converted to dose using the same calibration curve that was selected to obtain the prior yield. The standard deviation of this dose is then calculated. Limits that contain *e.g.*, 99.7% of the data can be obtained using the sorted results with the floor and ceiling R functions. As an example, with the following inputs, a prior yield of 0.00128, 1000 scored cells and curve parameters ($C=0.00128$, $\alpha=0.02103$, $\beta=0.06307$), a s_{ffp} value of 0.03 Gy is obtained. A yield of 0.001 represents a non-irradiated sample so laboratories reporting dose values higher than

$3 \times 0.03 = 0.09$ Gy were considered atypical results that require an action signal. The code for this approach is presented in supplementary file **s2**.

3. Results and Discussion

3.1 A simulation study with robust algorithms

The results simulating an ILC where the Expected Dose is 2.34 Gy and the *sd* of the results obtained by the laboratories is the 10% of the Expected dose, in this case 0.23 Gy, are shown in table 3. In the synthetic samples contaminated with almost no atypical values (*e.g.* 0.0001), the arithmetic algorithm shows the best performance, with the closest estimation to the Expected Dose and *sd*. This result agrees with the statement that the best estimation in the absence of atypical values is obtained by the arithmetic algorithm (ISO/IEC 13258). In the contaminated synthetic samples with 0.1 proportion at the lower tail the closest results to 2.34 Gy are obtained with the robust algorithms, especially with the Hampel estimate for the location (x^*) and algorithm B for scale (s^*).

<<table 3>>

The appearance of atypical results in one tail have occurred in some ILCs (*e.g.*, Lloyd *et al.* 1987), and in some cases a solution was the exclusion of the atypical value from the arithmetic when the data were summarized. The use of robust algorithms gives the opportunity to summarize the data without the exclusion of any atypical value. Dispersion increases notably when samples are contaminated at a 0.1 proportion. This affects specially the results obtained using the arithmetic algorithm. The algorithm B shows the best performance in the tested conditions. This algorithm was proposed by Rousseeuw and Verboven in 2002, and it is recommended for estimation in very small samples in the ISO 13258 standard. The algorithm B differs from iterative algorithm A and Q/Hampel by the logistic weighting function. There is a consensus that introducing smoothing functions improve the properties of robust M-estimators (Rousseeuw and Verboven 2002; Hampel *et al.* 2011). Specifically, the introduction of an algorithm with a logarithmic weighting function is

recommended in the ISO 13258 standard in cases of comparisons involving 4 or more laboratories (ISO 13258). The results of this simulation support the use of algorithm B over A to estimate the dispersion parameter in the case of regional comparisons with a small number of experienced laboratories.

The results simulating an ILC where a non-irradiated sample is evaluated, that considers a count of 1 dicentric in 1000 cells (IAEA 2011) are shown in Table 4. The *sd* value selected was obtained by simulation considering Poisson uncertainty. The value selected contain the 95% of the simulated rounds performed as described in section 2.3. The resultant dose and *sd* were 0.01948 and 0.0051 Gy. In the synthetic samples contaminated with almost no atypical values (*e.g.*, 0.0001), again the arithmetic algorithm shows the best performance as expected. The two simulations confirm that with almost no presence of atypical values, the mean obtained by the Algorithm A is the closest to the arithmetic mean that in this case is the more efficient estimator (ISO 13258). However, when the synthetic non-irradiated samples were contaminated with a 0.15 proportion (*e.g.* 6 of 39 laboratories) of atypical values with reported estimated doses higher than 0.1 Gy the Q/Hampel algorithm shows the best performance, with the closest estimation to the expected *mean* and the Algorithm B for the *sd* as in the first simulation performed. Inclusion of non-irradiated samples is quite important in ILCs because of their complexity and the probability that a laboratory receives samples close to the background is high (Sasaki *et al.* 2001; Sun *et al.* 2016).

<< table 4 >>

3.2 Algorithms performance using data from real ILC examples.

The performance of several algorithms with real ILC examples is presented in Table 5. The first example is from sample B that corresponds to a non-irradiated blood of a recent ILC (Gregorie *et al.* 2021). This dataset allows to test the performance of the robust algorithms in the presence of

several atypical values, some estimated doses higher than 0.1 Gy and with the additional complexity of having more than 50% of the identical reported values. The Algorithms A and B with iterative solutions are not designed to resist 50% or more identical values. It is also reported that the best performance of M-estimators, like algorithms A and B, is reached with fully iterative solutions (Rousseeuw and Verboven 2002), but as we appreciate in this example, they have limitations with a high proportion of identical values when the median and the median distance to the median (also known as MAD) is zero. On the other hand, the Q/Hampel algorithms and a non-iterative solution of Algorithm A are not affected when the median and the MAD are 0. The results in Table 5 show that the Q/Hampel algorithm resists the presence of the values higher than 0.1 Gy in a higher grade than A and arithmetic algorithms. The impact of the presence of some atypical values is evident when using the arithmetic algorithm that results in a *sd* estimation three times higher than the one obtained by Q/Hampel algorithm, and two times higher with respect to the one obtained using the algorithm A (Table 5).

<< table 5 >>

Table 6 presents a comparison of the *z* scores calculated with a prior mean, the physical dose, and a prior *sd* value. The last one, derived from a simulation considering the Poisson uncertainty (see section 2.3), with *z*-scores that were calculated with x^* and s^* values obtained by two robust algorithms, A and Q/Hampel, and the arithmetic mean and *sd*. The more sensitive approaches for assessment, those where estimated doses higher than 0.10 Gy obtained an unsatisfactory *z*-score, were the prior parameters and the Q/Hampel algorithm. This result agrees with the result observed in the simulation performed in 3.1.

<< table 6 >>

The analysis of non-irradiated samples is an important part of the routine analysis of biological dosimetry laboratories. Historically, an important proportion of the samples evaluated in a laboratory of biological dosimetry have been from people who have claimed to have been exposed to ionizing radiation, but without clear evidence, or from exposed workers who have hypothetically

received overexposure to low doses. In both cases the dicentric count will be close to that shown by unirradiated samples (Sasaki *et al.* 2001; Sun *et al.* 2016). So, the complexity in their assessment in the ILCs should not justify the exclusion of this type of samples in their planning. Statistical methods that deal with their complexity are necessary and the results with the sample B from Gregoire *et al.* 2021 support the Q/Hampel algorithm as a candidate to solve the problem.

Another example involves the analysis by 10 laboratories of a sample irradiated at 0.7 Gy (Lloyd *et al.* 1987). This ILC represents the problem of a common bias, an overestimation reported by all the laboratories. In this case the x^* obtained by the algorithms are between 0.87 and 0.88 Gy, far from the actual 0.7 Gy dose. Similarly, all the robust algorithms report a biased s^* because all the data in the sample are grouped around a biased value in the upper side of the distribution. This particular example justifies the application of a prior assigned value for the dose, as well as a prior defined s value. Table 7 presents the different results for the z-score when different algorithms are applied. This would be a representative example of how to detect objectively atypical values.

<< table 7 >>

3.3 Examples of prior s values for the assessment of ILC data.

The general consensus is that the physical dose should be the location parameter for assessment (Lloyd *et al.* 1987; Di Giorgio *et al.* 2011). However, the selection of standard deviation for the assessment is not an easy task (ISO/IEC 13528). Defining fitness for purpose deviations from the physical dose could be considered a key feature in the dicentric assay for future ILCs. In the case of a network of laboratories, and once several ILCs data are accumulated, it will be possible to compare prior s values with s^* . Then, robust estimation will be a permanent tool to summarize the data in ILCs.

<< table 8 >>

Table 8 summarizes examples of approaches that can be adopted. In fact, several approaches for the selection of *a priori* accuracy limits that derive in a dispersion parameter have been applied in past

comparisons or reviewed (Lloyd *et al.* 1987; ICRU 2019; Pan *et al.* 2019; Endesfelder *et al.* 2021). The Poisson uncertainty approach is suggested in the ISO standard for Poisson distributed data and it seems reasonable to be adopted for non-irradiated samples and for doses lower than 1 Gy. Poisson uncertainty predominates for doses lower than 1 Gy (Ainsbury *et al.* 2017; Gonzalez *et al.* 2020). Compared with the Poisson based uncertainty the 0.5 and 1 Gy accuracy limits appear too wide for the case of conventional scoring for doses lower than 1 Gy. On the other hand, the 0.5 Gy limit agrees with the 30% delivered dose deviation criteria for doses higher than 1 Gy but the Poisson approach appears too low in this range of doses (Table 8). This 0.5 Gy limit has been used in several triage comparisons and for conventional scoring (Endesfelder *et al.* 2021). It seems reasonable to adopt the 0.5 Gy limits for a range between 1 and 3 Gy. For doses greater than 3 Gy, the 1 Gy limit closely agrees with the 30% delivered dose deviation criteria and the rest of approaches seem too stringent. There are a number of unexplored scenarios that can be studied through simulation. For example, comparisons in which higher standard deviations are expected as in the simulation of non-homogeneous or heterogeneous irradiation scenarios. For these particular cases modifications of present algorithms including the uncertainty of the estimations are recommended (Uhlig, 2018). On the other hand, although the work focuses on the dicentric assay, the algorithms have the potential to be applicable to other assays used in ILCs.

4. Conclusions

As a result of the simulations performed when atypical results are expected, and under the conditions tested, the Q/Hampel algorithm seems a notable candidate to be applied for mean parameter estimation in the ILCs by dicentric assay. The Algorithm B should be considered as a candidate to estimate the standard deviation in the ILCs when the proportion of identical values is less than the half of the results. Without the presence of atypical values, the Algorithm A is the closest to the arithmetic. It is recommended that prior s values are defined before starting an ILC, these prior s values should consider the dose to be assessed. When prior s values are used for the

assessment, the robust algorithms are still necessary to summarize the data minimizing the influence of atypical values.

Acknowledgements

The authors wish to thank Eric Gregoire for the data analyzed in Table 6. We also thank the reviewer for their valuable suggestions. The opinions and assertions expressed herein are those of the authors and do not necessarily reflect the official policy or position of their institutions.

Disclosure statement

No conflict of interest was reported by the authors.

References

Ainsbury EA, Higuera M, Puig P, Einbeck J, Samaga D, Barquinero JF, Barrios L, Brzozowska B, Fattibene P, Gregoire E, et al. 2017. Uncertainty of fast biological radiation dose assessment for emergency response scenarios. *Int J Radiat Biol.* 93(1):127–135.

AMC. 1989. Analytical Methods Committee. Robust Statistics Part 2: Interlaboratory Trials. *Analyst*, 114 1699-1702.

AMC. 2015. TB No 68. Analytical Methods Committee. Fitness for purpose: the key feature in analytical proficiency testing. *Anal. Methods* 7: 7404

AMC. 2016. TB No 74. Analytical Methods Committee. z-Scores and other scores in chemical proficiency testing—their meanings, and some common misconceptions. *Anal. Methods* 8: 5553.

Beninson, DJ., Placer, A., Vander Elst, E. 1969. Study of a case involving accidental irradiation of a human being. In IAEA, *Handling of Radiation Accidents. Proceedings of a symposium organized by the IAEA in collaboration with the WHO, Vienna, 19-23 May 1969.* Vienna: IAEA, p. 415-429.

Chang W, Cheng J, Allaire JJ, Sievert C, Schloerke B, Xie Y, Allen J, McPherson J, Dipert A, and Borges B. 2021. shiny: Web Application Framework for R. R package version 1.7.0. <https://CRAN.R-project.org/package=shiny>

Di Giorgio M, Barquinero JF, Vallerga MB, Radl A, Taja AR, Seoane A, De Luca J, Oliveira MS, Valdivia P, Garcia O, Lamadrid A, Gonzalez JE, Romero I, Mandina T, Carvajal YC, Maldonado CA, Espinoza ME, Martinez-Lopez W, Mendez-Acuna L, Di Tomaso MV, Roy L, Lindholm C,

Romm H, Guclu I, Lloyd DC. 2011. Biological dosimetry intercomparison exercise: an evaluation of triage and routine mode results by robust methods. *Radiat Res* 175:638–649

DIN 38402-45:2014-06. German standard methods for the examination of water and sludge- General information (group A). Part 45: Interlaboratory comparisons for proficiency testing of laboratories (A 45).

Endesfelder D, Oestreicher U, Kulka U, Ainsbury EA, Moquet J, Barnard S, Gregoire E, Martinez JS, Trompier F, Ristic Y, *et al.* 2021. RENEB/EURADOS field exercise 2019: robust dose estimation under outdoor conditions based on the dicentric chromosome assay. *Int J Radiat Biol.* 97(9): 1181-1198.

Garcia OF, Ramalho AT, Di Giorgio M, Mir SS, Espinoza ME, Manzano J, Nasazzi N, Lopez I. 1995. Intercomparison in cytogenetic dosimetry among five laboratories from Latin America. *Mutat Res.* 327: 33–39.

Garcia O, Di Giorgio M, Vallergera MB, Radl A, Taja MR, Seoane A, De Luca J, Stuck Oliveira M, Valdivia P, Lamadrid AI, *et al.* 2013. Interlaboratory comparison of dicentric chromosome assay using electronically transmitted images. *Radiat Prot Dosim* 154:18–25.

González J. E., Barquinero J. F., Holladay B.A., Di Giorgio M., Higuera M. 2020. Uncertainty calculation methods in dose assessment for dicentric chromosome assay, *International Journal of Radiation Biology.* 96:5, 606-613.

Gregoire E, Barquinero JF, Gruel G, Benadjaoud M, Martinez JS, Beinke C, Balajee A, Beukes P, Blakely WF, Dominguez I, Pham ND, Monteiro Gil O, Güçlü I, Guogyte K, Hadjidekova SP, Hadjidekova V, Hande P, Jang S, Lumniczky K, Meschini R, Milic M, Montoro A, Moquet J,

Moreno M, Norton FN, Oestreicher U, Pajic J, Sabatier L, Sommer S, Testa A, Terzoudi G, Valente M, Venkatachalam P, Vral A, Wilkins RC, Wojcik A, Zafiropoulos D, Kulka U. RENEB Inter-Laboratory Comparison 2017: limits and pitfalls of ILCs. *Int J Radiat Biol.* 2021 May 10;1-52. doi: 10.1080/09553002.2021.1928782.

Hampel F, Hennig C, Ronchetti E. 2011. A smoothing principle for the Huber and other location M-estimators. *Comput Stat Data Anal.* 55(1):324–337

Huber P.J. 1964. Robust estimation of a location parameter. *Annals of Mathematical Statistics* 35, 73-101.

IAEA. 2011. Cytogenetic dosimetry: applications in preparedness for and response to radiation emergencies. Vienna (Austria): EPR Biodosimetry, International Atomic Energy Agency.

ICRU Report 94. 2019. Methods for Initial-Phase Assessment of Individual Doses Following Acute Exposure to Ionizing Radiation. Biodosimetry. *Journal of the ICRU.* 19: 26– 45.

ISO 13528. 2015. Statistical methods for use in proficiency testing by interlaboratory comparison. ISO, Geneva.

ISO 19238. 2014. Radiological protection -- Performance criteria for service laboratories performing biological dosimetry by cytogenetics. ISO, Geneva.

ISO 5725-5. Accuracy (trueness and precision) of measurement methods and results- Part 5: Alternative methods for the determination of the precision of a standard measurement method. ISO, Geneva.

Koch M. 2004. Regulatory aspects of proficiency testing: experience from the water sector. *Accred Qual Assur.* 9:684–687.

Kulka U, Abend M, Ainsbury E, Badie C, Barquinero JF, Barrios L, Beinke C, Bortolin E, Cucu A, De Amicis A, *et al.* 2017. RENEB – Running the European Network of biological dosimetry and physical retrospective dosimetry. *Int J Rad Biol.* 93(1):2–14.

Kuselman I, Fajgelj A. 2010. IUPAC/CITAC Guide: Selection and use of proficiency testing schemes for a limited number of participants—chemical analytical laboratories (IUPAC Technical Report). *Pure Appl Chem.* 82(5): 1099–1135.

Livingston GK, Wilkins RC, Ainsbury EA. 2011. Pilot website to support international collaboration for dose assessments in a radiation emergency. *Radiat Meas.* 46(9):912–915

Lloyd DC., Edwards AA, Prosser JS, Barjaktarovic N, Brown JK, Horvat D, Ismail SR, Koteles GJ, Almassy Z, Krepinsky A, Kucerova M, Littlefield LG, Mukherjee U, Natarajan AT, Sasaki MS. 1987. A collaborative exercise on cytogenetic dosimetry for simulated whole and partial body accidental irradiation. *Mutation Research.* 179:197-208.

Pan Y, Ruan J, Gao G, Wu L, Piao C, Liu J. 2019. Laboratory Intercomparison of Cytogenetic Dosimetry Among 38 Laboratories in China. *Dose-Response: An International Journal.* January-March, 1-7.

Rousseeuw P.J., Verboven S. 2002. Robust estimation in very small samples. *Computational Statistics and Data Analysis.* 40: 741-758.

Romm H, Ainsbury E, Bajinskis A, Barnard S, Barquinero JF, Barrios L, Beinke C, Puig-Casanovas R, Deperas-Kaminska M, Gregoire E, *et al.* 2014. Web-based scoring of the dicentric assay, a collaborative biodosimetric scoring strategy for population triage in large scale radiation accidents. *Radiat Environ Biophys.* 53(2):241–254.

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>

Sasaki Ms, Hayata I, Kamada N, Kodama Y, Kodama S. 2001. Chromosome aberration analysis in persons exposed to low-level radiation from the JCO criticality accident in Tokai-mura. *J Radiat Res.* 42: S107–S116.

Sun M, Moquet JE, Barnard S, Lloyd D C, Rothkamm K, Ainsbury E A. 2016. Doses in Radiation Accidents Investigated by Chromosomal Aberration Analysis XXV. Review of cases investigated, 2006–2015. *PHE-CRCE-025:* 1-25.

Szewczak E, Bondarzewski A. 2016. Is the assessment of interlaboratory comparison results for a small number of tests and limited number of participants reliable and rational? *Accred Qual Assur.* 21:91–100.

Uhlig S, Colson B, Gowik P. 2018. Taking laboratory uncertainties into account in the Hampel estimator. *Accred Qual Assur.* 16:483.

Venables WN, Ripley BD. 2002. *Modern Applied Statistics with S.* Fourth Edition. Springer, New York. ISBN 0-387-95457-0.

Wilkins RC, Beaton-Green L A, Lachapelle, S, Kutzner, B C, Ferrarotto C, Chauhan V., Marro L, Livingston G K, Greene HB, Flegal FN. 2015. Evaluation of the annual Canadian biodosimetry network intercomparisons. *Int. J. Radiat. Biol.* 91: 443–451.

Table 1. Approaches suggested to determine the scale parameter (s) (ISO/IEC 13528).

Approach	Comments
From perception of experts	Maximum permissible errors, <i>e.g.</i> 30% to 20%, of the physical dose. When the action signal is 3.0 times the s value the maximum permissible error is divided by 3.0 to obtain the s value.
From data obtained in the ongoing ILC	Robust algorithms are recommended to obtain an estimation of s from the data set when the differences between the mean from the robust algorithm and the physical dose have an explanation. The existence of a common bias between several laboratories can severely affect the result.
From previous ILC	The value of s can be determined by experience with previous ILC testing for the estimated dose comparable with the actual values, and where participants use compatible scoring criteria and procedures. This is a useful approach when there is no agreement among experts about fitness for purpose of the assay.

Table 2. Example of three simulated data sets obtained using parameters of a previously published ILC.

Reported doses for a 2.34 Gy irradiated sample (Lloyd <i>et al.</i> 1987) *	Simulated datasets obtained with <i>rnorm.contam</i> R function*		
	1	2	3
0.94	0.9509994	0.9493756	1.015508
1.98	1.6713652	2.0189598	2.137612
2.06	2.2982395	2.4518394	2.302099
2.36	2.3062237	2.4896372	2.405673
2.49	2.5493482	2.5403157	2.493992
2.55	2.5504352	2.6742096	2.565716
2.70	2.7578053	2.6931671	2.629180
2.72	2.8083212	2.7203480	2.636302
2.98	2.8786392	3.0709887	2.866302
3.12	3.1566504	3.2867899	2.975880

The ILC data set of a sample irradiated at 2.34 Gy was published by Lloyd *et al.* 1987. The three synthetic data sets are obtained with the *rnorm.contam* function with parameters *mean*=2.55 and *sd*=0.382 that are calculated from the original data set when the outlier result of 0.94 Gy is removed. Each synthetic dataset mimics the original sample (Lloyd *et al.* 1987). The numbers are ordered for the sake of comparison. * Data sets sorted for comparative purposes. All data is presented in Gray units.

Table 3. Results of a simulation of an ILC of 10 laboratories estimating a dose of (2.34 ± 0.23) Gy.

Algorithm	Contam.p ^a	Expected Dose (Gy)	Estimated Dose (Gy) <i>mean</i> or x^*	Distance from the expected dose	Expected sd	Estimated sd or s^*	Distance from the sd
Arithmetic			2.3430	0.05721		0.2283	0.04499
Alg.A	0.0001		2.3435	0.05795		0.2386	0.05090
Alg.B			2.3438	0.05784		0.2072	0.05908
QHampel			2.3444	0.05851		0.2516	0.06487
		2.34			0.23		
Arithmetic			2.2035	0.1581		0.4417	0.2303
Alg.A	0.1		2.2575	0.1155		0.3480	0.1391
Alg.B			2.2625	0.1079		0.2633	0.08697
QHampel			2.2988	0.09250		0.3118	0.1087

^a The samples are contaminated with outliers at a distance of 1.4 Gy in the lower tail at a contaminated proportion of 0.1 (one value of ten). All the statistics are the mean of 1000 ILC rounds simulated. Estimation by robust algorithms denoted as x^* and s^* . In bold the closest result from the expected. All data is presented in Gray (Gy) units.

Table 4. Results of a simulation of 1000 ILC rounds, each ILC sample of 39 laboratories estimating a dose of (0.01948 ± 0.00512) Gy.

Algorithm	Contam.p ^a	Expected Dose (Gy)	Estimated Dose (Gy) mean or x^*	Distance from the expected dose	Expected sd	Estimated sd or s^*	Distance from the sd
Arithmetic			0.01949	0.00067		0.00517	0.000570
Alg.A	0.0001		0.01947	0.00067		0.00515	0.00056
Alg.B			0.01947	0.00067		0.00499	0.00064
QHampel			0.01947	0.00067		0.00520	0.00057
Arithmetic		0.01948 ^b	0.03949	0.02001	0.00512	0.04825	0.04313
Alg.A	0.1500		0.02275	0.00333		0.00963	0.00455
Alg.B			0.02217	0.00272		0.00659	0.00160
QHampel			0.01946	0.00073		0.007298	0.00222

Table 5. Results of different algorithms to estimate mean and standard deviation in illustrative real examples of inter-laboratory comparisons with dicentric assay.

Example	Algorithm	Delivered Dose (Gy)	Mean or x^*	sd or s^*
Gregoire et al. 2021	Arithmetic	0	0.03600	0.05696
	Alg. A (ni) ^a		0.02302	0.03453
	Alg. B ^b		NA	NA
	Q/Hampel		0.01088	0.02340
Lloyd et al. 1987	Arithmetic	0.7	0.8811	0.1723
	Alg. A (iterative)		0.8798	0.1925
	Alg. B		0.8731	0.1476
	Q/Hampel		0.8748	0.1654

^a non-iterative (ni) solution proposed in epigraph 6.2.6 of ISO/IEC 5725-5 part 5 was implemented because the iterative algorithm does not converge. ^b the iterative solution fails because of the high proportion of identical values. In bold the closest result from the expected. All data is presented in Gray (Gy) units.

Table 6. Different z-scores calculated from the estimated doses (x_i) after the analysis of the sample B (0 Gy) non-irradiated sample included in Figure 3 of Gregoire et al. 2021.

x_i (Gy)	z score			
	Prior $x_{ref} = 0$ $s_{ffp} = 0.03^a$	Q/Hampel $x^* = 0.01088$ $s^* = 0.02340$	Alg.A $x^* = 0.02302^b$ $s^* = 0.03453^b$	Arithmetic $mean = 0.03600$ $sd = 0.05696$
0	0	-0.46	-0.67	-0.63
0	0	-0.46	-0.67	-0.63
0.004	0.13	-0.29	-0.55	-0.56
0	0	-0.46	-0.67	-0.63
0	0	-0.46	-0.67	-0.63
0	0	-0.46	-0.67	-0.63
0.16	5.33	6.37	3.97	2.18
0.09	3	3.38	1.94	0.95
0.02	0.67	0.39	-0.09	-0.28
0	0	-0.46	-0.67	-0.63
0.07	2.33	2.53	1.36	0.6
0	0	-0.46	-0.67	-0.63
0	0	-0.46	-0.67	-0.63
0	0	-0.46	-0.67	-0.63
0	0	-0.46	-0.67	-0.63
0	0	-0.46	-0.67	-0.63
0.14	4.67	5.52	3.39	1.83
0.05	1.67	1.67	0.78	0.25
0.02	0.67	0.39	-0.09	-0.28
0	0	-0.46	-0.67	-0.63
0.02	0.67	0.39	-0.09	-0.28
0	0	-0.46	-0.67	-0.63
0.14	4.67	5.52	3.39	1.83
0.15	5	5.95	3.68	2
0	0	-0.46	-0.67	-0.63
0.03	1	0.82	0.2	-0.11
0.07	2.33	2.53	1.36	0.6
0	0	-0.46	-0.67	-0.63
0	0	-0.46	-0.67	-0.63
0.00	0	-0.46	-0.67	-0.63
0	0	-0.46	-0.67	-0.63

0.16	5.33	6.37	3.97	2.18
0.01	0.33	-0.04	-0.38	-0.46
0.03	1	0.82	0.2	-0.11
0.03	1	0.82	0.2	-0.11
0	0	-0.46	-0.67	-0.63
0	0	-0.46	-0.67	-0.63
0.02	0.67	0.39	-0.09	-0.28
0.19	6.33	7.65	4.84	2.7

^a Prior dispersion (s_{ffp}) obtained by a Poisson sampling approach. ^b non-iterative solution of algorithm A, iterative algorithm does not converge. Marked in bold are the unsatisfactory estimations that could require actions from the labs. All x_i data is presented in Gray (Gy) units.

Table 7. Different z-scores calculated from the estimated doses after the analysis of the 0.7 Gy Gamma rays irradiated sample in Lloyd et al. 1987 laboratory comparison.

x_i (Gy)	z-score					
	Prior $x_{ref} = 0.7$ $s_{ffp} = 0.07^a$	Prior $x_{ref} = 0.7$ $s_{ffp} = 0.1^b$	Q/Hampel $x^* = 0.8748$ $s^* = 0.1654$	Alg.A $x^* = 0.8798$ $s^* = 0.1927$	Alg.B $x^* = 0.8731$ $s^* = 0.1476$	Arithmetic $mean = 0.8811$ $sd = 0.1723$
0.83	1.86	1.30	-0.27	-0.26	-0.29	-0.30
0.83	1.86	1.30	-0.27	-0.26	-0.29	-0.30
0.80	1.43	1.00	-0.45	-0.41	-0.5	-0.47
1.18	6.86	4.80	1.85	1.56	2.08	1.73
1.10	5.71	4.00	1.36	1.14	1.54	1.27
1.01	4.43	3.10	0.82	0.68	0.93	0.75
0.74	0.57	0.40	-0.81	-0.73	-0.9	-0.82
0.72	0.29	0.20	-0.94	-0.83	-1.04	-0.93
0.72	0.29	0.20	-0.94	-0.83	-1.04	-0.93

^a 30% permissible error with a 3.0 factor results in 10% of the Delivered dose (0.7 Gy). ^b Poisson sampling uncertainty approach. Marked in bold are the unsatisfactory estimations that could require actions from the labs. All x_i data is presented in Gray (Gy) units.

Table 8. Examples of Fitness for purpose or priors criteria and their resultant s values for the z -score calculation.

Approach	Dose (Gy)	s value (Gy)	$3 * s$ range (Gy) $ z\text{-score} \leq 3$
Poisson*	0	0.03	0 – 0.1
	0.5	0.1	0.2 – 0.8
	1.5	0.1	1.2 – 1.8
	3	0.1	2.7 – 3.3
30% of Delivered dose	0.5	0.05	0.35 – 0.65
	1.5	0.15	1.05 – 1.95
	3	0.3	2.10 – 3.90
20% of Delivered dose	0.5	0.0333	0.4 – 0.6
	1.5	0.10	1.2 – 1.8
	3	0.20	2.4 – 3.6
0.5 Gy limits**	0	0.1667	0 – 0.5
	0.5	0.1667	0 – 1.0
	1.5	0.1667	1.0 – 2.0
	3	0.1667	2.5 – 3.5
1 Gy limits**	0	0.3333	0 – 1.0
	0.5	0.3333	0 – 1.5
	1.5	0.3333	0.5 – 2.5
	3	0.3333	2 – 4

* maximum values obtained by a simulation based in 1 dicentric in 1000 cells for 0 Gy and 500 cells for the rest of doses with a yield from IAEA calibration curve example. The doses out of the range will receive scores higher than 3. **The proposed accuracy limits of 0.5 and 1 Gy taken from ICRU Report 94. All data is presented in Gray (Gy) units.

```
#####
###
# Functions for Simulation study of arithmetic and robust algorithms performance
#
#####
###

set.seed(234)

# Functions written by Bret A. Holladay

# Q/Hampel method for robust estimation of population mean and standard
deviation
# Takes as input vector of data values, 'y' and corresponding lab numbers,
'lab'.
# Allows for replicates.
# There is one accuracy tolerance level in the code tol.G1 which can be made
# smaller to yield slower but more accurate results.
# tol.G1 corresponds to the decimal place accuracy of the numerator of
s.star

#Check answers with
#https://quodata.de/en/web%C2%ADservices/QHampel.html#0

QHampel<- function(y, lab, tol.G1=0.000001){
#####
#####
#-----
#C.5.2.2 - Q method for robust estimation of standard deviation
#-----
#####
#####

#p: number of labs
p=length(unique(lab));p

#n[i]: number of replicates from lab i
n=NA
for(i in 1:p){
  n[i]=sum(lab==i)
}

#-----
#C.22 - Define H1, the cumulative distribution function of all absolute
between-labatory differences
#-----

#H1 function
H1 <-function(x){
  sum=0
  for(i in 1:(p-1)){
    for(j in (i+1):p){
```



```

        for(k in 1:n[i]){
            sum=sum+1/(n[i]*n[j])*sum(abs(y[lab==j]-(y[lab==i])[k])<=x)
        }
    }
}
return(2/(p*(p-1))*sum)
}

#-----
# find discontinuity points, x.discont[1],...,x.discont[r] of H1(x),
# where, x.discont[1]<...<x.discont[r]
# discontinuities occur at each unique between-labatory difference
#-----

#create vector of all between-labatory differences
x.discont=vector()
for(i in 1:(p-1)){
    for(j in (i+1):p){
        for(k in 1:n[i]){
            x.discont=append(x.discont,abs(y[lab==j]-(y[lab==i])[k]))
        }
    }
}
#sort & keep only 1 of each difference (i.e. delete duplicates)
x.discont=sort(unique(x.discont));x.discont

#keep only positive values
if(x.discont[1]==0){x.discont=x.discont[-1]}

#-----
#C.23 - Calculate G1(0) and G1 at discontinuity points of H1
#-----

#Evaluate H1 on discontinuity points
H1.discont=sapply(x.discont, H1)

#calculate G1(0) and G1(x1),...,G1(xr)
G1.discont=c(0,.5*H1.discont[1],
.5*(H1.discont[-1]+H1.discont[-length(H1.discont)]))

#-----
#linearly interpolate G1 between discontinuity points 0,
x.discont[1],...,x.discont[r]
#-----
# tolerance level for G1 inverse which is part of the numerator for calculation
of s*
tol.G1=tol.G1

#linearly interpolate G1 between discontinuity points 0,x1,...,xr
x.int=vector()
G1.int=vector()
for(i in 1:length(x.discont)){
    #for each i define two points of discontinuity (x1,y1) and (x2,y2)

```

```

x1=c(0,x.discont)[i]
x2=c(0,x.discont)[i+1]
y1=G1.discont[i]
y2=G1.discont[i+1]

#creates sequence of points between x1 and x2
x.temp=seq(x1,x2,by=tol.G1)

#linearly interpolates between the two points of discontinuity
y.temp=((y2-y1)/(x2-x1))*(x.temp-x1)+y1

#append interpolated values
x.int=c(x.int,x.temp[-length(x.temp)])
G1.int=c(G1.int,y.temp[-length(y.temp)])
}

#-----
#numerator of s* :
#-----
q.G1=0.25+0.75*H1(0)

#x value that results in smallest value of |G1-q.G1|
numerator=x.int[which.min(abs(G1.int-q.G1))]

#-----
#denominator of s*
#-----
p.z=0.625+0.375*H1(0)
denominator=sqrt(2)*qnorm(p.z, mean = 0, sd = 1, lower.tail = TRUE)

#-----
#C.24 - Calculate robust standard deviation s*
#-----
s.star=numerator/denominator

#####
#####
#-----
#C.5.3.3 - Hampel estimate of robust mean without iterative reweighting
# uses s.star, value of robust standard deviation calculate with above
# Q method
#-----
#####
#####

#calculate arithmetic mean for each lab
y.mean=NA
for(i in 1:max(lab)){
  y.mean[i]=mean(y[lab==i])
}
#redefine y as y.mean
y=y.mean

#-----

```

#C.26 - Psi function

```
#-----  
Psi <- function(q){  
  res=NA  
  if(q<=-4.5){  
    res=0  
  } else if(-4.5<q && q<=-3){  
    res=-4.5-q  
  } else if(-3<q && q<=-1.5){  
    res=-1.5  
  } else if(-1.5<q && q<=1.5){  
    res=q  
  } else if(1.5<q && q<=3){  
    res=1.5  
  } else if(3<q && q<=4.5){  
    res=4.5-q  
  } else if(q>4.5){  
    res=0  
  }  
  return(res)  
}
```

```
#-----  
#Calculate all interpolation nodes  
#-----  
d=vector()  
for(i in 1:length(y)){  
  d=c(  
    d,  
    y[i]-4.5*s.star,  
    y[i]-3.0*s.star,  
    y[i]-1.5*s.star,  
    y[i]+1.5*s.star,  
    y[i]+3.0*s.star,  
    y[i]+4.5*s.star  
  )  
}
```

```
#sort in ascending order  
d=sort(d,decreasing=FALSE)
```

```
#calculate p_m values  
p.vec=vector()  
for(m in 1:(6*p)){  
  p.vec[m]=sum(sapply((y-d[m])/s.star, Psi))  
}
```

```
#Find all solutions to equation C.25  
S=vector()  
for(m in 1:(6*p-1)){  
  if(p.vec[m]==0){S=c(S,d[m])}  
  if(p.vec[m+1]==0){S=c(S,d[m+1])}
```

```

if(p.vec[m]*p.vec[m+1]<0){S=c(S,d[m]-p.vec[m]/((p.vec[m+1]-p.vec[m])/(d[m+1]-d[m]
))))}
}

#determine x.star
x.star=NA
#if no solutions to C.25 s.star is the median (median of vector of lab means)
if(length(S)==0){
  x.star=median(y)
} else {
  #Else s.star is the solution closest to median
  x.star=S[abs(S-median(y))==min(abs(S-median(y)))]
}
#if there are two solutions nearest the median set s.star to be the median
if(length(x.star)>=2){x.star=median(y)}

#return Q/Hampel values of robust mean and standard deviation
return(data.frame(x.star=x.star,s.star=s.star))
}

#####
# Logistic M-estimate , algorithm B #
#####

# This code was translated from the Matlab functions 'mloclogist' and
'mscalelogist'
# which are part of LIBRA: the Matlab Library for Robust Analysis, available at:

# http://wis.kuleuven.be/stat/robust.html
# Written by S. Verboven with Revisions by N. Smets (Last update 28/08/03)

# x: vector of n observations
# iter_loc: number of iteration steps for location estimate (default=50)
# iter_scale: number of iteration steps for scale estimate (default=1000)

M_estimate <-function(x, iter_loc=50, iter_scale=1000){
  n=length(x)

  #starting value for the location estimate
  t_0=median(x)

  #starting value for the scale estimate
  s_0=sqrt((length(x)-1)/(length(x)-1.5))*mad(x);s_0

  #compute M-estimate of location
  if(n==1){
    loc_estimate=x #n=1 implies location estimate equals x
  }else if(n==2){
    loc_estimate=mean(x) #n=2 implies location estimate equals the mean
  }else{
    #see page 752 Rousseeuw, P.J. and Verboven, S. (2002) for
    #why denominator can be replaced by 0.4132
    alpha=0.413241928283814 #integrate(1/2*sech(x/2)^2*normpdf(x,0,1),-10,10)

```

```

tstep=t_0

if (s_0!=0){
  j=1
  while(j<=iter_loc){
    z=(x-tstep)/s_0;
    y=tanh(z/2);
    tstep=tstep+s_0*(sum(y)/(n*alpha)); #updating location estimate
    j=j+1;
  }
  loc_estimate=tstep
}
}

#Compute M-estimate of scale
if(n==1){
  #when x is of length 1, all scale estimators must equal to 0
  scale_estimate=0
}else{

  # b=0.3739 leads to a 50% breakdown
  # (see page 754 Rousseeuw, P.J. and Verboven, S. (2002))
  b=0.3739;
  beta=0.500038854875226 #integrate((tanh(x/(2*b)))^2)*normpdf(x,0,1),-5,5)
  sstep=s_0

  if (s_0!=0){
    j=1
    while(j<=iter_scale){
      u=(x-t_0)/sstep;u
      uu=tanh(u/(2*b))^2;uu
      sstep=sstep*sqrt(sum(uu)/(n*beta));sstep
      j=j+1;
    }
    scale_estimate=sstep
  }
}

return(data.frame(location=loc_estimate, scale=scale_estimate))
}

```

```

#####
# rnorm.contam generates a n x nsamp matrix containing nsamp random normal
# samples of size n each with a random number of contaminants/outliers.
# Each value in the sample has probability contam.p of being made a
# contaminant/outlier.

# rnorm.contam was adapted from code written By M.D. Edge, 9/19/2018.
#####
# n: sample size
# nsamp: number of samples
# mu: population mean
# sigma: population sd

```

```

# contam.p: probability of contaminant/outlier
# contam.sigma: sd for contaminant pop
# contam.dist: distance of mean of contaminant population from mu
# Each contaminants is drawn from either N(mu-contam.dist,contam.sigma)
# or N(mu+contam.dist,contam.sigma) with equal probability
rnorm.contam <- function(n, nsamp, mu = 0, sigma = 1, contam.p = 0.01,
contam.sigma=0, contam.dist=1){

  #preinitialize matrix to store samples
  sample_matrix=matrix(NA,nrow=n,ncol=nsamp)

  #generate sample matrix
  for(j in 1:nsamp){
    #random number of contaminants (different for each sample)
    ncontam <- rbinom(1, n, prob=contam.p)

    #number of contaminants in lower & upper tail respectively
    ncontam.lower=rbinom(1, ncontam, prob=0.9999)
    ncontam.upper=ncontam-ncontam.lower

    # random normal sample of size n, with number of contaminants ncontam
    # n-ncontam values from N(mu, sigma)
    # ncontam.lower contaminants from N(mu-contam.dist,contam.sigma)
    # ncontam.upper contaminants from N(mu+contam.dist,contam.sigma)
    sample_matrix[,j]=c(rnorm(n - ncontam, mu, sigma),
                        rnorm(ncontam.lower, mu-contam.dist, contam.sigma),
                        rnorm(ncontam.upper, mu+contam.dist, contam.sigma))
  }

  return(sample_matrix)
}

#####
#Example
#out.sample<-rnorm.contam(n=10, nsamp=1000, mu=0.85, sigma=0.085, contam.p=.25,
contam.sigma=0.01,contam.dist=0.255)
#mean(out.sample);sd(out.sample)

#####
# Compute robust estimates of location (x.star) and scale (s.star) for each
# sample in the sample matrix produced by rnorm.contam using algorithms:
# arithmetic mean/sd; Hubers; M-estimate; and QHampel
#####
# arithmetic mean/sd are in base R
library(MASS) # for Hubers (Algorithm A)
# Load function Qhampel for Qhampel
# Load function M_estimate for M-estimate (Algorithm B)

```

```
#####
#   Parameters input   #
#####

n=10 #sample size
nsamp=1000 #number of samples

mu=2.34 #population mean
sigma=0.23 #population sd

contam.p=0.1 #probability of contaminant/outlier
contam.sigma=0.09 #sd for contaminant pop
contam.dist=1.4 #distance of mean of contaminant population from mu

#####
#   End of Parameters input   #
#####

#create n random normal samples of size n each with a random number of
contaminant/outliers.
out.sample<-rnorm(contam(n=n, nsamp=nsamp,
                        mu=mu, sigma=sigma,
                        contam.p=contam.p,
contam.sigma=contam.sigma,contam.dist=contam.dist )
out.sample[out.sample<0]<-0
#preinitialize vectors to store robust estimates
x.star_Arithmetic=rep(NA,len=nsamp)
s.star_Arithmetic=rep(NA,len=nsamp)
x.star_Hubers=rep(NA,len=nsamp)
s.star_Hubers=rep(NA,len=nsamp)
s.star_Mestimate=rep(NA,len=nsamp)
x.star_Mestimate=rep(NA,len=nsamp)
s.star_Mestimate=rep(NA,len=nsamp)
x.star_QHampel=rep(NA,len=nsamp)
s.star_QHampel=rep(NA,len=nsamp)

# Compute robust estimates of location (x.star) and scale (s.star) for each
sample
# with each algorithm (arithmetic mean/sd, Hubers, M-estimate, Qhampel)
for(j in 1:nsamp){
  HubersTEMP=hubers(out.sample[,j]) #Hubers
  MestimateTEMP=M_estimate(out.sample[,j],iter_loc=50, iter_scale=1000)
#M-estimate
  QHampelTEMP=QHampel(y=out.sample[,j],lab=1:n)#QHampel

  x.star_Arithmetic[j]=mean(out.sample[,j])
  s.star_Arithmetic[j]=sd(out.sample[,j])
  x.star_Hubers[j]=HubersTEMP$mu
  s.star_Hubers[j]=HubersTEMP$s
  x.star_Mestimate[j]=MestimateTEMP$location
  s.star_Mestimate[j]=MestimateTEMP$scale
}
```

```

x.star_QHampel[j]=QHampelTEMP$x.star
s.star_QHampel[j]=QHampelTEMP$s.star
}

```

```

# statistics. More consistent run to run. the distance from the non contaminated
mean and sd.

```

```

D_mu_xstar<-c(mean(abs(mu-x.star_Arithmetic),na.rm=TRUE),mean(abs(mu-x.star_Hubers),na.rm=TRUE),mean(abs(mu-x.star_Mestimate),na.rm=TRUE),mean(abs(mu-x.star_QHampel),na.rm=TRUE))

```

```

D_sd_sstar<-c(mean(abs(sigma-s.star_Arithmetic),na.rm=TRUE),mean(abs(sigma-s.star_Hubers),na.rm=TRUE),mean(abs(sigma-s.star_Mestimate),na.rm=TRUE),mean(abs(sigma-s.star_QHampel),na.rm=TRUE))

```

```

mean_xstar<-c(mean(x.star_Arithmetic,na.rm=TRUE),mean(x.star_Hubers,na.rm=TRUE),mean(x.star_Mestimate,na.rm=TRUE),mean(x.star_QHampel,na.rm=TRUE))

```

```

mean_sstar<-c(mean(s.star_Arithmetic,na.rm=TRUE),mean(s.star_Hubers,na.rm=TRUE),mean(s.star_Mestimate,na.rm=TRUE),mean(s.star_QHampel,na.rm=TRUE))

```

```

Method<-c("Arithmetic","Alg.A","Alg.B","QHampel")

```

```

Diff.Alg<-data.frame(Method,mean_xstar,D_mu_xstar,mean_sstar,D_sd_sstar);
Diff.Alg

```

```

#####
#Change the ID run to run#
#####

```

```

ID<-c("2_34 Gy_p0_1_D1_4_n10_lowertail_r2")

```

```

#####
# Export results as a csv file #
#####

```

```

filename=paste("D:\\Diff-Alg-Rob", ID,"-",Sys.Date(), ".csv", sep = "")

```

```

write.csv(Diff.Alg, file = filename, row.names = FALSE)

```



```
#####
# Code for splitting a Poisson sample and to obtain an overall mean,  #
# sd and limits that contain a percentage e.g. 99.7% of the subsamples #
#####

set.seed(050222)

par(mfrow=c(2,1))

# the code to obtain the lambda from the calibration, d=dose in Gy units

d<-0

#####
# Data #
#####

# Set a number of cells (metaphases) in the non-irradiated or in the irradiated
sample to be splitted

N=2000000

# Yield of dicentrics(lambda) from a calibration function(Lc)

Lc<-function(d){0.00128 + 0.02103*d + 0.06307*d^2}

lambda=Lc(d)

#####
#number of scored cells# e.g. 500 or 1000 for conventional scoring
#####

n.samp=1000

#####
# Calibration (C,alpha,Beta) #
#####

Lc<-c(0.00128,0.02103,0.06307)

#####
#The end of data input#
#####

y=lambda

DEsp<-( -Lc[2] + sqrt( Lc[2]^2 + 4*Lc[3]*(y-Lc[1]) ) )/(2*Lc[3])
```

```

DEsp # Expected dose

# code written by Bret A. Holladay

# Sample of size N

x <- rpois(N, lambda = lambda)

# Convert x to an n.samp by N/n.samp matrix,
# each column representing a subsample of size n.samp
# use x[,i] to extract ith subsample

x <- matrix(x, nrow=n.samp)

dim(x)

# Compute and store the mean of each subsample,
y <- apply(x, 2, mean)

DE<-( -Lc[2] + sqrt( Lc[2]^2 + 4*Lc[3]*(y-Lc[1]) ) )/(2*Lc[3])

DE[DE<0]<-0

#####
#           Dose           #
# mean , sd and 99.7% limits #
#####

mean(DE)

sd(DE)

c(sort(DE)[floor(0.00136*length(DE))], sort(DE)[ceiling(0.99865*length(DE))]) #
99.7% of data

c(mean(DE)-3*sd(DE),mean(DE)+3*sd(DE))

#####
#           Yield          #
# mean , sd and 99.7% limits #
#####

mean(y)

sd(y)

c(sort(y)[floor(0.00136*length(y))], sort(y)[ceiling(0.99865*length(y))])

```

```
#####  
#####
```

```
# Graphical Outputs #
```

```
hist(y,main="yield")  
hist(DE,main="Dose")
```

```
# out of the limits -> flag with an action signal the laboratory result
```

```
#
```