Article title: Dose conversion in retrospective dosimetry: results and implications from an inter-laboratory comparison featuring a realistic exposure scenario
Authors: Jonathan Eakins[1], Michael Discher[2], Hyoungtaek Kim[3], Min Chae Kim[3], Hanjin Lee[4], Jungil Lee[3], Olivier Van Hoey[5], Hyungjoo Yu[4], Elizabeth Ainsbury[1], Celine Bassinet[6], Stephen McKeever[7], Sergey Sholom[7], Francois Trompier[6], Lovisa Waldner[8], Clemens Woda[9]
Affiliations: uk health security agency, rce, didcot, oxfordshire, united kingdom[1], paris-lodron-university of salzburg, salzburg, austria[2], korea atomic energy authority, daejeon, south korea[3], korea institute of nuclear safety, daejeon, south korea[4], belgian nuclear research center (sck cen), mol, belgium[5], institut de radioprotection et de surete nucleaire, fontenay-aux-roses, france[6], oklahoma state university, stillwater, ok, usa[7], lund university, malmö, sweden[8], helmholtz zentrum munchen, neuherberg, germany[9]
Orcid ids: 0000-0003-3566-7801[1]
Contact e-mail: jonathan.eakins@ukhsa.gov.uk

DOI: 10.14293/PR2199.001110.v1
Preprint first posted online: 02 October 2024
Keywords: Retrospective dosimetry, Emergency Dosimetry, Physical Dosimetry, Monte Carlo, Conversion coefficients, Dose uncertainty

Dose conversion in retrospective dosimetry: results and implications from an inter-laboratory comparison featuring a realistic exposure scenario

Jonathan Eakins[*,1], Michael Discher[2], Hyoungtaek Kim[3], Min Chae Kim[3], Hanjin Lee[4], Jungil Lee[3], Olivier Van Hoey[5], Hyungjoon Yu[4], Elizabeth Ainsbury[1], Céline Bassinet[6], Stephen McKeever[7], Sergey Sholom[7], Francois Trompier[6], Lovisa Waldner[8], Clemens Woda[9]

[1] *United Kingdom Health Security Agency (UKHSA), RCE, Didcot, Oxfordshire, United Kingdom*
[2] *Department of Environment and Biodiversity, Paris-Lodron-University of Salzburg, Salzburg, Austria*
[3] *Radiation Safety Management Division, Korea Atomic Energy Research Institute, Daejeon, South Korea*
[4] *Korea Institute of Nuclear Safety (KINS), Daejeon, South Korea*
[5] *Nuclear Medical Applications Institute, Belgian Nuclear Research Center (SCK CEN), Mol, Belgium*
[6] *Institut de Radioprotection et de Sûreté Nucléaire, Fontenay-aux-Roses, France*
[7] *Department of Physics, Oklahoma State University, Stillwater, OK, USA*
[8] *Department of Translational Medicine, Medical Radiation Physics, Lund University, Malmö, Sweden*
[9] *Institute of Radiation Medicine, Helmholtz Zentrum München, Neuherberg, Germany*

* *Corresponding author*: Jonathan.eakins@ukhsa.gov.uk

Abstract

Dose conversion coefficients attempt to harmonize the material-, location-, and exposure-dependent results from retrospective dosemeters. The issues and uncertainties arising from dose conversion are explored within the framework of an interlaboratory comparison exercise in which mobile phones were positioned around anthropomorphic phantoms and exposed to non-uniform photon fields, with the glass and resistors they contain employed as fortuitous dosemeters. The difficulties of adopting pre-calculated tables of generic conversion coefficients are evaluated first, and then compared against those arising through the use of bespoke data derived by Monte Carlo modelling, and also against not converting the doses measured by the phones. It is seen that the different subjective choices that users might make when selecting 'optimal' generic data can lead to a significant source of uncertainty (up to around 70 %), though may be improved (to around 30 %) by appropriate quality controls. Use of generic coefficients typically led to over-estimates of the organ doses: an average discrepancy of ca. a factor of 2 was found, but this is still better than the factor of around 3 observed when no conversion coefficients were applied. Use of bespoke conversion factors led to the best estimates of organ doses, although they still over-estimated by approximately 1.5 on average, and an uncertainty of around 20 % was associated with generating their values. Overall, applying bespoke conversion data improves but does not guarantee correct dose categorization of individuals, with the inconsistences in the measured results found generally to be the limiting factor in obtaining accurate dose assessments.

1. Introduction

Personal dosemeters are worn by individuals occupationally exposed to potentially hazardous sources of ionizing radiation, in order to estimate the doses that they receive. However, it is conceivable that people who do not have monitors might also be exposed to significant doses of ionizing radiation, due to accidental exposures from, say, the use of ionizing radiation in hospitals, the nuclear industry, or research, and in malicious attacks with radiological agents. For these reasons, considerable research effort has been expended to investigate the use of fortuitous retrospective dosemeters, i.e. common

materials that may be indicators of the absorbed doses received by exposed individuals (see e.g. Kulka *et al*, 2017; ICRU, 2019; McKeever *et al*, 2020). The idea is that such dosemeters may be analysed after a radiological incident as part of initial-phase ('triage') dosimetry of exposed individuals. The results may then add to the evidence base, and assist in deciding appropriate clinical treatment (if any) to mitigate against potentially adverse radiation effects, such as acute radiation syndrome (ARS).

Broadly, retrospective dosimetry techniques may be grouped into two complementary types: biological indicators, which rely on analyses of tissue samples taken from the exposed individual; and physical dosemeters, which rely on the dosimetric properties of certain objects carried by, or proximal to, an individual. Examples of biological methods include techniques such as dicentric scoring of chromosome aberration in blood samples (IAEA, 2011), methods based on DNA damage (Moquet *et al*, 2017), and gene expression analyses (Polozov *et al*, 2019). Physical dosimetry techniques include, but are not limited to, optically stimulated luminescence (OSL) or thermally stimulated luminescence (TL) of objects such as mobile phone electronic components (e.g. Inrig *et al*, 2008; Ekendahl and Judas, 2012; Lee *et al*, 2015; Eakins *et al*, 2016; Ademola *et al*, 2017; Geber-Bergstrand *et al*, 2018), display glass (e.g. Discher *et al*, 2013; 2016; Bassinet *et al*, 2014a; Mrozik *et al*, 2014; Chandler *et al*, 2019; Kim *et al*, 2019a), chip cards (e.g. Göksu, 2003; Sholom and McKeever, 2016; Mrozik *et al*, 2017; Kim *et al*, 2020), household salt (e.g. Bernhardsson *et al*, 2009; Ekendahl and Judas, 2011; Spooner *et al*, 2012), and other fortuitous materials carried by or near an individual. Electron paramagnetic resonance (EPR) on materials such as human tooth enamel (Fattibene and Callens, 2010) has also been studied.

Although retrospective dosimetry methods may prove invaluable in the event of a radiological incident, a thorough evaluation and understanding of their inherent limitations is fundamental to their success (or failure). One way of investigating these is by performing inter-technique and inter-laboratory comparison (ILC) exercises. The basic aim of such exercises is to expose various biological and/or physical dosemeters to controlled and well-characterized radiation fields, and then compare the results that are obtained. Within Working Group 10 (WG10, '*Retrospective Dosimetry*') of the European Radiation Dosimetry group (EURADOS), a number of inter-comparison exercises have been performed in the past (Bassinet *et al*, 2014b; Rojas-Palma *et al*, 2020; Discher *et al*, 2021a). Those exercises have helped to evaluate the state-of-the-art of retrospective dosimetry, both in terms of the performances of the dosemeters themselves and the laboratories that are conducting the techniques. An ILC was also recently performed by WG10 members in Malmö, Sweden. The set-up (Waldner *et al*, 2021) and results from the biological retrospective dosimetry (Endesfelder *et* al, 2021) are presented elsewhere, along also with details of some of the Monte Carlo modelling that was undertaken to support the reference dosimetry (Kim *et al*, 2024); a report summarizing the physical retrospective dosimetry is currently in production, led by the same authors as those of previous ILCs (Bassinet *et al*, 2014b; Discher *et al*, 2021a).

One common limitation of all physical dosemeters is that they can only ever provide estimates of the absorbed doses to themselves, whereas obviously it is the absorbed dose to the individual that is of clinical relevance. Whilst there is some debate (Eakins and Ainsbury, 2018a; 2018b) as to what is meant by 'absorbed dose to the individual' in the context of large whole-body exposures, it is nevertheless clear that further consideration is required of the relationship between the dose measured by the dosemeter and the actual 'biological' dose received by the individual. Such a need is especially important for physical dosemeters because they may be located at a place that receives a very different exposure from the majority of the body (e.g. a mobile phone kept in a pocket) and may rely on a sensitive material (e.g. aluminium oxide) that is not tissue-equivalent. Indeed, differences in location can mean that even two dosemeters of the same type carried by the same individual may report very different absorbed doses. Similarly, two different dosemeter materials may exhibit different results even if they are roughly co-located, an example being the absorbed doses to the

display glass and resistors in a phone. The question is, then: how should the 'true' dose to the individual be determined from the range of doses measured by the various retrospective dosemeters? Essentially, this becomes the question: how should the various location- and material-dependencies of the dosemeter results best be taken into account to arrive at a single assessment of "the dose" to the individual?

Previous work to investigate this issue was described by Eakins and Kouroukla (2015). Those authors derived sets of generic conversion factors that related doses to the resistors in mobile phones to organ doses within the body, as functions of the location of the phone relative to the person and the exposure. Other publications have since extended that endeavour, as discussed later. The current paper seeks to develop this work in the context of the recent ILC, in terms of exploring the impacts of applying the generic data and by presenting and comparing new data that are bespoke to its exposure conditions. The role of the laboratory performing the dose assessment, and the impact of the subjective choices made when selecting data, are also explored.

2. Dose conversion in retrospective dosimetry: Current Status

The immediate problem in attempting to account for variations in position and exposure is that there are clearly an infinite number of different possibilities and permutations. Work to-date has therefore sought to reduce these to a set of plausible 'standard' conditions, so that generic sets of conversion coefficient data may be generated; this is analogous to the ICRP approach for quantifying stochastic risks (ICRP, 2010), with fluence-to-effective dose conversion coefficient data tabulated at discrete energies for just a limited number of plane-parallel and isotropic exposure geometries.

Initially, Eakins and Kouroukla (2015) provided dose conversion coefficients between the aluminium oxide substrate of resistors in mobile phones and organs within the body, by performing sets of Monte Carlo calculations using realistic phone models positioned about the ICRP 110 voxel anthropomorphic phantom (ICRP, 2009). The authors considered iridium-192, caesium-137 and cobalt-60 radionuclide photon sources, and considered seven 'standard' exposure geometries: anterior-posterior (AP), posterior-anterior (PA), left-lateral (LLAT), right-lateral (RLAT), isotropic (ISO), floor (FLR, representing ground contamination), and quasi-rotational (ROT*, defined as the average of AP, PA, LLAT and RLAT data). Four positions of mobile phone about the body were considered: 'Chest', with the phone centred close to the location of the heart; 'Leg', with the phone centred just in front of the left thigh; 'Back', with the phone centred just behind the left buttock; and 'Hip', with the phone centred close to the left hip [Figure 1]. The outcome of this work was sets of tabulated coefficients that, for a given exposure (E), allowed doses measured by resistors in a phone to be converted to organ and whole body absorbed doses.

Since the above was completed, the work has been extended both by others and by the original authors to consider additional dosemeters, postures and exposure conditions (Discher, 2015; Kim, 2019b; Lopez *et al*, 2020; Rojas-Palma *et al.*, 2020; Discher *et al.*, 2021a; Discher *et al*, 2021b; Chandler *et al*, 2022). Under the auspices of EURADOS WG10, so-far unpublished datasets have also been produced for 100 keV exposures, for circuit board material within mobile phones, and for an additional phone location, 'In Use', which is 30 cm in front of the body. WG10 has also attempted to provide user-friendly software that combines the results to-date and performs the conversion. Specifically, Excel databases were created that allow users to input measured doses, select the appropriate measurement conditions, and obtain the desired organ doses. A flow-diagram schematic of this process and software is shown in Figure 2. Note that this software has not been published formally, but is available by request from the present authors.

Whilst considerable progress has been made, two overarching problems persist. Firstly, it is inevitable that data can only ever be calculated in advance for just a fraction of the possible exposure conditions that might occur in a real emergency scenario. Secondly, correct selection of the appropriate conversion coefficient presumes knowledge of the exposure conditions, but in practice users might have significant ignorance of exactly where the phone was located relative to the person, or how that individual was orientated relative to the source, as examples. Accordingly, users may be forced to 'best-guess' the most appropriate dataset, based on their judgement of available information and their interpretation of which of the pre-existing coefficients might most closely match them. These two factors therefore introduce potentially large uncertainties into the dose conversion process, some discussion of which has been reported previously (Eakins and Kouroukla, 2015; Van Hoey *et al*, 2021; Chandler *et al*, 2022).

Further exploration of these factors within the context of the recent ILC, and their implications for retrospective dosimetry, is a focus of the current paper. In particular, the objective uncertainties will be considered by calculating new conversion coefficients that are bespoke to the precise conditions of the ILC, and then comparing the results against those generated using the pre-existing datasets. The subjective uncertainties will be explored by presenting the results from a simple exercise that compared different individuals' different decisions based on their different interpretations of the same information.

Finally, mention must be made here of ICRU Report 94 (ICRU, 2019), which was published subsequent to the recent ILC and took the second of the above highlighted problems as a basis of discussion. The report recommended that "*For external exposures in large-scale events, the absorbed dose to the dosimeter material is the recommended dose to use in initial-phase dose assessment, whether this be determined by biological or physical dosimetry methods*", the primary reasoning being the very plausible assertion that the uncertainties and unknowns are likely too large for dose conversion to be workable. At first glance, this would seem to undermine the present endeavour, though it is remarked that the report did also permit exceptions when only a very few individuals were exposed (cf. the current ILC), in which case sufficient details may be known about the conditions to derive specific conversion data. Either way, it may be argued that the ICRU recommendations raise a critical question: is reporting relevant results with large uncertainties (e.g. estimated doses to individuals using dose conversion coefficients) better than reporting results that, by definition, are known not to have any biological basis at all (e.g. doses to fortuitous materials)? Or, more philosophically: is estimating the right quantity badly preferable to estimating a fundamentally different quantity well? In line with the call within ICRU 94 for further research on this issue, one additional aim of the current paper is therefore to investigate some of the uncertainties that arise through the use of conversion coefficients and compare them against those arising from their neglect.

3. The 2019 ILC in Malmö

The set-up and conditions of the 2019 ILC are described in detail elsewhere (Waldner *et al*, 2021), with just a brief summary provided here. Two exposure scenarios were conducted, labelled 'Setup 1' and 'Setup 2', schematics of which are shown in Figures 3 and 4 respectively. The exposures were performed outdoors in a large open space, above a concrete floor.

In Setup 1, an adult female anthropomorphic phantom (labelled 'P1', model ATOM 702, CIRS inc., VA) was placed facing a 1.36 TBq iridium-192 gamma-emitting source that was at a distance of 28 cm from it and 59.5 cm above the ground. Obliquely behind the female phantom and facing it at an angle of ~45° was an adult male Rando anthropomorphic phantom (labelled 'P2', Alderson Research Laboratories, USA), positioned at a distance of ~70 cm from the source; this male phantom was

therefore partially shielded from the source by the female phantom. In Setup 2, an adult male Rando anthropomorphic phantom (labelled 'P3', Alderson Research Laboratories, USA) was positioned facing the source at a distance of ~114 cm. A second adult male anthropomorphic phantom (labelled 'P4', model ATOM 701, CIRS inc., VA) was positioned at a 45° angle to the source, facing phantom P3, with its left side closest to the source and ~53 cm from it. None of the four phantoms had arms or legs, and all were placed on chairs at a height of ~45 cm. The exposures in Setup 1 and Setup 2 lasted approximately 1 hour and 2.5 hours, respectively.

Various dosimetric materials were located on the four phantoms, supplied by a number of retrospective dosimetry laboratories across Europe and from South Korea and the United States. These included mobile phones, both the display glass and resistors of which were used for the analyses. The locations of these are shown in Figure 5, where the labels signify the originating laboratory and the dosemeter number. Each laboratory provided two mobile phones, the locations of which were chosen to broadly match the areas where phones might most likely be found in a real emergency scenario, e.g. in front / back trouser pockets.

The arrangements shown in Figures 3-5 demonstrate how non-uniform exposure conditions were enforced explicitly, with attenuation, cross-scatter and inverse-square divergences of the photon field all significant. Such non-uniformity would have manifested through each individual phantom, from one phantom to another, and across the range of retrospective dosemeters distributed unevenly around them. Within the context of the current paper, the departure from 'standard' exposure conditions (e.g. plane-parallel) is therefore apparent, along with the expectation that the various dosemeters will experience very different exposures and hence record very different doses, even when located on the same 'individual'. This emphasizes the potential need for conversion factors that take these differences into account. It also provides a first test of the applicability, or otherwise, of the overall approach considered previously, and the potential use of standardized sets of generic data.

4. Participants' application of existing datasets to the ILC

4.1 Set-up

At the time of the Malmö ILC measurements, only the generic conversion coefficients were available (the bespoke data were generated subsequently). This therefore provided an opportunity to test their application and interpretation by participants, giving insight both into the uncertainties associated with their use and the impacts of the subjective decisions that users face. To that end, along with measuring the absorbed doses to their retrospective dosemeters, ILC participants were also invited to determine and report the concurrent 'body doses' for the phantoms upon which they were located. For this exercise, 'body dose' was associated with the quantity D_{GRB} that would have been received had the phantom been a real person, and defined previously (Eakins and Ainsbury, 2018b) for a given exposure as the average of the doses to the small intestine (SI) and stomach (referred to collectively as the Gastrointestinal (GI) tract), Red bone marrow (RBM) and Brain. Associated with this dose quantity is the corresponding conversion coefficient, C_{GRB}, given as:

$$C_{GRB} = {}^{D_{E,p,m}}\!/_{D_{GRB}} = {}^{D_{E,p,m}}\!/_{\{(D_{SI} + D_{Stomach} + D_{RBM} + D_{Brain})/4\}} \qquad [1]$$

where $D_{E,p,m}$ denotes the absorbed dose calculated for material m (either resistor (r) or glass (g)), for a phone located at nominal position p (i.e. Chest, Leg, Back or Hip) and in exposure condition E (e.g. AP, PA, LLAT, RLAT, ISO, ROT or FLOOR), and with the organ doses in the denominator.

The choice of dose quantity for emergency scenarios is problematic. D_{GRB} is a non-standard quantity, but conversely there is no alternative that is recommended for large whole body doses. The proposed D_{GRB} was therefore suggested as a means of accounting for average doses to the most important radiosensitive organs, useful to provide a rough estimate of overall risks to individuals. Such a single dose estimate is convenient for the current purposes, to provide consistent metrics for comparison. Nevertheless, the conversion software is able to provide data for all internal organs separately (Figure 2), which could have made for alternative analyses. However, that would have had the drawback of introducing additional variables into the discussions, in particular the differing locations of the organs; the stomach is an obvious example, adding an extra left-right bias into the results. Of course, that would in turn permit additional useful analyses and insight, but for brevity these are beyond the scope of the current paper, instead being left as an interesting consideration for future ILCs. For now, using the single quantity D_{GRB} is adequate for highlighting the impacts of the various choices and uncertainties in retrospective dosimetry.

The laboratories were provided with the EURADOS WG10 conversion software (Figure 2) for an iridium-192 source. Basic information on the locations of the various dosemeters were also provided in the form of the images reproduced here (Figure 5), along with the definitions of the phone positions presumed by the software (Figure 1). Assuming that some sort of 'questionnaire' process were in place to gather exposure information (e.g. Eakins *et al*, 2016), the information in Figure 5 may be argued to represent the maximum level of detail that an exposed individual might feasibly be able to provide in a real emergency scenario. That is, at best, they could likely only provide the rough vicinity of the phone during the exposure rather than its precise coordinates, with no more detail available than 'somewhere in a left trouser pocket', for example. As attendees of WG10 meetings, all participants had some knowledge of the ILC exposure configurations, although to different extents. Again, this situation is analogous to the likely range in the quality and quantity of information in a real emergency incident, resulting from an anticipated inability of individuals to always know exactly where they were positioned or in which precise direction they were facing.

Thus, participants had to decide which was the most appropriate location and exposure geometry for their phone from the set of options available. Best judgement was needed in this selection, which is an inherently subjective process; in some cases, this may have been tantamount to identifying a perceived least-worse option. To provide additional insight, participants were also invited to comment on why they made their particular selections. The organizer of the exercise also performed the procedure for each phone in order to yield a 'reference' set of solutions, with this process conducted 'blind' to the results submitted by the individual laboratories.

4.2 Results and feedback

Nine laboratories submitted results, though not all estimated doses using both resistors and glass. The results are provided in anonymized form for the resistors (*r*) in Table 1 and for the glass (*g*) in Table 2, labelled as $LX\text{-}Y_r$ and $LX\text{-}Y_g$ respectively, where *X* is an arbitrarily assigned laboratory number and *Y* denotes to which of the two phones from each laboratory the result relates. The choices made by the organiser (*O*) and participants (*L*) are shown in the tables, with differences highlighted in red, along with the corresponding conversion coefficients, C_{GRB}.

The impacts of differences in these choices on the subsequent estimates of D_{GRB} are evaluated as the ratios, $(R(L/O)_{GRB})$, of the conversion coefficients chosen by the participants to those of the organizer. As expected, there are large overlaps between the trends exhibited in Tables 1 and 2, but small differences still exist when like-for-like resistor and glass ratios are compared. This is caused by small location- and exposure-dependencies of the field's energy distribution and its impact on the differing responses of those two materials. Dosimetry on glass was not performed for phones *L*1-1, *L*1-2, *L*2-1 or *L*9-2; in principle, the data that would have been used for these phones could still have been

included in Table 2 but, notwithstanding the small differences noted above, such an exercise would not have greatly impacted the general observations discussed later.

Before analysing the data in Tables 1 and 2, it is worth noting the supplementary comments that were provided:

- *The organizer (mistakenly) thought that the source was located much closer to the ground than it actually was, and therefore chose the Floor geometry for P1 to best-represent a perceived upwards exposure. For the other phantoms, the organizer considered that their greater distances from the source might justify the assumption of a plane-parallel exposure; given their orientations, that meant AP for P3 and LLAT for P4. An AP exposure was similarly chosen for P2 as a least-worse option, despite its partial shadowing by P1, as that phantom was facing in the general direction of the source.*

- *Participant L2 reported that, because only the position of the phone was known and not the exposure scenario, AP, PA, LLAT, RLAT and ROT possibilities should all be considered. This gave a set of five results, which in turn was used to provide a dose estimate range corresponding to an interval [A,B], where A was defined as the smallest of the five values in the set minus their standard deviation, and B as the largest of the five values plus their standard deviation. The true value of D_{GRB} was then assumed by L2 to lie somewhere within that interval.*

- *L3 reported that a ROT geometry was chosen to mimic the anticipated movements of a person wearing the phone, which would likely occur in a real exposure scenario.*

- *L4 reported that AP and PA geometries were chosen because the orientation for each phone was not clear.*

- *L5 reported that they considered ISO to be the best choice because the location of the source was unknown, making then the assumption that the radiation field would surround the person and results may be determined be averaging.*

Comments were not received from *L1* or *L6-L9*. It is perhaps noteworthy that the choices made by these laboratories typically aligned with those of the organizer. It might plausibly be speculated that the decisions on the most appropriate generic conversion data for these phones may have been perceived as being 'more obvious'.

*L*2 also stated that for *L*2-2, for which both resistors and glass were analysed, the average of those results would have been reported as a single dose estimate. However, it was later noted that for the glass they had actually used the conversion coefficients for resistors, presumably by oversight. For this reason, their ratios $R(L/O)_{GRB}$ for glass contain an additional inherent bias. This highlights the propensity for human error in the process, and emphasizes that the mistakes inevitably made by users will represent sources of uncertainty in any real dosimetry system.

It is clear that a wide variety of personal interpretations occurred in this simple exercise, and that different participants adopted different tactics and approaches. This arises from ignorance of the precise exposure conditions, and hence divergent opinions on how best to mitigate for that and which conversion coefficient to apply. It is possible, therefore, that if two or more laboratories were to analyse even the same phone, their resulting estimates of the dose to the individual could differ significantly.

4.3 Analyses and discussion

The aim of the exercise was to test laboratories' use of the WG10 software, and provide a comparison of the various subjective choices. In that regard, it is emphasized that the choices made by the organizer are not claimed to be any more correct than those of the participants, with the same limited quantity and quality of data provided to both. It is also remarked that the ILC exposure conditions were highly specific and heterogeneous, which could make the choice of an appropriate generic conversion coefficient harder than in a real mass-casualty situation. Moreover, some of the participants had been involved in setting-up the ILC, and so would have been familiar with the layout, potentially introducing some bias. The findings from this simple scoping exercise should therefore be treated with some scepticism, and future work would aim for a larger cohort of participants and exclude those that had direct contact with the exposures.

One alternative to using the organizer's choices might have been to adopt the average of participants' results as the reference value at each location. This would have been possible in some cases, although not in others where there was only one phone present. A better way of conducting the exercise could therefore have been to request participants submit results for all phones instead of just their own; obtaining the reference set via averaging might then have been feasible. However, this would have a disadvantage that the choices themselves could no longer be contrasted so easily, only the conversion coefficients. In other words, whilst it might permit a better quantitative comparison of the numerical results, qualitative analyses in terms of "*Participant X made choice Y, whilst the organizer made choice Z*" would be more problematic. Nevertheless, it remains a methodology to consider in future studies.

Overall, the organizer and participants are seen (Tables 1 and 2) to have made the same choice of either location or exposure geometry in 19 out of 36 times. However, for only 4 out of 18 phones was the same choice made for both the location and the geometry. Most of the disagreements arose when choosing exposure geometries, for which the same selections were made in around only a quarter of the cases (5 out of the 18).

For the phone locations, the organizer and participants are seen to have made the same choices in over three quarters of the cases (14 out of 18). Generally, they agreed when the phone was located roughly in the Chest or Back positions, likely because there were no other plausible options in those instances. More disagreement was exhibited for the other two locations, potentially because it is not always clear from the photographs (Figure 5) whether the phone most closely related to the Leg or Hip locations (Figure 1). This type of ambiguity likely reflects the situation that would arise in a real emergency scenario.

The ratios, $R(L/O)_{GRB}$, may be used to provide some broad metric of the likely uncertainty that results from subjectivity within the conversion coefficient selection process. Specifically, it is noted that if the ranged data from *L*2 are excluded, then the average of the remaining sixteen $R(L/O)_{GRB}$ values for the resistors in Table 1 is 1.08, and the individual data points are distributed around this mean with a standard deviation of 0.77. For the glass in Table 2, the analogous data are 1.18 and 0.79, respectively. Of course, there is nothing to suggest that these ratios should follow anything like a Gaussian distribution, and different individuals might make choices that lead to vastly different results. It is also remarked that the current exercise is clearly not scientifically rigorous, and is instead intended just to provide some qualitative insight into the role of subjectivity. Finally, had a different organizer analysed the photographs, it is likely that a different set of 'reference' solutions (and hence ratios) would have been generated. Nevertheless, and with these caveats, the spread of ratios in Tables 1 and 2 might suggest something like a standard uncertainty of around $\pm$70 % for the subjective component of dose conversion, at least as a rough first approximation based on this limited scoping exercise.

The extremal values of $R(L/O)_{GRB}$ were for *L*4-1 and *L*9-2. Both of those phones were identified by the organizer and participants as being in the Back location. Phone *L*4-1 was assumed to be exposed AP by the organizer, but PA by the participant. However, that participant commented that the "*orientation for each phone is not clear*", so it is plausible that they might have flagged their result as questionable. A ground exposure of Phone *L*9-2 was assumed by the organizer (i.e. applying the conversion factors labelled 'FLOOR'), but AP by the participant; this misunderstanding of the point source location by the organizer is the cause of the large discrepancy. Such mistakes might be avoided in a real emergency situation by providing proper guidance and training, and by routinely implementing basic quality control procedures into the dosimetry service.

If *L*4-1 and *L*9-2 are excluded, the average of the remaining fourteen $R(L/O)_{GRB}$ values for the resistors is lowered slightly to 0.94, but the individual data points have a greatly reduced standard deviation of 0.21. Assuming that an operational retrospective dosimetry system had quality control checks sufficient for major mistakes and unknowns to be identified, such that outlier results were invalidated, these results would imply a more satisfactory standard uncertainty of <30 % for the subjective component of dose conversion.

Given the choices reported by the organizer and participants, it is easy to repeat the analysis using an alternative 'whole body' dose quantity; this is relevant, considering the lack of consensus within the community on any such quantity. For example, D_{GRB} may be replaced by the 'whole body average absorbed dose', D_B, defined (Eakins and Kouroukla, 2015) and used (e.g. Eakins and Kouroukla, 2015; Discher M, 2015; Kim M C, 2019b) elsewhere as the average of the doses to all radiosensitive organs listed in ICRP 103. The resulting conversion coefficients, $C_B = D_{E,p,m} / D_B$, are not included here for brevity, but their ratios, $R(L/O)_B$, which are defined and derived in the obvious way analogously to $R(L/O)_{GRB}$, are also shown in Tables 1 and 2. If the ranged data from *L*2 and the two outlier values for *L*4-1 and *L*9-2 are again excluded, the average of the remaining fourteen $R(L/O)_B$ values for the resistors is 0.91, and the individual data points are distributed with a standard deviation of 0.18. From this and Tables 1 and 2, it is clear that the choice of which of these 'whole body' dose quantities is used does not greatly impact the overall outcomes: the general patterns and trends in the datasets are similar in both cases. One implication is therefore that the precise choice of 'whole body' dose quantity is likely a less impactful factor than the uncertainty caused by ignorance of the phone location and exposure orientation, at least for the fairly uniform whole body exposures accounted for by the generic data.

Of course, it is reiterated that the current exercise is extremely limited in scope: comparing the decisions of just a small number of participants against those of one organizer cannot give any statistical power or rigour to the results. Nevertheless, even neglecting the organizer, it is observed that the choices made by different participants did not always agree. For example, phones *L*2-2 and *L*5-2 were located next to each other, but their respective laboratories still applied different conversion coefficients based on different approaches to interpreting the field conditions. It is therefore reasonable to suspect that the types of divergence highlighted in Tables 1 and 2 would likely have persisted had a different organizer conducted the test, even if the precise patterns of disagreement may not have been identical to those exhibited here.

5. Bespoke dose conversion coefficient data

5.1 Generation of data

In order to provide bespoke conversion coefficient data for the ILC, approximations of the exposure conditions were recreated within Monte Carlo models. The full details and results from this modelling

are presented and discussed in a sister publication (Kim *et al*, 2024), with just a brief summary provided below. In total, four groups used two Monte Carlo codes (GEANT4 (Agostinelli *et al*, 2003) or MCNP6.2 (Werner *et al*, 2018)) to model the scenarios: Korea Atomic Energy Research Institute (KAERI, labelled group 'Lab1' in [Kim *et al*, 2024]), United Kingdom Health Security Agency (UKHSA, 'Lab2'), Belgian Nuclear Research Center (SCK CEN, 'Lab3'), and Korea Institute of Nuclear Safety (KINS, 'Lab4'). The energy spectrum of the iridium-192 was taken from the Brookhaven National Laboratory resource (NNDC, 2020), with all groups adopting an idealized point source geometry.

Due to various factors, the four groups used slightly different approaches and approximations in their modelling. Those choices are summarized in Table 3, and result from considerations including: the respective limitations and abilities of GEANT4 vs. MCNP6.2 (but see (Yeom *et al*, 2019)); the availability of mesh (ICRP, 2020) or voxel (ICRP, 2009) phantoms; differing resource availabilities (e.g. Lab2 was restricted in computer memory, so only one voxel phantom could be modelled); different interpretations of the set-up conditions (e.g. exact phone and source locations, relative phantom positions etc.); and different approximations (e.g. dosemeters modelled fully vs. approximated as thin discs of material). Moreover, the physical phantoms used in the measurements were limbless (Waldner *et al*, 2021), but there are pros and cons of recreating this condition in the modelling: doing so provides better benchmarking against measured results, but not doing so provides a better comparison against both prior conversion coefficient data and the 'true' doses to people in such fields, as well as more accurate calculations of some organ doses (e.g. RBM, which is dependent on the long bones).
All dose results were provided with their associated statistical uncertainties, as outputted by the respective Monte Carlo codes. Those uncertainties were generally very low (typically <1 %), so have been omitted from much of the following work for clarity. Systematic uncertainties on the results, which were presumed to arise partly from the different ways in which the scenarios were modelled by the four groups, are harder to quantify and are discussed below.

Because of the presence/absence of the limbs, a direct comparison between the GEANT4 (Lab1 and Lab4) and MCNP (Lab2 and Lab3) results is problematic. This difficulty is obvious for the RBM, but is also relevant for other organs (e.g. testes) because attenuation and scatter by the limbs can drastically affect the doses received. The absorbed doses to the dosemeters are similarly affected. In order to derive meaningful dose conversion coefficients for the ILC exposures, along with their uncertainties, the following scheme was therefore implemented:

- The GEANT4 results from Lab1 have been adopted as the primary dataset for the organ doses, and hence for generating estimates for the aggregate 'whole body' doses used subsequently for the conversion coefficients. This choice was because limbed phantoms will give more accurate organ doses (especially RBM) than the truncated phantoms used by Lab2 and Lab3.
- The variations between the GEANT4 results from Lab1 and Lab4, which arise due to small differences in their approaches to the modelling, were used to provide estimates of the systematic uncertainty on each organ dose.
- The MCNP results from Lab3 have been adopted as the primary dataset for the dosemeter results, because the limbless phantoms were assumed better at recreating the attenuation and scatter conditions of the real measurements (Figure 5). Lab3 provided a fuller set of results than Lab2.
- Lab2 only provided results for P1. For each dosemeter on P1, the variation between the results from Lab2 and Lab3 was used to provide an estimate of its systematic uncertainty, which arises due to differences in the modelling and the shape of P2 on cross-scatter. In lieu of a better alternative, the average of those was then adopted as an estimate for the systematic uncertainty on the dosemeter results for the other phantoms.

- Lab2 and Lab3 used identical locations for their dosemeters: the impact of small variations in position cannot therefore be obtained from a comparison of their results. Instead, some indirect insight may be gained by comparing the results of Lab3 against Lab1 in cases where scatter and attenuation from the limbs is likely to have little impact. A predominantly frontal exposure of phones in the Chest location was assumed to meet this criterion, with those data for P1 and P3 therefore used as a first attempt at evaluating this variation.

For each dosemeter and phantom, the corresponding bespoke dose conversion coefficient, K_{GRB}, where the subscript again indicates the GI tract, Red bone marrow and Brain, was derived from the expression:

$$K_{GRB} = {}^{D_{X,Y,m}}\Big/_{D_{GRB}} = {}^{D_{X,Y,m}}\Big/_{\{(D_{SI} + D_{Stomach} + D_{RBM} + D_{Brain})/4\}} \qquad [2]$$

where $D_{X,Y,m}$ denotes the absorbed dose calculated for material *m* of phone *Y* from Laboratory *X*. For this exercise, the 'body dose' is associated with the aggregate quantity D_{GRB} that would have been received had the phantom been a real person, which again provides a useful metric for analyses and permits a direct comparison against the generic data discussed earlier. For completeness, however, bespoke conversion coefficients $K_B = D_{X,Y,m} / D_B$ to the alternative dose quantity D_B were also similarly defined, compared and discussed.

5.2 Results and Analyses

Using the raw data presented in (Kim *et al*, 2024), Table 4 shows the D_{GRB} and D_B evaluated for the four phantoms by Lab1 and Lab4; the statistical uncertainties on these doses were <2 % in all cases. Table 5 shows the resistor and glass doses calculated by Lab2 and Lab3 for phantom P1, and the resistor and glass doses calculated by Lab1 and Lab3 for those phones placed nominally in the Chest location on P1 and P3. Finally, bespoke dose conversion coefficients K_{GRB} and K_B are shown for resistors in Table 6 and for glass in Table 7, along with their uncertainties.

Table 4 demonstrates that Lab1 and Lab4 were generally consistent in their respective estimates of body doses (either D_{GRB} or D_B), differing on average by ~2 % and at most by less than 5 %, which was for P2. The exposure conditions of P2, with its partial shielding by P1, may anticipate the greatest sensitivity to the different choices and approximations that users might make when constructing their models. Overall, this average of 2 % may be associated with a rough assessment of the systematic uncertainty on whole body dose estimates due to user choices when nominally complete information is known about the exposure geometry.

Physical measurements of some of the organ doses were also achieved during the ILC by embedding sets of luminescence dosemeter pellets within them (Waldner *et al*, 2021; Kim *et al*, 2024). Not all organs could be included, so estimates of D_B are not possible, and organ doses in P3 were not measured due to limited resources during the field test, but from the remaining results it is nevertheless possible to derive measured values for D_{GRB} for P1, P2 and P4. These are also shown in Table 4, along with their negative and positive uncertainties. It is seen that the measured and modelled results broadly agreed for all three phantoms, though the uncertainties on the former are fairly large.

In general, the modelled values of D_{GRB} and D_B agreed fairly well with each other, with only small differences found in each case for P1, P3 and P4. However, larger differences were found for P2, with D_{GRB} around 40 % lower than D_B. One explanation for this is that a key organ for D_{GRB} is the stomach, which was chosen because it is sensitive for ARS. This organ is located towards the left side of the body, and hence was shielded by phantom P1 relative to those organs on its right side that do not

contribute to D_{GRB} but reduce the overall result when averaged for D_B. This hypothesis is further supported by the observation that the next greatest discrepancy between the two dose quantities was for P4: this phantom was predominantly LLAT exposed, so the stomach was a little closer to the source and less shielded than many of the other organs, raising D_{GRB} by 5 % relative to D_B. This observation indirectly supports earlier work advocating use of D_{GRB} in emergency dosimetry [Eakins *et al*, 2018b], rather than a more literal whole body dose quantity, due to its disregard for those organs that do not contribute significantly to ARS. That debate is far from resolved, however [ICRU, 2019].

Table 5 demonstrates that Lab2 and Lab3 were generally consistent in their respective estimates of doses for phones distributed around P1, with like-for-like data differing by ~3 % on average and by ~5 % at most. In all cases, the statistical uncertainties on the doses were small. This analysis provides an estimate of the systematic uncertainty within the modelling when factors such as the precise phone location and Monte Carlo code are harmonized.

The comparison between like-for-like resistor or glass doses ($D_{X,Y,m}$) calculated by Lab1 and Lab3 for those phones placed in the nominal Chest location on phantoms P1 and P3 found differences varying from ~9 % to ~21 %. Averaging all of those differences provides a general systematic uncertainty of ~13 % that may be attributed to factors such as divergencies in users' general approaches to the modelling, for example different approximations in defining the phone geometry or in choosing its precise position relative to the nominal location. This value is clearly not derived using any statistically rigorous approach, and will likely also be somewhat context dependent, so is difficult to extrapolate to other scenarios. The uncertainty will also include a component due to variations caused by using different Monte Carlo codes employing different physics data (e.g. cross-sections), though this contribution might reasonably be expected to be small.

Division of the data in Tables 4 and 5 is used to derive values for the bespoke conversion factors, i.e. $K_{GRB} = D_{X,Y,m} \,/\, D_{GRB}$ and $K_B = D_{X,Y,m} \,/\, D_B$. When the above analyses are taken together, overall uncertainties on these factors may also be obtained, as additionally shown in Tables 6 and 7. Specifically, combining the individual uncertainties in quadrature implies an average total uncertainty of ~14 %. So, if nominally complete information on the exposure geometry were available to Monte Carlo modellers, an uncertainty of 14 % might be appropriate for the resulting conversion factors, at least as a first approximation based on the current limited study. Of course, this figure represents a minimum: in reality, the available information will likely be incomplete, leading to higher uncertainties. The obvious statistical weakness of the analysis is also again emphasized (e.g. limited number of modellers, limited number of dosemeter locations etc.), so this figure should be interpreted only as broadly indicative.

6. Analyses of bespoke vs. generic datasets

6.1 Approach and raw results

Comparisons between K_{GRB} (Tables 6 and 7) and C_{GRB} (Tables 1 and 2), or between K_B and C_B, for a given material (resistor or glass) and phone, can be used to show the degree to which using generic conversion coefficients introduces errors into the dosimetry relative to the bespoke data, which are presumed to provide more accurate estimates of the nominal 'true' doses to individuals. These may be quantified by taking the ratios of the bespoke and generic data, defined as:

$$R(K/C)_{GRB} = {}^{K_{GRB}}/_{C_{GRB}} \qquad [3]$$

and similarly $R(K/C)_B = K_B / C_B$, where the values of C_{GRB} (and C_B) correspond to those resulting from the choices made by the laboratories during the exercise. The ratios are included in Tables 6 and 7.

6.2 Comparison of data

In most cases, $R(K/C)_{GRB}$ and $R(K/C)_B$ have similar values, as expected. Where larger divergencies occurred, it is likely again because the dose to a key organ (e.g. stomach) differed from those to other organs, such as is the case for P2. Because the phones are external to the body, the photons impinging on them are typically less attenuated than those through the internal organs, so $D_{X,Y,m} > D_{GRB}$ and hence generally $K_{GRB} > 1$; likewise $C_{GRB} > 1$, $K_B > 1$ and $C_B > 1$. The obvious exceptions are those few cases when the phone location 'opposed' the direction of the field, e.g. 'Back' for a frontal exposure.

Tables 6 and 7 show that $R(K/C)_{GRB}$ and $R(K/C)_B$ are also generally > 1, i.e., $K_{GRB} > C_{GRB}$ (or $K_B > C_B$). The explanation for this stems not just from the choices made by the participants, but also from the uniform exposure conditions underpinning the generic datasets compared to the non-uniformity within the bespoke modelling. Specifically, whilst both the generic and bespoke coefficients account for the variation in dose due to the location of the phone relative to the body, the latter also account for the additional variation in dose with position due to the divergence of the point source field. Accordingly, the range of differences between phone and body doses is typically larger for the bespoke data than for the generic data; and, because K_{GRB} and C_{GRB} are both > 1 in the majority of cases, the effect is that generally $K_{GRB} > C_{GRB}$, and hence $R(K/C)_{GRB} > 1$. Supporting this explanation is the observation that $K_{GRB} < C_{GRB}$ occurred in those few cases where the phone was maximally shielded by the body. An example of this was phone *L*5-2, which was located near the back right pocket of P4 and therefore shielded by the phantom when exposed from its left side. It received a dose $D_{X,Y,m}$ that was much lower than D_{GRB}, thereby giving a small value for K_{GRB}; but, that participant chose a generic conversion coefficient corresponding to an isotropic exposure, for which the phone dose is relatively independent of its location, leading to $D_{E,p,m} > D_{GRB}$ and hence $C_{GRB} > 1$. Taken together, the outcome was that $C_{GRB} > K_{GRB}$ for *L*5-2, and therefore $R(K/C)_{GRB} < 1$ as observed.

6.3 Implications for individual dose assessments

The values for $R(K/C)_{GRB}$ (and $R(K/C)_B$) indicate that, in general in the present exposures, use of the generic conversion coefficients instead of bespoke data would typically lead to over-estimates of the dose to the individual; in only two cases would an under-estimate instead have been caused, with one of those (*L*4-1) speculated previously as being a mistake. The over-estimates range from a few percent to a factor of 3.6 for resistors for D_{GRB} (up to 5.2 for D_B), with an average factor of 2.0 found for D_{GRB} (and 2.1 for D_B); similar values are found for glass.

These observations are somewhat open to interpretation. On the one hand, the argument for deriving bespoke conversion coefficients for retrospective dosimetry is clear: their application improves the accuracy of the dose estimates. On the other hand, the inaccuracy caused by using generic data may perhaps be acceptable in the wider context, especially considering the other large sources of uncertainty inherent to fortuitous dosimetry. Moreover, the use of the generic data would lead to conservative estimates of the doses to individuals in almost all of the cases considered here: whilst over-estimates are undesirable, they are preferable to the converse.

It is also evident that applying either set of conversion coefficients would lead to dose estimates that are closer to the organ doses than would be achieved by applying nothing: using the raw absorbed dose to the dosemeter as the only estimator of individual doses would lead to average and maximum over-estimates by factors of about 2.8 and 5.0 respectively for D_{GRB} (or 3.1 and 6.6 for D_B). More importantly perhaps, the dose to the person would be greatly under-estimated in some cases if only

the phone dose were reported without conversion. For example, phone *L*5-2 would assign a dose to an individual that is about half what they actually received.

6.4 Limitations of this study

The generic conversion coefficients were derived originally for use in large-scale scenarios, for which fairly uniform exposures are naturally presumed. Their application to a small-scale scenario featuring a highly heterogeneous exposure is therefore open to criticism. Nevertheless, it is plausible that laboratories might choose to apply generic data even in small-scale events, in lieu of anything else if the capability to generate bespoke data 'on the fly' was not available. Moreover, it is reasonable to question whether the above presumption and criticism are actually valid anyhow: is there even such thing as a 'typical' emergency dosimetry exposure scenario? Historically, large doses have arisen only in small-scale events affecting just a few unmonitored individuals; given that no analogous large-scale event has ever occurred, and the exact circumstances that might cause such a hypothetical scenario are unknown, it is not obvious how accurate any prediction could be regarding likely specifications.

7. Implications for measurements

7.1 Background and context

The goal of initial-phase ('triage') dosimetry is to support the making of decisions. In that vein, and putting it simplistically, the endeavour is to help make judgements of the general form: "*Result X may be associated with Person Y, therefore take course-of-action Z*". Nevertheless, there remains an open question of how to interpret and utilize the results from retrospective dosemeters. This is especially true when several dosemeters are available for use for the same individual. In such circumstances, the dosemeter results should ideally be consistent, in the sense that they would each imply the same general decision (e.g. reassure the individual and send home, versus refer onwards for urgent clinical attention). But what is the optimum interpretation when contradictions occur? Should perhaps their average be the result that is used to inform initial-phase dosimetry? Or, should the highest dose estimate be quoted: this may be more desirable as it has greater conservatism, but could also lead to false positives.

These questions cannot be answered by the current work. Indeed, for retrospective dosimetry using phones it is reasonable to assume that individuals would generally carry no more than one, so the above issue is relevant mainly with regards to what to do when resistor and glass results are both obtained from it. However, the ILC multiple-phone exposures may nevertheless usefully be interpreted in two different ways:

1. As a series of parallel scenarios, in which a phantom represents a set of individuals who received identical exposures (i.e. the same 'body dose') but who have different retrospective dosemeter estimates due to the different locations of their respective single phones;
2. As a scenario in which each phantom represents an individual who has several dosemeter estimates due to having several about their person.

The first of these has been the starting point so far in this paper, but some contributions towards the above general questions may be gained from also considering the latter.

Dose estimates accurate to just one significant figure are typically adequate for initial-phase dosimetry. In fact, the MULTIBIODOSE project went further by proposing the application of a simple colour-coded dose categorization scheme in the event of a mass-casualty scenario (Jaworska et al., 2014; ICRU, 2019):

- < 1Gy, "*unlikely to develop symptoms of acute radiation syndrome (ARS), no immediate care required"* (GREEN);
- 1-2 Gy, "*may experience mild or delayed ARS symptoms, follow-up care may be necessary"* (AMBER);
- >2 Gy "*moderate-to-urgent care may be required*" (RED).

It is therefore of interest to see how the converted and unconverted data from the ILC measurements (which will be summarized fully in an upcoming publication) might fare within these frameworks, especially from the perspectives of multiple dose estimates for the same individual or different single dose estimates for multiple individuals.

7.2 Scaled results and analyses

Previous results (Table 4) demonstrated that the values of D_{GRB} calculated by Lab1 were similar to the values calculated for D_B, as well as those measured using luminescence pellets embedded within the phantoms, especially if only nearest-gray doses are of interest. Those D_{GRB} values may therefore be interpreted as the nominally 'true' doses received by the individuals, at least for the purposes of the following analyses. These D_{GRB} doses would all fall trivially into the low-dose category, according to the MULTIBIODOSE scheme. To make the analysis more insightful, it is therefore imagined that the exposures lasted 5.5 times longer than they actually did, with all results hence multiplied by an artificial scaling factor of 5.5. The scaling factor was selected somewhat arbitrarily, to simulate an exposure situation in which the MULTIBIODOSE scheme would lead to resolvable differences in the dose categorizations. This choice was considered particularly interesting because multiplying all results by 5.5 means that P2 would then receive a low dose, P3 an intermediate dose, P4 only just a high dose, and P1 a dose that is well within the high dose category. In turn, it is insightful to see how well (or otherwise) the scaled phone doses then match these categorizations, and would hence lead to a 'correct triaging' of the individuals. Of course, higher doses would likely also reduce measurement uncertainties in practice, but this is ignored in this illustrative extrapolation.

Table 8 shows:

- The scaled body doses (D_{GRB}) calculated by Lab1;
- The scaled phone doses measured by the laboratories (i.e. the unconverted data);
- The scaled body dose estimates obtained by applying the laboratories' respective choices of generic conversion coefficient (Tables 1 and 2) to the scaled measured data;
- The scaled body doses implied by instead applying the bespoke conversion coefficients (Tables 6 and 7) to the scaled measured data.

Data for both resistors and glass are presented, where available. The results are colour coded, with green showing correct dose categorization, light/dark blue an over-estimate by one/two dose categories, and light/dark red an under-estimate by one/two categories; the last four are undesirable for initial-phase dosimetry, but the latter have more serious implications. The measurement uncertainties quoted by the laboratories are also shown; these sometimes imply a range that straddles more than one dose category. For clarity, the uncertainties discussed earlier for the dose conversion process are not shown; whilst insightful, it is presumed that these would not change dose categorization in a real scenario, so would complicate the analysis presented here. The statistical uncertainties on D_{GRB} (all < 1%) are also omitted.

At first glance, one observation seems to be that conversion coefficients do not necessarily improve correct dose categorization: the ratio of successes to failures is broadly similar across all columns. Where mis-categorizations have occurred, they cannot necessarily be attributed to errors in the conversion factors or the quoted measurement uncertainties. An example is phone *L8-1*, for which all converted and unconverted doses from resistors exceeded the body dose, whilst all results from glass under-estimated it; but, the measured resistor and glass doses were quoted with uncertainties of

~21 % and ~10 % respectively, which are insufficient to explain these discrepancies. The suggestion is therefore that there must be 'hidden' systematic uncertainties on some of the measured data, such as human error, misapplication of measurement protocols, unexpected signal loss during shipping, wide variations of sample response, etc. Some of the phones also received doses close to the detection limits of the measurement protocols. Scaling the results is hence also likely scaling biases that are greater than would have arisen had the doses genuinely been large. The data in Table 8 should therefore only be considered indicative, with more research required.

The unconverted data typically over-estimate the 'true' body doses by a much larger extent than the converted data. This is partly because, in almost all cases, $C_{GRB} > 1$ and $K_{GRB} > 1$. However, the effects of over-estimation are somewhat hidden in the dose categorization, especially in the highest level. For example, all the results for *L9-1* are in the correct high dose category, but the unconverted values are nearly 4 times larger than the body dose, whilst the estimates using bespoke conversion coefficients are within ~20 % of it. If the results from retrospective dosemeters are to be used only as binary 'flags' of whether an exposure has or has not occurred, serious over-estimates may not be considered a problem. Inevitably, however, such a scheme could lead to limited medical resources being directed towards individuals who do not need them. The pros and cons of this approach are therefore debates to be had (cf. ICRU, 2019).

For some of the results, in particular those close to 'boundaries', using a different scaling factor could have led to better or worse agreement with the nominally true dose category; the current data should therefore be considered just generally illustrative. However, because the unconverted doses typically over-estimated by the greatest extent, varying the scaling factor would have disproportionate effects on the different datasets. In all cases where over-estimates occurred, choosing an increasingly smaller scaling factor (i.e. a shorter exposure time) would have first made the bespoke results agree with the correct dose category, then the generic data, and finally the unconverted data. On the other hand, for *L6-2* for example, using a slightly larger scaling factor of 6.5 would have led to the converted results still matching the correct dose category for P3, but the unconverted results then over-estimating it.

Whilst the scaling factors were chosen somewhat artificially to imply a non-trivial range of body doses, they give insight into how often dose categorizations based on phones might match dose categorizations of individuals. This provides some indication of how effective phone dosimetry could be in leading to the 'correct' triage decision. Overall, applying bespoke factors may yield the most successes, followed by applying generic data, and then lastly using the unconverted doses.

7.3 Unscaled results and analyses

Table 9 shows the original unscaled data reordered by phantom, to better demonstrate the variations in dose assessments from the different phones. Specifically, it shows, for each set of four or five phones on each phantom:

- Their mean unconverted and converted (generic and bespoke coefficients) doses;
- The respective standard deviations of the individual dose estimates about those means;
- Their corresponding coefficients of variation (CoV);
- The mean responses, derived from the mean doses normalized to the nominally true D_{GRB} values calculated by Lab1.

Data obtained for resistors and glass are shown separately, as well as the arithmetic means of the results for both resistors and glass combined. Full uncertainty budgets have been omitted for clarity, noting again that some systematic uncertainties are likely hidden from those quoted.

Table 9 demonstrates that the mean response is generally improved by applying generic conversion coefficients, and improved even more by bespoke conversion coefficients. This is true for both

resistors and glass. However, it should be mentioned that the data used in Table 9 are somewhat arbitrary, especially the estimated values in the unconverted dataset, which can vary significantly depending on the selection of the ILC phone locations. On the other hand, the values of the generic and bespoke converted doses are less sensitive to this choice, leading (in principle, at least) to results closer to the 'true' body dose.

Ideally, all dosemeters located around the same 'person' would lead to consistent estimates of their dose. Instead, a spread of results is found for each phantom. For example, variations of several tens of percent are seen for each of P1, P2, P3 and P4, based on the CoVs of the individual dose estimates about their respective means. A large range is expected for the unconverted data, because the different locations of the phones typically lead to different exposures. But distributions are also seen in the converted data, for which attempts have been made to mitigate for this effect, although their CoVs are generally lower, indicating some success. The CoVs may be interpreted as providing a first handle on the minimum standard uncertainties from using multiple assays in assessing doses for the four 'individuals'. These range from ~20% (for glass, bespoke, P1) to ~80% (resistors and glass, unconverted, P2).

7.4 Observations and implications

Spreads in the converted data do not automatically imply that the approach is flawed: the large uncertainties on the measured data are significant factors. As an illustration, it is recalled that measured dose estimates from phones can still vary greatly when located adjacently, and even resistor and glass doses from the same phone measured by the same laboratory are not always consistent (e.g. *L8-1* (Table 8)). Indeed, the spreads in the datasets caused by inaccuracies in the measurement techniques are typically larger than those arising due to the spatial distribution of phones about the phantoms, and cannot therefore be removed by using conversion coefficients. This is an important observation for retrospective dosimetry endeavours: whilst dose conversion would generally lead to more accurate assessments of individuals' true doses, even with its application the results obtained from phones do not always provide reliable estimates of the exposures actually received. On the other hand, it is again recalled that some of the measured doses were close to the detection limits of the samples, so some variance is potentially due to particularly large measurement errors.

The large standard deviations, and the potential for dose averaging, have other implications. Taking the mean can hide instances where under- and over-responses both occur within the same dataset; the consequences of this may differ depending on whether multi-assay analyses are available, and could have either positive or negative impacts. For example, consider that the unconverted result for resistors in *L9-2* was roughly half the body dose for P1, whilst the *L3-1* result was roughly twice it, so neither provides reliable dosimetry on its own. If these results corresponded to two different individuals, each exposed identically and each with their own single phone dose assessment, the two estimates could lead to both individuals receiving incorrect dose categorizations: the first an under-estimate, the second an over-estimate. Conversely, if the results instead related to one individual with two dose estimates, obtained independently using two different phones, the advantage of averaging to inform the subsequent decision making process is clear.

Even when a multi-assay approach is an option, increasing the size of the data pool does not necessarily improve the quality of inferences drawn from it. Whilst some outlier results may be mitigated by averaging, the inclusion of extreme values with large systematic uncertainties may be detrimental elsewhere. For example, averaging all five unconverted resistor results for P1 leads to a mean dose that is closer to the body dose than four of those individual results were; but, extending the averaging to also include the P1 glass results leads to a worse estimate. This leads naturally to the question: could a real retrospective dosimetry system incorporate a reliable method for excluding

outliers? Or, instead of adopting a 'the-more-data-the-better' perspective, might it be preferable to disregard some retrospective dosimetry methods in favour of those with greater potential for accuracy? These issues are left open , as part of wider questions regarding the optimal use of fortuitous dosimetry results during real emergencies.

8. Summary and Conclusions

The main results and outcomes from this work may be summarized as follows:

- Ignorance of the exposure conditions can significantly impact results when conversion coefficients are applied. Moreover, even if nominally complete information is available, the choice of phone position and exposure type is still subjective: in the simple comparison exercise, the organizer and participants came to the same decision in only a minority of cases.

- An uncertainty of up to ~70 % might be appropriate due to subjectivity when choosing generic coefficients. This could potentially be reduced to <30 % if adequate systems of quality control were in place. Of course, these estimates are highly heuristic and based on just one small study.

- For the uniform exposures for which the generic coefficients were derived, the choice of 'whole body' dose quantity had only modest impact (Tables 1-2). Conversely in the bespoke modelling, whilst D_{GRB} and D_{B} showed some agreement when the phantoms were exposed fairly homogeneously, they differed by several tens of percent when part of the body was shielded (Tables 6-7). The need for further study on this topic is thus reinforced.

- When nominally complete information is available on the exposure conditions, the uncertainty on a modelled phone dose was found to be a few percent. However, factors such as the modelling approach and exact placement of the phone inevitably varied from one modeller to another, leading to a larger average systematic uncertainty of about 13 %.

- Taken together, the combined systematic uncertainty on the bespoke conversion factors was about 14 %. However, this represents a minimum: in reality, the available information will likely be incomplete, leading to higher uncertainties.

- The ratios between the bespoke conversion coefficients and the generic data implied that, in almost all cases, the nominally 'true' dose to the individual would be over-estimated by using the latter, by an average factor of about 2. This compares with an average over-estimate of about 3 if no conversion coefficient were applied. In some cases, however, use of the wrong or no conversion coefficient led to serious under-estimates.

- The above Monte Carlo analyses suggest a critical need for dose conversion in retrospective dosimetry. However, the advantages were less clear-cut when applied to measured data from the ILC. Specifically, whilst conversion still led to better dose assessments, and therefore improved dose categorization of individuals, those benefits were eroded by the large systematic uncertainties associated with the measurements themselves.

- Ideally, all dosemeters distributed around the same 'person' would report an identical dose. However, variations of several tens of percent were instead seen. Again, inconsistencies in the measured data were the likely cause. Underestimating the uncertainties on the bespoke

conversion coefficients may also be a factor, especially in highly inhomogeneous radiation fields.

Some of the results presented in this study are specific to the current ILC scenario, with the findings from the simulations applicable just to the small-scale exposure conditions modelled. In many cases, however, they may still be considered extrapolatable, if not numerically then at least heuristically. That is, if the work were repeated for an alternative set-up, the types of difficulties and divergences found here would likely still persist, even though the absolute doses and dose ratios could change. As examples: different people in different laboratories would likely still make different subjective choices; the phone location would still be a key factor for the dose received; doses from resistors and glasses in the same phone may not always agree with each other; and phones located around the same individual would not necessarily harmonize to a consistent single dose estimate for them, although the use of conversion coefficients would likely still improve that. Nevertheless, the current study would clearly benefit from repeating, and the results specific to this ILC would be enhanced by considering additional exposure scenarios and conditions.

As a first exploration of any of these issues, the current study provides some useful insight into some of the sources of 'human' and other errors that would arise in practice from using dose conversion coefficients in retrospective dosimetry. In that regard, the value of 14 % found for the uncertainty on the bespoke conversion factors is too specific to generalize, with a one significant figure estimate of 20 % perhaps representing a more realistic estimate. Such 'rounding up' is valid given that the largest component of this uncertainty, which originated from ambiguity in the precise phone position relative to its nominal location site (e.g. 'Chest'), is likely somewhat context dependent and therefore hard to extrapolate.

The uncertainty of ~20 % on the bespoke conversion coefficients, and ~30 % from the subjective selection of generic data (assuming suitable quality control checks etc.), corresponded to a situation in which modellers had complete knowledge about the exposure conditions. These values therefore provide a first estimate of likely minimum uncertainties. In more realistic circumstances, where only incomplete information is available, these uncertainty components will inevitably be higher. However, their values were small compared to the systematic uncertainties induced by applying generic data to this highly bespoke exposure, which could cause individuals' doses to be mis-assessed by several hundreds of percent. This observation reinforces the need for bespoke conversion coefficients in small-scale scenarios (ICRU, 2019). Moreover, the values are even less significant relative to not making any dose conversion at all, and instead just reporting the physical dosemeter result as the nominal dose to the individual, which in this case led to larger errors than those arising through the use of generic data.

The outcomes discussed above were based on analyses of modelled data. A different pattern of results was found (Tables 8 and 9) when the conversion coefficients were instead applied to the real phone measurements from the ILC. Overall, the unconverted measured data over-estimated the nominally true body doses by an average factor of 3.0, the generic-converted measured data over-estimated by an average of 1.9, and the bespoke-converted measured data over-estimated by an average of 1.5. The analogous mean CoVs, which were interpreted as first estimates of the minimum uncertainties on the individuals' dose assessments, were about 60 % for the unconverted data and about 50 % for both the generic- and bespoke-converted data. The potential advantages of dose conversion are therefore apparent. However, such direct comparisons neglect the various statistical and systematic uncertainties, as well as other human-related factors , that arise from using conversion coefficients. When those are also included, the overall uncertainty budgets are broadly similar in all three approaches.

So, whilst in principle individuals' dose assessments are improved by using conversion factors, in practice the overall accuracy is not greatly enhanced, because ultimately the reliability of phone dosimetry appeared dominated by the uncertainties and inconsistencies in the measurements. Of course, this conclusion is based solely on the current ILC, which featured a small-scale and greatly heterogeneous exposure scenario, so caution should be applied before extrapolating to phone dosimetry in the wider context. Conversely, however, the current study was performed under idealized conditions of maximal information about the exposure scenarios, which is an unlikely luxury in many real emergency circumstances (Eakins and Kouroukla, 2014; ICRU, 2019). Moreover, over-estimates of any magnitude may still be considered acceptable if it were decided that only dose/no-dose binary flagging is required for initial-phase dosimetry. The ultimate question of whether highly uncertain data are better than no data is therefore left open for future debate.

Acknowledgments

This international cooperation was conducted within task group 10.8 of Working Group 10 of the European Radiation Dosimetry Group (EURADOS). The authors are grateful for the support of that group, especially those involved with setting-up the ILC and those who contributed data for the results shown in Sections 4 and 7.

References

Ademola J and Woda C (2017). Thermoluminescence of electronic components from mobile phones for determination of accident doses. Radiat. Meas., 104, 13–21. http://dx.doi.org/10.1016/j.radmeas.2017.07.002

Agostinelli S et al. (2003). GEANT4 – A Simulation Toolkit. Nucl. Instrum. Meth. A 506, 250-303.

Bassinet C, Pirault N, Baumann M, Clairand I (2014a). Radiation accident dosimetry: TL properties of mobile phone screen glass. Radiat Meas., 71, 461–5. http://dx.doi.org/10.1016/j.radmeas.2014.03.025

Bassinet C, Woda C, Bortolin E, Della Monaca S, Fattibene P, Quattrini MC, et al (2014b). Retrospective radiation dosimetry using OSL of electronic components: Results of an inter-laboratory comparison. Radiat Meas., 71, 475–9.

Bernhardsson C, Christiansson M, Mattsson S, Rääf CL (2009). Household salt as a retrospective dosemeter using optically stimulated luminescence. Radiat Environ Biophys., 48(1), 21–8.

Chandler JR, Sholom S, McKeever SWS, Hall HL (2019). Thermoluminescence and phototransferred thermoluminescence dosimetry on mobile phone protective touchscreen glass. J Appl Phys., 126, 74901.

Chandler JR, Sholom S, McKeever SWS, Bakhanova E, Chumak V, Velásquez D, Hall HL (2022). Dose conversion factors for absorbed dose in a mobile phone to absorbed dose in critical organs in an anthropomorphic phantom for emergency dosimetry applications: OSL and TL experimental results, and Monte Carlo simulations. Radiat. Meas., 154, 106781.

Discher M, Woda C (2013). Thermoluminescence of glass display from mobile phones for retrospective and accident dosimetry. Radiat Meas., 53–54, 12–21. http://dx.doi.org/10.1016/j.radmeas.2013.04.002

Discher M (2015). MCNP simulations of a glass display used in a mobile phone as an accident dosimeter. Radiat. Meas. 75, 21 – 28.

Discher M, Bortolin E, Woda C. (2016). Investigations of touchscreen glasses from mobile phones for retrospective and accident dosimetry. Radiat. Meas., 89, 44–51.

Discher M, Woda C, Ekendahl D, Rojas-Palma C, and Steinhäusler F. (2021a). Evaluation of physical retrospective dosimetry methods in a realistic accident scenario: results of the CATO field test. Radiat. Meas. 142, 106544

Discher M, Eakins J, Woda C and Tanner R (2021b). Translation of the absorbed dose in the mobile phone to organ doses of an ICRP voxel phantom using MCNPX simulation of an Ir-192 point source. Radiation Measurements, 146, 106603. Doi: https://doi.org/10.1016/j.radmeas.2021.106603.

Eakins J and Kouroukla E. (2015). Luminescence-based retrospective dosimetry using Al2O3 from mobile phones: a simulation approach to determine the effects of position. Journal of Radiological Protection 35, 343-381. https://doi.org/10.1088/0952-4746/35/2/343

Eakins J, Hager L, Kouroukla E, Smith R and Tanner R (2016). The PHE fortuitous dosimetry capability based on optically stimulated luminescence of mobile phones. Radiation Protection Dosimetry Vol. 170 No. 1-4, 412-415. https://doi.org/10.1093/rpd/ncv520

Eakins J and Ainsbury E (2018a). Quantities for assessing high doses to the body: a short review of the current status. Journal of Radiological Protection 38, 731-742. https://doi.org/10.1088/1361-6498/aabffe

Eakins J and Ainsbury E. (2018b). Quantities for assessing high photon doses to the body: a calculational approach. Journal of Radiological Protection 38, 743-762. https://doi.org/10.1088/1361-6498/aabffc

Ekendahl D and Judas L (2011). NaCl as a retrospective and accident dosemeter. Radiat Prot Dosimetry, 145(1), 36–44.

Ekendahl D. and Judas L. (2012). Retrospective dosimetry with alumina substrate from electronic components. Radiation Protection Dosimetry Vol. 150 No. 2, 134-141. https://doi.org/10.1093/rpd/ncr380

Endesfelder D, Oestreicher U, Kulka U, Ainsbury E, Moquet J, Barnard S, Gregoire E, Martinez J, Trompier F, and Ristic Y (2021). RENEB/EURADOS Field Exercise 2019: Robust Dose Estimation under outdoor Conditions based on the Dicentric Chromosome Assay. Int. J. Radiation Biology, 1-48.

Fattibene P and Callens F (2010). EPR dosimetry with tooth enamel: A review. Appl Radiat Isot., 68(11), 2033–116. http://dx.doi.org/10.1016/j.apradiso.2010.05.016

Geber-Bergstrand T, Bernhardsson C, Christiansson M, Mattsson S, Rääf CL (2018). Optically stimulated luminescence (OSL) dosimetry in irradiated alumina substrates from mobile phone resistors. Radiat Environ Biophys., 57(1), 69–75.

Göksu HY (2003). Telephone chip-cards as individual dosemeters. Radiat Meas., 37(6), 617–20.

Inrig EL, Godfrey-Smith DI, Khanna S. (2008). Optically stimulated luminescence of electronic components for forensic, retrospective, and accident dosimetry. Radiat Meas., 43(2–6), 726–30.

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

International Commission on Radiological Protection (ICRP), (2007). The 2007 Recommendations of the International Commission on Radiological Protection. ICRP Publication 103.

International Commission on Radiological Protection (ICRP), (2009). Adult Reference Computational Phantoms. ICRP Publication 110.

International Commission on Radiological Protection (ICRP), (2010). Conversion Coefficients for Radiological Protection Quantities for External Radiation Exposures. ICRP Publication 116.

International Commission on Radiological Protection (ICRP), (2020). Adult mesh-type reference computational phantoms. ICRP Publication 145.

International Commission on Radiation Units (ICRU) (2019). Methods for Initial-Phase Assessment of Individual Doses Following Acute Exposures to Ionizing Radiation. ICRU Report No. 94, Bethesta, MD.

Jaworska A, Ainsbury E A, Fattibene P, Lindholm C, Oestreicher U, Rothkamm K, Romm H, Thierens H, Trompier F, Voisin P, Vral A, Woda C and Wojcik A (2014). Operational guidance for radiation emergency response organisations in Europe for using biodosimetric tools developed in EU MULTIBIODOSE project, Radiat. Prot. Dosim. 164, 165–169.

Kim H, Kim MC, Lee J, Chang I, Lee SK, Kim JL (2019a). Thermoluminescence of AMOLED substrate glasses in recent mobile phones for retrospective dosimetry. Radiat Meas., 122, 53–6. https://doi.org/10.1016/j.radmeas.2019.01.004

Kim M C, Kim H, Han H, Lee J, Lee S K, Chang I, Kim J-L and Kim C H (2019b). A study on dose conversion from a material to human body using mesh phantom for retrospective dosimetry. Radiation Measurements 126, 106126. https://doi.org/10.1016/j.radmeas.2019.106126

Kim H, Kim MC, Lee J, Discher M, Woda C, Lim S, et al (2020). Characterization of thermoluminescence of chip cards for emergency dosimetry. Radiat Meas., 134, 106321. https://doi.org/10.1016/j.radmeas.2020.106321

Kim H *et al* (2024). Monte Carlo reference dosimetry for the EURADOS WG 10 and RENEB field test of retrospective dosimetry techniques in a realistic exposure scenario. SUBMITTED TO RADIATION MEASUREMENTS, 2024. DOI: 10.14293/PR2199.001104.v1. https://www.scienceopen.com/document/read?id=13ec5637-5f5a-4d5d-b487-aef9df7f2eee

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

Lee JI, Chang I, Pradhan AS, Kim JL, Kim BH, Chung KS (2015). On the use of new generation mobile phone (smart phone) for retrospective accident dosimetry. Radiat Phys Chem. 116, 151–4.

López M, Berkovskyy V, Ratia G, Challeton-de-Vathaire C, Davesne E, Eakins J, Franck D, Giussani A, Gregoratto D, Hernandez C, Kulka U, Mafodda A, Marsh J, Navarro J, Oestreicher U, Pérez B, Sierra I and Woda C. (2020) Improvements in individual dose measurement techniques following nuclear emergencies. Radioprotection, 55(HS1), S89–S93. https://doi.org/10.1051/radiopro/2020017 . See also CONFIDENCE project website: https://resy5.iket.kit.edu/CONFIDENCE/index.php

McKeever S, Sholom S and Chandler J (2020). Developments in the use of thermoluminescence and optically stimulated luminescence from mobile phones in emergency dosimetry. Radiation Protection Dosimetry, Vol. 192, No. 2, pp. 205–235 (doi:10.1093/rpd/ncaa208).

Moquet J, Barnard S, Staynova A, Lindholm C, Monteiro Gil O, Martins V, et al (2017). The second gamma-H2AX assay inter-comparison exercise carried out in the framework of the European biodosimetry network (RENEB). Int J Radiat Biol., 93(1), 58–64.

Mrozik A, Marczewska B, Bilski P, Kłosowski M. (2014). Investigation of thermoluminescence properties of mobile phone screen displays as dosimeters for accidental dosimetry. Radiat Phys Chem., 104, 88–92.

Mrozik A, Marczewska B, Bilski P, Książek M (2017). OSL signal of IC chips from mobile phones for dose assessment in accidental dosimetry. Radiat Meas., 98, 1–9.

National Nuclear Data Center (NNDC) (2020). NNDC Chart of nuclides. Brookhaven National Laboratory website: https://www.nndc.bnl.gov/nudat2/

Polozov S, Cruz-Garcia L, Badie C. (2019). Rapid Gene Expression Based Dose Estimation for Radiological Emergencies. Radiat. Prot. Dosim., 186(1), 24–30.

Rojas-Palma C, Woda C, Discher M, and Steinhäusler F. (2020). On the use of retrospective dosimetry to assist in the radiological triage of mass casualties exposed to ionizing radiation. J. Radiol. Prot. 40. 1286–1298.

Sholom S and McKeever SWS (2016). Integrated circuits from mobile phones as possible emergency OSL/TL dosimeters. Radiat Prot Dosimetry, 170(1–4), 398–401.

Spooner NA, Smith BW, Creighton DF, Questiaux D, Hunter PG (2012). Luminescence from NaCl for application to retrospective dosimetry. Radiat Meas., 47(9), 883–9. http://dx.doi.org/10.1016/j.radmeas.2012.05.005

Van Hoey O, Römkens D, Eakins J, Kouroukla E, Discher M and Vanhavere F. (2021). Uncertainty evaluation for organ dose assessment with optically stimulated luminescence measurements on mobile phone resistors after a radiological incident. Rad. Meas. 141. 106520. https://doi.org/10.1016/j.radmeas.2021.106520

Waldner L, Woda C, Abend M, Bassinet C, Bernhardsson C, Discher M, Eakins J, Endesfelder D, Gregoire E, Van Hoey O, Kim H, Kim M C, Kulka U, Lee J, Oestreicher U, Ristic Y, Rääf C, Trompier F, Wojcik A, Yu H and Ainsbury E. (2021). The 2019-2020 EURADOS WG10 and RENEB field test of retrospective dosimetry methods in a small-scale incident involving ionising radiation. Radiat Res. 195(3), 253-264. Doi: 10.1667/RADE-20-00243.1

Werner C J, *et al*. (2018), MCNP6.2 Release Notes, Los Alamos National Laboratory, report LA-UR-18-20808.

Yeom YS, Han MC, Choi C, Han H, Shin B, Furuta T, and Kim CH (2019). Computation speeds and memory requirements of mesh-type ICRP reference computational phantoms in Geant4, MCNP6, and PHITS. Health Physics, 116(5), 664-676.

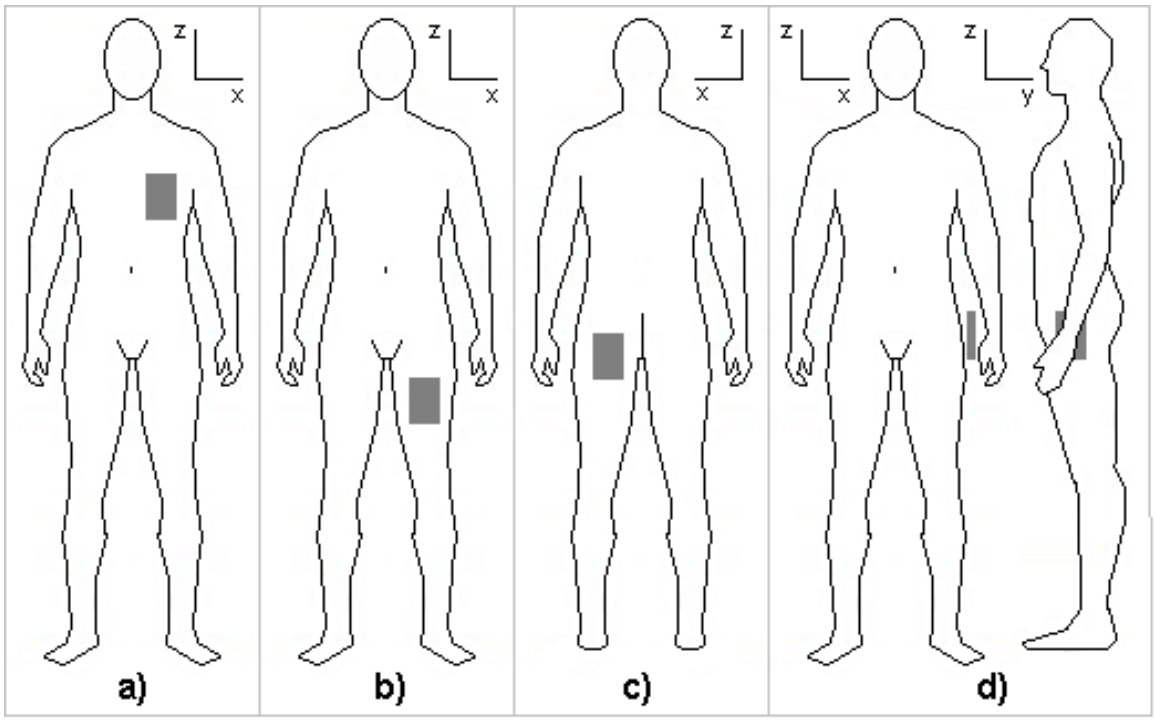

Figure 1 Illustration of the approximate locations of the phones (grey rectangles) in the four geometries of interest: *a*) Chest, *b*) Leg, *c*) Back and *d*) Hip. *(Figure reproduced from (Eakins and Kouroukla, 2015))*

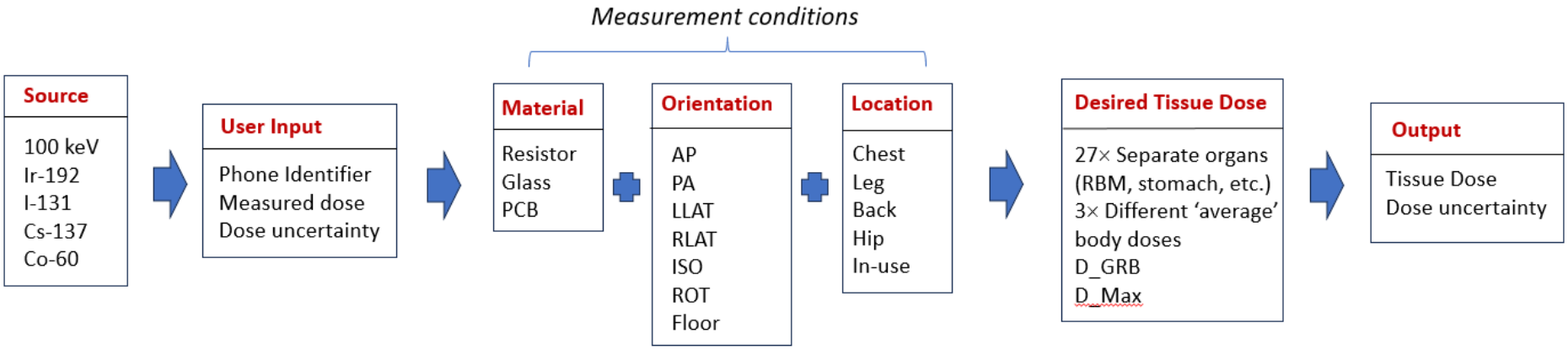

Figure 2 Flow-diagram schematic of the dose conversion spreadsheet developed within EURADOS WG10.

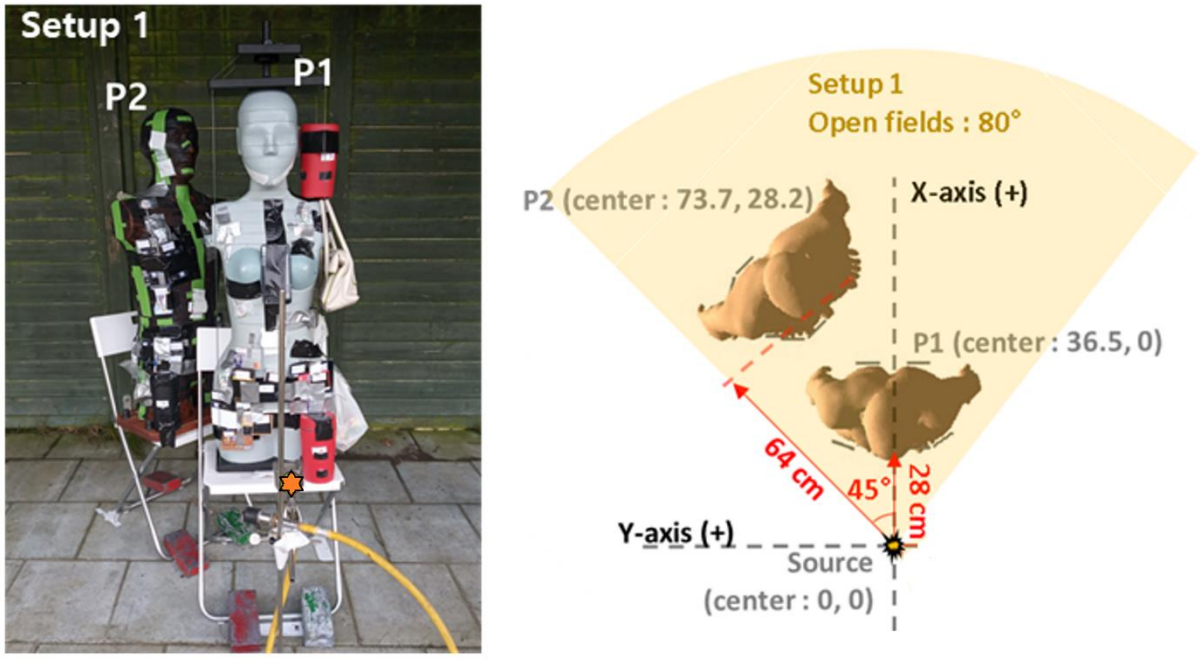

Figure 3 Schematic of Setup 1. Phantom P1 was a female phantom, and P2 was a male phantom. The star indicates the position of the ^{192}Ir source. Phantom P1 was facing the source, whilst P2 was facing P1. *(Figure reproduced from (Kim et al, 2024))*

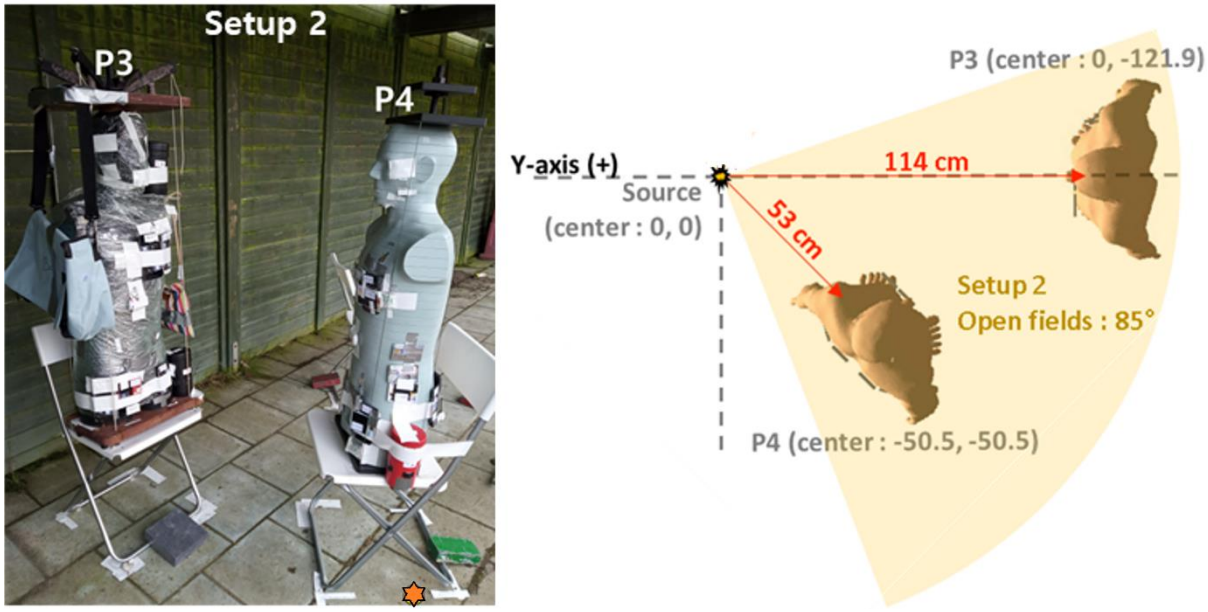

Figure 4 Schematic of Setup 2. Phantoms P3 and P4 were both male phantoms. The star indicates the position of the ^{192}Ir source. Phantom P3 was facing the source, whilst P4 was facing P3. *(Figure reproduced from (Kim et al, 2024))*

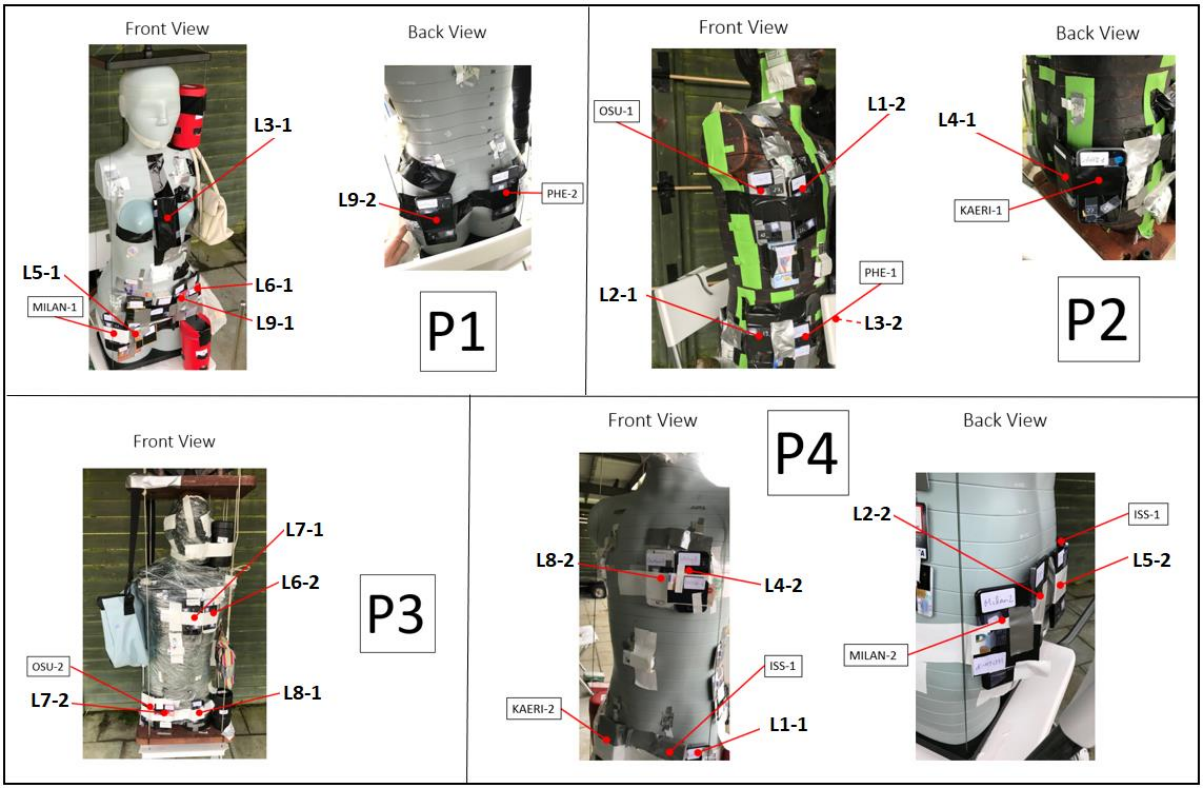

Figure 5 Locations of the retrospective dosemeters around the four phantoms. All phones from the ILC are shown for completeness, but not all laboratories also participated in the dose conversion exercise, so some do not have the specific identifiers used in the current paper.

Table 1 Resistor-based conversion coefficients. Phone location and exposure geometries chosen by the organizer (O) and nine participating laboratories (L), and the ratio ($R(L/O)_{GRB}$) of their associated conversion coefficients (C_{GRB}) for resistors (r). Analogous ratios ($R(L/O)_B$) using alternative conversion coefficients to the dose quantity D_B are also shown. The uncertainties on the data are small (few %), and have been omitted.

Phone Identifier	Organizer (O)			Laboratory (L)			Ratio $R(L/O)_{GRB}$	Ratio $R(L/O)_B$
	Exposure Geometry	Phone Location	C_{GRB} (Gy/Gy)	Exposure Geometry	Phone Location	C_{GRB} (Gy/Gy)		
L1-1_r	LLAT	Leg	1.31	LLAT	Hip	1.68	1.28	1.29
L1-2_r	AP	Chest	1.37	AP	Chest	1.37	1.00	1.00
L2-1_r	AP	Hip	1.08	A,P,L,R,Rot	Hip	0.1-1.68	0.09-1.56	0.09-1.98
L2-2_r	LLAT	Back	1.47	A,P,L,R,Rot	Back	0.45-1.71	0.31-1.16	0.24-0.93
L3-1_r	FLOOR	Chest	1.59	ROT	Chest	1.17	0.74	0.82
L3-2_r	AP	Leg	1.35	ROT	Hip	1.11	0.82	0.88
L4-1_r	AP	Back	0.45	PA	Back	1.71	3.80	3.90
L4-2_r	LLAT	Chest	1.34	AP	Chest	1.37	1.02	0.80
L5-1_r	FLOOR	Leg	1.88	ISO	Hip	1.22	0.65	0.70
L5-2_r	LLAT	Back	1.47	ISO	Back	1.21	0.82	0.67
L6-1_r	FLOOR	Leg	1.88	AP	Leg	1.35	0.72	0.74
L6-2_r	AP	Chest	1.37	AP	Chest	1.37	1.00	1.00
L7-1_r	AP	Chest	1.37	AP	Chest	1.37	1.00	1.00
L7-2_r	AP	Leg	1.35	AP	Leg	1.35	1.00	1.00
L8-1_r	AP	Leg	1.35	Floor	Hip	1.69	1.25	1.20
L8-2_r	LLAT	Chest	1.34	Floor	Chest	1.59	1.19	0.89
L9-1_r	FLOOR	Leg	1.88	AP	Leg	1.35	0.72	0.74
L9-2_r	FLOOR	Back	1.77	AP	Back	0.45	0.25	0.26

Table 2 Glass-based conversion coefficients. Phone location and exposure geometries chosen by the organizer (O) and nine participating laboratories (L), and the ratio ($R(L/O)_{GRB}$) of their associated conversion coefficients (C_{GRB}) for glass (g). Analogous ratios ($R(L/O)_B$) using alternative conversion coefficients to the dose quantity D_B are also shown. Dosimetry on glass was not performed for phones L1-1, L1-2, L2-1 or L9-2; L2-2 used a range of conversion data for resistors (shown in italics). The uncertainties on the data are small (few %), and have been omitted.

Phone Identifier	Organizer (O)			Laboratory (L)			Ratio $R(L/O)_{GRB}$	Ratio $R(L/O)_B$
	Exposure Geometry	Phone Location	C_{GRB} (Gy/Gy)	Exposure Geometry	Phone Location	C_{GRB} (Gy/Gy)		
L1-1_g	LLAT	Leg	-	-	Hip	-	-	-
L1-2_g	AP	Chest	-	-	Chest	-	-	-
L2-1_g	AP	Hip	-	-	Hip	-	-	-
L2-2_g	LLAT	Back	1.12	A,P,L,R,Rot	Back	*0.45-1.71*	*0.40-1.53*	*0.31-1.21-*
L3-1_g	FLOOR	Chest	1.68	ROT	Chest	1.26	0.75	0.84
L3-2_g	AP	Leg	1.47	ROT	Hip	1.15	0.78	0.83
L4-1_g	AP	Back	0.48	PA	Back	1.78	3.71	3.78
L4-2_g	LLAT	Chest	1.27	AP	Chest	1.52	1.20	0.93
L5-1_g	FLOOR	Leg	2.01	ISO	Hip	1.38	0.69	0.75
L5-2_g	LLAT	Back	1.12	ISO	Back	1.32	1.18	0.96
L6-1_g	FLOOR	Leg	2.01	AP	Leg	1.47	0.73	0.76
L6-2_g	AP	Chest	1.52	AP	Chest	1.52	1.00	1.00
L7-1_g	AP	Chest	1.52	AP	Chest	1.52	1.00	1.00
L7-2_g	AP	Leg	1.47	AP	Leg	1.47	1.00	1.00
L8-1_g	AP	Leg	1.47	Floor	Hip	1.82	1.24	1.19
L8-2_g	LLAT	Chest	1.27	Floor	Chest	1.68	1.32	0.99
L9-1_g	FLOOR	Leg	2.01	AP	Leg	1.47	0.73	0.76
L9-2_g	FLOOR	Back	-	AP	Back	-	-	-

Table 3 Modelling choices, setups and applications

Laboratory	MC Software	Phantom	Body	Dosemeters	Main Use
Lab1	GEANT4	Mesh	Limbed	Full model	Organ dose primary dataset
Lab2	MCNP	Voxel	No limbs	Disc approx.	Dosemeter uncertainties
Lab3	MCNP	Voxel	No limbs	Disc approx.	Dosemeter primary dataset
Lab4	GEANT4	Mesh	Limbed	Full model	Organ dose uncertainties

Table 4 The whole body dose quantities D_{GRB} and D_B calculated for the phantoms by Lab1 and Lab4, and by measurements. Statistical uncertainties on the modelled data are <2 %, so are omitted. Upper and lower uncertainties on the measured data are shown as super- and sub-scripts respectively.

Data Source	P1		P2		P3		P4	
	D_{GRB} (Gy)	D_B (Gy)	D_{GRB} (Gy)	D_B (Gy)	D_{GRB} (Gy)	D_B (Gy)	D_{GRB} (Gy)	D_B (Gy)
Lab1	0.544	0.534	0.056	0.090	0.206	0.207	0.401	0.384
Lab4	0.552	0.542	0.059	0.094	0.209	0.206	0.402	0.378
Measurements	$0.44_{(-0.19)}^{(+0.27)}$	-	$0.12_{(-0.06)}^{(+0.14)}$	-	-	-	$0.28_{(-0.18)}^{(+0.21)}$	-

Table 5 Resistor and glass doses calculated by Lab2 for P1, by Lab3 for all phantoms, and by Lab1 for phones placed nominally in the Chest location for phantoms P1 and P3.

Phantom	Phone	Resistor, $D_{X,Y,r}$ (Gy)			Glass, $D_{X,Y,g}$ (Gy)		
		Lab1	*Lab2*	*Lab3*	*Lab1*	*Lab2*	*Lab3*
P1	*L3-1*	0.58	0.77	0.73	0.62	0.74	0.71
	L5-1	-	2.44	2.35	-	2.38	2.30
	L6-1	-	2.51	2.42	-	2.47	2.38
	L9-1	-	2.51	2.42	-	2.47	2.38
	L9-2	-	0.26	0.26	-	0.27	0.27
P2	*L1-2*	-	-	0.37	-	-	0.37
	L2-1	-	-	0.26	-	-	0.27
	L3-2	-	-	0.11	-	-	0.11
	L4-1	-	-	0.10	-	-	0.10
P3	*L6-2*	0.31	-	0.35	0.32	-	0.35
	L7-1	0.32	-	0.36	0.33	-	0.37
	L7-2	-	-	0.37	-	-	0.38
	L8-1	-	-	0.38	-	-	0.39
P4	*L1-1*	-	-	1.90	-	-	1.96
	L2-2	-	-	1.39	-	-	1.39
	L4-2	-	-	0.88	-	-	0.91
	L5-2	-	-	0.24	-	-	0.24
	L8-2	-	-	0.88	-	-	0.91

Table 6 The dose conversion coefficients K_{GRB} and K_B and their uncertainties for resistors, along with their corresponding ratios $R(K/C)_{GRB}$ and $R(K/C)_B$ to the generic coefficients C_{GRB} and C_B chosen by the nine participating laboratories (Table 1).

Phone	Phantom	K_{GRB}	*Unc.*	K_B	*Unc.*	$R(K/C)_{GRB}$	$R(K/C)_B$
L1-1	P4	4.95	*0.67*	4.74	*0.64*	2.95	2.35
L1-2	P2	4.16	*0.56*	6.64	*0.90*	3.03	5.15
L2-1	P2	2.85	*0.38*	4.56	*0.62*	-	-
L2-2	P4	3.63	*0.49*	3.48	*0.47*	-	-
L3-1	P1	1.38	*0.19*	1.35	*0.18*	1.18	1.14
L3-2	P2	1.22	*0.16*	1.94	*0.26*	1.09	1.73
L4-1	P2	1.10	*0.15*	1.76	*0.24*	0.64	1.07
L4-2	P4	2.30	*0.31*	2.20	*0.30*	1.68	1.70
L5-1	P1	4.40	*0.59*	4.31	*0.58*	3.60	3.59
L5-2	P4	0.61	*0.08*	0.59	*0.08*	0.51	0.49
L6-1	P1	4.54	*0.61*	4.45	*0.60*	3.36	3.50
L6-2	P3	1.70	*0.23*	1.71	*0.23*	1.24	1.32
L7-1	P3	1.76	*0.24*	1.76	*0.24*	1.28	1.37
L7-2	P3	1.81	*0.24*	1.81	*0.24*	1.34	1.43
L8-1	P3	1.85	*0.25*	1.86	*0.25*	1.10	1.21
L8-2	P4	2.30	*0.31*	2.20	*0.30*	1.44	1.53
L9-1	P1	4.54	*0.61*	4.45	*0.60*	3.36	3.50
L9-2	P1	0.50	*0.07*	0.49	*0.07*	1.10	1.16

Table 7 The dose conversion coefficients K_{GRB} and K_B and their uncertainties for glass, along with their corresponding ratios $R(K/C)_{GRB}$ and $R(K/C)_B$ to the generic coefficients C_{GRB} and C_B chosen by the nine participating laboratories (Table 2).

Phone	Phantom	K_{GRB}	*Unc.*	K_B	*Unc.*	$R(K/C)_{GRB}$	$R(K/C)_B$
L1-1	P4	5.12	*0.69*	4.90	*0.66*	-	-
L1-2	P2	4.12	*0.56*	6.58	*0.89*	-	-
L2-1	P2	2.99	*0.40*	4.78	*0.64*	-	-
L2-2	P4	3.63	*0.49*	3.48	*0.47*	-	-
L3-1	P1	1.33	*0.18*	1.31	*0.18*	1.09	1.03
L3-2	P2	1.27	*0.17*	2.03	*0.27*	1.06	1.75
L4-1	P2	1.14	*0.15*	1.81	*0.24*	0.62	1.07
L4-2	P4	2.38	*0.32*	2.28	*0.31*	1.51	1.59
L5-1	P1	4.31	*0.58*	4.22	*0.57*	3.19	3.11
L5-2	P4	0.62	*0.08*	0.59	*0.08*	0.47	0.45
L6-1	P1	4.46	*0.60*	4.37	*0.59*	3.09	3.14
L6-2	P3	1.71	*0.23*	1.71	*0.23*	1.12	1.20
L7-1	P3	1.77	*0.24*	1.78	*0.24*	1.16	1.24
L7-2	P3	1.84	*0.25*	1.85	*0.25*	1.23	1.33
L8-1	P3	1.88	*0.25*	1.88	*0.25*	1.02	1.14
L8-2	P4	2.38	*0.32*	2.28	*0.31*	1.37	1.50
L9-1	P1	4.46	*0.60*	4.37	*0.59*	3.09	3.14
L9-2	P1	0.51	*0.07*	0.50	*0.07*	-	-

Table 8 Scaled body doses (D_{GRB}) from the calculations by *Lab1*, and unconverted and converted (generic and bespoke coefficients) scaled measured doses for the 18 phones; data for both resistors and glass are shown, where available. Uncertainties on the measured data are shown in brackets according to the manner provided by the laboratories: either as a single figure representing one standard uncertainty, or as the lower and upper boundaries of a range. Green shows nominally correct dose categorization, light / dark blue an over-estimate by one / two dose categories, light / dark red an under-estimate by one / two categories. Phone locations are as shown in Figure 5.

Phantom	Phone	D_{GRB} (Gy)	Dose from resistors (Gy)			Dose from glass (Gy)		
			Meas.	Generic	Bespoke	Meas.	Generic	Bespoke
P1	*L3-1*	3.0	6.5 (1.1)	5.5 (0.9)	4.7 (0.8)	3.8 $_{(2.6)}^{(5.1)}$	3.0 $_{(2.1)}^{(4.0)}$	2.8 $_{(1.9)}^{(3.8)}$
	L5-1	3.0	9.8 (1.2)	8.1 (1.0)	2.2 (0.3)	12.4 $_{(10.1)}^{(14.7)}$	9.0 $_{(7.3)}^{(10.6)}$	2.9 $_{(2.3)}^{(3.4)}$
	L6-1	3.0	7.6 (1.2)	5.6 (0.9)	1.7 (0.3)	7.8 $_{(5.8)}^{(9.7)}$	5.3 $_{(4.0)}^{(6.6)}$	1.7 $_{(1.3)}^{(2.2)}$
	L9-1	3.0	11.7 (1.9)	8.7 (1.4)	2.6 (0.4)	10.7 $_{(8.9)}^{(12.9)}$	7.3 $_{(6.1)}^{(8.8)}$	2.4 $_{(2.0)}^{(2.9)}$
	L9-2	3.0	1.5 (0.2)	3.3 (0.4)	3.0 (0.3)	-	-	-
P2	*L1-2*	0.3	0.8 (0.4)	0.6 (0.3)	0.2 (0.1)	-	-	-
	L2-1	0.3	3.2 (0.7)	-	1.1 (0.3)	-	-	-
	L3-2	0.3	1.7 (0.4)	1.5 (0.3)	1.4 (0.3)	0.6 $_{(0.1)}^{(0.9)}$	0.5 $_{(0.1)}^{(0.8)}$	0.5 $_{(0.1)}^{(0.7)}$
	L4-1	0.3	0.4 (0.2)	0.2 (0.1)	0.4 (0.2)	0.7 $_{(0.2)}^{(0.9)}$	0.4 $_{(0.1)}^{(0.5)}$	0.6 $_{(0.1)}^{(0.8)}$
P3	*L6-2*	1.1	1.7 (0.7)	1.2 (0.5)	1.0 (0.4)	1.9 (0.3)	1.2 (0.2)	1.1 (0.2)
	L7-1	1.1	4.3 (0.9)	3.1 (0.7)	2.4 (0.5)	4.6 $_{(2.5)}^{(6.7)}$	3.0 $_{(1.6)}^{(4.4)}$	2.6 $_{(1.4)}^{(3.8)}$
	L7-2	1.1	5.8 (0.8)	4.3 (0.6)	3.2 (0.4)	3.0 $_{(1.6)}^{(4.5)}$	2.1 $_{(1.1)}^{(3.1)}$	1.6 $_{(0.9)}^{(2.4)}$
	L8-1	1.1	9.2 (1.9)	5.5 (1.1)	5.0 (1.0)	0.6 (0.1)	0.3 (0.1)	0.3 (0.1)
P4	*L1-1*	2.2	4.2 (0.9)	2.5 (0.5)	0.8 (0.2)	-	-	-
	L2-2	2.2	7.2 (1.3)	-	2.0 (0.4)	4.3 $_{(2.0)}^{(5.9)}$	-	1.2 $_{(0.6)}^{(1.6)}$
	L4-2	2.2	6.4 (1.8)	4.7 (1.3)	2.8 (0.8)	5.2 $_{(4.0)}^{(7.0)}$	3.4 $_{(2.6)}^{(4.6)}$	2.2 $_{(1.7)}^{(3.0)}$
	L5-2	2.2	1.8 (0.3)	1.5 (0.2)	2.9 (0.4)	2.6 (0.5)	2.0 (0.4)	4.2 (0.8)
	L8-2	2.2	5.2 (1.1)	3.3 (0.7)	2.3 (0.5)	2.7 (0.7)	1.6 (0.4)	1.1 (0.3)

Table 9 Body doses (D_{GRB}) from the calculations by *Lab1*, and unconverted and converted (generic and bespoke coefficients) data from resistors, glass, and resistors and glass averaged, for each set of phones on each phantom: mean doses, standard deviations across each set, mean responses (i.e. mean dose divided by 'true' D_{GRB}), and coefficients of variation (i.e. standard deviation as percentage of mean).

		Resistors				Glass				Resistors & Glass			
***Phantom,* D_{GRB} (Gy)**	**Dataset**	**Mean (Gy)**	**St.Dev. (Gy)**	**Mean Resp.**	**CoV (%)**	**Mean (Gy)**	**St.Dev. (Gy)**	**Mean Resp.**	**CoV (%)**	**Mean (Gy)**	**St.Dev. (Gy)**	**Mean Resp.**	**CoV (%)**
P1, 0.544	Meas.	1.35	0.63	2.48	47	1.57	0.59	2.89	38	1.45	0.62	2.66	43
	Generic	1.14	0.35	2.09	31	1.11	0.40	2.05	36	1.13	0.38	2.07	33
	Bespoke	0.517	0.19	0.95	36	0.448	0.083	0.82	19	0.486	0.15	0.894	32
P2, 0.056	Meas.	0.275	0.20	4.91	73	0.115	0.005	2.05	4.3	0.222	0.18	3.96	81
	Generic	0.138	0.097	2.46	70	0.082	0.014	1.46	17.3	0.115	0.080	2.06	70
	Bespoke	0.138	0.091	2.46	66	0.096	0.010	1.72	10	0.124	0.077	2.21	62
P3, 0.206	Meas.	0.955	0.50	4.64	52	0.458	0.27	2.22	60	0.706	0.47	3.43	67
	Generic	0.642	0.28	3.12	44	0.301	0.18	1.46	61	0.472	0.293	2.29	62
	Bespoke	0.529	0.26	2.57	49	0.256	0.15	1.24	60	0.392	0.254	1.90	65
P4, 0.401	*Meas.*	0.900	0.35	2.24	38	0.673	0.20	1.68	30	0.799	0.31	1.99	39
	Generic	0.540	0.21	1.35	39	0.422	0.14	1.05	34	0.490	0.20	1.22	40
	Bespoke	0.391	0.13	0.97	34	0.395	0.23	0.99	57	0.393	0.18	0.980	46