Monte Carlo reference data sets for imaging research: Executive summary of the report of AAPM Research Committee Task Group 195

Ioannis Sechopoulos[a)]
Department of Radiology and Imaging Sciences, Emory University, Atlanta, Georgia 30322

Elsayed S. M. Ali[b)]
Carleton Laboratory for Radiotherapy Physics, Department of Physics, Carleton University, Ottawa, Ontario K1S 5B6, Canada

Andreu Badal and Aldo Badano
Food and Drug Administration, Silver Spring, Maryland 20993-0002

John M. Boone
Departments of Radiology and Biomedical Engineering, University of California, Davis, California 95817

Iacovos S. Kyprianou
Food and Drug Administration, Silver Spring, Maryland 20993-0002

Ernesto Mainegra-Hing
National Research Council Canada, Ottawa, Ontario K1S 5B6, Canada

Kyle L. McMillan and Michael F. McNitt-Gray
Department of Biomedical Physics and Department of Radiology, University of California, Los Angeles, California 90024

D. W. O. Rogers
Carleton Laboratory for Radiotherapy Physics, Department of Physics, Carleton University, Ottawa, Ontario K1A 0R6, Canada

Ehsan Samei
Carl E. Ravin Advanced Imaging Laboratories, Department of Radiology, Duke University, Durham, North Carolina 27705

Adam C. Turner[c)]
Department of Biomedical Physics and Department of Radiology, University of California, Los Angeles, California 90024

(Received 4 August 2014; revised 27 July 2015; accepted for publication 4 August 2015; published 8 September 2015; corrected 28 September 2015)

The use of Monte Carlo simulations in diagnostic medical imaging research is widespread due to its flexibility and ability to estimate quantities that are challenging to measure empirically. However, any new Monte Carlo simulation code needs to be validated before it can be used reliably. The type and degree of validation required depends on the goals of the research project, but, typically, such validation involves either comparison of simulation results to physical measurements or to previously published results obtained with established Monte Carlo codes. The former is complicated due to nuances of experimental conditions and uncertainty, while the latter is challenging due to typical graphical presentation and lack of simulation details in previous publications. In addition, entering the field of Monte Carlo simulations in general involves a steep learning curve. It is not a simple task to learn how to program and interpret a Monte Carlo simulation, even when using one of the publicly available code packages. This Task Group report provides a common reference for benchmarking Monte Carlo simulations across a range of Monte Carlo codes and simulation scenarios. In the report, all simulation conditions are provided for six different Monte Carlo simulation cases that involve common x-ray based imaging research areas. The results obtained for the six cases using four publicly available Monte Carlo software packages are included in tabular form. In addition to a full description of all simulation conditions and results, a discussion and comparison of results among the Monte Carlo packages and the lessons learned during the compilation of these results are included. This abridged version of the report includes only an introductory description of the six cases and a brief example of the results of one of the cases. This work provides an investigator the necessary information to benchmark his/her Monte Carlo simulation software against the reference cases included here before performing his/her own novel research. In addition, an investigator entering the field of Monte Carlo simulations can use these descriptions and results as a self-teaching tool to ensure that he/she is able to perform a specific simulation correctly. Finally, educators can assign

these cases as learning projects as part of course objectives or training programs. © *2015 American Association of Physicists in Medicine.* [http://dx.doi.org/10.1118/1.4928676]

Key words: Monte Carlo simulation, diagnostic imaging, PENELOPE, GEANT4, EGSnrc, MCNP, x-rays, test cases

TABLE OF CONTENTS

1. INTRODUCTION

The use of Monte Carlo methods in diagnostic medical imaging research is an attractive option because of the relative ease with which many different calculations can be performed. These simulations sometimes span large parameter spaces and/or obtain estimates of quantities that are challenging to measure, e.g., absorbed dose or x-ray scatter. The ever continuing reduction in cost of computing power has also helped in increasing the use of this research methodology.

As with all other types of experimentation, Monte Carlo simulations need to be validated before their results can be trusted. In many cases, the validation process for a Monte Carlo simulation is not a simple endeavor. Ideally, Monte Carlo-based computer programs are validated by generating a simulation that replicates an experimental result and then comparing simulation and measured results. However, this requires replicating the conditions of a physical experiment to a high level of accuracy and detail in the computer simulation to minimize differences from the measurement conditions. This level of replication of a physical experiment can be extremely challenging even for simple experimental conditions.

One possible step in the validation process of Monte Carlo simulations is the replication of published (and previously validated) simulations. However, this approach also presents challenges. In the first place, a relevant previously published simulation has to exist. Once one is identified, a common pitfall is the lack of sufficient details in the description of the simulation, making its replication challenging. In addition, many times the published results are provided in graphs or summarized form, so the possibility of performing an appropriate comparison with the results of the simulation being validated can be limited and the effort laborious due to the need to obtain numeric values from graphs.

To aid researchers in their efforts to validate their Monte Carlo simulations in imaging research or to assist in the validation process of their use of a general purpose code already available, the American Association of Physicists in Medicine (AAPM) created Task Group (TG) 195 with the following charge:

"The primary charge of this TG is to develop a set of Monte Carlo simulation geometries which are relevant to diagnostic imaging, perform benchmark Monte Carlo studies on a number of the geometries using a variety of established code systems such as EGSnrc, MCNP, PENELOPE, and GEANT4, and report the findings. Findings will include Monte Carlo measured parameters such as dose, variance in dose, and scatter to primary relationships. For scatter assessment, dose in the mathematical phantoms and in realistic imaging geometries to the detector will be evaluated."

The resulting TG report,[1] approved by the Science Council of the AAPM, provides six different reference sets of simulations that include a complete description of the simulation

TABLE I. List of the cases included in the TG report.

Case number	Modality/description	Quantities scored
1	Half value layer	Energy fluence
2	Radiography and body tomosynthesis	(a) Total dose (b) X-ray primary and scatter photon energy
3	Mammography and breast tomosynthesis	(a) Total dose and glandular dose (b) X-ray primary and scatter photon energy incident on a scoring plane
4	Computed tomography (simple volumes)	Dose
5[a]	Computed tomography (voxelized volumes)	Dose
6	X-ray production	Photon energy fluence and fluence spectra

[a]Case 5 may also be useful for dosimetry simulations with other imaging modalities (e.g., radiography and body tomosynthesis) that involve voxelized models of patients.

TABLE II. Monte Carlo packages used to generate the results included in the TG report, in alphabetical order.

Package name	Version	Photon cross-sections	References	Comments
EGSnrc	V4 r2.4.0	EGSnrc default (XCOM: Rayleigh, photoelectric, and pair production. RIA: Incoherent scattering.)	12–15	Used in all cases except for Case 1, where NIST XCOM cross sections were used for all interactions.
GEANT4	9.6 patch 2	Electromagnetic physics option 4 package	16 and 17	Used in all cases except for Case 6, where the Livermore electromagnetic physics were used.
MCNPX	2.7a	ENDF/B-VII	18	Used in all cases.
PENELOPE	2006	PENELOPE default (XCOM and EPDL)	19–21	Used in all cases except Case 6. Used with the penEasy main program adapted for parallel simulation with MPI library.
	2008	PENELOPE default (XCOM and EPDL)		Used in Case 6, with standard penEasy main program.

conditions. In addition, the simulation results for all sets as generated by four commonly used Monte Carlo packages are provided in tabulated form along with their statistical uncertainty. It is the intent of this TG that this information allows researchers performing Monte Carlo simulations to have a standardized methodology to benchmark their simulations before embarking on research. We believe that providing these data will avoid the need to search the literature for potential comparable simulations, try to obtain numerical data from published graphs, and contact the author of the publications in hope of gaining more information on the relevant inputs and results of a simulation performed in some cases many years ago. Although the results provided are based on the use of four common Monte Carlo packages, the information provided may be used to benchmark simulations implemented with any Monte Carlo code.

These data sets allow the reader to accurately replicate the geometry, source properties, and scoring. In this way, the reader can have confidence that his/her simulation is providing results comparable to those of these thoroughly vetted reference cases. The agreement amongst the four codes used here indicates the level of accuracy to be expected as these codes have all been benchmarked against experimental data in different situations.[2–11] It is worthwhile to emphasize that as our results were not based on experimental measurements, the validation achieved through this benchmarking is only relative to these calculations with widely used codes and not directly to measurements. But the concordance between the four Monte Carlo results provides strong confidence for the value and the validity of this exercise. It is in that sense that we use the term "validation" throughout the body of this document. In publications and presentations involving the code, a reference to the TG report and the obtained level of agreement will be sufficient to communicate the performance of this component of a validation effort.

Another often complicating factor when performing Monte Carlo simulations is that learning to perform these simulations involves a steep learning curve. It takes a substantial amount of time and effort to become comfortable in the programming method for using any of the publicly available Monte Carlo packages. To help address this, it is hoped that these reference simulation sets can also be used as a teaching tool, either for a scientist entering the field of Monte Carlo simulations in diagnostic imaging himself/herself, or for a supervising educator training students and/or junior researchers in the field. The availability of the complete descriptions of relevant simulations and their expected results when using different well-established Monte Carlo codes may be used to check one's own skills or as an assignment for trainees.

The diagnostic imaging simulation sets included in the TG report are listed in Table I. They span a number of x-ray based imaging modalities and common quantities scored with Monte Carlo simulations. While this is a limited set that does not encompass all modalities and tests that are commonly investigated using Monte Carlo methods, the defined conditions provide a meaningful representation of Monte Carlo-based research.

This paper is organized as follows. In Sec. 2, a description of all the parameters common to the simulation sets is provided. This includes the Monte Carlo packages used to generate the results and the material compositions and the x-ray spectra used throughout the simulation sets. In addition, this section provides an abbreviated description of each of the simulation cases included in the TG report. Section 3 provides some of the results obtained by the TG members for one of the simulation cases described in Sec. 2. The complete description and all results for all six cases are available in the TG report. Section 4 includes a discussion on the common pitfalls encountered in the implementation of the simulations ("lessons learned") and a discussion on the differences in the

TABLE III. Characteristics of x-ray spectra used for the simulations included in the TG report. Half value layer (HVL) and quarter value layer (QVL) values are in terms of air kerma determined using planar energy fluence.

IEC 61267 name	RQR-8	RQR-9	RQR-M3
Target/filter elements	W/Al	W/Al	Mo/Mo
Tube voltage (kVp)	100	120	30
Anode angle (deg)	11	11	15
Ripple (%)	0	0	0
Filter thickness (mm)	2.708	2.861	0.0386
Mean energy (keV)	50.6	56.4	16.8
HVL (mm Al)[a]	3.950	5.010	0.3431
QVL (mm Al)[b]	9.840		0.7663

[a]The values in the table are not those reported by IPEM Report 78 software, but were calculated by the TG using the NIST XCOM dataset (Ref. 22).
[b]Provided for the two spectra used in Case 1 only.

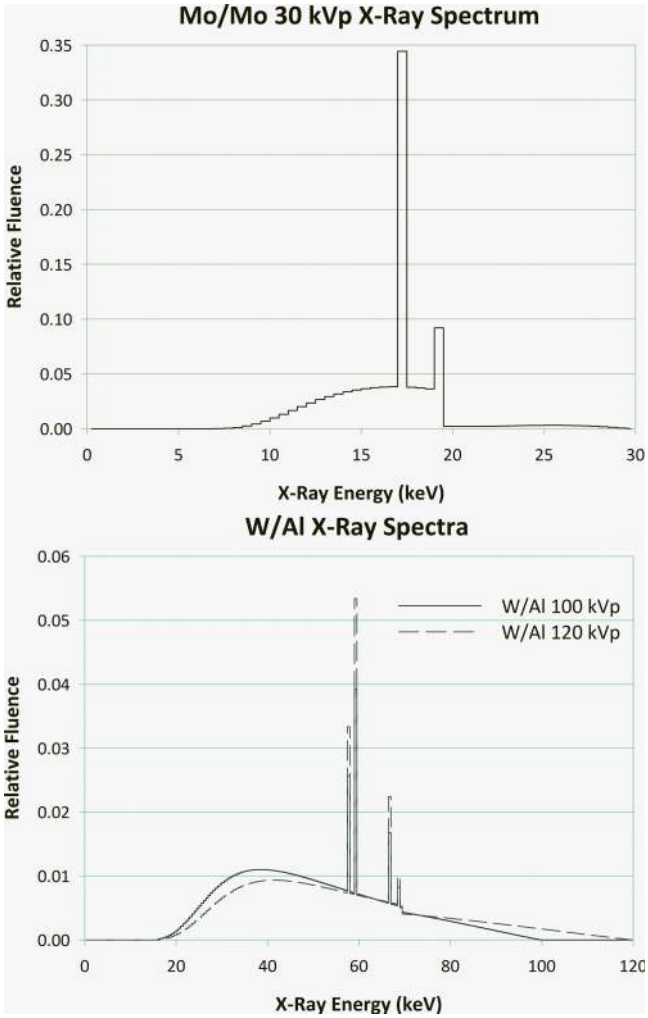

FIG. 1. The x-ray spectra, normalized to unit fluence under the curve, used in the Monte Carlo simulations included in the TG report.

results obtained with the Monte Carlo packages used in the TG report.

2. COMMON PARAMETERS

2.A. Monte Carlo packages

All simulation conditions described in the TG report were implemented with four commonly used and well-established

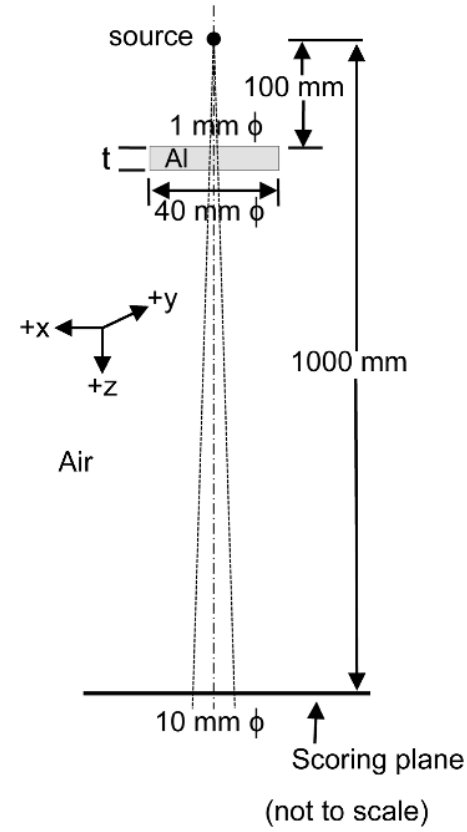

FIG. 2. The geometry setup for Case 1.

Monte Carlo packages, listed in Table II. Results from all four implementations of each simulation are provided in the TG report. These packages were selected for inclusion in the TG report due to their being commonly used by the diagnostic imaging research community, their availability and continued development and maintenance, and the expertise of the TG members. Inclusion of these Monte Carlo packages in the TG report should not be interpreted as any form of endorsement by the TG or by the AAPM, and exclusion of any other Monte Carlo package should not be interpreted as disapproval by the same.

The results and performance of the four Monte Carlo packages used for the TG report should not be necessarily construed to be the best achievable with each package. The results and performance obtained with these packages reflect

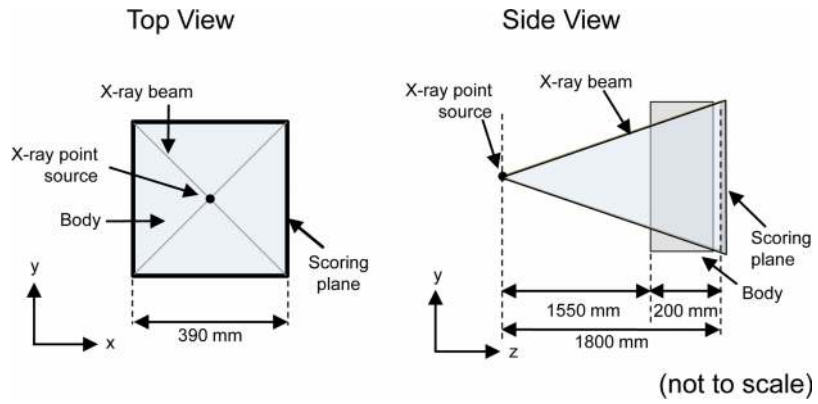

FIG. 3. The geometry setup for Case 2 with the x-ray source positioned for the radiography acquisition (equivalent to the tomosynthesis 0° angle).

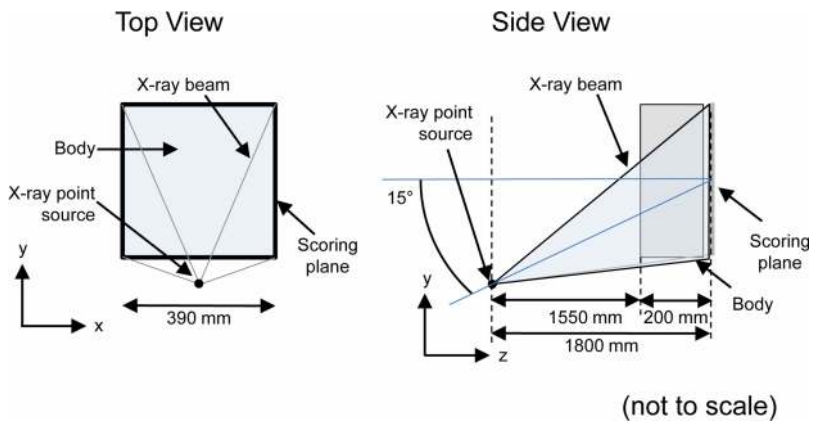

FIG. 4. The geometry setup for Case 2 with the x-ray source positioned for the tomosynthesis 15° angle acquisition.

the best attempts of the TG members to match the conditions specified in each case description, with logical choices for each code-specific parameter. The simulations were not necessarily optimized for maximum efficiency and/or for maximum level of accuracy beyond that of typical use. In addition, errors could have been made during the implementation of the simulations.

2.B. Monte Carlo parameters

2.B.1. Simulated material compositions

The composition of materials used in all simulations is provided in two separate electronic files available for download with the TG report. One file includes material compositions for all cases except those used in Case 5, which are defined in a separate file.

The densities specified for the elements, as well as the material composition for the simpler compounds were those provided by the National Institute of Standards and Technology (NIST).[22] The composition of soft tissue used is that provided by the International Commission on Radiation Units and Measurements (ICRU).[23] The material compositions related to mammography and breast tomosynthesis are those provided by Hammerstein *et al.*[24] due to their common use in breast imaging research.

2.B.2. X-ray spectra

The probability distribution functions of the x-ray spectra used as inputs in the simulations included in the TG report were obtained using Report 78 of the Institute of Physics and Engineering in Medicine (IPEM).[25] They were defined so as to approximate certain x-ray spectra defined in the International Electrotechnical Commission (IEC) Publication 61267.[26] The spectra definitions do not comply with the IEC 61267 requirement for their homogeneity coefficient, the ratio of the first to the second half value layer. However, all other requirements are met, and this has no effect on the validity of this benchmarking process. Table III lists the x-ray spectra used and the parameters specified for the generation of their distribution functions using the IPEM Report 78 software, with the resulting spectra plotted in Fig. 1. The distribution functions are available for download in the electronic resources included with the TG report. The energy bin width for all spectra is 0.5 keV, starting at 0 keV, and the energy of the center of the bin is listed. During the Monte Carlo simulations, the

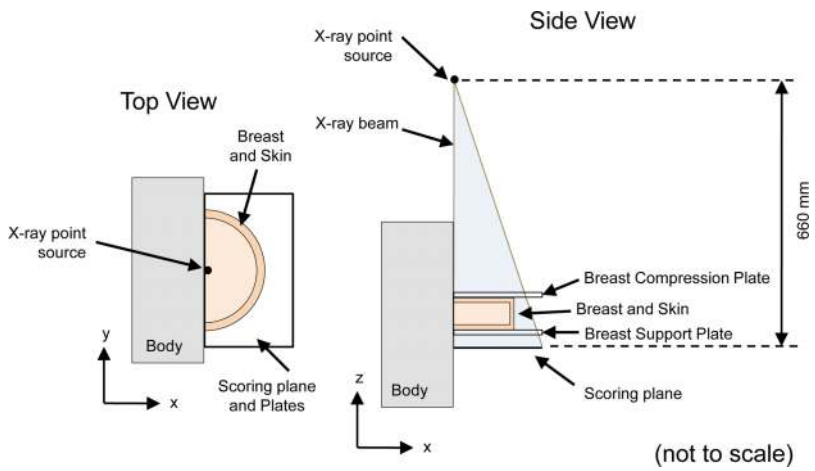

FIG. 5. The geometry setup for Case 3 with the x-ray source positioned for the mammography acquisition (equivalent to the tomosynthesis 0° angle).

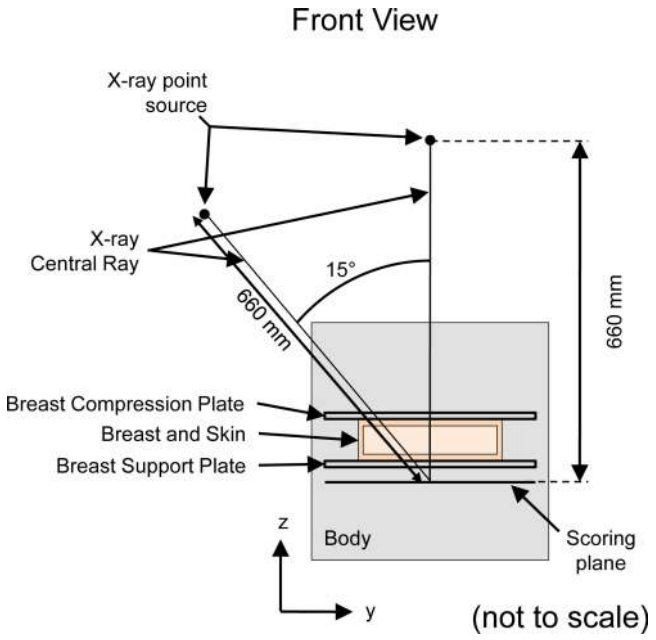

FIG. 6. The front view of the geometry setup for Case 3 with the x-ray source positioned for both the mammography and the breast tomosynthesis 15° angle acquisitions.

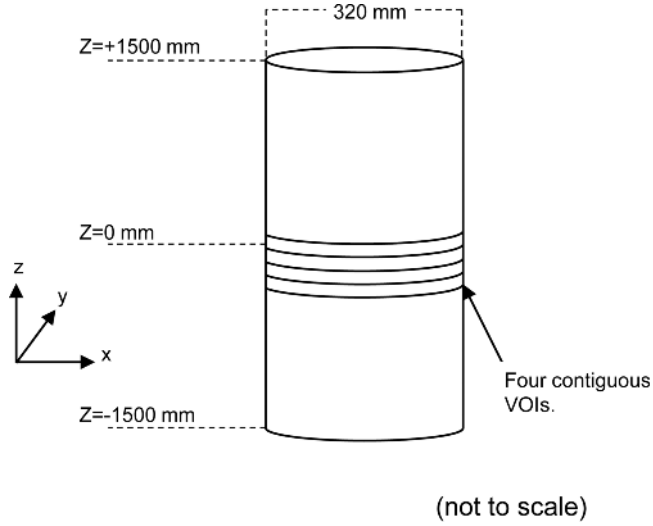

FIG. 8. Locations of the four cylindrical VOIs within the simulated body phantom in Case 4 where the energy deposition is scored for the first test of this CT dosimetry simulation.

provided distribution functions should be sampled uniformly within each bin.

3. SIMULATION DESCRIPTIONS

3.A. Case 1: Half value layer (HVL)

This case allows verification of input x-ray spectrum sampling, basic material attenuation, and HVL calculations. The simulation consists of an x-ray source (Fig. 2) emitting a circular beam toward an aluminum filter. The energy fluence transmitted through the filter is scored so HVLs can be calculated off-line. These HVLs are compared with the expected first half value and quarter value layers for both monoenergetic and spectral sources.

3.B. Case 2: Radiography and body tomosynthesis

This case allows validation of x-ray transport and interaction characteristics in general radiography and whole body tomosynthesis simulations. The simulation consists of an isotropic x-ray source collimated so that the beam coincides with the borders of the square scoring plane. The x-ray source is positioned to simulate acquisition of a radiographic image (Fig. 3), equivalent to the central projection of a body tomosynthesis exam, and to one side to simulate a noncentral tomosynthesis projection (Fig. 4). A simple homogeneous box composed of soft tissue represents the patient body. The simulation involves the scoring of both absorbed dose in the body and the energy of both x-ray primary and scatter components incident upon the detection plane. For the dosimetry simulation, both the doses deposited in the whole body and in a number of volumes-of-interest (VOIs) are scored. For the scatter simulation, the scatter energy incident on the scoring plane at a number of regions-of-interest (ROIs) for both full field beams and zero-area beams is scored. Depending on the application, the dose and/or the fluence component results may be of interest.

3.C. Case 3: Mammography and breast tomosynthesis

This case allows verification of x-ray transport and interaction characteristics in mammography and breast tomosyn-

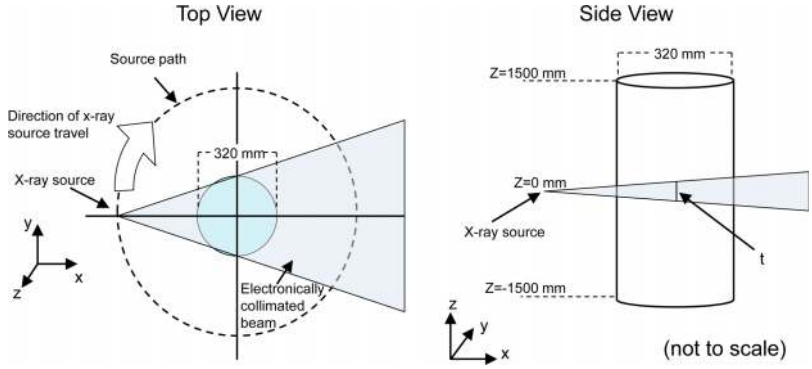

FIG. 7. The geometry setup for Case 4.

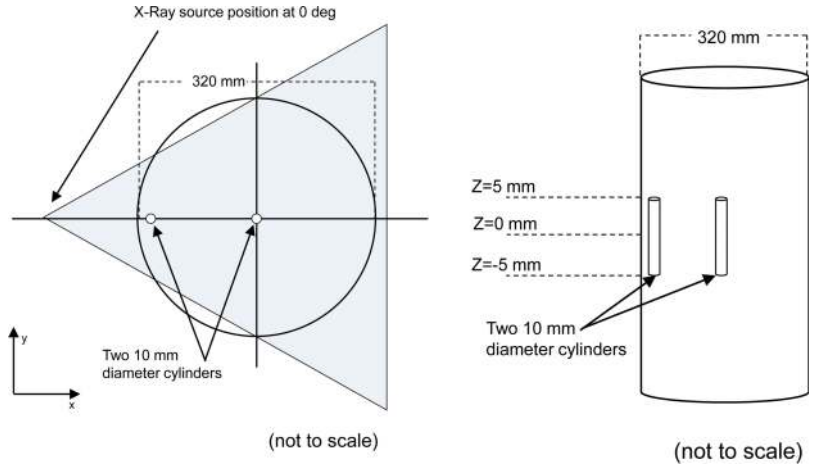

FIG. 9. Locations of the two cylindrical VOIs within the simulated body phantom in Case 4 where the energy deposition is scored for the second test of this CT dosimetry simulation.

thesis simulations, resulting in the validation of estimates of absorbed dose in the breast glandular tissue and x-ray scatter incident upon the scoring plane. This case is similar to Case 2 but adapted for simulation of breast imaging. This involves the use of a semicircular solid representing the compressed breast instead of the box representing the body in Case 2, and the inclusion of the breast compression and support plates. To include the effect of backscatter, a box representing the body is included posterior to the breast. As in Case 2, the x-ray source involves an isotropic source limited to the scoring plane borders, and is positioned for both mammography (Fig. 5) and breast tomosynthesis imaging (Fig. 6). Scoring includes energy deposition in the whole breast and in specific VOIs, average glandular dose to the whole breast, and x-ray scatter energy incident on the scoring plane at specific ROIs. Depending on the application, the dose and/or the fluence component results may be of interest.

3.D. Case 4: Computed tomography with simple solids

This case allows validation of x-ray transport and interaction characteristics in computed tomography in addition to x-ray source rotation, resulting in the validation of estimates of absorbed dose in a simple CT phantom. The simulated phantom is similar to the computed tomography dose index (CTDI) phantom, with a diameter of 32 cm (Fig. 7). For the first test, the scoring is the energy deposited in four contiguous cylindrical segments from a single projection, as shown in Fig. 8. For the second test, the scoring is the energy deposited in two 10 mm diameter PMMA cylinders with a height of 100 mm (in the z-direction from −50 to +50 mm), as shown in Fig. 9. In both cases the source is an isotropic fan beam, and the simulations are tested with two fan beam thicknesses.

3.E. Case 5: Computed tomography with a voxelized solid

This case allows validation of voxel-based x-ray transport and interaction characteristics in computed tomography, in addition to x-ray source rotation, resulting in the validation of estimates of absorbed dose in a complex, voxelized CT phantom. The simulation geometry is the same as that defined for Case 4, but with a voxelized box replacing the cylindrical body phantom (Fig. 10). The three dimensional image with the information for the material content of the voxelized volume is available with the TG report. This reference case is the XCAT model, courtesy of Ehsan Samei and Paul Segars of Duke University, to serve as a reference platform for Monte Carlo

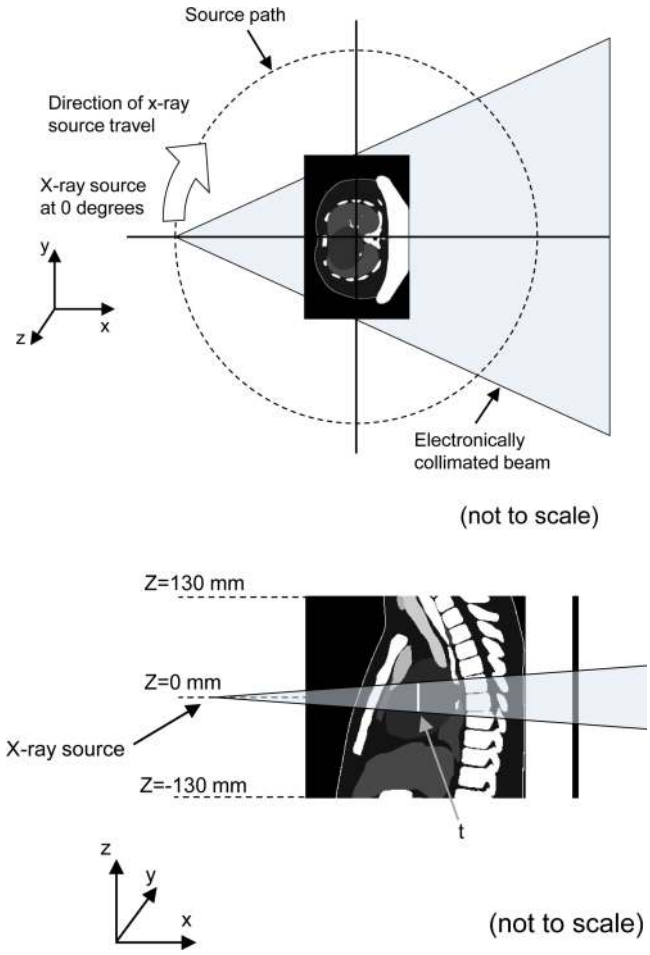

FIG. 10. The geometry setup for Case 5.

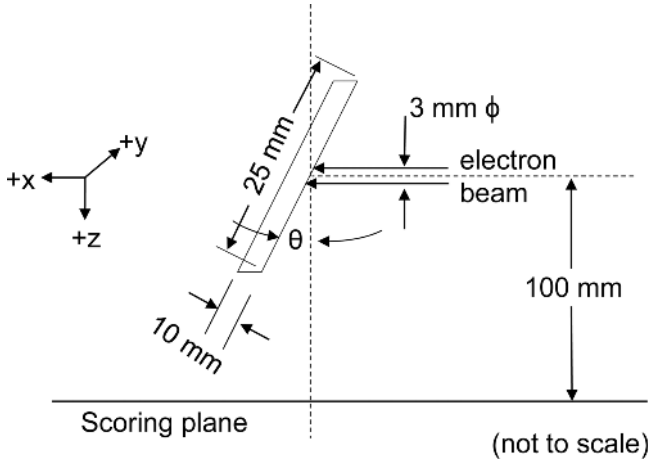

FIG. 11. The geometry setup for Case 6. Electron energy and target angle are 30 keV/15° and 100 keV/11° for molybdenum and tungsten targets, respectively.

simulations. The voxels in the image contain values ranging from 0 to 19 that correspond to the different organs of the body. Material definitions for each organ are provided in the TG report. The scoring is the energy deposited in 17 of the organs of the body phantom. The source is defined as an isotropic fan beam that rotates around the voxelized volume.

Even though this simulation uses a relatively thin fan beam, this case may also be useful for verification of dosimetry simulations involving voxelized solids in other modalities such as radiography and body tomosynthesis. For this, comparison of the results for a single or a limited number of projection angles may be sufficient.

3.F. Case 6: X-ray production

This case allows validation of electron transport and x-ray generation in idealized tube targets in mammography and radiography. The simulation involves a beam of electrons aimed at a target slab composed of either molybdenum or tungsten to simulate mammography and radiography conditions, respectively (Fig. 11). The x-ray fluence at a scoring plane 100 mm away from the center of the target is scored in five ROIs, positioned so as to study both symmetry and the heel effect. The results depend on the physical models used to simulate bremsstrahlung emission and inner shell ionizations, as well as the internal tables on characteristic x-ray emission energies.

4. RESULTS

Detailed tabular results for all six cases simulated with the Monte Carlo packages listed in Table II are available for download in the electronic resources included with the TG report, along with some tables and graphs for some specific simulation conditions included in the TG report itself. As an example of the results obtained and included in the report, a small portion of the results obtained for Case 2 are shown here. As can be seen in Fig. 12, all four packages provided very consistent results for the energy deposited in the whole body, with the variation in results being in the range of 0.1%–0.2%. In Fig. 13, the comparisons of primary and scatter x-ray energy incident on the ROIs at the scoring detector plane are shown. These graphs show a larger variation in the primary incident energy (up to ~2%) and a very small variation in the scatter results. The variation in the results for the primary energy is explained at length in the full report. Briefly, the use of different Compton scattering models by the different Monte Carlo codes has an impact on the resulting primary x-ray energy incident on the detector. The monoenergetic primary photon beam reaching the scoring plane is very sensitive to differences in the photon cross sections since it is directly proportional to $e^{-\mu x}$, where μ is the total photon cross section and x is the photon path length in the attenuating material. For

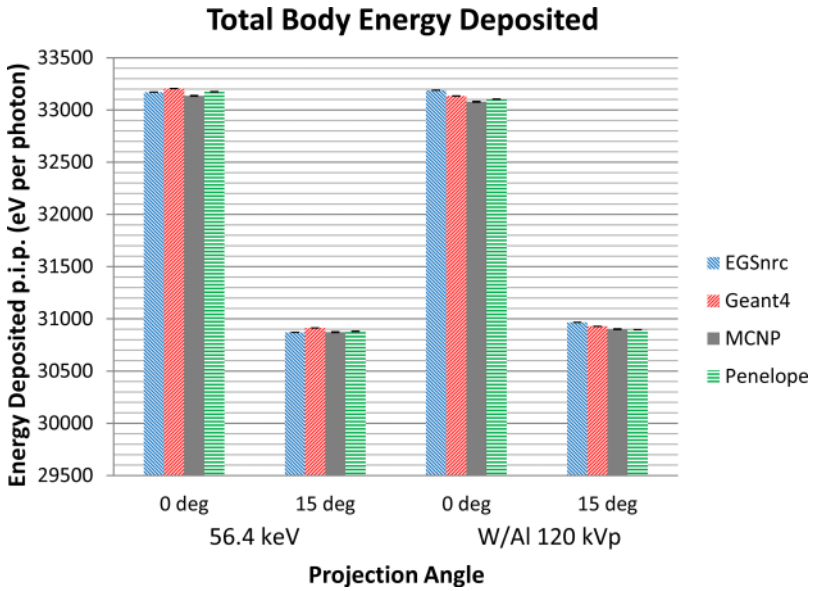

FIG. 12. Energy deposition in the whole body per initial photon (p.i.p) in Case 2 for the four Monte Carlo packages included in Table II. Note the narrow range of the vertical scale, and the presence of the statistical uncertainty bars in each simulation.

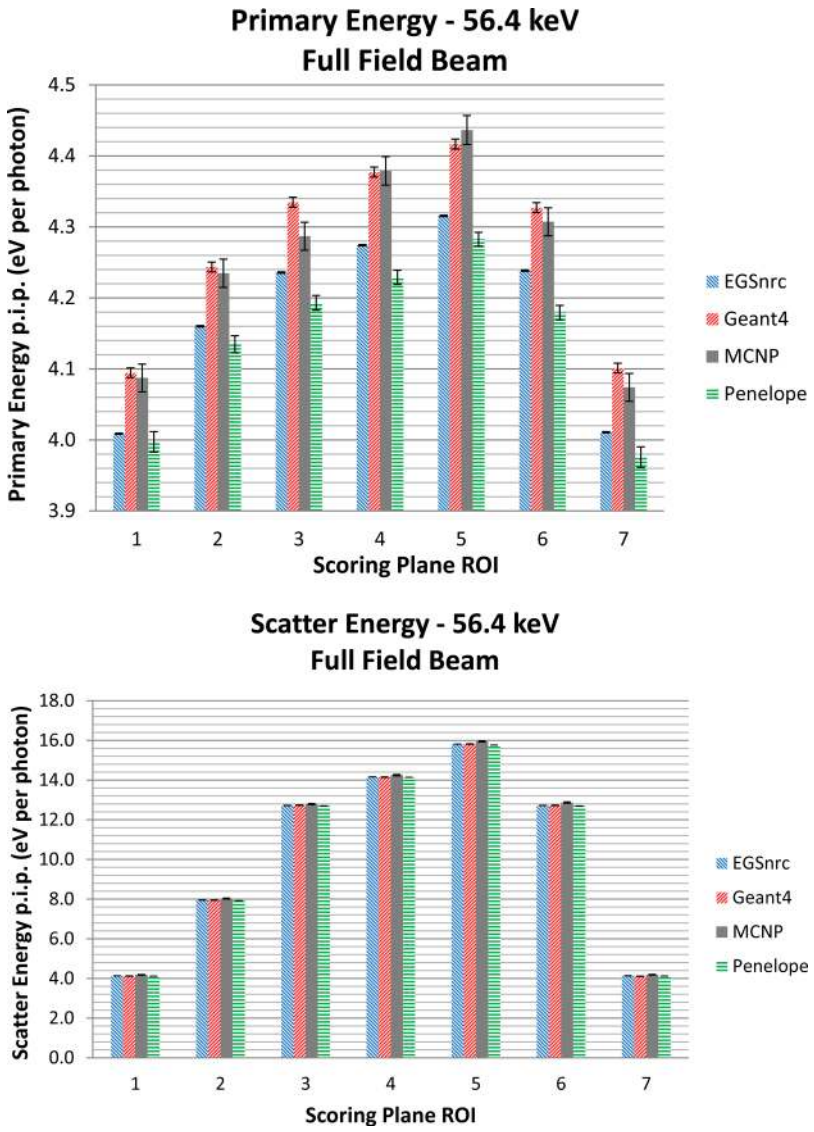

FIG. 13. Primary component (top) and scatter component (bottom) of energy incident on the scoring plane ROI per initial photon (p.i.p.) with the full field x-ray beam for the radiography (0° projection). 56.4 keV x-ray energy simulations of Case 2 with the four Monte Carlo packages included in Table II.

this 20 cm thick phantom, a 1% change in μ translates into a greater than 4% change in $e^{-\mu x}$.

Additional graphs are included in the electronic spreadsheets that provide the simulation results. All uncertainties shown in tables and graphs are due only to the statistical uncertainties from the Monte Carlo simulations and do not include any uncertainties or limitations of the underlying physics models or cross sections.

5. DISCUSSION

The comparison of simulated results with experimental measurements is still an essential step in the process of validating a new simulation algorithm. However, the possibility to benchmark simulation results among different codes using clearly defined geometries, material composition, and scoring methods eliminates many of the sources of variability and error found in experimental measurements and provides an insight into the inner workings of the simulation code at a precision level beyond the accuracy of standard experimental measurements.

5.A. Comparison of results among Monte Carlo packages

In general, the implementations of all six cases with the four Monte Carlo packages included in the TG report resulted in estimates with very good concordance, with most of the results being within the statistical uncertainty of the simulations. For those differences that exceeded the statistical uncertainty, many differences are within ~5% and most within 10% of the mean of the results. Some, if not all, of the differences beyond statistical uncertainty could be due to small differences in interpretation of the simulation conditions or even errors in implementation of the cases by the TG members. Every attempt was made to minimize this possibility. In some of the cases, the differences were also clearly produced by the underlying differences in the physical models used by the multiple codes, such as small differences in the energy of

the emitted characteristic x-rays or the modeling of Compton scattering events.

It should be noted that the implementations of Case 6 (Sec. 3.F) resulted in somewhat larger differences among the codes than those seen for the other cases. This simulation case is more sensitive to subtle differences in physics models compared to the other simulations because it is sensitive to electron transport details. Thus, it is to be expected that the comparison among the different Monte Carlo simulation packages would be the most challenging. Differences among the predicted energy-discriminated fluence were found to be approximately ±10%–20%, except for the very low-energy bins below approximately 3 keV for the molybdenum spectrum and 12 keV for the tungsten spectrum, respectively. At these very low energies, where the electron interaction physics models are known to have larger uncertainties, the discrepancies between codes were considerably larger. At these energies, however, diagnostic imaging conditions normally result in these x-rays being fully attenuated by added filtration, so these discrepancies are not expected to impact realistic imaging research. MCNP produced characteristic emission peaks for the molybdenum target that were substantially higher than those predicted by the other codes. While the total energy fluence estimates for the tungsten target spanned a range of approximately ±1%–±2% with only one difference being 4%, the molybdenum target spectrum simulations demonstrated much larger differences, mostly due to the substantially higher estimates in characteristic emission by the MCNP simulations. These calculations are sensitive to the electron impact ionization cross sections used.

5.B. Lessons learned

The exercise performed for this TG report, specifically the design and implementation of several simulation test cases with a number of different Monte Carlo simulation packages, resulted in a number of problems that were identified and corrected along the way. These problems and inconsistencies in interpretation now provide information that we hope will help readers when implementing their own Monte Carlo simulations.

Chief among the problems that could be encountered is that the specifications of the simulations may not be followed accurately. Preliminary results of some of the cases included in the TG report resulted in large differences amongst the codes, which had to be identified and corrected until the TG was confident that the remaining discrepancies were present either due to statistical uncertainties or real differences in either the physics tables of each package or some other aspect inside each code. An example of actual differences in the codes is the small difference in the energy of some characteristic emission x-rays, which resulted in associated peaks corresponding to different energy bins depending on the Monte Carlo package used. Therefore, great care should be taken in implementing these simulations with the correct input parameters, in terms of geometry, materials, source, and scoring. Errors observed during this exercise and which may affect others include the following.

1. Incorrect source type (e.g., isotropic versus nonisotropic point source): For all simulations except Case 6 and the point spread function sources (Cases 2 and 3), the source should be isotropic.
2. Incorrect scoring metric (e.g., fluence versus planar fluence versus energy fluence): For comparison to the present data, all fluence scorings should be planar fluence, i.e., the factor for the cosine of the incidence angle does not apply.
3. Incorrect scoring units (e.g., total x-rays incident in a ROI versus x-rays per unit area).
4. Incorrect binning (e.g., whether the energy of a bin represents the midlevel, the minimum, or the maximum of the bin): All spectral information, i.e., the x-ray spectra provided for the simulations and the results for Cases 1 and 6, specify the midlevel of the energy bin. When sampling, energies are sampled uniformly across each bin.
5. Incorrect material composition and density definition: Care should be taken when defining the materials.
6. Direction of travel in rotating/translating sources: The directions in which rotations or translations are positive are marked in the figures.
7. X-ray source translation versus rotation: Note that to more accurately reflect the way the corresponding imaging systems function, the simulation of body tomosynthesis calls for the x-ray source to translate parallel to the scoring plane to an angle of 15° while the simulation of breast tomosynthesis involves the source rotating about a point at the scoring plane to an angle of 15°.
8. Incorrect normalization: All results normalized "per photon history" are normalized to only the acceptable histories that are emitted toward the area specified (in most cases, the entire scoring plane). The results should not be normalized to all histories potentially emitted in all directions covering the 4π sphere. The EGSnrc collimated source gives quantities normalized to the incident real fluence, not the number of histories. To convert to a quantity normalized per number of particles, one has to calculate the solid angle subtended by the point source and the scoring plane.
9. Excluding the surrounding air was found to have a small but significant effect on the results in Cases 1 and 4.

One issue that can introduce differences in the results is the setting of the energy or range threshold/cut. This is a very important parameter in many simulations and can have a substantial impact on the results in many types of Monte Carlo simulations. Among the cases included in the TG report, this is especially true for Case 6, the simulation of the generation of x-rays.

An interesting issue was observed with MCNP for Case 4. For all simulations for this case, a void collimator (material of the collimator set as a void region where photons have no importance) was used to properly collimate the isotropic x-ray point source. When the collimator and point source were rotated in 10° increments about the phantom, large differences between MCNP results and the other codes were observed

between 60°–80° and 280°–300°. These differences were eliminated when the phantom inserts, not the collimator and point source, were rotated in 10° increments. It appears that the rotation of a rigid shape (like the collimator assembly) in an environment with some finite spatial resolution may result in unexpected behavior when photons interact with the edges of the rotated, rigid structure. Although no clear explanation is available to account for this observation, care should be taken, at least with MCNP, when using a rotated collimator assembly.

Finally, care should be taken when incorporating variance reduction techniques in the simulations. During the early stages of obtaining the results for the TG report, a bug in the implementation of the variance reduction technique of "particle splitting" in GEANT4 version 9.5 and 9.5 patch 1 was found. Due to this bug, the reduced weight given to each split photon was not inherited by any subsequently created secondary particle, resulting in biased results. This was only identified during the comparison of the results of the different packages performed during this work, but could have gone unnoticed in regular research work. The bug has been corrected in later versions of GEANT4, including the one used for the final simulations of all cases included in the TG report (version 9.6 patch 2).

5.C. Simulation times

All the results provided in the electronic spreadsheets include timing information for the simulations implemented with the four Monte Carlo packages listed in Table II. The timing information for each simulation is presented in two forms: (1) the mean time per history simulated and (2) the total simulation time needed to achieve an uncertainty of 1% for a specific scored quantity. Given the possible use of different variance reduction techniques, the latter quantity in general provides more complete information regarding simulation times than the former. The timing results are provided based on the execution time of each simulation with the computational resources used by each TG member, assuming a sequential execution of the simulation on a single core. Therefore, the timing differences include variations in the type and clock speed of the processors used.

It was not the goal of this TG to compare the efficiency of the Monte Carlo packages tested, but rather to provide consistent results among them, with the priority being that the simulation conditions be as similar to each other as possible. Therefore, the use of the most efficient methods to perform these simulations, including the use of variance reduction techniques, varied considerably among the groups implementing the simulations. For example, the PENELOPE simulations were not run with any variance reduction techniques, and the PENELOPE and GEANT4 simulations included the tracking of electrons, although their inclusion is not expected to affect the results. Although it was not always the case, some of the other implementations did include variance reduction and did not include tracking of electrons. This resulted in the PENELOPE simulation times being larger in general than those for the other codes. The specific simulation parameters used in the implementation of all cases for each Monte Carlo package are provided in the Appendix of the report.

The scored quantity that the timing to 1% precision is referred to is clearly marked in the results spreadsheets. Useful observations can be made from the timing information. In various situations, variance reduction techniques have the ability to very substantially improve the efficiency of calculations. For example, a factor of 2000× improvement was obtained in the efficiency of the EGSnrc calculations in Case 6 (x-ray spectra) by using variance reduction techniques. When using variance reduction methods which are part of general purpose code systems, one must take great care when implementing them in a new code and verify their accurate implementation by comparison to runs with the technique turned off.

6. CONCLUSION

This TG developed and implemented six simulations of typical experiments performed for diagnostic imaging research with x-rays. These simulations were implemented with four different well-known and publicly available Monte Carlo packages and the results compared. All details of the inputs and the outputs to the simulations are provided in the TG report and in associated electronic files.

It should be noted that each research investigation has different requirements and therefore the Monte Carlo simulations needed may include higher levels of complexity compared to those presented in this report. For example, the simulations described in Case 3 (mammography and breast tomosynthesis) may be sufficient for investigations of dose in a homogeneous breast phantom. On the other hand, Case 5 (computed tomography dosimetry with a voxelized patient model) may not be sufficient to estimate dose to patients from helical CT scans using automatic exposure control techniques and/or in the presence of a complex bow-tie filter. If higher levels of complexity are required, it is up to the investigator to determine what additional validation steps would be needed (e.g., comparisons with physical measurements). Even for those situations in which the cases presented here are not sufficient for complete validation, comparisons to the reference cases described in this report are strongly recommended. In addition, to consider a new simulation validated, the results obtained do not necessarily have to be within the ranges presented here. It is up to the investigator's discretion to decide when his/her results are considered to be consistent with those found by this TG, and whether his/her code can be considered validated.

Therefore, this TG recommends the following.

1. That all Monte Carlo methods used for medical imaging research for publication in scientific journals be analyzed by comparing results obtained with the proposed methods to results of the relevant case(s) included in the full report.
2. That the analyses be included in published and presented works, with a reference to the specific case

used, and a description of the relevant details of the case including x-ray spectra, imaged objects, tallies performed, and statistics of the Monte Carlo results, as well as the relative percentage difference found between the results obtained and the corresponding mean of the results in this report. If the results obtained are outside the ranges of the results presented in this report, the relative percentage deviations from these ranges also need to be provided and justified.

3. This Task Group suggests the following language to describe the comparison: "The Monte Carlo methods used in this work were tested by performing case(s) XX–YY of the AAPM TG report 195 (*ref*). The simulations were performed using... (*provide details of simulation as listed in point 2 above*). The results of our simulations agree to within *X%* of the mean results published by TG 195 and they (*either*) fall within the range of the results published by TG 195 (*or*) are *Y%* from the lower/upper limit of the range of results published by TG 195. The statistical uncertainty obtained for these simulations was *Z%*." A table listing the detailed results of this analysis with appropriate uncertainty estimates is preferred.
4. That in published and presented works, the Monte Carlo code and methods being used be described in detail including software and hardware, total computation time, and any optional parameters being used (e.g., variance reduction techniques, energy cutoff values). Any relevant modification of the standard code release (such us a new scoring routine) should be described with enough detail that it could be reproduced by independent investigators.
5. That in published works, reference to this report should be made by including a citation to the paper describing this TG report in the journal Medical Physics.

The TG report can also be useful as an educational tool for trainees and scientists gaining experience in the field of Monte Carlo simulations, either as a self-teaching tool to ensure that he/she is able to correctly perform a specific simulation or as assigned learning projects as part of course objectives or training programs.

Of course, many imaging modalities and potential simulation types are not included in the TG report. This first undertaking had to be limited to a manageable number of cases to be completed in a realistic time frame. In the future, similar reports that include simulations of other types of measurements and/or other imaging modalities may be considered by the AAPM.

ACKNOWLEDGMENTS

The TG would like to thank Samir Abboud, Cecilia Marini Bettolo, Bahaa Ghammraoui, Sebastien Incerti, Vladimir Ivantchenko, Alfonso Mantero, Luciano Pandola, Joseph Perl, and Yakun Zhang for their contributions to this work, Eric Jablonowski for his contributions to the figures, and Karen MacFarland, Lisa Giove, Corbi Foster, and Debbie Gilley from the AAPM for assisting with administrative issues throughout the TG's work. Work by I.S. supported in part by the U.S. Department of Energy under Contract No. DE-AC02-76SF00515.

[a]Author to whom correspondence should be addressed. Present address: Department of Radiology and Nuclear Medicine, Radboud University Medical Center, P.O. Box 9101, 6500 HB Nijmegen (766), The Netherlands; Electronic mail: ioannis.sechopoulos@radboudumc.nl
[b]Present address: The Ottawa Hospital Cancer Centre, 501 Smyth Road, Ottawa, Ontario K1H 8L6, Canada.
[c]Present address: Ironwood Cancer and Research Centers, 5810 W Beverly Lane, Glendale, AZ 85306.

[1]The full Task Group report is available at http://www.aapm.org/pubs/reports/RPT_195.pdf.
[2]J. R. Mercier, D. T. Kopp, W. D. McDavid, S. B. Dove, J. L. Lancaster, and D. M. Tucker, "Modification and benchmarking of MCNP for low-energy tungsten spectra," Med. Phys. **27**, 2680–2687 (2000).
[3]G. A. P. Cirrone *et al.*, "Precision validation of GEANT4 electromagnetic physics," in *IEEE Nuclear Science Symposium Conference Record* (IEEE, New York, NY, 2003), Vol. 481, pp. 482–485.
[4]D. Lazaro, I. Buvat, G. Loudos, D. Strul, G. Santin, N. Giokaris, D. Donnarieix, L. Maigne, V. Spanoudaki, S. Styliaris, S. Staelens, and V. Breton, "Validation of the GATE Monte Carlo simulation platform for modelling a CsI(Tl) scintillation camera dedicated to small-animal imaging," Phys. Med. Biol. **49**, 271–285 (2004).
[5]C. M. Salgado, C. C. Conti, and P. H. B. Becker, "Determination of HPGe detector response using MCNP5 for 20–150 keV x-rays," Appl. Radiat. Isot. **64**, 700–705 (2006).
[6]S. España, J. L. Herraiz, E. Vicente, J. J. Vaquero, M. Desco, and J. M. Udias, "PeneloPET, a Monte Carlo PET simulation tool based on PENELOPE: Features and validation," Phys. Med. Biol. **54**, 1723–1742 (2009).
[7]A. Loehr, J. Durst, T. Michel, G. Anton, and P. Geithner, "Comparison of recent experimental data with Monte Carlo tools such as RoSi, GEANT4 and PENELOPE," in *IEEE Nuclear Science Symposium Conference Record (NSS/MIC)* (IEEE, New York, NY, 2009), pp. 3904–3908.
[8]M. Batic, G. Hoff, M. G. Pia, and P. Saracco, "Photon elastic scattering simulation: Validation and improvements to GEANT4," IEEE Trans. Nucl. Sci. **59**, 1636–1664 (2012).
[9]K. Sangroh, T. Y. Terry, Y. Fang-Fang, and J. C. Indrin, "Spiral computed tomography phase-space source model in the BEAMnrc/EGSnrc Monte Carlo system: Implementation and validation," Phys. Med. Biol. **58**, 2609–2624 (2013).
[10]G. Weidenspointner, M. Batic, S. Hauf, G. Hoff, M. Kuster, M. G. Pia, and P. Saracco, "Validation of Compton scattering Monte Carlo simulation models," in *IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)* (IEEE, New York, NY, 2013), pp. 1–5.
[11]R. Schmidt, J. Wulff, and K. Zink, "GMctdospp: Description and validation of a CT dose calculation system," Med. Phys. **42**, 4260–4270 (2015).
[12]D. W. O. Rogers, B. A. Faddegon, G. X. Ding, C. M. Ma, J. We, and T. R. Mackie, "BEAM: A Monte Carlo code to simulate radiotherapy treatment units," Med. Phys. **22**, 503–524 (1995).
[13]I. Kawrakow, "Accurate condensed history Monte Carlo simulation of electron transport. I. EGSnrc, the new EGS4 version," Med. Phys. **27**, 485–498 (2000).
[14]I. Kawrakow, E. Mainegra-Hing, D. W. O. Rogers, F. Tessier, and B. R. B. Walters, "The EGSnrc code system: Monte Carlo simulation of electron and photon transport," NRC Technical Report PIRS-701 v4-2-3-2 (Ottawa, 2011).
[15]D. W. O. Rogers, B. Walters, and I. Kawrakow, "BEAMnrc users manual," NRC Technical Report PIRS-509(A) rev. L (Ottawa, 2011).
[16]S. Agostinelli *et al.*, "GEANT4—A simulation toolkit," Nucl. Instrum. Methods Phys. Res., Sect. A **506**, 250–303 (2003).
[17]J. Allison *et al.*, "GEANT4 developments and applications," IEEE Trans. Nucl. Sci. **53**, 270–278 (2006).
[18]D. B. Pelowitz, J. S. Hendricks, J. W. Durkee, M. R. James, M. L. Fensin, G. W. McKinney, S. G. Mashnik, and L. S. Waters, MCNPX 2.7.A Extensions–Report LA-UR-08-07182, Los Alamos, NM, 2008.
[19]J. Sempau, E. Acosta, J. Baro, J. M. Fernández-Varea, and F. Salvat, "An algorithm for Monte Carlo simulation of coupled electron-photon transport," Nucl. Instrum. Methods Phys. Res., Sect. B **132**, 377–390 (1997).

[20]F. Salvat, J. M. Fernández-Varea, and J. Sempau, PENELOPE-2011: A Code System for Monte Carlo Simulation of Electron and Photon Transport, OECD-NEA, Issy-les-Moulineaux, France, 2011, available in PDF at http://www.oecd-nea.org/lists/penelope.html.

[21]J. Sempau, A. Badal, and L. Brualla, "A PENELOPE-based system for the automated Monte Carlo simulation of clinacs and voxelized geometries-application to far-from-axis fields," Med. Phys. **38**, 5887–5895 (2011).

[22]J. H. Hubbell and S. M. Seltzer, Tables of X-Ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients from 1 keV to 20 MeV for Elements Z = 1 to 92 and 48 Additional Substances of Dosimetric Interest, July 2004 [accessed December 2007], available at http://physics.nist.gov/PhysRefData/XrayMassCoef/cover.html.

[23]International Commission on Radiation Units and Measurements, "Photon, electron, proton, and neutron interaction data for body tissues," ICRU Report No. 46 (International Commission on Radiation Units and Measurements, Bethesda, MD, 1992).

[24]G. R. Hammerstein, D. W. Miller, D. R. White, M. E. Masterson, H. Q. Woodard, and J. S. Laughlin, "Absorbed radiation dose in mammography," Radiology **130**, 485–491 (1979).

[25]K. Cranley, B. J. Gilmore, G. W. A. Fogarty, and L. Desponds, "Catalogue of diagnostic x-ray spectra and other data," Report No. 78 (Institute of Physics and Engineering in Medicine, York, 1997).

[26]International Electrotechnical Commission, IEC 61267-Medical diagnostic x-ray equipment—Radiation conditions for use in the determination of characteristics, Geneva, Switzerland, 2005.