

# Validation of the MC-GPU Monte Carlo code against the PENELOPE/penEasy code system and benchmarking against experimental conditions for typical radiation qualities and setups in interventional radiology and cardiology

David Fernández<sup>1,\*</sup>, Victor García<sup>1</sup>, Clara Delgado<sup>2</sup>, Sara Principi<sup>1</sup>, Maria Amor Duch<sup>1</sup>, Mercè Ginjaume<sup>1</sup>

<sup>1</sup> Universitat Politècnica de Catalunya, Avda. Diagonal 647, 08028 Barcelona, Spain

<sup>2</sup> Physics and Radiation Protection Department, Hospital Universitari Vall d'Hebron, Passeig de la Vall d'Hebron, 119, 08035 Barcelona, Spain

**Abstract. Introduction:** Interventional procedures are associated with potentially high radiation doses to the skin. Monte Carlo codes of radiation transport are considered to be one of the most reliable tools available to assess doses. However, they are usually too time consuming for use in clinical practice. This work presents the validation of the fast Monte Carlo code MC-GPU for application in interventional radiology. **Methodologies:** MC-GPU calculations were compared against the well-validated Monte Carlo simulation code PENELOPE/penEasy by simulating the organ dose distribution in a voxelized anthropomorphic phantom. In a second phase, the code was compared against thermoluminescent measurements performed on slab phantoms, both in a calibration laboratory and at a hospital. **Results:** The results obtained from the two simulation codes show very good agreement, differences in the output were within 1%, whereas the calculation time on the MC-GPU was 2500 times shorter. Comparison with measurements is of the order of 10%, within the associated uncertainty. **Conclusions:** It has been verified that MC-GPU provides good estimates of the dose when compared to PENELOPE program. It is also shown that it presents very good performance when assessing organ doses in very short times, less than one minute, in real clinical set-ups. Future steps would be to simulate complex procedures with several projections.

**KEYWORDS:** Monte Carlo, Interventional procedures, Dose distribution, Execution time, Voxel phantom.

## 1 INTRODUCTION

Nowadays interventional radiology and cardiology are among the medical specialties that are most commonly used because of their minimally invasive procedures. However, the X rays employed during fluoroscopically guided procedures in order to generate medical images entail high exposure to the patient and the possibility of suffering deterministic effects. In this regard, the 2013/59/EURATOM Directive [1] establishes basic safety standards for protection against dangers arising from exposure to ionizing radiation. In article 60, it is specified that the equipment used for interventional radiology must have a device or a feature informing the practitioner of relevant parameters for assessing patient dose at the end of the procedure. Thermoluminescence dosimeters (TLDs) are widely used in various patient dosimetry applications because they are nearly tissue equivalent and have good dosimetry performance. However, they have strong limitations in monitoring dose. In general, it is not possible to predict the locations of the highest dose, so a great number of detectors are needed. In addition, the results are only known a few days after delivery of the doses and they only provide information for very specific regions of tissues. This highlights their limitation in order to deliver data on the dose distribution of extended organs, such as the skin [2]. Different monitoring software packages have been developed to tackle these limitations and also to fulfil the need for automatic recording and evaluation of patient-relevant data [3]. In particular, due to the risk of skin injuries [4], special attention has been paid to skin dose and several codes claim to be able to estimate it or are even specially designed to assess it [5]. In the vast majority of cases these codes apply

---

<sup>1\*</sup> Corresponding author, e-mail: david.fernandez.bosman@upc.edu

mathematical models and the patient is either described in a simplistic way [6, 7, 8, 9] or more complex models are used [10, 11].

Another possible solution may be the development of a dosimetry tool based on Monte Carlo (MC) simulations, for which the main challenge is to obtain results in relatively short computation times. MC codes, which are considered to be the gold standard for computing radiation doses, are generally time-consuming and would not lead to a quick estimation of the requested output. The available MC codes which are able to provide results in short computation times are very few: MC-GPU, FDEIR, ARCHER and irtGPUMCD [12, 13, 14, 15] and all of them are based on the computational power of Graphics Processing Units (GPU). MC-GPU handles voxelized geometries and provides both mean and maximum organ doses for typical diagnostic imaging procedures. FDEIR is based on MC-GPU and focuses on the calculation of skin doses in interventional radiology. The other two codes address other health applications. ARCHER is designed for dose assessment in CT imaging and radiotherapy, while irtGPUMCD is used for dose calculation in nuclear medicine. The Universitat Politècnica de Catalunya (UPC) is participating in Work-Package 2 of MEDIRAD, a H2020 European Project aimed to enhance the scientific bases and clinical practice of radiation protection in the medical field. The task of UPC consists of validating MC-GPU code for patient dose monitoring during interventional procedures. MC-GPU has been validated against measurements in clinical CT [16], however, at present, there are no published validation studies on the reliability of MC-GPU for interventional radiology. Therefore, the purpose of the present work is first to compare MC-GPU with a well-validated MC simulation code and secondly to compare it against measurements carried out in a calibration laboratory and under realistic conditions in an operating room.

## 2 MATERIALS AND METHODS

The MC-GPU code is in the public domain and is freely distributed. It implements a massively multi-threaded MC simulation algorithm for the transport of X rays in voxel geometries and for applications in diagnostic imaging. The publicly available version of MC-GPU uses the X-ray interaction models and cross sections from PENELOPE, 2006 [17]. However, the physics models in PENELOPE 2006 have been updated in the code version PENELOPE, 2014 [18]. Consequently, in this work the physics models in MC-GPU have been changed accordingly. For the diagnostic energy range taken into consideration (20 to 120 keV), the main change between the 2014 and 2006 versions mainly consisted of new cross sections implemented for Rayleigh and photoelectric effects. Differences in the total mean free paths are small at low energies (2-3% depending on the material of interest) and almost negligible for high energies. However, updating the materials description is relevant for this study since photoelectric interactions are predominant at the typical energy range of interventional cardiology. In addition, a set of functions were developed to automatically set the optimal values for the GPU calculations (number of blocks per kernel, number of threads per block, number of histories per thread to be simulated by the GPU). Finally, the MC-GPU libraries were modified to increase the maximum number of different materials that could be taken into account in the calculations. Initially, the maximum number was set to 15 and it was increased to 30. It was tested that, with the computational resources described in 2.1.3, one could simulate up to 70 different materials. The impact on the efficiency of the simulations [19] was negligible, lower than 1%.

The code package used for the MC-GPU validation was PENELOPE (version 2014) / penEasy (version 2015). PENELOPE is one of the reference codes for Monte Carlo simulation of electron, positron and photon transport. penEasy [20] is a general-purpose main program defined specifically for PENELOPE. It provides the user with a set of source models and tallies which are invoked from a structured code.

For benchmarking of the simulations against clinical measurements, a series of irradiations was carried out at the UPC Calibration Laboratory under controlled conditions and subsequently other irradiations were performed in a realistic environment at the Hemodynamic Department of the Hospital Universitari Vall d'Hebron (HUVH), simulating a clinical scenario. The experimental measurements were performed with thermoluminescent dosimeters. Details of both simulations and experiments are provided below.

## 2.1 MONTE CARLO SIMULATION

### 2.1.1 Comparison against PENELOPE/penEasy

The Duke anthropomorphic phantom from the Virtual family was used. The Duke voxel phantom has  $122 \times 62 \times 372$  voxels of dimensions  $0.5 \times 0.5 \times 0.5 \text{ cm}^3$ . It represents a 34-year-old-male and is one of the four whole body models from the Virtual family [21]. For the sake of simplicity, in this verification, 9 materials were taken into account among all the initially available segmented tissues: muscle, soft tissue, bone, cartilage, adipose tissue, blood, skin, lung and air.

The aim was to calculate and compare the doses both in a whole organ or tissue (made up of thousands of voxels) and at an individual voxel level. The dose in each material is calculated by the program as the average of the dose across all voxels belonging to the same material. The program also provides the value and location of the maximum dose.

Due to its energy range of application and to speed calculations, MC-GPU does not simulate the transport of charged particles. For comparison with PENELOPE, the energy absorption parameters of electrons and positrons in PENELOPE/penEasy were set to very high energies (e.g.  $1 \text{E}9 \text{ MeV}$ ). In this way, once the secondary particles are generated, they are absorbed locally in the material and are no longer followed. This approximation is valid for the energy range of interventional radiology [22, 23, 24]. The requested outputs for the PENELOPE/penEasy code were the dose in each voxel and the energy deposited in each labeled material. Doses are given in  $\text{eV}/(\text{g} \cdot \text{history})$  ( $\text{eV}/\text{g}$  per simulated history), while the energy deposition units are  $\text{eV}/\text{history}$ . Then, both the value and position of the maximum dose are easily retrieved from the dose distribution. On the other hand, the dose in each tissue can be straightforwardly calculated by dividing the deposited energy by the corresponding tissue mass.

The X-ray source used in the comparison is a typical diagnostic X-ray spectrum defined in the International Standard IEC 61267 [25], and is one of the reference radiation qualities in the field of X-ray diagnostics when dealing with radiation emerging from the X-ray source. Calculations were performed for RQR6 quality [25], generated by a tube voltage of 80 kV (see details in Table 1). The energy distribution of the X-ray spectrum for the simulations was generated by using XCOMP5R [26]. The source was defined in order to completely irradiate the chest of the phantom. The postero-anterior (PA) projection was simulated so that the radiation strikes the back of the phantom. The Duke phantom was positioned at 30.5 cm from the beam source. A pyramidal field was defined in order to completely irradiate the Duke phantom chest and had dimensions of  $61 \times 61 \text{ cm}^2$  at a distance of 30.5 cm from the source. 61 cm is just the width of the Duke. This beam size is much larger compared to the ones present in interventional radiology, which are aimed at irradiating a specific organ. The reason for choosing this size was to ensure that all the different organs in the Duke phantom chest were irradiated.

### 2.1.2 Comparison against experimental measurements

Different experimental measurements were carried out at the UPC Calibration Facility and in an operating room of the HUVH hospital. Details on the geometry of the experimental irradiations and the voxelized representation with MC-GPU are given in sections 2.2.3.1 and 2.2.3.2.

For comparison of simulated doses with experimental values in a clinical environment, a normalization factor  $N$  was used to convert the calculated dose values per simulated history (eV/(g·history)) to absolute values (Gy). This factor is the ratio between experimental ( $K_{\text{air,exp}}$ ) and simulated ( $K_{\text{air,sim}}$ ) air kerma at a reference point:

$$N = \frac{K_{\text{air,exp}}}{K_{\text{air,sim}}} \quad (1)$$

The experimental air kerma ( $K_{\text{air,exp}}$ ) value was obtained at the reference point from the kerma area product (KAP). The KAP value is included in the radiation dose structured report (RDSR) defined by DICOM (DICOM PS3.16 2020b) [27] and is obtained with a KAP-meter located on the housing of the gantry at a fixed distance from the focal spot. The reference point is a virtual position defined at a specified distance from the isocenter of the C-arm; the field size at the reference point can be obtained using the inverse square law from the source-to-detector distance and the collimated area at the detector, both included in the RDSR. In order to obtain the simulated air kerma ( $K_{\text{air,sim}}$ ), a box of air of dimensions 1 x 0.5 x 1 cm<sup>3</sup> is simulated, mimicking the clinical air kerma at the reference point.

Regarding the measurements in the Calibration Laboratory, instead of comparing values in units of dose, the comparison was carried out in terms of relative values obtained as the quotient of air kerma within tissue-equivalent materials and air kerma in free air. Therefore, the above-mentioned normalization factor  $N$  was not needed.

The energy distribution of the X-ray beams defined in Tables 1 and 3 was generated by using XCOMP5R.

### 2.1.3 Computational resources

The simulations presented in this work were run on the UPC Argos Cluster. PENELOPE/penEasy runs on any of the 15 available nodes and runs on two CPU Intel Xeon E5520 Processors, 8x2GB Dual Rank UDIMMs 1066MHz, 1 x 160GB SATA 7200 3,5". The compiler used in PENELOPE/penEasy was gfortran 4.8.5. Although the simulations were carried out on a computer cluster, it is worth pointing out that each simulation ran on a single CPU, since PENELOPE cannot be parallelized.

## 2.2 MEASUREMENTS

### 2.2.1 Thermoluminescent dosimetric system

The detector of the thermoluminescent dosimetric system is a LiF:Mg,Cu,P crystal manufactured by Conqueror Electronics Technology Co.Ltd. (Beijing, China) under the trade name TLD-2000C. They are 4.5-mm-diameter circular pellets, 0.8 mm thick, with a mass density of 2.65 g/cm<sup>3</sup>. The annealing process previous to the detectors exposure to radiation was performed in a PTW TLDO oven, and consisted of 240 °C for 10 min followed by a fast quench on an aluminium block. The readout after the exposure was done on a semiautomatic Harshaw 5500 reader (Thermo-Electron) in a two-step heating cycle: a preheat process at a temperature of 160 °C for 10 s, followed by a reading phase of 26 s up to 250 °C with a linear heating rate of 4 °C/s. The readout was carried out within 24 h after irradiation.

Instead of using batch calibration, detectors were individually calibrated. To determine the response of a single detector with respect to the whole set, individual calibration factors were determined by irradiating all detectors using a <sup>137</sup>Cs source, at a value of air kerma of 3 mGy.

The use of an individual correction factor is a standard procedure to reduce uncertainty in the measurement [28]. Several background dosimeters were used for each measurement and the background signal was subtracted from the measured reading. In order to provide the output in terms of absorbed dose (Gy), calibration factors were calculated in order to convert reader arbitrary units to air kerma (Gy) in the material in which the detectors are embedded. It is worth mentioning that in the energy range of interest differences between air kerma and absorbed dose in air are negligible. These calibration factors were determined under reference conditions at the UPC facility.

## 2.2.2 Energy calibration for the TLDs used in the hospital

The calibration factors are both energy and material dependent. In order to have appropriate calibration factors to measure the air kerma when the TLDs were irradiated in the hospital, the energy calibration of the dosimetric system was obtained by irradiating a group of 10 dosimeters at the UPC calibration laboratory. The laboratory has been accredited by the Spanish Accreditation Board and is traceable to the PTB (German National Metrology Institute) for IEC61267 qualities [25]. The detectors were irradiated with two IEC61267 qualities: RQR5 and RQR6. The main features of the beams are presented in Table 1. These qualities were chosen as they have similar X-ray characteristics to the beams encountered in the hospital [29].

## 2.2.3 Experimental set-up

Different set-ups were tested. Slab phantoms with surface dimensions of 30x30 cm<sup>2</sup> and made of different tissue-equivalent materials were used. Some of these plates have small cavities where the detectors can be accommodated. The materials of the plates were plastic water (PW®,  $\rho=1.03$  g/cm<sup>3</sup>), lung tissue-equivalent material ( $\rho=0.199$  g/cm<sup>3</sup>) and bone tissue-equivalent material ( $\rho=1.93$  g/cm<sup>3</sup>) from Computerized Imaging Reference Systems (CIRS). Lung and bone equivalent materials match the linear attenuation coefficients of their reference tissues [30] to within +/- 2.5%. The materials were simulated according to the atomic composition specified by the manufacturer. Nine detectors were irradiated at each measurement position and the measured dose to be compared to the simulation is then obtained as the average of the nine readings.

### 2.2.3.1 Experiment under controlled conditions

The first validation was performed under controlled conditions at the UPC Calibration Laboratory. Sandwich-like phantoms made of the above-mentioned tissue-equivalent materials were used for the irradiations. The TLD detectors were placed at different depths (Table 2) defined with respect to the point where the photons enter the phantom. RQR5 and RQR6 qualities were used for the irradiations (Table 1). In all cases the imparted air kerma was 5 mGy at 2 m from the focal spot. The radiation field at the irradiation plane (defined by the plane where the photons enter the material) had a diameter of 32 cm.

**Table 1:** Features of the beam qualities used for irradiation under controlled conditions at the UPC laboratory.

Beam quality	Anode angle	Voltage (kV)	Inherent filtration	Added filtration	1st HVL
RQR5	18°	70	7 mm Be	2.5 mm Al	3.07 mm Al
RQR6	18°	80	7 mm Be	3.0 mm Al	3.45 mm Al

**Table 2:** Phantom thickness and dosimeter position inside the phantom for irradiation under controlled conditions at the UPC laboratory.

Phantom	Total thickness (cm)	Depth (cm)
PW®	1.5	0.5
Lung	4	2
Bone	2	1

The irradiations were also carried out free in air. In this case, the detectors were sealed in small opaque polyethylene plastic bags, with a density thickness of  $11 \text{ mg/cm}^2$ . A sketch of the irradiation configurations at the UPC laboratory is shown in Figure 1.

Three voxelized geometries were implemented in MC-GPU to reproduce the PW®, lung and bone phantoms. The voxel size dimension was 3.33 cm, as the nine detectors in the experimental set-up were arranged over an area of 3.33 cm x 3.33 cm, whereas the voxel thickness was 0.1 cm. Therefore, TLDs were modeled as a single 0.1-cm-thick voxel made of air. In order to reproduce the irradiation of the TLDs free in air (without attenuation and scattering material), the same geometries were used, but were considered as being all made of air.

### 2.2.3.2 Experiment in hospital

The second part of the measurements was performed at the HUVH Department of Hemodynamics in a real operating room equipped with a Philips Allura Clarity X-ray system. A sketch of the irradiation configuration at the HUVH hospital is shown in Figure 2. In order to simulate the patient during an interventional procedure for thorax irradiation, the phantom was made up of three different materials in the following order from bottom to top: 2 cm of PW®, 2 cm of bone, 9 cm of lung and 2 cm of PW® (Figure 3 a). Hereafter it is referred as *phantom 1*. The detectors were located at 1, 3 and 6 cm depth in PW®, bone and lung plates respectively. A similar irradiation was repeated for a phantom made of PW® layers only, in which the detectors were placed at 1, 7 and 13 cm depth (Figure 3 b). This phantom is, hereafter, referred to as *phantom 2*.

Simulations were designed and carried out for the two experimental configurations. The nine TLDs at each depth level were modeled as a single voxel made of air. Independent simulations were carried out for each depth and each material.

For the purposes of benchmarking, the geometry set-up of the Monte Carlo simulation should be as close as possible to the realistic scenario. In the clinical practice, the table pad where the patient lies reduces the beam intensity to the patient for under-table projections, i.e. PA. In order to quantify this attenuation, two measurements were performed in the clinical room with a semiconductor detector, the RaySafe X2, located first on the table and, subsequently, under the same irradiation conditions, beneath the table. The ratio between the two measurements (given in terms of dose rate) gave the attenuation provided by the treatment couch. These two last experiments, both with and without the table, were reproduced using Monte Carlo simulations to

calculate the Aluminum thickness that would produce the same attenuation. Based on this result, a 1.8-mm Al layer was finally introduced in the voxel simulation geometry of the phantoms.

Irradiations were carried out in the conditions described in Table 3. The radiation field size was 25 cm x 25 cm at a distance of 100 cm between the source and the image detector.

Table 3: Features of the radiation beams used for irradiations with *phantom 1* (Figure 3 a) and *phantom 2* (Figure 3 b) under clinical conditions.

Phantom type	Anode angle	Voltage (kV)	Inherent filtration	Added filtration
Phantom 1	11°	62	2.5 mmAl	1 mmAl+0.1 mmCu
Phantom 2 (PW®)	11°	65	2.5 mmAl	1 mmAl+0.1 mmCu

### 2.3 UNCERTAINTY EVALUATION

Uncertainties of measurements are calculated according to the recommendations of the Guide to the Expression of Uncertainty in Measurement (GUM) [31]. All sources of uncertainty associated to a given measurement are evaluated and combined to obtain the range of dispersion of the measurand. In order to have a confidence level of approximately 95%, an expanded uncertainty is calculated, multiplying the standard uncertainty by a coverage factor,  $k$ , equal to 2.

For the calculations, when only Monte Carlo results are compared, the statistical standard deviation,  $sd$ , is considered. In order to have a confidence level of approximately 95%, uncertainties are expressed as 2  $sd$ . When Monte Carlo calculations are normalized using equation (1), the uncertainty associated to the experimental air kerma is also computed following [31] recommendations. An uncertainty of about 10% for  $k=2$  for the experimental KAP measurement was considered and has been included in the calculation of the uncertainty of the simulation results.

## 3 RESULTS

### 3.1 VALIDATION WITH THE MONTE CARLO SIMULATION CODE PENELOPE/PENEASY

Figure 4 and Table 4 show the dose distribution for chest irradiation of the anthropomorphic phantom and the average dose values for each tissue, respectively. The codes showed very good agreement with differences below 1%. These small differences can be attributed to the different algorithms that the codes use to obtain photon cross sections. PENELOPE v2014 photon cross sections are tabulated on a denser grid of energies to describe the structure of the cross section near photo-electric absorption edges and interpolated on a linear log-log scale, whereas MC-GPU applies tabulated values on a linear grid.

The PENELOPE/penEasy code took around 9.3 hours of execution time to complete the simulation and achieve standard errors in the materials of about 0.1% ( $k=2$ ). On the other hand,

MC-GPU execution time was 13 seconds, reaching standard errors of the same order while being about 2500 times faster than PENELOPE/penEasy. Furthermore, both programs identified the maximum voxel dose at the same position, that is, in the air at the entrance of the patient. High doses were also observed on the bone and on the skin on the most exposed side (on the back of the patient). The difference between the Peak Skin Dose (PSD) reported by both codes was only 0.6% with an associated statistical uncertainty ( $k=2$ ) of 1% for MC-GPU and 2 % for PENELOPE/penEasy. Both codes reported the same location of the PSD. In contrast, the reported dose differences between FDEIR and PHITS [32] were up to 6 % in the validation study of FDEIR [13]. FDEIR is based on the publicly available version of MC-GPU, which in turn, as has been pointed out previously, uses relevant photon interaction data from PENELOPE 2006. Both PHITS and PENELOPE 2006 are based on the LLNL Evaluated Photon Data Library [33], but PENELOPE uses different numerical algorithms.

**Table 4:** Type of tissue, number of voxels, computed dose from PENELOPE/penEasy and MC-GPU, uncertainty of the simulation (U) with ( $k=2$ ), ratio between the two codes for the anthropomorphic phantom simulation and uncertainty of the ratio ( $U_{\text{ratio}}$ ).

Material/Tissue	#voxels	PENELOPE dose eV/(g·history)	U (%)	MC-GPU dose eV/(g·history)	U (%)	$\frac{\text{MC-GPU}}{\text{PENELOPE}}$	$U_{\text{ratio}}$ (%)
Air	2273413	1.38E-01	0.16	1.38E-01	0.16	1.00	0.23
Muscle	278365	2.27E-01	0.01	2.27E-01	0.01	1.00	0.02
Soft tissue	34535	1.37E-01	0.05	1.38E-01	0.04	1.01	0.07
Bone	31577	6.30E-01	0.02	6.32E-01	0.02	1.00	0.02
Cartilage	11840	6.34E-02	0.12	6.39E-02	0.11	1.01	0.16
Adipose	104890	1.66E-01	0.02	1.66E-01	0.02	1.00	0.03
Blood	8945	1.07E-01	0.10	1.07E-01	0.10	1.00	0.15
Skin	39449	2.68E-01	0.03	2.67E-01	0.03	1.00	0.04
Lung	30794	2.93E-01	0.04	2.95E-01	0.04	1.01	0.06

In order to verify the agreement of the dose distribution across all voxels between the two codes, the dose-volume histograms (DVH) were obtained for each tissue. The cumulative DVH relates radiation dose to tissue volume in order to summarize 3D dose distributions in a graphical 2D format (Figure 5). When the material is distributed in the human body, such as the skin, the voxel dose varies within a broad range (back skin directly exposed to the radiation beam vs. skin on the front of the patient). This makes it also difficult to appreciate the differences between the two outputs. This is not the case of tissues related to specific limited organs, such as the lungs. Figure 5 shows the cumulative DVH for skin and lung, showing very good agreement between MC-GPU and PENELOPE.

### 3.2 BENCHMARKING AGAINST EXPERIMENTAL MEASUREMENTS

#### 3.2.1 Experiment in controlled conditions

Table 5 shows the results of the measurements and simulations given as the ratio of the detector reading within the different materials ( $L_{\text{mat}}$ ) and in air ( $L_{\text{air}}$ ), both for RQR5 and RQR6 qualities, corresponding to the experiments performed at the UPC calibration lab. The simulation time,  $t_{\text{sim}}$ ,



and the ratio between experimental and simulation results are also provided together with the expanded uncertainty of the ratio for a coverage factor ( $k=2$ ). The uncertainty associated with TLD measurements corresponds to the combination of the standard deviation of the different dosimeter readings and the uncertainty associated with the experimental set-up, and is of the order of 5%. Simulations were performed in MC-GPU. 5E10 histories were generated, resulting in an execution time of 1.5 minutes and standard deviation between 2% and 4% (2 sd). Since data are normalized to air kerma, only the statistical uncertainty is considered in the calculation in this verification.  $U_{\text{ratio}}$  is the expanded uncertainty of the ratio between simulated and measured results and is obtained as a combination of the experimental and Monte Carlo uncertainties (for  $k=2$ ).

**Table 5:** Experimental and simulated ratios  $L_{\text{mat}}/L_{\text{air}}$  for PW<sup>®</sup>, Lung-, Bone- equivalent phantoms and RQR5 and RQR6 qualities. Time of simulation  $t_{\text{simul}}$  (s) and ratio between the experimental and simulated output with associated expanded uncertainties ( $U_{\text{ratio}}$ ) (for  $k=2$ )

Phantom	Radiation Quality	TLD		Simulation (5E10 histories)		
		$L_{\text{mat}}/L_{\text{air}}$	$L_{\text{mat}}/L_{\text{air}}$	$T_{\text{simul}}$ (s)	Sim/TLD	$U_{\text{ratio}}$
PW	RQR5	0.953	0.975	70	1.02	5%
Lung	RQR5	1.087	1.083	46	1.00	6%
Bone	RQR5	0.279	0.270	90	0.97	7%
PW	RQR6	1.070	1.024	71	0.96	6%
Lung	RQR6	1.122	1.058	47	0.94	6%
Bone	RQR6	0.367	0.366	97	1.00	7%

In column 6 of Table 5 the ratios between the simulated and experimental values are reported. In general, differences are below 5%, except for the lung with RQR6, where a difference of 6% was found. Simulation were also performed with 2E11 histories. In this case, the execution time was up to 7 minutes and standard deviation was below 2%. However, there was no significant improvement with respect to the simulation shown in Table 5. Thus, the faster simulations were preferred because of computational time saving. In all cases, these differences are within the statistical expanded uncertainty of the ratio,  $U_{\text{ratio}}$ .

### 3.2.2 Experiment in clinical conditions

Table 6 compares the results of the measurements performed at the hospital and the corresponding simulations implemented on MC-GPU with 1E10 histories. TLD readings are corrected by the energy calibration factor described in 2.2.2. Thus, the uncertainty in measurements is higher (8%) than the uncertainty of the measurements presented in section 3.2.1, as, in addition to the uncertainties considered in Table 5, the difference between the calibration beams and the hospital beams, and the reference air kerma uncertainty, must be taken into account. Simulation data are given in units of absorbed dose in air by applying the normalization factor,  $N$ , defined in equation (1). Thus, the uncertainty associated with the simulated values includes both the statistical standard deviation and the uncertainty associated with the KAP recorded in the RDSR. Column 8 shows the ratio between simulated and measured doses. The last column shows the expanded uncertainty associated with the ratio. All uncertainties are given with a 95 % confidence level ( $k=2$ ).

**Table 6:** Comparison of experimental and simulated absorbed dose in a clinical set-up (expanded uncertainties are given for  $k=2$ )

Phantom Type	Material (depth)	Experimental		Simulated		$t_{\text{simul}}$ (s)	$\frac{\text{Simulated}}{\text{Experimental}}$	$U_{\text{ratio}}$ (%)
		D (mGy)	U(%)	D (mGy)	U(%)			
1	PW® (1 cm)	2.61	8%	2.46	10%	37	0.95	13%
1	Bone (3 cm)	0.67	8%	0.70	11%	37	1.06	13%
1	Lung (6 cm)	0.21	8%	0.24	12%	37	1.12	14%
2	PW® (1 cm)	4.44	8%	3.99	10%	55	0.90	13%
2	PW® (7 cm)	1.30	8%	1.24	12%	55	0.95	14%
2	PW® (13 cm)	0.18	8%	0.19	16%	55	1.07	18%

Good agreement is observed between the simulated and experimental doses. Differences are below 15% and within the associated uncertainty.

#### 4 DISCUSSION AND CONCLUSIONS

The comparison of MC-GPU and PENELOPE/penEasy for the simulation on the Duke anthropomorphic phantom shows very good agreement in the calculation of the mean doses in the different tissues, with differences below 1%. Furthermore, both codes identified the maximum dose at the same voxel position, which is at the entrance surface of the patient near the center of the beam. The most relevant difference between the two codes is the simulation time. MC-GPU implements a massively multi-threaded Monte Carlo simulation algorithm for the transport of X-rays on NVIDIA GPUs, which considerably speeds up the simulation with respect to execution on a single CPU. Execution time for PENELOPE running on a single CPU took approximately 2500 times longer than the same simulation implemented on MC-GPU. As one of the most relevant issues of the study was the comparison of the computing times for completing the simulations by the two algorithms, and agreement between both outputs, the results can be considered widely satisfactory. We can conclude that it has been verified that MC-GPU provides good estimates of the dose when compared with the PENELOPE code in a much shorter execution time.

Comparison with the experimental set-up is more challenging. One of the main issues is the model of the geometry for the simulated scenario. The differences between the MC-GPU estimates and the doses measured at the UPC Calibration Laboratory were within 5%, whereas they increased to 15 % for the hospital set-up. In all cases, differences were within uncertainties. IEC60580 [34] tolerance for displayed KAP is specified up to  $\pm 35\%$  and manufacturer's specifications are typically  $\pm 25\%$ . In addition, the energy dependence of KAP-meters is often around 10% [35]. Therefore, the differences between calculations and TLD measurements obtained in this study can be considered excellent given the uncertainties found in this field. Also, for the implementation of the set-up at the HUVH hospital, MC-GPU execution time was less than one minute. This shows that MC-GPU overcomes the time limitations of CPU-based MC codes, and is a reliable and near real-time tool for dose estimation in clinical practice, provided that relevant parameters of the procedure and the KAP value for dose normalization are available on-line.

The results can also be considered satisfactory when comparing with other software performances [36], e.g. some studies reported organ dose differences up to 17 % when estimations from Dose (QAELUM NV, Leuven, Belgium) were compared against

VirtualDoseCT (Virtual Phantoms Inc., Albany, New York, USA). In the specific domain of interventional radiology, VirtualDoseIR (Virtual Phantoms Inc., Albany, New York, USA) organ dose results were compared with measurement data previously reported in literature, with reported differences of 40% [37]. Regarding skin dose estimations, the validation studies reported a wider range of differences [38, 39, 40] between 15 % and 76%.

The different MC-GPU approach used to calculate doses when compared with other software tools now available on the market should be commented on. The approach used by other software packages is the same and consists of the correction of the displayed air kerma given by the system by a set of factors accounting for different phenomena such as table attenuation or backscattering of the photons [5]. MC-GPU calculates mean doses in tissues and the peak doses for different organs as well as other complex analyses, such as cumulative dose volume histograms.

This paper shows the high performance and accuracy of MC-GPU for dose assessment when compared to standard Monte Carlo codes and TLD measurements. In addition, it should be highlighted that MC-GPU provides not only the dose values at specific positions, as in the case of the TLDs, but also the dose distribution, the position of maximum dose and the organ doses.

At this point, the following steps for general implementation in clinics are the following: simulation of complex procedures with different projections (not only postero-anterior) in order to reproduce more realistic interventional cardiology and radiology procedures and, benchmarking against measurements in clinical scenarios with anthropomorphic phantoms of different sizes, as in this work only the 34-year-old Duke phantom was computationally tested.

### ***Acknowledgments***

The authors gratefully acknowledge the interesting and helpful assessment and discussions with the author of MC-GPU, Andreu Badal, and the valuable contributions of the UPC master students Anna Torras and Alvaro Merino, as well as the collaboration of staff at the UPC Laboratory and HUVH hospital where experimental work was carried out.

This project has received funding from the Euratom research and training programme 2014-2018 under grant agreement No 755523.

## **5 REFERENCES**

- [1] European Society of Radiology (ESR), Akata D, Adam EJ, Damilakis J, Ducou le Pointe H, Loose R, et al. Summary of the European Directive 2013/59/Euratom: essentials for health professionals in radiology. *Insights Imaging* 2015;6:411–7. <https://doi.org/10.1007/s13244-015-0410-4>.
- [2] Bjelac OC, Farah J, Malchair F, Siiskonen T. EURADOS Report: Patient maximum skin dose in interventional procedures in radiology and cardiology. 2019.
- [3] Boos J, Meineke A, Bethge OT, Antoch G, Kröpil P. Dose Monitoring in Radiology Departments: Status Quo and Future Perspectives. *RoFo Fortschritte Auf Dem Gebiet Der Rontgenstrahlen Und Der Bildgeb Verfahren* 2016;188:443–50. <https://doi.org/10.1055/s-0041-109514>.
- [4] Balter S, Hopewell JW, Miller DL, Wagner LK, Zelefsky MJ. Fluoroscopically Guided Interventional Procedures: A Review of Radiation Effects on Patients' Skin and Hair. *Radiology* 2010;254:326–41. <https://doi.org/10.1148/radiol.2542082312>.
- [5] Malchair F, Dabin J, Deleu M, Merce MS, Bjelac OC, Gallagher A, et al. Review of skin dose calculation software in interventional cardiology. *Phys Medica* 2020;80:75–83. <https://doi.org/10.1016/j.ejmp.2020.09.023>.
- [6] Bordier C, Klausz R, Desponds L. Patient dose map indications on interventional X-ray systems and validation with gafchromic XR-RV3 film. *Radiat Prot Dosimetry*

- 2014;163:306–18. <https://doi.org/10.1093/rpd/ncu181>.
- [7] Magnier F, Poulin M, Van Ngoc Ty C, Osmond E, Bonniaud G, Coulot J, et al. Comparison of Patient Skin Dose Evaluated Using Radiochromic Film and Dose Calculation Software. *Cardiovasc Intervent Radiol* 2018;41:762–71. <https://doi.org/10.1007/s00270-018-1888-1>.
- [8] Habib Geryes B, Hadid-Beurrier L, Waryn MJ, Jean-Pierre A, Farah J. Benchmarking the DACS-integrated Radiation Dose Monitor® skin dose mapping software using XR-RV3 Gafchromic® films. *Med Phys* 2018;45:4683–92. <https://doi.org/10.1002/mp.13125>.
- [9] Hintenlang D, Little K, Jiang X, Yang X, Fitousi N. 2018. Evaluation of fluoroscopic dose metrics predicted by dose management software. 60th annual meeting of the American Association of Physicists in Medicine, AAPM 2018, 29 July – 2 August 2018, Nashville, Tennessee, USA.
- [10] Rana VK, Rudin S, Bednarek DR. A tracking system to calculate patient skin dose in real-time during neurointerventional procedures using a biplane x-ray imaging system. *Med Phys* 2016;43:5131–44. <https://doi.org/10.1118/1.4960368>.
- [11] Gardavaud F, Tavolaro S, Gussenmeyer-Mary N, Cornelis F, Boudghène F. 56 Peak Skin Dose evaluation for vascular clinical procedures in interventional radiology: a comparison between three computation numerical solutions. *Phys Medica* 2018;56:32. <https://doi.org/10.1016/j.ejmp.2018.09.069>.
- [12] Badal A, Badano A. Accelerating Monte Carlo simulations of photon transport in a voxelized geometry using a massively parallel graphics processing unit. *Med Phys* 2009;36:4878–80. <https://doi.org/10.1118/1.3231824>.
- [13] Takata T, Kotoku J, Maejima H, Kumagai S, Arai N, Kobayashi T, et al. Fast skin dose estimation system for interventional radiology. *J Radiat Res* 2018;59:233–9. <https://doi.org/10.1093/jrr/rrx062>.
- [14] Peng Z, Fang X, Yan P, Shan H, Liu T, Pei X, et al. A method of rapid quantification of patient-specific organ doses for CT using deep-learning-based multi-organ segmentation and GPU-accelerated Monte Carlo dose computing. *Med Phys* 2020;47:2526–36. <https://doi.org/10.1002/mp.14131>.
- [15] Frezza A, Joachim-Paquet C, Chauvin M, Després P. Validation of irtGPUMCD, a GPU-based Monte Carlo internal dosimetry framework for radionuclide therapy. *Phys Medica* 2020;73:95–104. <https://doi.org/10.1016/j.ejmp.2020.04.010>.
- [16] Sharma S, Kapadia A, Fu W, Abadi E, Segars WP, Samei E. A real-time Monte Carlo tool for individualized dose estimations in clinical CT. *Phys Med Biol*. 2019;64(21):215020. <https://doi.org/10.1088/1361-6560/ab467f>.
- [17] Salvat F, Fernandez-Varea JM, and Sempau J. PENELOPE 2006: A code system for monte carlo simulation of electron and photon transport. Workshop Proceedings Barcelona, Spain, 4-7 July 2006, OECD 2006 NEA No.6222, 2006.
- [18] NEA (2015), PENELOPE 2014: A code system for Monte Carlo simulation of electron and photon transport: Workshop, Barcelona, Spain, 29 June - 3 July 2015, OECD Publishing, Paris, <https://doi.org/10.1787/4e3f14db-en>.
- [19] Saidi P, Sadeghi M, Tenreiro C. Variance Reduction of Monte Carlo Simulation in Nuclear Engineering Field. *Theory Appl Monte Carlo Simulations* 2013. <https://doi.org/10.5772/53384>.
- [20] Sempau J, Badal A, Brualla L. A PENELOPE-based system for the automated Monte Carlo simulation of clinacs and voxelized geometries-application to far-from-axis fields. *Med Phys* 2011;38:5887–95. <https://doi.org/10.1118/1.3643029>.
- [21] IT'IS Foundation. Virtual population, <https://itis.swiss/virtual-population/virtual-population/overview/>
- [22] Daures J, Gouriou J, Bordy JM. Monte Carlo determination of the conversion coefficients Hp(3)/Ka in a right cylinder phantom with “PENELOPE” code. Comparison with “MCNP” simulations. *Radiat Prot Dosimetry* 2011;144:37–42. <https://doi.org/10.1093/rpd/ncq359>.

- [23] Principi S, Guardiola C, Duch MA, Ginjaume M. Air kerma TO Hp(3) conversion coefficients for IEC 61267 RQR X-ray radiation qualities: Application to dose monitoring of the lens of the eye in medical diagnostics. *Radiat Prot Dosimetry* 2016;170:45–8. <https://doi.org/10.1093/rpd/ncv435>.
- [24] W. G. Cross, J. Böhm, M. Charles, E. Piesch, S. M. Seltzer. Report 56 Dosimetry of External Beta Rays for Radiation Protection. *Journal of the International Commission on Radiation Units and Measurements*, Volume os29, Issue 1, 5 January 1997, <https://doi.org/10.1093/jicru/os29.1.Report56>.
- [25] International Electrotechnical Commission, Medical diagnostic X-ray equipment - radiation conditions for use in the determination of characteristics, IEC 61267, IEC, Geneva (2005).
- [26] Nowotny R, Höfer A. Ein Programm für die Berechnung von diagnostischen Röntgenspektren. *RoFo Fortschritte Auf Dem Gebiete Der Rontgenstrahlen Und Der Nukl* 1985;142:685–9. <https://doi.org/10.1055/s-2008-1052738>.
- [27] NEMA PS3 / ISO 12052, Digital Imaging and Communications in Medicine (DICOM) Standard, National Electrical Manufacturers Association, Rosslyn, VA, USA (available free at <http://www.dicomstandard.org/>).
- [28] European Commission. Ambrosi P, Garcia-Alves J, Bartlett D. Technical Recommendations for Monitoring Individuals Occupationally. <https://doi.org/10.2768/19817>.
- [29] Chiriotti S, Ginjaume M, Vano E, Sanchez R, Fernandez JM, Duch MA, et al. Performance of several active personal dosimeters in interventional radiology and cardiology. *Radiat. Meas.*, vol. 46, Pergamon; 2011, p. 1266–70. <https://doi.org/10.1016/j.radmeas.2011.05.073>.
- [30] White DR, Booz J, Griffith RV, Spokas JJ, Wilson IJ, Report 44. Tissue substitutes in radiation dosimetry and measurement. *Journal of the International Commission on Radiation Units and Measurements*. 1989, 23, 1, <https://doi.org/10.1093/jicru/os23.1.Report44>.
- [31] Joint Committee for Guides in Metrology. Evaluation of measurement data — Guide to the expression of uncertainty in measurement. JCGM 100:2008 GUM 1995 with minor corrections. Available in: [https://www.bipm.org/utis/common/documents/jcgm/JCGM\\_100\\_2008\\_E.pdf](https://www.bipm.org/utis/common/documents/jcgm/JCGM_100_2008_E.pdf).
- [32] Sato T, Iwamoto Y, Hashimoto S, Ogawa T, Furuta T, Abe S, Kai T, Tsai P, Ratliff HN, Matsuda N, Iwase H, Shigyo N, Sihver L and Niita K. Features of particle and heavy ion transport code system (PHITS). version 3.02. *J. Nucl. Sci. Technol.* 2018;55(6):684-690. <https://doi.org/10.1080/00223131.2017.1419890>.
- [33] Cullen D, Hubbell J and Kissel L. (1997). EPDL97: The evaluated photon data library.
- [34] International Electrotechnical Commission, Medical electrical equipment - Dose area product meters, IEC 60580, IEC, Geneva (2019).
- [35] Toroi P, Komppa T, Kosunen A, Tapiovaara M. Effects of radiation quality on the calibration of kerma-area product meters in x-ray beams. *Phys Med Biol* 2008;53:5207–21. <https://doi.org/10.1088/0031-9155/53/18/024>.
- [36] Tsapaki V, Fitousi N, Caracappa P, Jacobs J, Papailiou I. CT chest and abdomen organ doses estimation using a dose management tool 59 th annual meeting of the American Association of Physicists in Medicine, AAPM 2017, 30 July – 1 August 2017, Denver, Colorado, USA.
- [37] Huo W, Pi Y, Feng M, et al. VirtualDose-IR: a cloud-based software for reporting organ doses in interventional radiology. *Phys Med Biol*. 2019;64(9):095012. <https://doi.org/10.1088/1361-6560/ab0bd5>.
- [38] Järvinen J, Sierpowska J, Siiskonen T, Husso M, Järvinen H, Kiviniemi T, et al. Contemporary radiation doses in interventional cardiology: A nationwide study of patient skin doses in finland. *Radiat Prot Dosimetry* 2020;188:181–90. <https://doi.org/10.1093/rpd/ncz273>.

- [39] Greffier J, Van Ngoc Ty C, Bonniaud G, Moliner G, Ledermann B, Schmutz L, et al. Assessment of peak skin dose in interventional cardiology: A comparison between Gafchromic film and dosimetric software em.dose. *Phys Medica* 2017;38:16–22. <https://doi.org/10.1016/j.ejmp.2017.05.044>.
- [40] Bordier C, Klausz R, Desponds L. Accuracy of a dose map method assessed in clinical and anthropomorphic phantom situations using gafchromic films. *Radiat Prot Dosimetry* 2015;165:244–9. <https://doi.org/10.1093/rpd/ncv034>.