

# FINAL PUBLISHABLE REPORT

Grant Agreement number 15HLT06

Project short name MRTDosimetry

Project full title Metrology for clinical implementation of dosimetry in molecular radiotherapy

Project start date and duration:		01 June 2016 (36 months)
Coordinator: Andrew Robinson, NPL		Tel: +44 208 943 7174
Project website address: <a href="http://mrt dosimetry-empir.eu/">http://mrt dosimetry-empir.eu/</a>		E-mail: <a href="mailto:andrew.robinson@npl.co.uk">andrew.robinson@npl.co.uk</a>
Internal Funded Partners:	External Funded Partners:	Unfunded Partners:
1 NPL, United Kingdom	7 ASUL RE, Italy	13 AOSP, Italy
2 BEV-PTP, Austria	8 Christie, United Kingdom	14 BRFAA, Greece
3 CEA, France	9 INSERM, France	15 CARD, United Kingdom
4 CMI, Czech Republic	10 LUND, Sweden	16 OPBG, Italy
5 ENEA, Italy	11 THG, Greece	17 OUHT, United Kingdom
6 SCK•CEN, Belgium	12 UKW, Germany	18 RSCH, United Kingdom
RMG: -		

## TABLE OF CONTENTS

1	Overview .....	3
2	Need .....	3
3	Objectives .....	3
4	Results .....	5
4.1	Objective 1: To determine branching ratios and emission probabilities for $^{90}\text{Y}$ and $^{166}\text{Ho}$ in order to enable improved QI accuracy and dose estimation for these radionuclides, and to exploit new technologies in order to develop a suitable transfer instrument optimised for accuracy of measurements of the activity of MRT agents in clinics and radiopharmaceutical companies. ....	5
4.2	Objective 2: To develop 3D printing methods in order to generate a range of quasi-realistic anthropomorphic phantoms containing compartments fillable with known activities of radioactive liquid or standardised sealed radioactive test sources, having a range of geometrical complexity for validation of multimodal QI or absorbed dose measurement, and estimation of the uncertainties of measurement. In addition, to expand the protocol developed in JRP HLT11 for traceable calibration of SPECT QI for $^{177}\text{Lu}$ activity to include PET-CT QI of $^{90}\text{Y}$ and SPECT QI of $^{131}\text{I}$ , validated by measurements using the quasi-realistic anthropomorphic 3D printed phantoms. ....	8
4.3	Objective 3: To generate multimodal images either from SPECT or PET-CT phantom measurements or MC simulations to provide material for an open-access database of reference images to be used as reference data for commissioning and QC of QI using SPECT or PET-CT. In addition, to develop an architecture for and host the open-access database. ....	15
4.4	Objective 4: To improve the accuracy and metrological traceability in the calculation of dose from time-sequences of QI measurements by optimisation of the time points (i.e. obtaining cumulated activity from a TAC), choice of measurement modality (imaging or non-imaging), refinement of absorbed dose standards, and validation of alternative absorbed dose calculation methods in phantoms using physical measurement techniques such as MR sensitive gel-based and film-based dosimetry and MC simulations. ....	21
4.5	Objective 5: To determine uncertainties in relation to the full MRT dose measurement chain from a primary standard to a range of commercial and non-commercial dosimetry calculation platforms. This includes image quantification (such as uncertainties in the selection of volumes of interest (VOI) and image reconstruction); integration of TACs, propagation of uncertainties in NTCP models, and determination of the overall evaluated uncertainty in the absorbed dose quantification process. ....	28
5	Impact .....	33
6	List of publications .....	35

## 1 Overview

The overall aim of this project was to provide the metrology for the clinical implementation of absorbed dose calculations in Molecular Radiotherapy (MRT). The project built on the results and outputs from the preceding EMRP project HLT11 MetroMRT, which took the first steps towards providing data, methods, protocols and guidance for MRT dosimetry in collaboration with many European MRT clinics as well as radiopharmaceutical companies and camera manufacturers. The focus of this follow-on project was “clinical implementation” and it is strongly directed by the involvement of leading MRT clinics across Europe as well as building on metrology expertise.

## 2 Need

Over the past few years there has been an increase in Europe in the development and use of radiopharmaceuticals for treating cancer as well as an increase in the number of MRT clinical trials. However, in spite of the growing acceptance that an accurate knowledge of the radiation absorbed dose to critical tissues would provide a more effective targeted use of MRT, most patient treatments still follow the historical practice of administering a nominal activity of the radiopharmaceutical.

It is well known that the administered activity is not a good predictor of tissue dose and hence the outcome of patient treatment, due to individual variation in uptake and retention. However, one of the main reasons for a reluctance to perform individual patient dose measurements is that the process is complicated and there are no standard methods for calibrating or implementing MRT dosimetry in clinics. Therefore, prior to this project the MRT community had an urgent need for dosimetry calibration standards, validation methods, and clear guidance on how to implement MRT dosimetry in every European clinic offering MRT. As without this, it would not be possible to comply with EC Directive 2013/59/EURATOM, Article 56, which states that individual dose planning for radiotherapy patients (including MRT) must be enforced in legislation by EU member states.

The preceding EMRP project HLT11 MetroMRT clearly identified the key needs for obtaining dose measurements for MRT patients. These are: (1) measurement of the administered activity, (2) quantitative imaging (QI) of the activity localised in the patient using Single Photon Emission Computed Tomography (SPECT) or Positron Emission Tomography (PET), (3) integration of activity measurements over the time of treatment, (4) calculation of the dose from activity measurements and (5) estimation of the overall uncertainty of the measurement. Each of these needs is addressed in the MRTDosimetry project's objectives. The first objective addresses a specific need for more accurate QI measurement of the administered activity, for emerging beta-emitters therapies with  $^{90}\text{Y}$  and  $^{166}\text{Ho}$ , which are used with microspheres for liver cancer treatment. The remaining objectives are focused on meeting the generalised needs of MRT dosimetry.

Prior to this project the main sources of uncertainty in MRT dosimetry were in taking the step from dose measurements on simple reference geometries to QI measurements of the complex and varying geometries of the activity localised in real patients, as well as activity measurements over the time of treatment. All of these issues were addressed within this project for SPECT and PET based imaging, through the development of 3D printed quasi-realistic anthropomorphic phantoms and by creating a database of reference images of geometries covering typical clinical situations.

Dosimetry for MRT, as routinely performed, has no traceability to primary standards of absorbed dose. Therefore, prior to this project there was an urgent need to achieve traceability and to validate the dose calculation methods. Further to this, and central to any recommendations for dosimetry methods, is knowledge of the overall uncertainty associated with any particular method. Hence, the uncertainties in relation to the full MRT dose measurement chain (i.e. from a primary standard to a dosimetry calculation platform) also needed to be determined.

## 3 Objectives

The overall aim of this project was to provide the metrology for the clinical implementation of absorbed dose calculations in MRT. In order to do this, the specific objectives of this project were:

1. To determine branching ratios and emission probabilities for  $^{90}\text{Y}$  and  $^{166}\text{Ho}$  in order to enable improved quantitative imaging (QI) accuracy and dose estimation for these radionuclides, and to exploit new

technologies in order to develop a suitable transfer instrument optimised for accuracy of measurements of the activity of MRT agents in clinics and radiopharmaceutical companies.

2. To develop 3D printing methods in order to generate a range of quasi-realistic anthropomorphic phantoms containing compartments fillable with known activities of radioactive liquid or standardised sealed radioactive test sources, having a range of geometrical complexity for validation of multimodal QI or absorbed dose measurement, and estimation of the uncertainties of measurement. In addition, to expand the protocol developed in JRP HLT11 for traceable calibration of SPECT QI for  $^{177}\text{Lu}$  activity to include PET-CT QI of  $^{90}\text{Y}$  and SPECT QI of  $^{131}\text{I}$ , validated by measurements using the quasi-realistic anthropomorphic 3D printed phantoms.
3. To generate multimodal images either from SPECT or PET-CT phantom measurements or Monte Carlo (MC) simulations to provide material for an open-access database of reference images to be used as reference data for commissioning and Quality Control (QC) of QI using SPECT or PET-CT. In addition, to develop an architecture for and host the open-access database.
4. To improve the accuracy and metrological traceability in the calculation of dose from time-sequences of QI measurements by optimisation of the time points (i.e. obtaining cumulated activity from a time-activity-curve (TAC)), choice of measurement modality (imaging or non-imaging), refinement of absorbed dose standards, and validation of alternative absorbed dose calculation methods in phantoms using physical measurement techniques such as Magnetic Resonance (MR) sensitive gel-based and film-based dosimetry and MC simulations.
5. To determine uncertainties in relation to the full MRT dose measurement chain from a primary standard to a range of commercial and non-commercial dosimetry calculation platforms. This includes image quantification (such as uncertainties in the selection of volumes of interest (VOI) and image reconstruction); integration of TACs, propagation of uncertainties in NTCP models, and determination of the overall evaluated uncertainty in the absorbed dose quantification process.
6. To facilitate the take up by healthcare professionals (clinical centres) and industry (scanner manufacturers and software developers) of the technology and measurement infrastructure developed by the project.

## 4 Results

**4.1 Objective 1:** *To determine branching ratios and emission probabilities for  $^{90}\text{Y}$  and  $^{166}\text{Ho}$  in order to enable improved QI accuracy and dose estimation for these radionuclides, and to exploit new technologies in order to develop a suitable transfer instrument optimised for accuracy of measurements of the activity of MRT agents in clinics and radiopharmaceutical companies.*

The aim of this objective was to determine branching ratios and emission probabilities for  $^{90}\text{Y}$  and  $^{166}\text{Ho}$  to enable improved QI accuracy and dose estimation for these radionuclides. In addition, to exploit new technologies to develop transfer instruments optimised for accuracy of measurements of the activity of MRT agents in clinics and radiopharmaceutical companies.

### Activity measurements and determination of nuclear data for $^{166}\text{Ho}$

Holmium-166 is a high-energy  $\beta^-$ -emitter radionuclide ( $\sim 1.8$  MeV) with a short half-life ( $\sim 26.8$  h) that offers great potential as an alternative to  $^{90}\text{Y}$  for the treatment of liver cancer based on radioembolisation using microspheres. The possibility of quantitative SPECT imaging of the main  $\gamma$ -ray emission at 80.6 keV, in addition to strong paramagnetic properties suitable for Magnetic Resonance Imaging (MRI), complement this therapeutic potential. The establishment of a primary standard of radioactivity for  $^{166}\text{Ho}$  is essential to support its clinical use and in addition accurate nuclear data is required.

Primary measurements and new determinations of nuclear data of  $^{166}\text{Ho}$  were carried out at CMI, CEA and NPL. The primary measurements were independently performed using different coincidence counting techniques: Triple to Double Coincidence Ratio (TDCR) method,  $4\pi\beta\text{-}\gamma$  and  $4\pi(\text{LS})\text{-}\gamma$  coincidence counting using solid and liquid scintillation sources. As  $^{166}\text{Ho}$  is a high-energy  $\beta^-$ -emitting radionuclide high detection efficiencies were obtained in proportional counters or TDCR apparatus leading to low uncertainties associated to the activity concentrations ( $\sim 0.2 - 0.4$  %).

The newly standardised solutions at CMI, CEA and NPL were used to determine new precise absolute intensities for ten of the 17 existing evaluated photon emissions. The values determined in this work were typically in internal agreement and with the existing evaluation, with the standard uncertainties bettering those of the currently recommended values. The three 80.6 keV absolute intensity values determined in this work showed a relative increase of approximately 1 % over the current recommended value. Five independent half-life measurements were also carried out at CMI, CEA and NPL using ionisation chambers and gamma - spectrometers. In addition, ENEA performed measurements of  $^{166}\text{Ho}$  linked to these measurements, in the framework of the EURAMET.RI(II)-K2.Ho-166 comparison, by using Ionisation Chambers, a NaI(Tl) 5"x5" well- type detector and high-energy resolution high-purity germanium (HPGe) spectrometer.

### New measurements of the intensity of the internal pair production intensity of $^{90}\text{Y}$

In Selective Internal Radiotherapy (SIRT) based on radioembolisation using  $^{90}\text{Y}$  microspheres, the accuracy of activity measurements based on Positron-emission tomography (PET) QI strongly depends on the knowledge of the intensity of internal positron/electron ( $e^+/e^-$ ) pair production. Up to now, this nuclear data relies on a single published measurement based on high-resolution gamma spectrometry using HPGe detectors. New measurements of the internal  $e^+/e^-$  pair production branching ratio for  $^{90}\text{Y}$  were carried out at CEA and CMI by means of high-resolution gamma spectrometry using HPGe detectors and at ENEA using two cylindrical NaI(Tl) 5"x5" detectors working in coincidence coupled with a TDCR system.

At CMI, the count rates of the 511 keV full-energy peak (positron annihilation) were measured using two calibrated HPGe detectors. The measurements were performed with two sources of different constructions. One source was a 1 mL ampoule filled with  $^{90}\text{Sr}$  solution (in equilibrium with  $^{90}\text{Y}$ ) inserted into an aluminium absorber. The second source was the evaporation residue of another  $^{90}\text{Sr}$  solution (in equilibrium with  $^{90}\text{Y}$ ) in a polyethylene absorber. In both cases, the annihilation of positrons occurred in the source materials. The detection efficiency of 511 keV photon emission was determined by MC calculation, where the source was defined as a theoretical continuous positron spectrum with a maximum energy of 738 keV. At CEA a HPGe detector was first calibrated in the 511 keV photon energy region using the  $\beta^+$  emission of  $^{65}\text{Zn}$ . The measurements were carried out with an ampoule filled with a standardised radioactive solution. The intensity of the 511 keV annihilation photon emission resulting from the  $^{90}\text{Y}$  decay, was measured using an ampoule filled with a radioactive solution of  $^{90}\text{Y}$  standardised by means of the TDCR method. At ENEA the two cylindrical

Nal(Tl) 5"x5" detectors were first calibrated by using long-lived gamma-emitters ( $^{241}\text{Am}$ ,  $^{60}\text{Co}$ ,  $^{137}\text{Cs}$ ) and short-lived gamma-emitters ( $^{177}\text{Lu}$ ,  $^{64}\text{Cu}$ ), in point-like geometry. These measurements allowed the study of the energy resolution of the two Nal(Tl) detectors as a function of the gamma energy. The  $^{64}\text{Cu}$  source was used to define a Region-Of-Interest (ROI) around the 511 keV peak for each gamma-detector as well as to tune the algorithm for the gamma-gamma coincidences used to select the photons coming from the  $\beta^+$  branch of  $^{90}\text{Y}$  source and recorded in coincidences in the two gamma detectors. The activity of the  $^{90}\text{Y}$  source was measured by the TDCR method.

NPL's approach to determining the internal  $e^+/e^-$  pair production was to develop a method to measure directly the number of positrons emitted per decay of  $^{90}\text{Y}$ . Given the small branch, the challenge is to detect positrons in an overwhelming background of electrons. The method developed is based on a magnetic spectrometer, which allows to filter out the large number of electrons. Energy and efficiency calibration are paramount for a precision measurement, and this is challenging because a monoenergetic source of positrons of different energy is not available as opposed to electron measurement where electrons from internal conversion (IC) can be used. For this measurement, the spectrometer hardware was upgraded with the addition of a thick silicon detector and a new digital acquisition system. The Si detector, calibrated with IC sources, was then used to determine the relationship between the magnetic field (i.e., coil electrical current) and the positron energy. This method was validated using a source of  $^{22}\text{Na}$ , and the method demonstrated good agreement between the experimental spectrum shape and the shape expected as a beta-allowed transition. The stability of the system was studied over several months and determined to be sufficient for the measurement of  $^{90}\text{Y}$  that given the small branch will have to be a long measurement. Having established the feasibility of this novel technique the first measurements of a  $^{90}\text{Y}$  source is now ongoing with the aim of publishing the result by the end of 2019.

**Table 1:  $^{90}\text{Y}$  position branching ratio measurement results**

<i>Positron branching ratio for <math>^{90}\text{Y}</math></i>	
CMI	$3.26 (4) \times 10^{-5}$
CEA	$3.27 (4) \times 10^{-5}$
ENEA	$3.20 (4) \times 10^{-5}$

## Development of a transfer instrument for high-energy beta-emitters

### Portable TDCR system

The ENEA TDCR portable system consists of three small square PhotoMultipliers (PMTs - Hamamatsu R7600U-200) arranged around a prismatic optical chamber, made from white PTFE (Teflon), in a planar 120° geometry. Three compact on-board High Voltage (HV) power supplies provide 900 V to each PMT. Vials to be measured can be placed in the optical chamber by a lift; a shutter allows this operation to be performed without turning off the PMTs. The detector is also equipped with a self-contained electronics unit (CAEN Digitiser DT572). The weight of the complete system is < 6 kg and it can be easily transported to a clinical site where it is necessary to calibrate devices such as Ionisation Chambers and PET or SPECT scanners.

Using this system, a new activity primary standard of  $^{90}\text{Y}$  was developed with an uncertainty of 0.57 %. The activity value, determined by measuring sources prepared using Ultima Gold (UG) as a liquid scintillator, was also useful for determining with low uncertainty the counting efficiency for Double and Triple coincidences for Cherenkov light emissions, by measuring  $^{90}\text{Y}$  in water contained in anti-static plastic vials. With this new standard it was then possible to perform calibration on-site of the Ionisation Chamber of the Children Hospital Bambino Gesù (partner OPBG) in Rome. In this case the sources were prepared directly at OPBG and the  $^{90}\text{Y}$  activity measurements were carried out at OPBG simultaneously in the ENEA TDCR portable counter and in the OPBG ionisation chamber by using the same  $^{90}\text{Y}$  master solution, in liquid chloride form. The ENEA TDCR portable system was also used to perform the on-site calibration of the PET device used at a stakeholder site at Gemelli Hospital in Rome. In this case different anthropomorphic phantoms were filled with high-activity (about 1 GBq) of  $^{90}\text{YCl}_3$  and a set of  $^{90}\text{Y}$  sources in UG liquid scintillator were also prepared at Gemelli and measured in the ENEA TDCR detector. Measurements were also performed at a second stakeholder site the Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori (Scientific Institute in Emilia Romagna region for the Study and Care of Tumors) (IRST) where the ENEA TDCR portable system was used to perform the calibration of PET and SPECT devices for  $^{90}\text{Y}$  by anthropomorphic phantoms. In all the measurements

performed the combined standard uncertainty achieved on the calibration factor of the different devices was lower than 2 %.

### ***Development of novel Cherenkov counting system***

The design of classical ionisation chambers consists of a gas-filled cylindrical vessel with a well at the centre in which containers (e.g. ampoules, vials, syringes) filled with radioactive solutions are placed to obtain measurements in a quasi  $4\pi$  geometry. The activity is determined by measuring an ionisation current proportional to the energy deposition rate of ionising radiation in the gas medium. In the case of high-energy pure  $\beta^-$ -emitters such as  $^{90}\text{Y}$  ( $E_{\text{max}} \sim 2280$  keV), the response of ionisation chambers predominantly relies on the detection of bremsstrahlung emitted along the track of beta particles in the container and its surroundings (e.g. source holder). The reproducibility of the measurements is then impaired by their higher sensitivity to variations in the geometrical characteristics of those components, e.g. container wall thickness. In order to diminish the impact of such problems, the new transfer instrument investigated at CEA is based on the implementation of pulse counting instead of measuring an ionisation current. Additionally, the possibility of counting saturation due to the high-activity of radioactive solutions must be prevented (e.g. up to 10 GBq for  $^{90}\text{Y}$ -microspheres in aqueous suspensions such as (selective internal radiation) SIR-Spheres® (Sirtex, Australia)).

The aim of the new transfer instrument is to take advantage of the Cherenkov threshold to obtain a low-sensitivity detector to prevent counting saturation for high-activity sources. The design was investigated and validated at CEA by means of MC simulations using the Geant4 code. The simulated geometry was based on a quartz-transparent well sensitive part coupled to a photomultiplier to enable the measurements of the same containers (ampoules, vials, syringes, etc.) used in hospitals. MC calculations have shown that detection efficiencies lower than  $10^{-4}$  can be obtained thanks to the Cherenkov threshold, thus enabling the measurements of sources without counting saturation with activities up to a few GBq in the case of  $^{90}\text{Y}$ .

Regarding the counting hardware, a new digital system was designed to implement list-mode data format acquisition using a fast digitiser (TI ADC32RF45EVM,  $\sim 2.4$  GHz, 14 bits) to enable an accurate sampling of photomultiplier pulses (rise time  $\sim 3$  ns). The ADC mezzanine card is coupled to a field-programmable gate array (FPGA) circuit (Arria 10 GX, Alaric host card). Specific front-end electronics were also developed to optimise the coupling between the photomultiplier output and the ADC input. A FPGA program was implemented to register pulse trains given by the new transfer instrument. First acquisition tests indicated that a good signal to noise ratio was obtained on single photon pulses used to set the detection threshold. Initial counting tests based on a Plexiglas sensitive part using gamma-emitting sources ( $^{137}\text{Cs}$ ,  $^{60}\text{Co}$ ) confirmed the need to use quartz material to optimise the production of Cherenkov photons (spectral bandwidth, spectral transmittance) to obtain the expected detection efficiency with the new transfer instrument. CEA was not able to implement further tests during the project due to a problem with the provision of the quartz sensitive part with the expected well geometry.

### **Summary of outcomes**

The establishment of the first primary standard activity measurements for  $^{166}\text{Ho}$ , coupled with significant improvements to the nuclear data and half-life will enable traceability and improved precision for measurements of  $^{166}\text{Ho}$  made by clinical centres and radiopharmaceutical companies. Similarly, for  $^{90}\text{Y}$  new precise measurements of the  $e^+/e^-$  branching ratio will improve confidence in clinical quantitative PET measurements for this isotope – which are particularly challenging. The data for both radionuclides has been submitted for publication in peer-reviewed journals ensuring that these important new results are made available to all stakeholders.

The use of transfer instruments for in a clinical setting has been demonstrated at several different clinical sites. This work has led to reduced uncertainties for the calibration of ionisation chambers and imaging at all centres, highlighting the benefits of reducing the traceability chain for beta-emitting radionuclide. The exciting prospect of a new design of transfer instrument which can be operated without specialised radionuclide metrology training has also been developed in this project. The promising results on this prototype device provide an excellent platform for its continued development.

4.2 *Objective 2: To develop 3D printing methods in order to generate a range of quasi-realistic anthropomorphic phantoms containing compartments fillable with known activities of radioactive liquid or standardised sealed radioactive test sources, having a range of geometrical complexity for validation of multimodal QI or absorbed dose measurement, and estimation of the uncertainties of measurement. In addition, to expand the protocol developed in JRP HLT11 for traceable calibration of SPECT QI for  $^{177}\text{Lu}$  activity to include PET-CT QI of  $^{90}\text{Y}$  and SPECT QI of  $^{131}\text{I}$ , validated by measurements using the quasi-realistic anthropomorphic 3D printed phantoms.*

The aim of this objective was to produce a range of quasi-realistic anthropomorphic phantoms containing compartments fillable with known activities of radioactive liquid or standardised sealed radioactive test sources, having a range of geometrical complexity for validation of multimodal QI or absorbed dose measurement, and estimation of the uncertainties of measurement. In addition, to define an expanded protocol for traceable calibration of SPECT QI and PET- Computed tomography (CT) QI, validated by measurements using the quasi-realistic anthropomorphic 3D printed phantoms.

### Generation of Anthropomorphic Phantoms

Conventional phantoms for validation of quantitative SPECT/CT imaging typically consist of regular geometric shapes such as spheres or cylinders. 3D printed phantoms can provide unique, patient-representative, activity distributions that can be used to validate quantitative SPECT/CT imaging. Specifically, these phantoms provide a more realistic partial volume effect in reconstructed images similar to those found in clinical images.

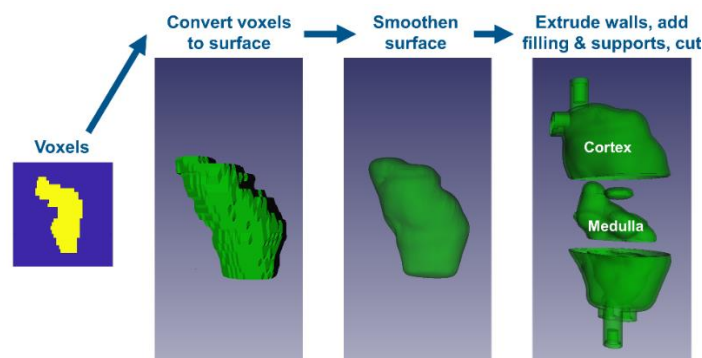
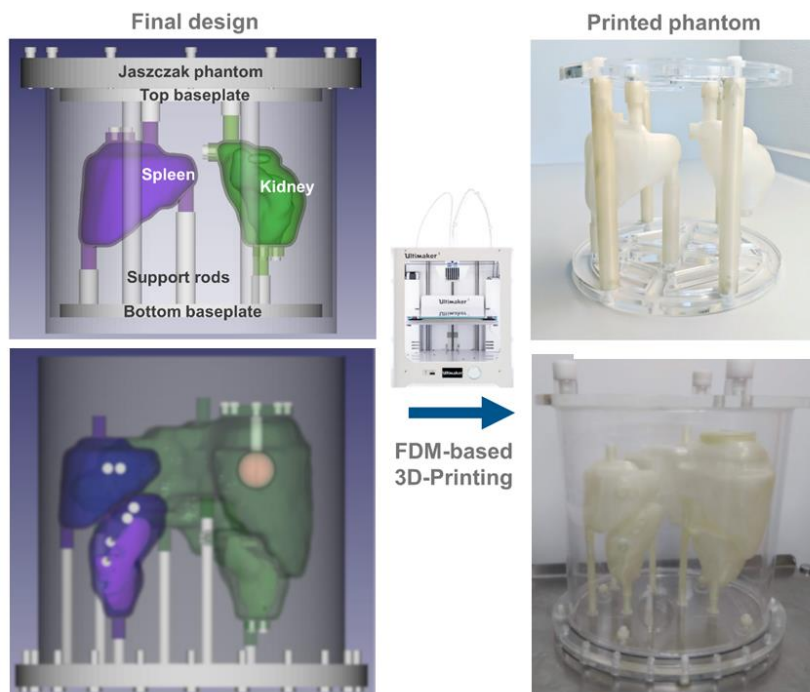


Figure 1: Process for conversion from voxelised phantom to CAD model

Three anthropomorphic phantoms were designed as part of this project. The first, the *Two-Organ Phantom*, was designed to provide a realistic geometry to be used in a QI comparison exercise and to allow validation of the sphere-geometry based partial volume correction. The model is based on International Commission on Radiological Protection's (ICRP) Publication 110 ("Adult Reference Computational Phantoms"), which has been widely accepted as dosimetry standard. The female phantom was used as it is based on a higher resolution CT scan than the male. The two-organ phantom contains a spleen and a 2-compartment right kidney (Cortex and Medulla). The second, the *Four-Organ Phantom* is also based on the ICRP 110 publication model and contains a liver with a sphere-shaped tumor, a spleen and left and right kidneys with 2 compartments. In addition, an anthropomorphic neck phantom was produced to simulate thyroid cancer treatment.

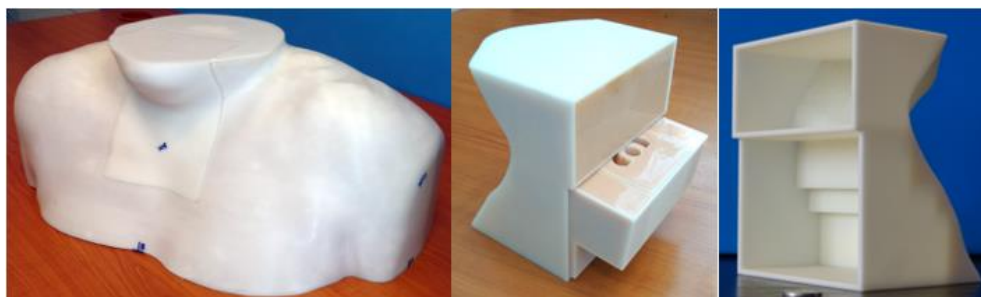
A method for converting the organs in the computational phantom to a set of triangular mesh volumes was developed using in-built MATLAB functions (see Figure 1 **Error! Reference source not found.**). To remove the voxelisation that was visible in the meshes, each surface was smoothed using a Laplacian smoothing algorithm. After smoothing, the surface models were hollowed out to create a set of hollow objects and cut into parts so that they could be 3D printed. CAD features were added to allow filling and to reproducibly position in a phantom. The phantoms were printed using fused deposition modeling (FDM) with an Ultimaker 3 extended 3D printer. The phantoms were printed in polylactic acid, which has a CT number of approximately 130 HU which is similar to Perspex (110-130). Phantom parts were sealed using chloroform and then further sealed with Araldite 2020 low viscosity glue. Laser cut perspex baseplates and top plates were also produced, designed to be easily attached to the holes in the base of the standard Jaszczak phantom. The computer-aided designs and the assembled 3D printed phantom are shown in Figure 2 **Error! Reference source not**

**found..** The four-organ phantom design was too large to fit into the standard Jaszczak phantom and required a bespoke Perspex elliptical phantom (see Figure 2Error! Reference source not found.).



**Figure 2: Design and printed phantoms including baseplate for attachment**

CMI developed a solid anthropomorphic phantom for use in thyroid cancer treatment based on CT images of a real patient. The CT images were processed and segmented for 3D printing. The phantom consists of two main parts: a body part and a neck part. The body part was 3D-printed as a 2.5 mm-thick one-piece shell that also included an outer shell of a simplified spine. The material used was Acrylonitrile Butadiene Styrene (ABS) plastic. The spine was later filled with a bone-equivalent material made of silicone resin “lukopren” (57 % by mass) with the addition of  $\text{Al}_2\text{O}_3$  (36 %) and  $\text{TiO}_2$  (7 %) to increase proton number and density. The rest of the phantom was filled with a water-equivalent material made of polyurethane rubber (97.2 % of mass) with the addition of  $\text{CaCO}_3$  (2.8 %). The neck part is a removable object that fits into the front middle cut-off space of the body part. It was also 3D-printed as a 3 mm-thick one-piece shell. The neck part contains empty space (8 cm x 11 cm x 4.5-6.0 cm) where various insets can be inserted. All materials used in the phantom were characterised using photon attenuation and density.



**Figure 3: Anthropomorphic phantom for validation of QI of the thyroid. Left: Full view. Centre: Neck part with inserts. Right: Neck part empty**

## Development of sealed cylinder surrogate sources

### Production of sources

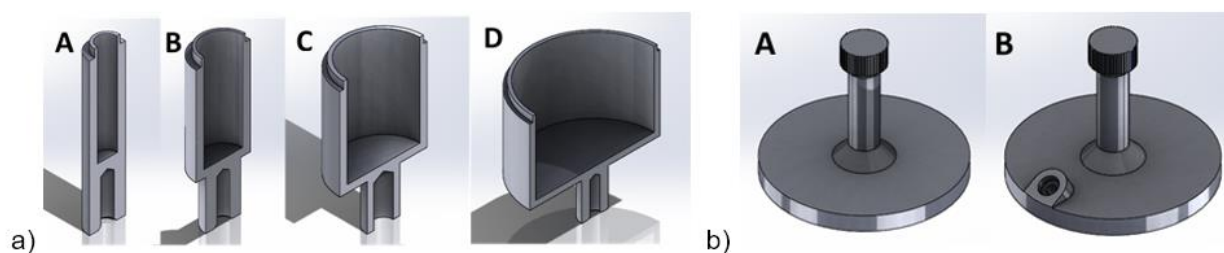
In nuclear medicine imaging, commissioning and quality control (QC) is typically performed based on hollow, fillable containers, which are manually filled with radionuclide solutions. This leads to a multitude of problems such as ensuring that the filling volume and activity concentration are determined with an acceptably small error and that similar activities are used for repeated measurements to be comparable in terms of counting statistics. Alternative, but costlier, QC could be performed with sealed, long-lived (i.e. useable for more than one year) surrogate radioactive test sources containing radionuclides of comparable energies and emission probabilities to the ones to be used for imaging (e.g.  $^{57}\text{Co}$  for  $^{99\text{m}}\text{Tc}$  or  $^{133}\text{Ba}$  for  $^{131}\text{I}$ ). However, the activity of these reference sources must be traceable to a primary activity standard.

In this project, two sets of traceable  $^{133}\text{Ba}$  sources (as a surrogate for  $^{131}\text{I}$ ) were produced at CEA and CMI. Both sets were encapsulated in 3D-printed cylinders printed by NPL. To assess the feasibility of using the solid sources as surrogates, identical containers to be filled with liquid  $^{131}\text{I}$  were sent to partner sites (ASMN, Christie, Lund, NPL, OUHT, RSCH, THG) for a quantitative SPECT/CT imaging comparison exercise to be conducted. A third set of traceable sources filled with  $^{68}\text{Ge}$  (as surrogate for positron emission from  $^{90}\text{Y}$ ) were produced for PET/CT imaging. Details on the geometry of the sources and the respective activities are provided in Table 2.

CAD designs were produced for cylinders with four volumes ( $1.7\text{ cm}^3$ ,  $6.7\text{ cm}^3$ ,  $26.9\text{ cm}^3$ ,  $107.4\text{ cm}^3$ ). To facilitate the assessment of the partial volume effect, the height of the active volume was set to 38 mm in all cases, while the diameter was changed (7.5 mm, 15 mm, 30 mm, 60 mm). Two types of cylinder cap were designed, one for containing a resin and one for injection of solution. The cylinder and cap designs are shown in Figure 4. Cylinders and caps were produced using a stereolithography (SLA) 3D printing system (Formlabs Form 2) using a photopolymer resin resulting in a watertight model. Following the ISO 9978 standard, leakage and contamination tests of the sealed sources were conducted. The wipe tests detection limit was below 4 Bq and no activity was detected. For the immersion test, a closed geometry with 400 kBq of  $^{68}\text{Ge}$  spiked resin inside was immersed in water at room temperature for 48 h. The remaining water was measured by gamma spectrometry and no activity was detected (detection limit of 0.7 Bq).

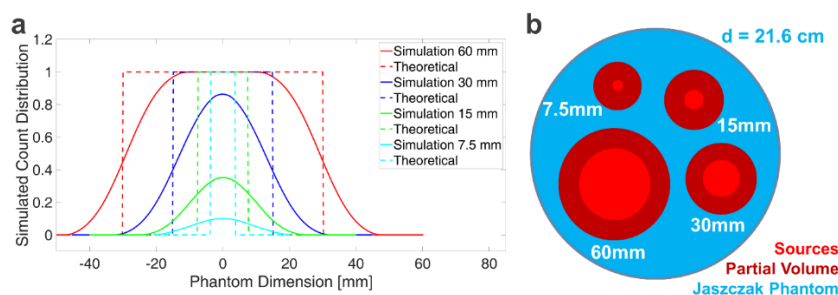
**Table 2: Specifications of the cylindrical solid sources produced within this project**

Diameter (mm)	Length (mm)	Volume ( $\text{cm}^3$ )	Activity $^{133}\text{Ba}$ CMI (kBq)	Active Vol. ( $\text{cm}^3$ )	Activity $^{133}\text{Ba}$ CEA (kBq)	Active Vol. ( $\text{cm}^3$ )	Activity $^{68}\text{Ge}$ CEA (kBq)
7.5	38.0	1.68	$161.9 \pm 1.8$	$1.66 \pm 0.02$	$341.8 \pm 8.9$	$1.68 \pm 0.02$	$184.6 \pm 6.3$
15.0	38.0	6.72	$674.2 \pm 7.4$	$6.68 \pm 0.07$	$1377.0 \pm 35.8$	$6.51 \pm 0.07$	$715 \pm 24.3$
30.0	38.0	26.9	$2685 \pm 30$	$25.6 \pm 0.3$	$5280 \pm 137$	$26.3 \pm 0.3$	$2890 \pm 104$
60.0	38.0	107.4	$10710 \pm 118$	$101.1 \pm 1.0$	$20860 \pm 542$	N/A	N/A



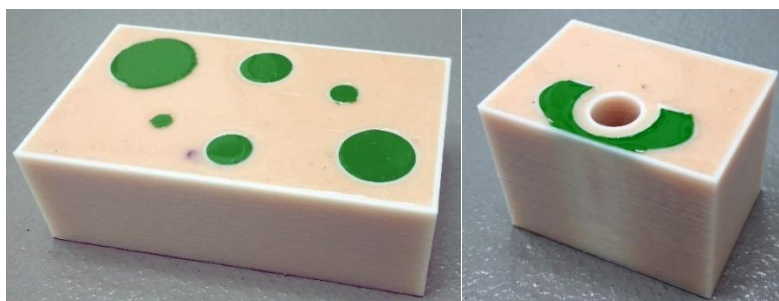
**Figure 4: (a) Cross-sections of the computer-aided designs for A: Ø7.5 mm, B: Ø15 mm, C: Ø30 mm, and D: Ø60 mm active diameter cylinders. (b) Designs of cylinder caps. A: resin cylinder cap (surrogate production), B: injection cylinder cap (liquid measurement)**

The placement of the sources in a standard Jaszczak phantom was optimised by simulating the energy- and collimator-dependent spill-out of counts (Gaussian convolution with 20 mm full width at half maximum (see Figure 5). A laser-cut baseplate for attaching the sources according to this optimal positioning was also produced.



**Figure 5: (a) Simulation of partial volume effect (red: 60 mm, blue: 30 mm, green: 15 mm, cyan: 7.5 mm). (b) Optimised source placement inside the Jaszczak cylinder**

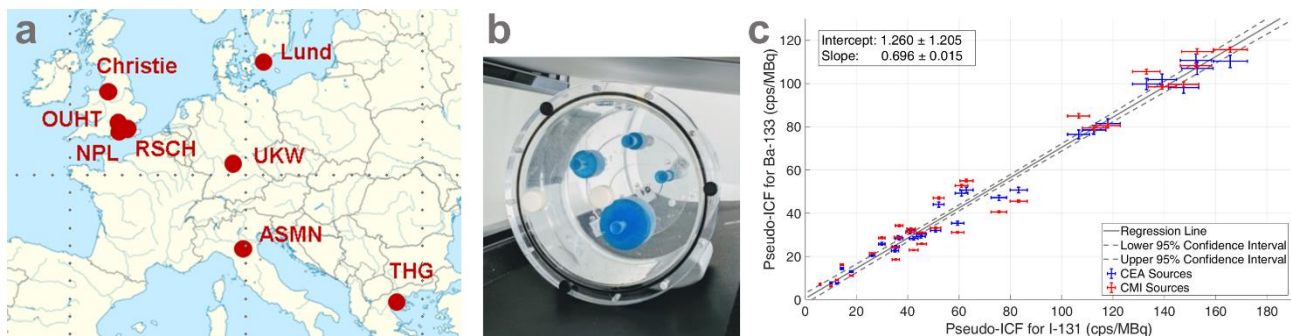
An additional set of  $^{131}\text{I}$  radioactive test sources was produced at CMI for inclusion in a head and neck torso. From one original  $^{131}\text{I}$  solution, three different sources were produced: (i) thyroid gland, (ii) cylindrical sources, and a (iii) point source. The thyroid gland source and cylindrical sources consisted of activity homogeneously distributed within silicone resin lukopren which was poured into appropriate vessels. The vessels were designed at CMI and printed on a commercial 3D-printer, such that they fitted into a neck part of the CMI phantom (see Figure 3). The 3D-printed shell of the thyroid gland took into account the gland's shape and average volume and its dimensions were approximately 4.2 cm in height, 4.0 cm in width and 1.6 cm in depth with the total active volume of 8.5 ml (Figure 6 on the right-hand side). The total activity was 11.71(16) MBq ( $k=1$ ) at the reference time. This source was used for testing of SPECT measurement and reconstruction settings usable for treatment planning of benign thyroid diseases. The cylindrical source consisted of six cylinders filled with radioactive solution with the same specific activity as the thyroid source (1.411(20) MBq/g,  $k=1$ ). The height of all cylinders was 3.2 cm and the diameters were 0.4, 0.8, 1.2, 1.6, 2.0, and 2.4 cm (Figure 6 on the left-hand side). This source was used for testing of general SPECT measurements and reconstruction setting, primarily spatial resolution and partial volume effect. The volume with activity was covered by a thin ( $\sim 2$  mm) protective layer of non-radioactive, coloured lukopren.



**Figure 6:  $^{131}\text{I}$  sources in 3D-printed cylinders (left) and 3D-printed thyroid gland (right) for use in neck phantom**

### **Solid source SPECT/CT measurements ( $^{133}\text{Ba}$ )**

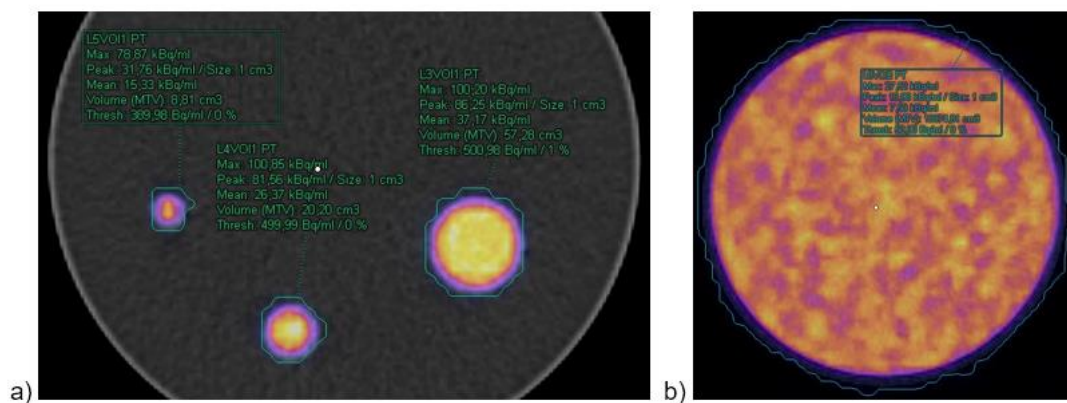
The solid surrogate sources were sent around 8 different sites of the MRTDosimetry partners (Christie, OUHT, LUND, NPL, RSCH, UKW, ASMN and THG) to perform an inter-comparison exercise between the solid  $^{133}\text{Ba}$  surrogate sources and sources of the same geometry and material filled with liquid  $^{131}\text{I}$ . The participating sites are shown in Figure 7 (a). This exercise required the sites to attach all 4 sources in a water-filled Jaszczak cylinder and perform SPECT/CT measurements based on a high-energy collimation protocol (Figure 7 (b)). Subsequently, a set of fillable cylinders with liquid  $^{131}\text{I}$  was attached in the same order in the Jaszczak cylinder and imaged using the same protocol. A VOI analysis was then carried out to compare differences between  $^{131}\text{I}$  and the  $^{133}\text{Ba}$  surrogates in image calibration factor (cps / MBq). (see Figure 7 (c)). A regression analysis yielded a cross-calibration factor of  $0.70 \pm 0.02$ , which can be explained by the emission probabilities.



**Figure 7:**  $^{133}\text{Ba}$  SPECT/CT comparison exercise. (a) Map of participating sites. (b) SPECT/CT acquisition of the  $^{133}\text{Ba}$  sources mounted in the Jaszczak cylinder. (c) Image calibration factors (ICF) for the two sets of  $^{133}\text{Ba}$  sources (CEA: blue, CMI: red) plotted against the  $^{131}\text{I}$  values for all SPECT/CT systems used in the exercise

### Solid Sources PET/CT Measurements ( $^{68}\text{Ge}$ )

PET/CT measurements of the solid  $^{68}\text{Ge}$  sources attached to a Jaszczak cylinder were carried out at partner UKW (Siemens Biograph mCT S(64), standard clinical imaging protocol). To compare homogeneity, accuracy, and partial volume effects of the 3 cylindrical sources, their performance was compared against a standard commercial QC phantom (i.e. a large cylinder filled with  $^{68}\text{Ge}$  in epoxy). The acquisition for both phantoms was set to stop after 10 million counts had been collected to achieve comparable count statistics. Standard reconstructions were performed using the  $^{68}\text{Ge}$  calibration pre-set provided directly by the vendor. A VOI analysis (enlarged volumes to account for partial volume) was performed in Siemens syngo via (MM Oncology), the results of which are shown in Figure 8 and Table 3. While the agreement between traceable activity and PET/CT-based activity is 0.1 % for the standard QC phantom, a difference of ~6 % was found for all analysed VOIs of the sources produced by the consortium.



**Figure 8:**  $^{68}\text{Ge}$  PET/CT measurements. Fusions of both acquisitions including drawn VOIs. (a) 3 sources produced by CEA. (b) QC phantom

**Table 3: Results of  $^{68}\text{Ge}$  PET/CT measurements and VOI analysis**

	7.5-mm Cylinder	15-mm Cylinder	30-mm Cylinder	Large VOI (All Cylinders)	QC Cylinder
Calibrated Activity [MBq]	145.2 $\pm$ 4.9	562.4 $\pm$ 19.1	2273.2 $\pm$ 81.8	3789.6 $\pm$ 104	77.7 $\pm$ 3.9
PET/CT-based Activity (exact VOI) [MBq]	91.6	436.5	1906.4	2434.5	76.5
Difference [%]	-36.9	-22.4	-16.1	-35.8	-1.6
PET/CT-based Activity (enlarged VOI) [MBq]	135.1	532.7	2129.1	2815.5	77.8
Difference [%]	-7.0	-5.3	-6.3	-5.6	0.1

## Expanded Calibration Protocol for Quantitative SPECT/CT and PET/CT Imaging

The calibration protocol for quantitative SPECT imaging developed in the preceding project HLT11 MetroMRT was expanded to provide a method for traceable quantitative calibration for any radionuclide. Commercially available SPECT/CT systems include corrections for attenuation, scatter and point spread function, and return a 3D volume of counts or count rate. The expanded calibration protocol was developed such that it could describe how to obtain an image calibration factor, that relates the measured count rate to activity in the field-of-view (for a specific set of acquisition and reconstruction parameters). Based on this factor, any SPECT/CT acquisition data (provided that the same hardware such as detector and collimators, but the same reconstruction algorithm and parameters have been used) can be converted from counts to activity. The expanded protocol determines the image calibration factor based on a traceable amount of activity dissolved in a water-filled large-volume phantom (e.g. Jaszczak cylinder). It describes in detail the phantom geometry to use, how to perform the activity measurement in a radionuclide calibrator, prepare the radioactive solution, position the phantom on the SPECT/CT patient bed, acquire and reconstruct a SPECT image volume, and draw a volume-of-interest based on which the image calibration factor is determined. In addition, the expanded calibration protocol describes how to validate the calibration procedure using a test phantom and makes suggestions on how to assess quantification errors related to the partial volume effect.

In contrast to SPECT/CT which reports a count rate, PET/CT systems typically automatically convert counts to activity concentration or standardised uptake value (SUV). This calibration is often performed using NIST-traceable surrogates (e.g. large cylinder sources containing  $^{68}\text{Ge}$  in epoxy), the results of which are then cross-calibrated to clinically relevant radionuclides such as  $^{18}\text{F}$  or  $^{68}\text{Ga}$ . Although this approach is likely to provide acceptably low uncertainties, the exact implementation is usually proprietary, impeding a closer examination. Therefore, the PET/CT section of the expanded protocol mainly describes how to use traceable liquid sources to either assess the quantitative accuracy for existing radionuclide calibration factors or calibrate the system for new, not automatically pre-calibrated radionuclides.

### Validation SPECT/CT Measurements with Two-Organ Phantom

After determination of sensitivity and partial volume effects at the participating sites, the Two-Organ phantoms were used for validation of the QI setup at each individual site. Multiple sets of phantom inserts and attachment baseplates were sent to participating sites. Using a QI Commissioning Protocol (based on the expanded calibration protocol) the phantoms were filled with a solution of traceable activity, attached to the Jaszczak cylinder, and measured at each site's SPECT/CT system. Sensitivity and partial volume correction factors determined at each site were applied to quantify the activity in each organ. An example of the VOI analysis at UKW is shown below. VOIs were drawn based on a threshold on the count rate until the nominal volume of the phantom inserts (cortex and medulla combined to one kidney compartment) was reached. To reach the exact volume, one slightly smaller and one slightly larger volume were outlined and a linear interpolation was performed to reach the exact filling volume.

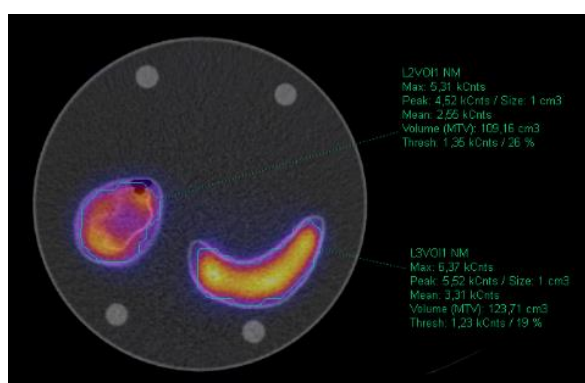
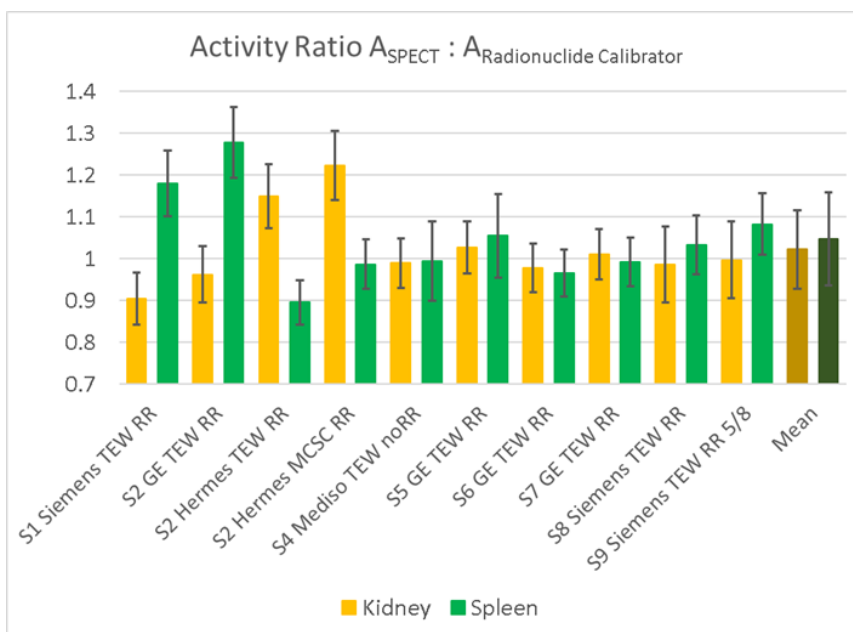


Figure 9: VOIs for Two-Organ phantom imaged on a Siemens Intevo Bold system



**Figure 10: Ratio between SPECT-based activity and radionuclide calibrator based activity for kidney (medulla + cortex, orange) and spleen (green). The mean and standard deviation is indicated by a darker colour**

The ratio between SPECT-based and radionuclide calibrator based activity for the SPECT systems used in the comparison exercise is shown in Figure 10. The indicated errors include the uncertainties of activity measurement, SPECT counts, acquisition duration, sensitivity, and partial volume correction. Although the average activity recovery in both phantom inserts is relatively close to unity ( $1.02 \pm 0.09$  for the kidney and  $1.05 \pm 0.11$  for the spleen) there is significant variation between systems and reconstructions, highlighting the crucial role of reconstruction algorithms in QI.

### **Summary of outcomes**

A new range of andromorphic phantoms suitable for 3D printing were developed. These phantoms are partly based on the ICRP 110 dosimetry standard model, and the use of 3D printing for production means that they can be made available to all clinical sites at a fraction of the cost of conventional commercial phantoms. To compliment the phantoms a range of traceable solid source has been produced which can be used for QC of SPECT and PET systems.

Coupling these phantoms and sources with the newly expanded calibration protocol allows quantitative SPECT/CT imaging to be traceably calibrated. This allows the harmonisation of imaging across multiple centres, systems and countries – a major step towards personalised therapies. The successful multi-site inter-comparison exercise (covering 8 clinical sites in 5 countries) has demonstrated the feasibility of clinical application of this protocol and the new phantoms.

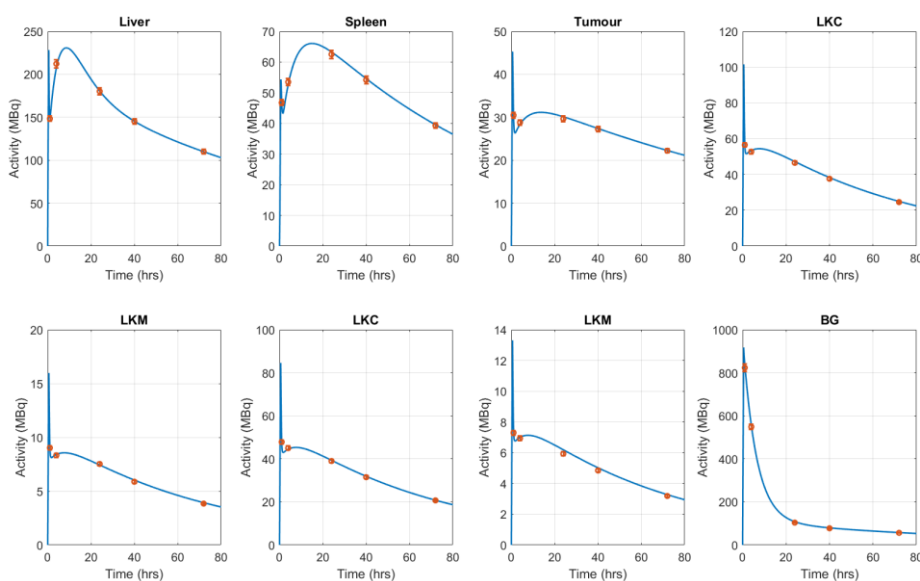
**4.3 Objective 3:** To generate multimodal images either from SPECT or PET-CT phantom measurements or MC simulations to provide material for an open-access database of reference images to be used as reference data for commissioning and QC of QI using SPECT or PET-CT. In addition, to develop an architecture for and host the open-access database.

The 3<sup>rd</sup> objective of the project was to provide a database of reference SPECT/CT images for a range of geometries covering typical clinical situations that can be used for commissioning and QC of QI and MRT dosimetry platforms. Experimental SPECT/CT images of the 3D-printed anthropomorphic phantoms (described in Objective 2) were generated on different clinical systems. Complementary MC simulated SPECT images and simulated CT images were also provided using the digital models of the same anthropomorphic phantoms.

### Experimental imaging of Four-Organ phantom

A series of SPECT/CT images of the Four-Organ 3D phantom (described in Objective 2), filled to demonstrate time points of a representative MRT therapy, was produced at partner Christie. This required the use of an appropriate model to determine the activity in each organ as a function of time. There have been several theoretical studies on the biokinetics of  $^{177}\text{Lu}$  Dotatate therapy and the digital biokinetic phantom established by partner LUND was chosen as the basis for the phantom activities; as it contained a comprehensive multi-compartmental mathematical model, based on clinical measurements. The LUND model was adapted to reflect the volumes of the individual phantom organs and used to generate a set of TACs. The determination of activities for each organ and background was based on the starting point of a 7.4 GBq administration to the whole body of the patient and scaled accordingly, the medulla to cortex ratio was set at 3:1.

The phantom was filled with  $^{177}\text{Lu}$  on 6 occasions to reproduce 6 time points chosen to best represent typical clinical imaging time points. The time points chosen were 1, 4, 24, 40, 72 and 144 hrs post administration. These scans were then split into 2 datasets of 5 points, one including an earlier time point 4hr, and one a late time point 144hr reflecting the conclusion of the time point optimisation work of the project (see objective 5). The phantoms were filled with stock solutions of  $^{177}\text{LuCl}$  in a 0.4M HCl buffer, where the activity concentration was determined by measurements on a clinical calibrator using factors referenced to NPL's existing primary activity standard. Figure 11 shows the accuracy of filling the phantom organs, tumour and background at each time point, representing the theoretical TAC curve to better than 2.0 %



**Figure 11: TAC (solid line) and phantom filling activities (filled circles) for the Four-Organ phantom**

SPECT/CT imaging was acquired on a GE Discovery 670 SPECT/CT camera at partner Christie. The images were taken using an acquisition protocol matching the calibration protocol (which in turn was fitted to the clinical settings). In addition to SPECT/CT scans, planar scans were taken for each time point to enable assessment of systems using planar or planar/PECT hybrid dosimetry. In order to control some of the variability in the

system two sets of images were prepared. The first set was SPECT/CT reconstructed with the same protocol as the calibration, and also preregistered. The raw tomographic scans were also made available. To ensure the data could be read by a variety of software systems, (i.e. both vendor neutral and vendor specific), the image sets were adapted by careful manipulation of the Digital Imaging and COmmunications in Medicine (DICOM) header. Three versions were produced to cover GE, Siemens and Hermes DICOM formats.

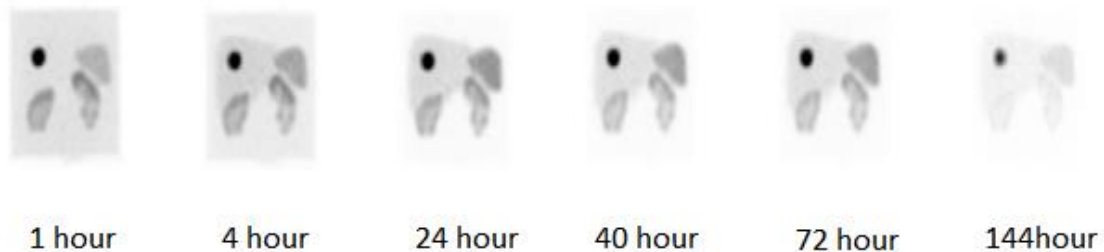


Figure 12: SPECT image sequence generated by phantom imaging (6 time points)

### MC simulation of SPECT images of the Two-Organ and Four-Organ phantoms

MC SPECT simulations were performed to generate realistic images from activity distributions in complex geometries. Simulated images were created with two different MC codes: SIMIND and GATE. The added value compared to the experimental imaging from generating MC based images is that each contribution in image formation can be simulated separately to provide the following images:

- Total images: these include all events from primary photon as well as events from scattered photons. Total images represent the real measured projections.
- Perfect images: no scatter and attenuation effects i.e. a situation where the activity distribution can be regarded as being in vacuum without any photon interactions in the phantom. Perfect images are not possible to obtain experimentally but can be useful since they represent a real study corrected by a perfect attenuation and scatter correction method.
- Primary images: the image includes events from photon attenuation only
- Scatter images: contain only those events coming from photons scattered in the phantom

In all cases, different types of camera effects such as collimator resolution, scatter in the crystal and in the collimator and septal penetration were included and 6 energy windows were created. The data sets produced consisted of raw data meaning that the files need to be reconstructed for use in a clinical work station.

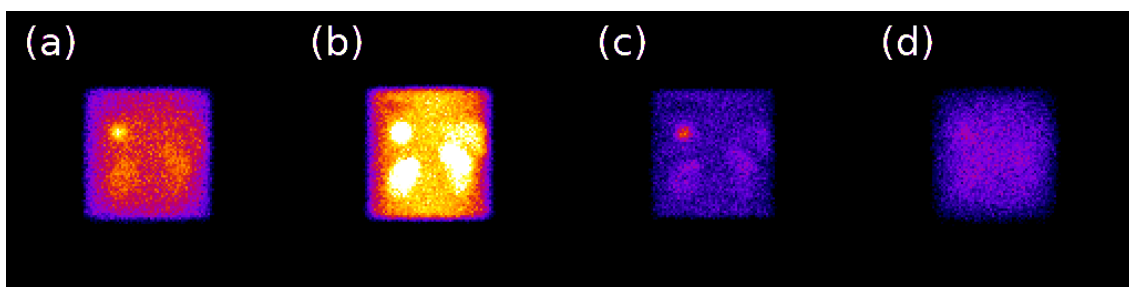
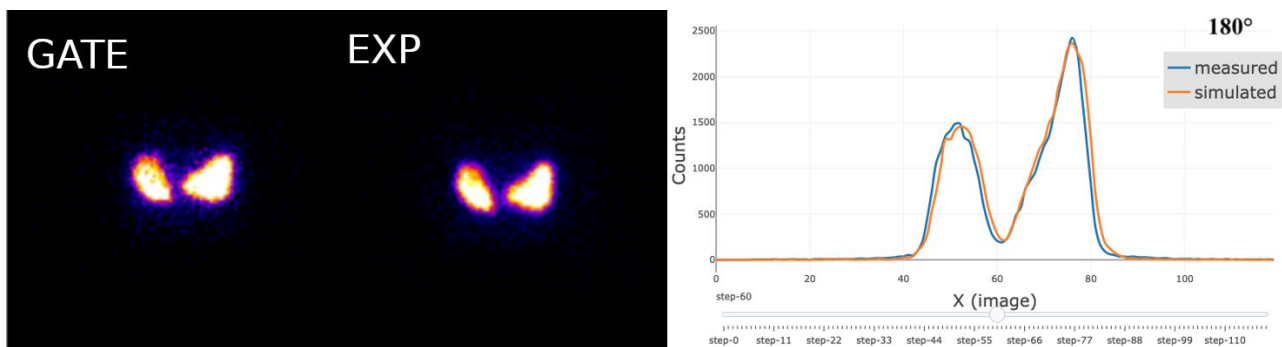


Figure 13: Simulated SPECT projections for the Four-Organ phantom corresponding to (a) Total - all photons, (b) Perfect - no scatter and attenuation effects, (c) Primary – attenuation only and (d) Scatter – scattered only

The MC SIMIND simulations used a CT based voxel-based phantom to define the activity distribution and a density distribution, providing a direct link to the physical Two-Organ and Four-Organ phantoms (described in Objective 2). SPECT data sets were produced for circular orbits (diameter 27.5 cm) and non-circular orbits, for which a routine is used that calculates the distance to the boundary from the centre of rotation for each projection angle. Images were created for a generic GE model (3/8" thick crystal) and a Siemens Symbia T2 (5/8" crystal). In the MC GATE simulations, at the start of the project, generating projections simulating the auto-contour acquisition mode was not possible, only the movement of the detector with a fixed radius was possible. However, new developments in GATE have been made to address the 2 critical issues preventing

auto-contouring: a) Implementing non-circular motion of the detector head; and b) Large distance between the phantom model and gamma camera due to voxelised geometries. These new developments use tessellation modelling to allow the movement of each camera head independently, for which the positions and rotation angles were extracted from the experimentally acquired DICOM images. Thus, now using GATE, simulated SPECT images have been created by the project for the Two-Organ and Four-Organ phantom (with tumour spheres) for a Siemens Symbia T2 (5/8") system.

To assess the similarity between the simulated and experimental images, a common approach is to compare profiles across the images at different projection angles. In this project additional metrics were tested: Mean Squared Error (MSE) and Peak Signal to Noise Ratio (PSNR). MSE and PSNR can be used to estimate the absolute errors, and the normalised MSE can evaluate preservation of border gradients, whereas the Structural Similarity Index (SSIM) is a perception-based model similar to the Human Vision System (HVS) which can quantify the visual similarity. The project's initial results are very encouraging, with a SSIM of 0.95 comparing GATE simulated and experimental images of the Four-Organ phantom for a Siemens Symbia T2 system. As it is not easy to perfectly position the experimental phantom during repeated imaging acquisition, an important contribution of the difference between measured and simulated images is due to differences in phantom/model positioning. In Figure 14 a GATE simulated (left) and experimental (middle) Four-Organ phantom image is shown, as well as one of the X-profiles (summed over the Y-axis) at 180° for both images (right).



**Figure 14: (left) GATE simulated image of four-organ phantom and corresponding experimental image. (right) Horizontal profiles (summed over the Y-axis) at 180° for both images**

### Simulation of CT images of the Two-organ and Four-organ 3D printed phantoms

In patient dosimetry for radionuclide therapy CT images are used to quantify the physical properties of patient tissues as needed for attenuation and scatter corrections in the quantitative SPECT reconstruction, and for quantification of the mass density distribution for subsequent voxel based absorbed-dose calculations. However, the contribution from the CT image values to the combined uncertainty in absorbed doses estimated from SPECT/CT is often not considered. One possible way for undertaking uncertainty analyses is to produce simulated CT images where the true underlying properties of the materials are known by definition. In this project, the feasibility of two approaches to address this issue, were investigated for the simulation of CT images: the first is based on stoichiometric calibration and noise characterisation, the second on CT modelling with the GATE MC software, including a similar noise characterisation.

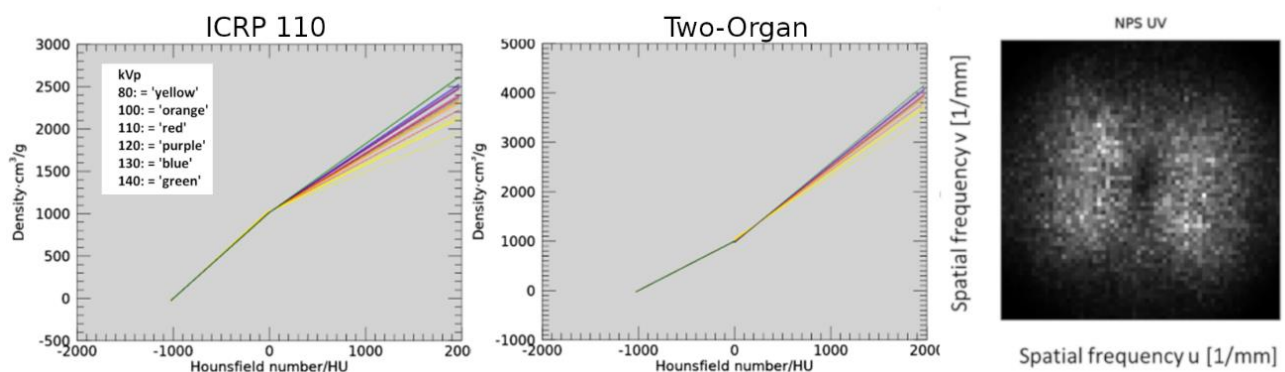
#### ***Stoichiometric calibration and noise characterisation***

Experimental Measurements: CT measurements of a CIRS calibration phantom were made at partners LUND, UKW, and Christie, for different settings of the known tube voltage (kVp), mAs (milliampere-second), and tomographic reconstruction methods. This resulted in 28 combined sets of Hounsfield Unit (HU) data for the inserts of the CIRS phantom. The same centres also performed measurements of a uniform water phantom.

Stoichiometric calibration: A mathematical expression, deduced from first physical principles, was used to describe the HU for different materials. The expression has one parameter whose value is determined by fitting the expression to measured HU data and is used to characterise the particular scanner. The value of this parameter was determined for all sets of the measured HU data, acquired at the 3 different centres at LUND, UKW, and Christie.

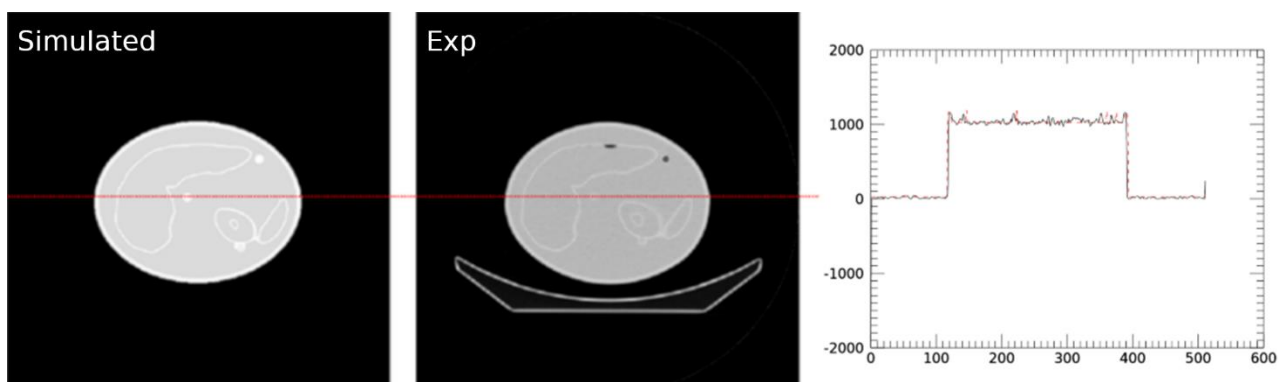
**Determination of the noise-power spectrum (NPS):** The noise in 3D VOIs placed symmetrically in a difference image showing the change in two consecutive scans of the same uniform phantom was analysed in Fourier space. A normalised NPS was calculated from this data which described the frequency dependence of the noise in 3D Fourier domain. By inverse Fourier transformation and adjustment of the overall noise magnitude for different HN, a probability density function was created from which random components of the HN could be sampled. The addition of the noise to the CT image was performed with the assumption that the imaging system behaves as a linear system.

**Results:** Figure 15 shows the calculated HN values for materials of different chemical composition and mass density, for human tissues from ICRP 110 (left) and the Two-organ phantom (middle). The slopes of the segment for higher CT image values differ because of the different photon interaction probabilities for the different types of materials, and for different settings of the voltage during CT acquisition. The right figure below shows one of the 3 orthogonal planes in the frequency domain, through the three-dimensional NPS determined for the CT system in LUND, for 120 kVp. This NPS was then used to mimic the noise for the simulated data.



**Figure 15: Calculated HN values for materials of different chemical composition and mass density, for (Left) human tissues from ICRP 110 and (Middle) the Two-Organ phantom. (Right) An orthogonal plane in the frequency domain, through the three-dimensional NPS determined for the CT system in Lund, for 120 kVp**

A comparison between simulated (left) and measured (middle) CT images of the two-organ phantom for a GE 670 CT system is shown in Figure 16. The right panel shows a profile through the simulated and measured images, at the vertical position indicated by the red line. As can be seen there is an excellent agreement between the simulated and measured images. Preliminary data suggest that the noise modelling manages to capture the noise level of the measured image, however these noise properties need to be evaluated further.



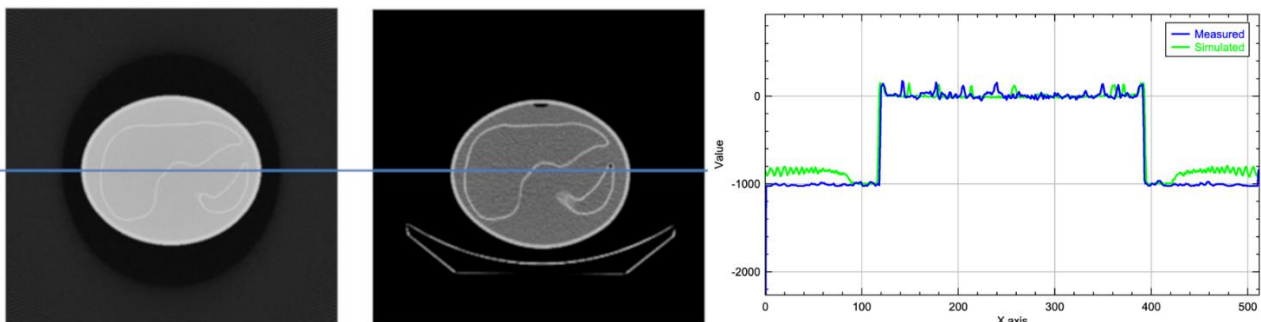
**Figure 16: (Left) Comparison between simulated and measured CT images of the Two-organ phantom for a GE 670 CT system. (Right) profile through the simulated and measured images, at the position indicated by the red line. (Stoichiometric method)**

### CT simulation with GATE and noise characterisation

GATE was chosen as a MC code in the second approach as it is optimised for image simulation and has a dedicated template for CT system simulations. This template allows the simulation of the interaction of the photons with a real detector, as long as the characteristics of the detector are known. To begin, the CT projection (or sinograms) are simulated. These can then be split into 2 parts: (i) the interaction of the X-ray beam with the phantom and (ii) the interaction of the photons with the detector.

Interaction with the phantom: The X-ray spectra used in the simulations were based on the kVp, the other components determining the spectrum are the flat and the bowtie filters of which the designs are proprietary information of manufacturers. Therefore, a single filter was applied for which the thickness was analysed to achieve a Half Value Layer (HVL) similar to the values available in the International Commission on Radiation Units and Measurements (ICRU)12 report on CT X-ray spectrum characterisation. As a CT image reflects the attenuation of the radiation within the patient this can be obtained directly using the Fixed Force Detector Actor in GATE, which is a deterministic computation of the probability of photons reaching each specific pixel of a perfect detector (ray tracing technique). The contribution of scattered photons to the primary image was investigated by performing full MC simulations with GATE, scoring only the contribution of scattered photons reaching the detector. It was found that the scatter-to-primary ratio was less than 5. The obtained CT projections were then able to be reconstructed using standard techniques.

Interaction with the detector: The noise characteristics of the CT system was determined with a similar approach as previously described for the stoichiometric calibration method: i.e. the determination of the NPS was from images acquired with a homogeneous water phantom on a system that is under simulation and with similar acquisition settings as those used for the 3D printed phantoms. The magnitude of the overall noise level applied on the normalised noise image (mean 0 and standard deviation 1) was assessed by determining the standard deviation in pixel values in a homogeneous phantom. The obtained noise images were then added on the reconstructed CT slices. The figure below shows a comparison between simulated (left) and measured (middle) CT images.



**Figure 17: (Left) Comparison between simulated and measured CT images of the two-organ phantom for a GE 670 CT system. (Right) profile through the simulated and measured images, at the position indicated by the blue line. (MC method)**

In conclusion, both approaches for CT simulation combine measured data (characteristics for a specific CT system) with in one method advanced analysis methods for the stoichiometric calibration approach and in the other method with fast ray tracing techniques using GATE. Both approaches appear to be promising in the production of simulated CT images. The data also suggests that the overall level of noise compares well with the measured level. Although the realism of the simulated CT images could be improved by also including effects of limited spatial resolution and effects such as cupping. Furthermore, the GATE approach is limited by missing proprietary system-specific information, such as filter information and the reconstruction methods applied.

### Development of a tool to read and reconstruct simulated images

A software tool was needed to enable clinical sites to read and reconstruct the reference simulated SPECT CT images (SPECT projections and CT slices) on clinical workstations. Representative SPECT/CT DICOM files from Siemens, GE, Mediso and Philips systems were compared by SCK.CEN to identify those tags necessary for compatibility with clinical workstations. However due to the lack of harmonisation among the systems, a system-specific approach was chosen by the project.

Copies of existing files were used as system-specific templates for converting the simulated images into useable DICOM files. A python-based script was written to import simulated SPECT/CT images in existing DICOM files and enable their use in clinical practice. For each SPECT/CT system, a list of specific DICOM tags to modify was found and was subsequently modified by the python-based script. Simulated images were successfully incorporated in files from 5 systems (Siemens: Symbia T2 and Symbia Intevo Bold; GE: Discovery 670 and Optima 640; Philips: Brightview XCT) and could be read on the manufacturer's respective workstations, (Siemens Syngo, GE Xeleris and Philips Jetstream), or on a vendor-independent workstation (Hermes). A python-based script was also written to modify the radial and angular position of the detector heads to allow examination-specific body contouring.

### **Development of a design for a database for MRTDosimetry reference data**

Ensuring that the datasets generated in this project are widely available to clinical centres will maximise their impact. It is envisaged that the datasets will be used in conjunction with both the QI (Objective 2) and dosimetry commissioning protocols (Objective 5) to benchmark existing sites and for testing new image-based dosimetry calculation methods. The adoption of the datasets (and future datasets based on the methodologies developed in this project) will be accelerated by the promotion of the datasets from appropriate standards bodies. To support this a design for a database to host the data from this project and an internal example implementation were developed. The project partners are currently working on the database being made publicly available through an appropriate international standards body following the completion of the project.

### **Summary of outcomes**

The imaging data produced in this project provides a unique dataset that has uses in both QI and dosimetry. The design of the project's phantoms provides a realistic set of activity distributions to validate a number of parts of the QI chain. The reconstructed images may be used for validation of partial volume correction methods by comparing the recovered activity with the (known) filled activities. Different reconstruction methods can also be validated by using the projection dataset, while dosimetry platforms can be commissioned by performing dose calculations on the reconstructed data. The open-source availability of the images edited such that they can be used in the majority of clinical camera systems in Europe, and together with the QI calibration protocols developed by the project, will also allow ready calibration of clinical systems by hospital physicists.

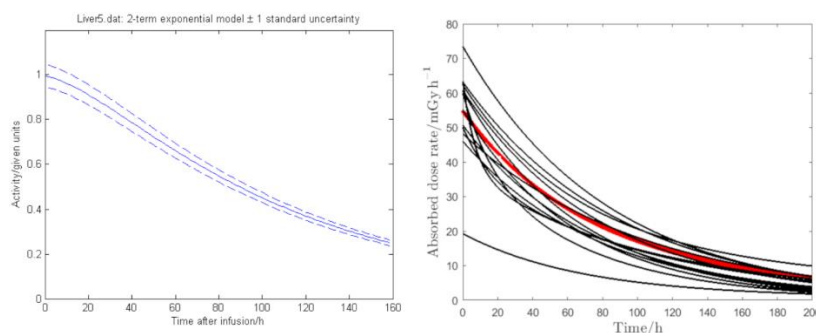
**4.4 Objective 4:** *To improve the accuracy and metrological traceability in the calculation of dose from time-sequences of QI measurements by optimisation of the time points (i.e. obtaining cumulated activity from a TAC), choice of measurement modality (imaging or non-imaging), refinement of absorbed dose standards, and validation of alternative absorbed dose calculation methods in phantoms using physical measurement techniques such as MR sensitive gel-based and film-based dosimetry and MC simulations.*

The aim of this objective was to improve the accuracy and metrological traceability in the calculation of dose from time-sequences of QI measurements through the optimisation of the time points and choice of measurement modality (imaging or non-imaging). In addition, to support this by refining absorbed dose standards, and validation of alternative absorbed dose calculation methods in phantoms (e.g. using physical measurement techniques such as MR sensitive gel-based and film-based dosimetry and MC simulations).

### Modelling methods for the determination of optimal scan times

One of the major sources of uncertainty identified for the determination of absorbed dose for MRT is the fitting of the TAC data. Therefore, a theoretical study was undertaken to determine if the choice of optimal scan times within a dosimetry study could influence the uncertainty of the TAC. A pilot method was developed in the preceding project HLT11 MetroMRT for obtaining optimal scan times. However, although this method was tested on suites of data sets provided by clinical partners, it was not validated. In this project modifications were made to this pilot method and the way the method was validated. The pilot method was extended to consider that activity values derived from a succession of quantitative images have associated correlation. As part of this a review of current recommendation from The European Association of Nuclear Medicine (EANM), ICRU and Committee on Medical Internal Radiation Dose (MIRD) and several clinical groups for the choice of time points within dosimetry was performed. The conclusions made were that these techniques relied on population biokinetic modelling and not individual patients. It was also noted that compliance with some suggestions would be hard due to the number of scans required, particularly if outside normal departmental hours.

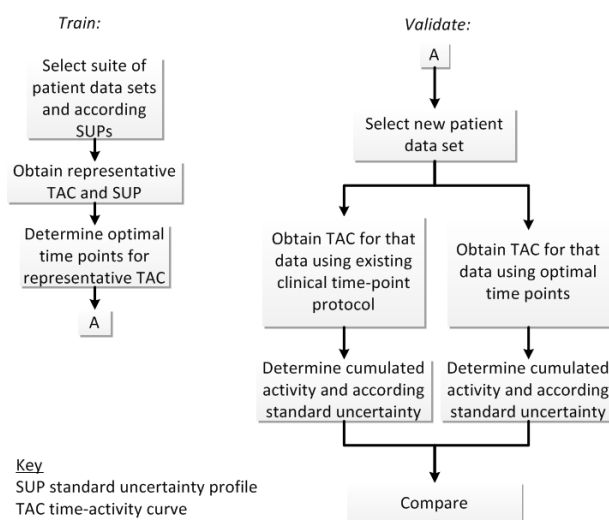
Taking this into consideration the proposed technique for the determination of optimal scan times relies on the definition of an activity standard uncertainty profile (SUP) a representation of the standard uncertainty as a function of activity (e.g. proportional to the activity). The activity SUP is superimposed on a representative TAC as  $\pm 1$  standard uncertainty. To do this, firstly, a representative TAC had to be determined for a particular isotope and therapy. This TAC needs to be based on a suitably large number of TACs of a specific type of pharmacokinetic model (e.g. bi-exponential) obtained for patients with the same disease and treated with the same pharmaceutical at the same clinical centre. Thus, clinical TACs were obtained from anonymised patient records for  $^{177}\text{Lu}$  Dotatate therapies. A representative “training” TAC was then created by defining a median of fits of 13 clinical TACs. This choice of median served to reduce the effects of outlying curves away from a generalised group of curves. The median TAC also generated an activity SUP which can be defined as above.



**Figure 18: (Left) Plot showing activity standard uncertainty superimposed on a TAC. (Right) Plot showing generation of median TAC (red) from clinical data**

The system was then trained to determine the optimal time points for this median TAC, defined as having the minimum of standard uncertainty for the area under TAC. An arbitrary set of 5 time points was chosen together with the activities and activity uncertainties from the SUP for those time points. The integral of the TAC is then evaluated together with the standard uncertainty for the estimation of the area under the curve. The value of the integral is unchanged since the TAC is fixed. The process is repeated many times with the arbitrary time

points selected each time. After extensive sampling a solution will be determined that indicates the time points that represent the optimal solution as a minimum of uncertainty. The optimisation process was controlled by a discrete stipulation of starting candidate time points that (i) increased in time (as in a dosimetry exercise) and (ii) included time points used in clinical protocols. Time points could also be excluded, such as night time points, and those that were too close together to be clinically viable. The minimisation process controlling the optimisation produced a Jacobean matrix of partial derivatives of the linearisation of the fit. This matrix plays a key role establishing variances and covariances associated with the estimated values of the model parameters. A similar process was followed to identify time points for whole body measurements of  $^{131}\text{I}$  mIBG therapies. Once the optimal time points for the training TAC had been achieved a planned validation of the technique was trialled. This involved taking individual patient data and then comparing the cumulated activity achieved with the clinical time points against that determined by using the optimal time points derived by the training set.



**Figure 19: Training and validating the TAC time-point optimisation process**

It was found that for patients with TACs sufficiently close to the representative TAC, the method could be used to yield scan times that offered a valuable improvement. However, the method gave only a small improvement for other patients. After extensive consideration the manner of selecting the representative TAC and the presence of extreme TACs were ruled out as the causes of this, and the major effect was attributed to the considerable variation in the response of patients to the therapy. Therefore, in order to achieve better performance, the most promising method would appear to be the use of pre-therapeutic (tracer) data on an individual patient basis. Several recommendations have also been deduced from this work which can be used clinically. Firstly, that it is highly desirable to include a time point that is as late as possible in the set of time points to capture more faithfully the tail of the TAC. This is because the tail of the TAC often makes an appreciable contribution to the cumulated activity. Secondly, the concept of an activity SUP provides knowledge of how the quality of measured activity values can be quantified in terms of the activity itself.

### Choice of measurement modality (imaging or non-imaging)

In the last few years, two main methods have been described in the literature for obtaining absorbed doses to organs and to the bone marrow in radionuclide therapies:

1. Non-imaging based methods (e.g. blood or whole-body dosimetry),
2. Image based methods (e.g. planar or SPECT/CT imaging).

In this project the potential differences between 3D imaging, planar whole-body scans and whole-body scans with measured or calculated attenuation correction and external probes with respect to the integration of the TAC were assessed. Since dosimetry is not performed routinely in radionuclide therapies, no comprehensive clinical datasets comprising non-imaging data (e.g. blood data or external probe measurements) or imaging data (e.g. planar or SPECT/CT imaging) were available for an inter-comparison in any of the partner or collaborating clinical centres. Therefore, separate comparison exercises for non-imaging and imaging based

methods were performed and then from the results of these recommendations for future considerations were derived.

For the study, 10 datasets for non-imaging scenarios ( $^{131}\text{I}$  mIBG: whole-body TAC,  $^{131}\text{I}$  therapy of differentiated thyroid cancer: blood activity measurement data) and two patient data sets with multiple post-therapeutic SPECT/CTs (after  $^{177}\text{Lu}$  Dotatate therapy of neuroendocrine tumours) were distributed to project partners (UKW, THG, LUND, ASMN, Christie, INSERM, OUHT, RSCH and THG) so that they could use their available software to integrate the respective TACs and assess the respective uncertainties. Although the number of ~7-30 **non-imaging** measurements was considered to be sufficient for deriving the TAC with a sufficiently low uncertainty, some software solutions provided area-under-the-curve values with differences of up to 21 % to the mean value of all results. In addition, only two results included the covariance term when calculating the error propagation.

For the **imaging**-based TAC integration, the majority of sites relied on in-house solutions. One of the major difficulties encountered during this exercise was to prepare the image data including the VOIs in such a way that it could be read by all types of software. Unfortunately, none of the sites provided an estimate of the uncertainty of the TAC integration, as their software solutions could not process uncertainties. Overall, the results differed considerably, with differences to the mean value of up to 50 %. Although all the images provided for the exercise were given in the DICOM format, there is still no standardised way to transmit VOIs. In addition, many software solutions do not read the DICOM-RT format used to distribute a set of pre-delineated VOIs (i.e. to be then applied in this exercise's datasets).

To try and address these issues the project concluded that, software users need to be provided with sufficient information for accepting or rejecting fitting functions. Also, the resulting uncertainties need to be reported, including the covariance term of the error propagation. This would lower the inter-observer variability and should, therefore, be an essential part of any dosimetry software package. Finally, as there is currently no standardised way of registering and integrating voxel-based TAC, all results derived from voxel-based codes should be considered preliminary and should be reported only in conjunction with the time integrated activity curve (or absorbed dose) of the whole organ or lesion including the respective uncertainty of the whole-organ or lesion values.

## Refinement of absorbed dose standards for radionuclides

In the preceding project HLT11 MetroMRT a primary standard instrument was developed to realise the quantity absorbed dose to water for a given radionuclide in solution. This was achieved using an extrapolation chamber to detect the ionisation in air at the surface of a radionuclide solution. This work was the first application of the technique to unsealed radionuclides and the first proof of its feasibility.

Measurements made at NPL during the preceding HLT11 MetroMRT project using a solution of  $^{90}\text{Y}$  were published during this current project. This work demonstrated that it is possible to make a primary standard measurement of absorbed dose to water within a radioactive solution with a standard uncertainty of 1.4 % ( $k=1$ ). The measurement agreed well with values from the standard dose calculation methods with an expanded uncertainty of 2.8 % ( $k=2$ ). This (in the case of  $^{90}\text{Y}$ ) effectively achieves traceability of the calculations to a primary standard of absorbed dose.

For the current project the work on the primary standard instrument was taken forward and developed in the following ways: (1) A re-analysis of the uncertainty budget of the previous measurements was performed including the uncertainties in the MC modelling used to determine correction factors. The limiting value in the measurement uncertainty budget was found to be the determination of the volume of the radioactive solution in the well of the measurement phantom. (2) In order to reduce the uncertainty of the measure of solution volume, a technique was introduced using a peristaltic pump to inject the solution in the chamber that could be used to determine the volume of the radionuclide solution. The pump used had a good repeatability and an accuracy of 0.21 % on injecting a mass of approximately 62 g.

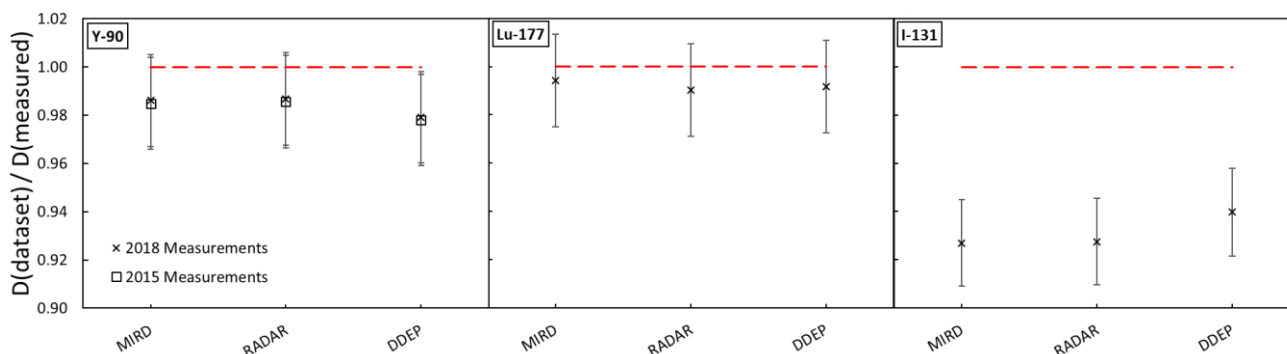
In the case of the MC correction factor  $k_{MC}$  which converts (corrected) measured absorbed dose to air to the absorbed dose to water at centre of large phantom (sufficiently large to establish emitted particle equilibrium for a specific radionuclide), the factors re-estimated for the uncertainty budget of the MC simulation were: (i) the uncertainty in the diameter of the collecting electrode in the extrapolation chamber, and (ii) the uncertainty due to the choice of beta spectrum of the  $^{90}\text{Y}$  used in the MC simulation. The effective electrode diameter was re-measured at NPL and was determined to be 30.302 mm  $\pm$  0.011 mm ( $k=2$ ), a value which agreed with the previous capacitance method within 0.1 %. The beta spectra used in the initial study in preceding project

HLT11 MetroMRT were taken from the MIRD and RADAR data tables. In this project the spectra were rebinned to improve accuracy by minimising binning artefacts inherent in the previous spectra. The combinations of these improvements resulted in an improvement of the Type B uncertainty of  $k_{MC}$  which was updated to 0.30 % ( $k = 2$ ).

### New Results

Following all these improvements the measurement of dose in  $^{90}\text{Y}$  solution was repeated and after uncertainty analysis was found to have improved with expanded uncertainty 1.98 % ( $k=2$ ). The uncertainty on volume of solution was now  $<1$  % ( $k=1$ ) compared to 1.24 % previously. Preliminary analysis also showed an agreement, with the previous  $^{90}\text{Y}$  measurement of 0.1 %. Agreement with analytic data from MIRD, RADAR and the Decay Data Evaluation Project (DDEP) evaluation are in the order of 98 % (see Figure 20).

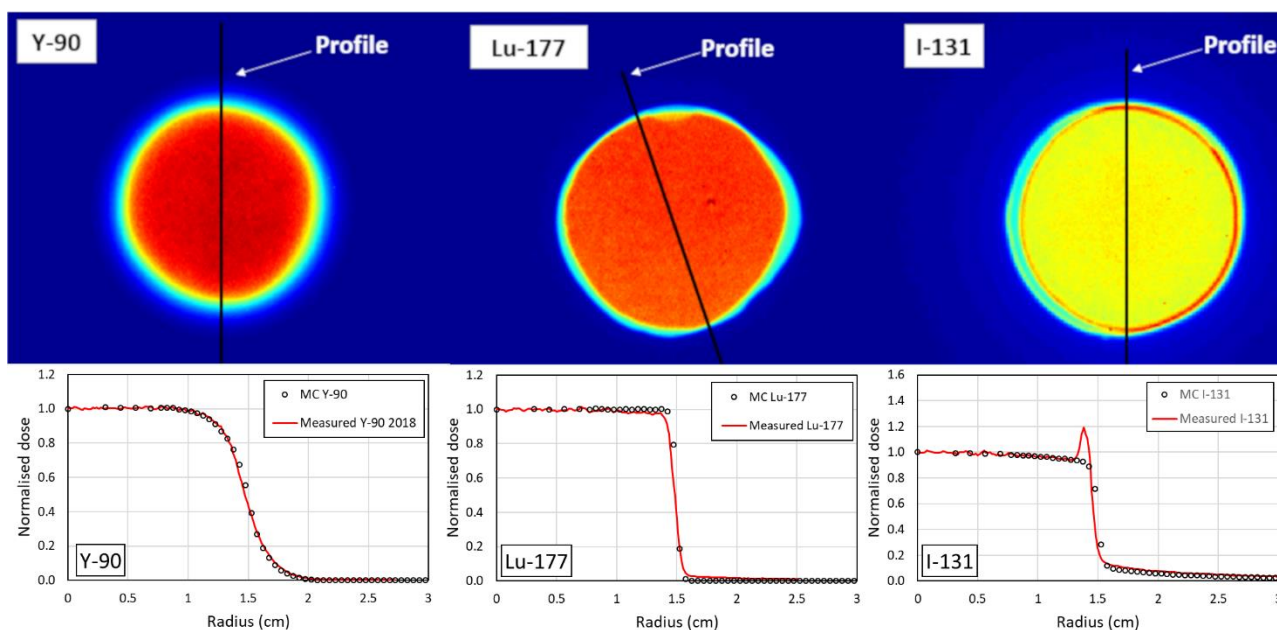
The same technique was then used to determine the chamber response with solutions of  $^{177}\text{Lu}$  and  $^{131}\text{I}$ . A depth range of 0.25 mm to 2.5 mm was covered with 0.25 mm steps. The current at each depth was corrected for, temperature, pressure, humidity, leak current, ion recombination and source decay using LabVIEW software remotely controlling the measurements. New measurements with a  $^{177}\text{Lu}$  solution gave results that again showed agreement with MIRD and RADAR data with a measurement uncertainty of 1.92 % ( $k=2$ ) (see Figure 20). However, for  $^{131}\text{I}$  solution the measurements show a significant disagreement when compared to the MIRD and RADAR data of +7%. This inconsistency has been identified as most likely to be a problem with simulation of the photons in the laboratory environment, and due to backscatter not being accounted for during the simulation.



**Figure 20: Comparison of measured absorbed dose to water for  $^{90}\text{Y}$ ,  $^{177}\text{Lu}$  and  $^{131}\text{I}$  to reference values. Datasets from this work (2018 - cross) and previous measurement (2015 - square) are shown for  $^{90}\text{Y}$ . Red dashed line indicate line of unity**

### Film measurements of absorbed dose

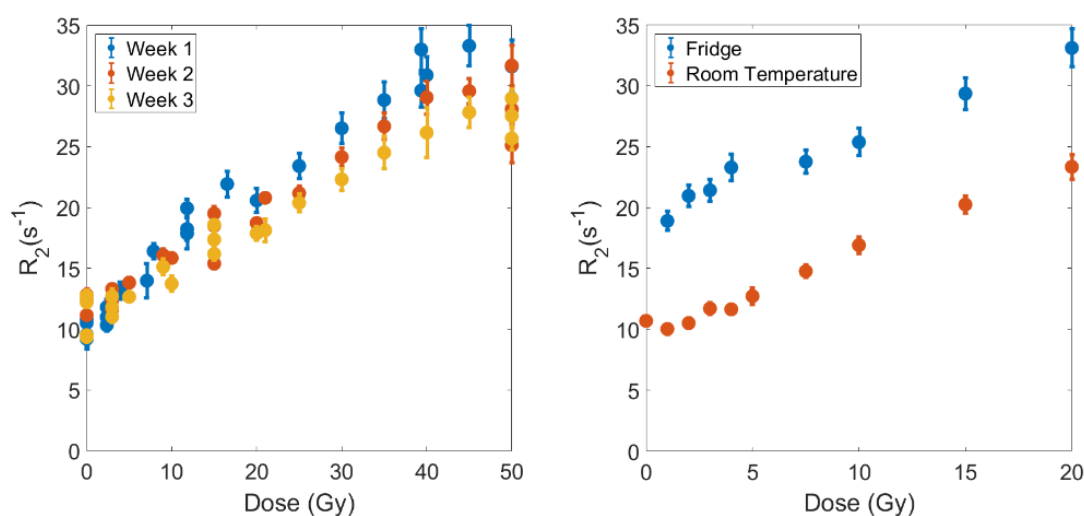
The preceding HLT11 MetroMRT project established the use of radiochromic film measurements as a way of verifying the MC codes used for simulation of beta radiation transport. This was achieved with the measurement of dose in  $^{90}\text{Y}$  solution in a cylindrical phantom geometry at NPL. Therefore, this project aimed to expand on this and further verify the technique for both  $^{177}\text{Lu}$  and  $^{131}\text{I}$  solution in the same geometry where a film is suspended between mylar layers within the radionuclide solution. MC calculations were performed in EGSnrc software to determine the absorbed dose profiles from all emitted particles at the edge of the film and into the surrounding phantom wall. Measurements were made for all three isotopes (i.e.  $^{90}\text{Y}$ ,  $^{177}\text{Lu}$  and  $^{131}\text{I}$ ). Profiles across the dose map of the three radionuclide solutions were used to compare with MC simulations. Unfortunately, the measured profiles suffered from the effects of uneven thickness of the glue distribution around the edges of the containers. Therefore, the least affected profile was chosen for comparison with MC calculations in each case. Figure 21 shows a good agreement between the radial profiles determined with MC simulations for  $^{90}\text{Y}$ ,  $^{177}\text{Lu}$  and  $^{131}\text{I}$ . Please note that the anomaly in the  $^{131}\text{I}$  measurement is most likely to have been caused by the effect of the glue.



**Figure 21: Comparison of film measurement with MC simulation for  $^{90}\text{Y}$ ,  $^{177}\text{Lu}$  and  $^{131}\text{I}$ . Profiles are shown from the center of the phantom (0) as indicated by the black lines**

### MRI Sensitive Gel Measurements of Absorbed Dose

The physical response in a gel from a radionuclide solution was investigated for the use of MR responsive (Methacrylic and Ascorbic acid in Gelatin Initiated by Copper) MAGIC gel as a potential independent dose measurement. MAGIC gels have been shown to provide a relaxation rate ( $R_2$ ) response to absorbed dose. MAGIC gels may be manufactured under normal atmospheric conditions; however prolonged exposure to oxygen causes some inhibition and artefacts. The gels also have advantages for measuring dose in MRT situations. For example, the elemental composition means that the electron-density and mean Z are closer to muscle than water, and as well as having an MR response the effects of irradiation can also be observed optically.



**Figure 22: (Left) Dispersion of response curve with delays between manufacture and irradiation. (right) Effect of cooling rate on dose response**

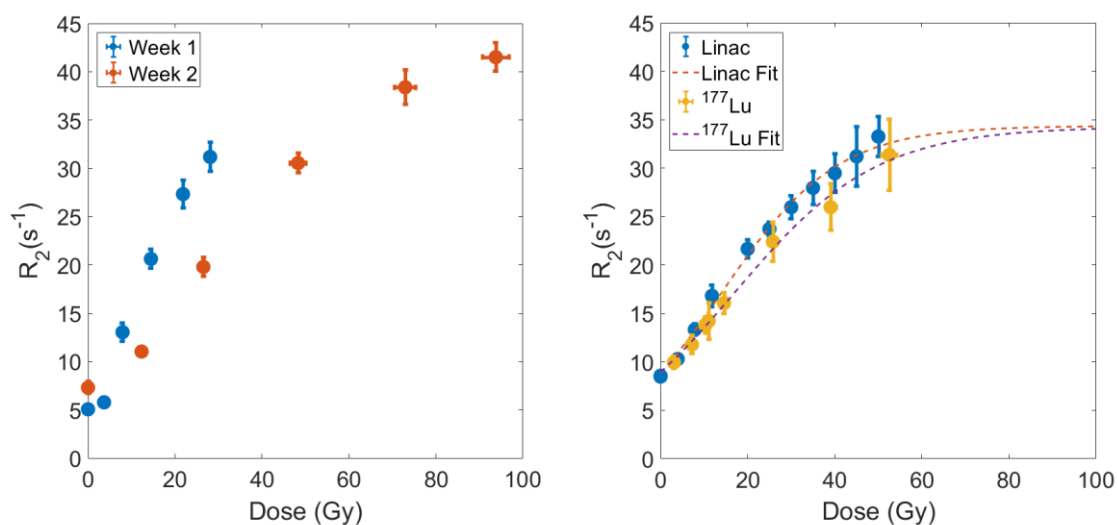
A provisional set of experiments to determine optimal manufacture for gels was undertaken at the University of Manchester (a collaborator) and NPL. Gels were irradiated by the 6MV linear accelerator at NPL to establish an initial calibration under a beam related to a primary standard. Irradiation was controlled to provide doses

from 0 to 50 Gy in 5 Gy steps This work was then repeated, irradiating the gels using a 6MV beam on a clinical linear accelerator at partner Christie. All imaging was performed on a clinical 1.5T MR scanner at Christie. The dominating uncertainty in determining dose response was the uncertainty in measuring the  $R_2$  MR signal which was limited by noise. This was improved by adapting the MR sequence to allow multiple averages.

A full study of a set of MAGIC gels was made, investigating the effect of time of irradiation after manufacture and the stability of measured response after linear particle accelerator (LINAC) irradiation. It was determined that the response of the gels became more disperse, dependent on the time of irradiation after manufacture. Therefore, subsequent irradiations were performed immediately after manufacture (see Figure 22). The effect of controlling the temperature of polymerisation was also investigated, with identical batches of gel split and cooled at room temperature or at 4°C in a refrigerator. A significant difference was seen between the  $R_2$  values of the different sets of gels (Figure 22). These results clearly show that the method of cooling the gels has a large influence on the response under MR imaging and should be held constant across different experiments.

### Isotope doped gels

The project continued its work on gels with a study of the calibration of gels manufactured with a radioisotope added to the gel. This work was performed in sealed nitrogen filled P6 vials. 10 mls of irradiated gels were manufactured in each vial, with 10 % (1 ml) of the water replaced by 1 ml of isotope solution in water. Production was controlled such that each vial contained components with an uncertainty <1 %. All gels were left to polymerise at room temperature. MC modelling (GEANT4) was performed to determine the dose factors for the isotope filled gels thus ensuring composition and density matched the formulation. The activities used were such that doses of 0 to 50 Gy could be achieved over a range of time points to match the LINAC range.

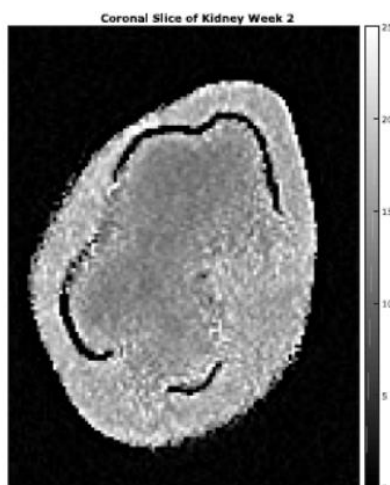


**Figure 23: (left) Effect of leaving a gel to decay and build up dose: Blue week1; Red week2. (right) Comparison of linac (blue) and <sup>177</sup>Lu (red) irradiated gel dose response**

The first isotope tested was <sup>177</sup>Lu, in the form of <sup>177</sup>LuCl. Once produced, the isotope irradiated gels were imaged at several time points in order to determine any changes of response with time. This was done by examining the response to the same dose at different times and by matching activities to give the same dose at different times. The results show that at later times the  $R_2$  response appeared to reduce with the same dose, indicating that the gel response is dose rate and/or time dependent (see Figure 23). Therefore, this must be taken into account when designing experiments as the time point for reaching dose should be considered. A comparison was also made of isotope dose response to that of LINAC irradiated gels and the results showed good agreement, albeit a bigger discrepancy in uncertainties for gels where the dose was realised later.

A further investigation was continued with gels mixed with <sup>131</sup>I, in the form of NaI. For <sup>131</sup>I a response consistent with the <sup>177</sup>Lu measurements was not observed and responses were either missing, reduced and delayed. In addition, vials showed separation into regions of further reduced  $R_2$  which grew in proportion to the amount of isotope solution added. The <sup>131</sup>I came in a stock solution containing Sodium Iodide, Sodium Thiosulphate, Sodium Chloride, Sodium Dihydrogen Phosphate, Sodium Hydroxide, Disodium Hydrogen Phosphate, and Water. Sodium Thiosulphate has been identified as a strong reducing agent, and thus this may be the reason

for the results as it could be interfering with the polymerising process or the normoxic conditions required for the gel.



**Figure 24: 3D printed phantom R2 dose response showing lower response (dose) in the medulla**

A further study was made of the response of a  $^{177}\text{Lu}$  doped gel in a 3D printed phantom. Initial tests in a Polylactic acid (PLA) printed kidney phantom showed severe artefacts caused by oxygen at boundaries. A 3D kidney was also printed in transparent resin on an SLA Formlabs 2 printer at NPL. Both medulla and cortex regions were filled with the same activity concentration. When imaged, the  $R_2$  in inner Medulla compartment appears to be lower than the  $R_2$  of the gel in the outer, Cortex compartment (see Figure 24 **Error! Reference source not found.**). One explanation for this could be that the gel in the Medulla is kept warmer for longer as it is surrounded by the gel in the Cortex. This may slow the cooling rate of the gel in the Medulla which will affect the  $R_2$  response, as has been previously observed. The large oxygen contamination that was present in the previous 3D printed phantom is also exhibited in this phantom due to the different printing technique and material used. In addition, doses were more consistent with those modelled by MC, but not as comparable as in the vials. These results show the potential of MAGIC gels to be used in realistic phantoms, but issues such as a temperature control, oxygen contamination, supporting chemistry and container material still require solutions before this technique can become generally applicable.

### **Summary of outcomes**

A method for the determination of optimal scan times for an example MRT was achieved, based on an activity standard uncertainty for a reference TAC. This optimisation was based on minimisation of uncertainties. The expected improvements were limited to patient data that closely matched the reference TAC. Further approaches based on pre-therapeutic tracers may prove more personalised.

The problems encountered with the exercise on comparison of choice of measurement techniques highlighted two serious problems within the clinical practice of dosimetry. Firstly, there is no common standard for defining and transfer of VOI information (as there is in external beam therapy). Although, the external beam format is not accepted by commercial nuclear medicine systems without a great deal of work if at all. Secondly, it highlighted the paucity of software available to clinical groups that was able to perform advanced uncertainty analysis on the fits it produced. Both these problems indicate a clear need for software solutions to be developed so that clinical groups can take advantage and use this information.

The use of the primary standard of absorbed dose (developed in the preceding project HLT11) was extended to other isotopes with improved uncertainties. Good results with reduced uncertainties were achieved for  $^{90}\text{Y}$  and  $^{177}\text{Lu}$ , albeit with reduced success for  $^{131}\text{I}$ . The technique depends on the accuracy of the MC simulation used and relies on validation.

The work on the development of MR dose sensitive gels also showed promise. However widespread use will require problems with temperature control and oxygen sensitivity to be addressed. As this currently limits their use to specialist laboratories.

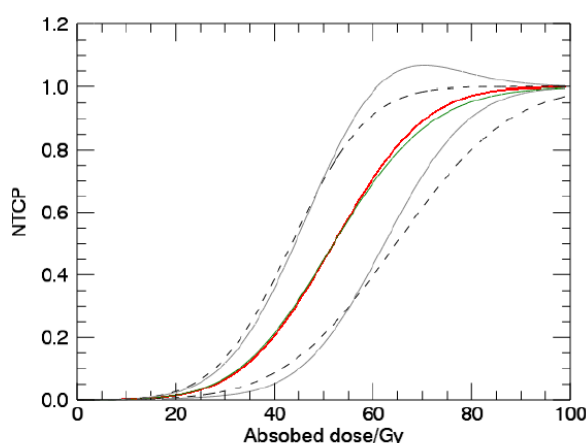
**4.5 Objective 5:** To determine uncertainties in relation to the full MRT dose measurement chain from a primary standard to a range of commercial and non-commercial dosimetry calculation platforms. This includes image quantification (such as uncertainties in the selection of volumes of interest (VOI) and image reconstruction); integration of TACs, propagation of uncertainties in NTCP models, and determination of the overall evaluated uncertainty in the absorbed dose quantification process.

The aim of this objective was to determine the uncertainties in relation to the full MRT dose measurement chain from a primary standard to a range of commercial and non-commercial dosimetry calculation platforms. This includes image quantification, integration of TACs, propagation of uncertainties in Normal Tissue Complication Probability (NTCP) models, and determination of the overall evaluated uncertainty in the absorbed dose quantification process.

### Propagation of uncertainties in NTCP models

NTCP is the parameter used in other radiation therapy modalities for estimating the risk of inducing harmful side effects. The evaluation of NTCP uncertainty has the potential to indicate an acceptable level of uncertainty for absorbed dose for MRT. NTCP curves are traditionally characterised by two parameters:  $D_{50}$  (absorbed dose corresponding to a probability of 50%) and  $\gamma$  (maximum slope). As absorbed dose estimates are sometimes claimed with 10 % uncertainty there is a need for methods that propagate uncertainties in accordance with internationally accepted principles, and also lead to feasible solutions, giving probabilities between zero and one both for the values on the NTCP curve and also the range of values lying in the 95 % ( $k=2$ ) uncertainty swathe. The choice of model is key; as different NTCP models can give very different results.

Following the performing of a literature search comparing empirical models used to describe NCTP curves, the empirical model of a logistic function was chosen. An exploration was made as to the propagation of uncertainties using the traditional GUM (Guide to the Expression of Uncertainty in Measurement, JCGM 100) approach. This contains two assumptions, firstly that the probability function of the NTCP can be linearised. This tends to be true for the central regions of the NTCP curve, but not at the low or high dose regions where linearisation cannot be assumed. Secondly the traditional model assumes that the probability distribution for NTCP is gaussian in the neighbourhood of the estimate of probability measured. However, the standard uncertainty associated with D propagates to a much larger value in the steeper parts of the curve.



**Figure 25: NTCP as function of absorbed dose. Red line: original NTCP curve. Grey lines: 2.5th and 97.5th percentiles of NTCP probability (dash - from MC approach, solid from GUM approach. Green line: mean NTCP from uncertainty propagation**

Traditional calculations of the uncertainty propagation using the GUM technique were performed and showed that the lower and high dose ends of the curves could produce uncertainties that resulted in the probability being unfeasible, either below 0 or above 1. This highlighted the dangers of the extrapolation of clinical data, particularly at either end of the curve. Therefore, an alternative approach was taken by the project, using the techniques recommended in GUM supplement 1. This is a proponent for using MC methods as an implementation of the propagation of distributions. Within this technique, the absorbed dose is assumed to have a Gaussian probability distribution with mean and standard deviation at each input dose. A probability

distribution of NTCP is then determined by sampling from the individual Gaussian distribution, giving an estimate of standard uncertainty and coverage interval at that point. This technique leads to very different probability distributions than the one assumed in the general GUM, particularly at the low and high end of the NTCP curve. Figure 25 shows an example of the “band” of possible NTCP values that can arise for these two differing approaches to uncertainty propagation.

The results proved that the summary statistics obtained by the general GUM method in such a case are unreliable. The estimate of NCTP uncertainty provided by the GUM is optimistically small and that provided by the MC method of GUM supplement 1 is higher. The coverage interval provided by the GUM is also unfeasible as the lower range becomes  $< 0$ . That provided by MCM is feasible (contained within 0-1). With the MGM the confidence interval is kept within the range of feasible results but with higher upper levels than that predicted by GUM. The current NTCP model is non-linear and the standard uncertainty associated with an absorbed dose value is large. In such cases, the law of propagation of uncertainty in the GUM may provide poor results and MC approaches are recommended. Finally, by looking at the ratio of absorbed dose relative standard uncertainty to NTCP relative standard uncertainty, an amplification factor can be defined,  $K = \tilde{u}(\hat{p}) / \tilde{u}(\hat{D})$ .

The current analysis for MRT data gives  $K \sim 10$ . This has two important clinical implications: firstly, to obtain an NTCP value with a relative standard uncertainty of 50 % would require an absorbed dose uncertainty of approximately 5 %. In comparison for external beam radiotherapy NTCP curves  $K \sim 5$ . Secondly, it may become feasible in the future to define an acceptable maximum standard uncertainty in absorbed dose  $D$  to be that which would yield a NTCP having an associated standard uncertainty of at most 10 %.

## Comparison of MRT Dosimetry calculations

Dosimetry calculations in MRT can be described theoretically in a relative straightforward manner by the MIRD schema. However, in reality dosimetry calculations are a complex series of functions governed by practical considerations which vary between centres that may use different imaging systems, calibration systems, analysis software and dosimetry systems. As such it has proved very challenging for the project to compare dosimetry techniques in MRT dosimetry. The doses calculated rely on a whole chain of actions, each of which contributes to the overall accuracy and uncertainty in the final calculation.

### *Theoretical Determination of Uncertainty Budget*

Several theoretical studies on uncertainty propagation were made by NPL working in collaboration with clinical groups during the preceding HLT11 MetroMRT project. This resulted in establishing a framework for estimating absorbed dose and the associated uncertainty evaluation in the measurement processes constituting the dosimetry chain in internal absorbed dose calculations. The framework is based on the basic MIRD model where an absorbed dose in a VOI is product of cumulated activity and a dose factor, (where cumulated activity is given by the area under a TAC derived from a time sequence of activity values). Each activity value is itself obtained in terms of a count rate, calibration factor and recovery factor. Propagation of these estimates of the quantities concerned and their associated uncertainties through the dosimetry chain leads to an estimate of mean absorbed dose in the VOI and associated uncertainty.

This study as part of the preceding HLT11 MetroMRT project applied internationally acknowledged uncertainty principles but did so based on the clinical measurements to be made at each stage and considering the correlation and covariance of uncertainties from one measurement to another. It considered each part of the dosimetry chain from administration of a pharmaceutical to absorbed dose determination. The result is a methodology for estimating mean absorbed dose in a VOI and evaluating its associated uncertainty. During the current project this methodology has been accepted published as an EANM guide covering theory and practice (see the publications in Section 6). This work proposes a framework for modelling the uncertainty in the measurement processes constituting the dosimetry chain that are involved in internal absorbed dose calculations. The propagation of estimates of the quantities concerned and their associated uncertainties through the dosimetry chain to obtain an estimate of mean absorbed dose its associated uncertainty is demonstrated using a clinical example.

***MC modelling of dose and Imaging***

To establish a ground truth for the absorbed doses in each organ in the Four-Organ phantom (from objective 2) the doses from the activity distributions in the voxel phantoms were generated by MC modelling using different codes (GATE, GEANT, EGSnrc and MCNPX). The dosimetry results of the 4 MC codes agreed, showing the reliability of different codes in calculating dose. Self-dose S-Factors match within 2-3 %. These results were used to generate a ground truth of dose with which to compare with clinical dose calculations from experimental and simulated imaging.

***Cross comparison of dosimetry calculations with uncertainty analysis***

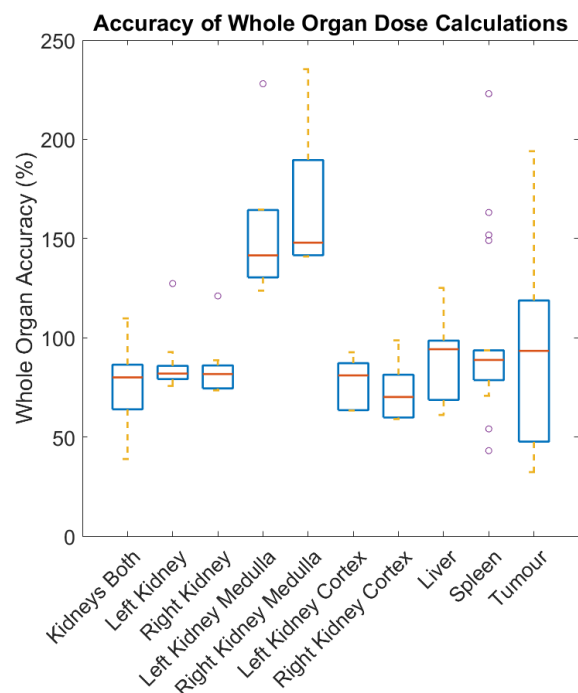
Although relative differences in the dosimetry calculations in a set of patient scans have been studied by various groups, including in this project, the comparisons have only shown relative differences of dose by each technique. This is essentially because the comparison has been made on an unknown distribution of activity (i.e. the patient itself).

Therefore, it was the aim of this project to try and create an acceptable system to test the accuracy and uncertainty of clinical dosimetry systems through a series of measurements that had traceability to a ground truth dose and that allowed the propagation of uncertainties throughout. The experimental and simulated Four-Organ phantom imaging (Objective 2) was combined with dose calculations to produce a unique dataset for validating dosimetry calculations systems.

A Standard Operating Procedure (SOP) was created to guide clinical users through the series of operations of the dosimetry chain using their own platform to determine the absorbed doses from the phantom (and MC) images. All factors needed for using the reconstructed images such as the calibration factor and partial volume recovery coefficients were given along in the SOP with their uncertainties.

Further to the SOP, a reporting spreadsheet was developed to record the results of the key steps of the dosimetry chain and to establish estimates of uncertainties for each step. The reporting spreadsheet allows the user to enter information about choice of technique used for each step of the calculation such as the method of defining VOI or choice of fit for the TAC etc., and also key data such as volumes, activities, TAC fits and S factors used. The SOP and spreadsheet also importantly allowed users to determine the uncertainties associated with these measurements to then be able to estimate uncertainties on calculated dose values with emphasis placed on following the newly accepted EANM guidelines. An exemplar uncertainty analysis, following these guidelines, was also provided to guide users.

Clinical groups within the project and also commercial software companies were invited to take part in a cross-comparison exercise to compare the accuracy and uncertainty of different dosimetry techniques. A total of 12 groups using commercial and non-commercial systems participated in the comparison exercise. The results of the comparison were compared in terms of both accuracy and the chain of uncertainties and a report has been produced which highlights the key areas where accuracy and uncertainty could be improved.



**Figure 26: Plots of whole organ dose comparisons with spread and uncertainties from the comparison exercise**

The key areas that showed the major influence on the accuracy were; VOI determination (particularly for small volumes), activity recovery in regions with a high background and how these influence the overall accuracy and uncertainty. Another key area is the use of TAC analysis in voxel-based dosimetry, where uncertainties are very high.

An important result from the exercise was that none of the commercial systems and only some of the clinical groups were able to determine a full uncertainty analysis (as per previous other comparisons). This has highlighted two important areas of note.

1. By following a specified system of uncertainty analysis, the dosimetry results can be shown to have general uncertainties ranging from a few percent to 40 % depending on the organ/tumour being analysed. This is the first time a range of dosimetry systems has been able to be compared with uncertainties analysis and gives valuable information as the advantages of the techniques used.
2. Only the commercial systems have Federal Drug Administration (FDA) approval for clinical work and yet these do not have the facilities for giving uncertainties on their key operating stages. Therefore this is an important area that needs addressing.

**Table 4: Comparison of uncertainties in complete absorbed dose calculation chain by technique**

Organ	Local Energy Deposition	MIRD Schema	MC
Both Kidneys	9.97%	11.58%	-
Left Kidney	6.86%	7.72%	5.50%
Right Kidney	7.23%	8.77%	5.96%
Liver	4.58%	5.00%	4.35%
Spleen	7.48%	7.74%	6.00%
Tumour	37.59%	39.03%	34.88%

The procedure for the comparison exercise has been produced as a protocol for commissioning a dosimetry platform to allow the users to use the datasets and ground truth to determine both the accuracy and uncertainties of the techniques used for their dosimetry. In addition, this system can be used as a training set for performing dosimetry on images with known dose and identifying sources of inaccuracy. Users can

compare the performance of their system when changing techniques, in turn leading to a better outcome of MRT Dosimetry. However, if this technique is to be used to help dosimetry for different therapies then more datasets are needed for other isotopes, pharmacokinetic models, disease groups, phantoms and especially tumours.

### **Summary of Outcomes**

The first key achievement for this objective is the acceptance of the procedure to produce an uncertainty budget for MRT dosimetry by EANM. This is a key document to be embedded in future dosimetry systems and lays the foundations for a method of uncertainty comparison and improvement (See Section 6 Publications). A further significant achievement is the implementation of a “ground truth” exercise to establish uncertainties and accuracy for a given clinical dosimetry system. This is the first comparison using “known dose” between centres and with commercial partners. These results highlight key areas to be addressed to play a significant role in improving MRT dosimetry.

Following these achievements, is the preparation of an SOP for commissioning of a dosimetry platform, from camera calibration to the final dose estimate, where uncertainty budget is considered at every step. The SOP is designed to allow users to determine the accuracy and uncertainties within their dosimetry technique and compare the output of the clinical system with the ground truth. The SOP will be submitted to bodies at IEAE, EANM and EFOMP. It is intended that this will engage users in employing full uncertainty analysis in commissioning a dose platform and embed the reporting of uncertainties as recommended in EANM guidelines.

Finally, it must be remarked that this work was also limited by the problems found in other objectives- namely the extreme difficulty in providing DICOM formats to suit all commercial and academic dosimetry systems. However, the most concerning outcome is that although it has been demonstrated that an uncertainty budget can be generated, no commercial dosimetry systems have the capability to do so, along with many academic systems. This should be addressed as a matter of urgency. and if the IAEA accepts the project's SOP for commissioning of a dosimetry system then this should also be raised with the FDA and other licencing agencies.

## 5 Impact

The project output has been disseminated through journal publications and presentations at a range of national and international meetings. Highlights of this engagement included; an invited presentation at the European Association of Nuclear Medicine (EANM) congress 2017 – the world’s leading nuclear medicine meeting with over 6000 delegates, and a recent EANM Guideline on “Practical guidance on uncertainty analysis for molecular radiotherapy absorbed dose calculations”. In addition, the project has hosted four very well attended European workshops covering, “Quantitative Imaging for molecular radiotherapy – metrology for clinical practice”, “The principles and clinical implementation of dose calculation in molecular radiotherapy” and “Molecular Radiotherapy Dosimetry: Clinical Implementation of Personalised Dosimetry in Molecular Radiotherapy”. A final workshop presenting the output of the project was attended by 80+ representatives from clinical centres, academia, industry and international standards bodies. Further details and presentations from these workshops can be found on the project website [http://mrt-dosimetry-empir.eu/?page\\_id=210](http://mrt-dosimetry-empir.eu/?page_id=210).

### *Impact on industrial and other user communities*

The key to the long-term success of this project will be in ensuring that the methods and tools developed are taken up by the European MRT community. In addition to the 12 clinical consortium members the project had 22 collaborators representing commercial, clinical and academic organisations.

There are many tangible outputs from this project which can be of direct use to the MRT community, including not only MRT clinics, but manufacturers of software, imaging equipment, and radiopharmaceuticals. The establishment of a primary standard for  $^{166}\text{Ho}$  (Objective 1) will allow Quirem Medical (a manufacturer of  $^{166}\text{Ho}$  microspheres) to standardise activity measurements for their product across multiple sites in different countries. Over the course of the project Quirem Medical have increased their engagement with the metrology community, becoming a collaborator in the project and presenting at the final project workshop.

A major route for maximising clinical uptake of the output from this project is engagement with SPECT/CT camera manufacturers. Representatives from all three major commercial SPECT camera vendors (GE Healthcare, Siemens Healthineers and Mediso Medical Imaging Systems) attended the public workshops held by the project. Mediso Medical Imaging Systems is also a project collaborator.

The calibration protocol from objective 2 and commissioning guidance from objective 5 has recently been incorporated into prototype clinical nuclear medicine dosimetry software of a commercial vendor Mirada Medical Ltd who collaborated on the project. The software has been publicly demonstrated at the British Nuclear Medicine Society (BNMS) meeting in 2019.

The designs for the 3D printed quasi-realistic anthropomorphic phantoms produced by the project (Objective 2) will be made available to end-users (via scientific publications), as will the open-access database of test images (Objective 3), which is intended to be readable by contemporary clinical camera systems and accessible to most commercial software packages, so that it can be used for commissioning dosimetry systems. The project’s protocols for the clinical use of these phantoms and the database of test images for the commissioning of quantitative SPECT imaging and dosimetry calculation platforms were also presented to the MRT dosimetry community at the project’s final workshop in [Teddington, UK in May 2019](#) and are currently being developed into eLearning teaching modules.

### *Impact on the metrology and scientific communities*

The project’s continuation of the development of a primary standard of absorbed dose to water from a radionuclide solution (Objective 4) is an essential part of establishing traceability of MRT dosimetry to primary standards. This is the only primary standard of this type worldwide and is of considerable interest to the ionising radiation metrology community.

The project has also produced new nuclear data to enable more accurate PET imaging of  $^{90}\text{Y}$  and new activity standards and measurements of the decay scheme and half-life for  $^{166}\text{Ho}$  (Objective 1). This data has been submitted to peer-reviewed scientific journals, thus making it available to the radionuclide metrology, nuclear structure and the nuclear medicine communities. Nuclear data published in this way will consequently be evaluated by the Decay Data Evaluation Project (DDEP) for inclusion in their internationally-recognised database of evaluated nuclear data.

#### *Impact on relevant standards*

The project has engaged with a variety of standards bodies throughout the project including the:

**BIPM (Bureau International des Poids et Mesures):** where primary standards activity measurements of  $^{166}\text{Ho}$  made during this project have been submitted to BIPM under the EURAMET.RI(II)-K2.Ho-166 comparison. This comparison exercise will demonstrate equivalence between primary standard measurements at 4 NMIs for this important emerging medical radionuclide.

**EANM (European Association of Nuclear Medicine):** A key achievement of the project is the adoption and publication of its practical guidance on uncertainty analysis for MRT dosimetry by the EANM as an EANM guideline (objective 5). Furthermore, reports developed in this project on “Recommendations for the Best Imaging and Non-Imaging Methods for Obtaining Cumulated Activity from a Time-Activity-Curve (TAC), and the Associated Uncertainties” (objective 4) and a “Procedure for Commissioning a Clinical Dosimetry Platform” (objective 5) have recently been submitted to the EANM Dosimetry committee. In response the EANM has scheduled a special Pre-Congress Symposium at the EANM Annual Congress 2019 dedicated to “European Projects for Clinical Implementation of Dosimetry in Molecular Radiotherapy” to present the protocols developed during this project.

**IAEA (International Atomic Energy Agency):** A Working Group has been set up by the Dosimetry and Medical Radiation Physics section of the Division of Human Health at the IAEA, to produce a publication on MRT dosimetry in the Human Health Series and this Working Group includes 4 of the project partners. The IAEA is considered to be the main international authority on medical radiation dosimetry protocols and has also agreed in principle to host the open-access database of test images produced by the project and to publish the SPECT calibration and dosimetry commissioning protocols developed in this project (in objectives 3 and 5). The IAEA have invited the project to present its final results on the development of a primary standard for absorbed dose to water from a radionuclide at the International Symposium on Standards, Applications and Quality Assurance in Medical Radiation Dosimetry (IDOS) to be held in Vienna in June 2019. The procedure for commissioning a clinical dosimetry platform (objective 5) has also been submitted to the IAEA.

**MIRD (Committee on Medical Internal Radiation Dose):** The project's results have been cited in and influenced the upcoming update to the MIRD primer of absorbed dose calculations. The schema set out in previous versions of this MIRD publication has been adopted as a de facto standard for clinical absorbed dose calculations. In addition, the project's recommendations for the best imaging and non-imaging methods for obtaining cumulated activity from a TAC (objective 4), and the associated uncertainties have also been submitted to the MIRD.

Members of the consortium are also members of several ISO committees such as (ISO/TC 85 and ISO/TC 69) and have liaised with the European Federation of Organisations in Medical Physics (EFOMP) to provide information on the project's results to them. Finally, the upcoming ICRP (International Commission on Radiological Protection) publication “ICRP 140: Radiological protection in therapy with radiopharmaceuticals” has also cited and been influenced by this project's results.

#### *Longer-term economic, social and environmental impacts*

The overall goal of this project was to encourage and assist European MRT clinics, as well as those worldwide, to adopt dosimetry as a routine part of patient treatment. To this end the project has helped support compliance with the EC Directive 2013/59/EURATOM primarily through the production of guidance which can be used in the implementation of the directive within member states and through the educational workshops delivered by the project. This work will provide a significant step in bringing MRT into line with other radiotherapy modalities. Adoption of the techniques and protocols developed during the project should also allow clinical sites to use MRT with greater confidence and help predict patient treatment outcomes.

Clinical trials play a major role in the development of standardised dosimetry (including MRT). Absorbed dose is a critical parameter in both treatment effectiveness and harmful side-effects, therefore a reduction in the uncertainty of absorbed dose determination will give a greater statistical power to clinical trials. In turn, this should support the incorporation of standardised dosimetry into clinical trials and hence lead to the adoption of MRT in routine clinical treatment. The final results from this project are also being disseminated to national radiotherapy clinical trials quality assurance organisations, such as IAEA, EANM and MIRD, an important step towards widespread adoption.

The long-term results of the MRTDosimetry project will be a significant contribution to delivering more effective, better targeted treatments, improved outcome for the patients receiving them, and savings to national and European health systems providing this care.

## 6 List of publications

- [1]. Billas, I., Shipley, D., Galer, S., Bass, G., Sander, T., Fenwick, A., & Smyth, V. (2016). "Development of a primary standard for absorbed dose from unsealed radionuclide solutions". *Metrologia*, 53(6), 1259–1271. <http://doi.org/10.1088/0026-1394/53/6/1259>;
- [2]. "Reply to Comment on "Development of a primary standard for absorbed dose from unsealed radionuclide solutions". *Metrologia*, 54, 615–616. <https://doi.org/10.1088/1681-7575/aa78ff>
- [3]. Villoing, D., Marcatili, S., Garcia, M.-P., & Bardiès, M. (2017). "Internal dosimetry with the Monte Carlo code GATE: validation using the ICRP/ICRU female reference computational model". *Physics in Medicine and Biology*, 62(5), 1885–1904. <http://doi.org/10.1088/1361-6560/62/5/1885>
- [4]. D'Arienzo, M. and Cox, M. (2017). "Uncertainty Analysis in the Calibration of an Emission Tomography System for Quantitative Imaging". *Computational and Mathematical Methods in Medicine*, Volume 2017, Article ID 9830386, <https://doi.org/10.1155/2017/9830386>
- [5]. Solc, J., Vrba, T., & Burianova, L. (2018). "Tissue-equivalence of 3D-printed plastics for medical phantoms in radiology". *Journal of Instrumentation*, 13(09), P09018–P09018. <https://doi.org/10.1088/1748-0221/13/09/P09018>
- [6]. Gear, J. I. *et al*, (2018). "EANM practical guidance on uncertainty analysis for molecular radiotherapy absorbed dose calculations". *Eur. J. Nucl. Med. Mol. Imaging*, 1–12. <https://doi.org/10.1007/s00259-018-4136-7>
- [7]. Tran-Gia, J and Lassmann, M., (2019). "Characterization of Noise and Resolution for Quantitative <sup>177</sup>Lu SPECT/CT with xSPECT Quant". *J. Nucl. Med.*, 80(1), 50-59. <https://doi.org/10.2967/jnumed.118.211094>