



Quantitative assessment of radionuclide production yields in in-beam and offline PET measurements at different proton irradiation facilities

Julia Bauer, Meret Hildebrandt, Michael Baumgartl, Fine Fiedler, Charlotte Robert, Irène Buvat, Wolfgang Enghardt, Katia Parodi

► To cite this version:

Julia Bauer, Meret Hildebrandt, Michael Baumgartl, Fine Fiedler, Charlotte Robert, et al.. Quantitative assessment of radionuclide production yields in in-beam and offline PET measurements at different proton irradiation facilities. *Physics in Medicine and Biology*, 2022, 67 (15), pp.155001. 10.1088/1361-6560/ac7a89 . inserm-03872932

HAL Id: inserm-03872932

<https://inserm.hal.science/inserm-03872932>

Submitted on 26 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Quantitative assessment of radionuclide production yields in in-beam and offline PET measurements at different proton irradiation facilities

Julia Bauer^{1,2}, Meret Hildebrandt^{2,3,4*}, Michael Baumgartl^{4,5**}, Fine Fiedler⁶, Charlotte Robert⁷, Irène Buvat⁸, Wolfgang Enghardt^{6,9} and Katia Parodi^{2,4***}

¹ Department of Radiation Oncology, Heidelberg University Hospital, Heidelberg, Germany

² Heidelberg Ion Beam Therapy Center; Heidelberg University Hospital, Heidelberg, Germany

³ German Cancer Research center, Heidelberg, Germany

⁴ Ludwig Maximilian University Munich, Chair of Experimental Physics - Medical Physics, Munich, Germany

⁵ Department of Radiation Oncology, University Hospital Zurich, Zurich, Switzerland

⁶ Helmholtz-Zentrum Dresden-Rossendorf., Institute of Radiation Physics, Dresden, Germany

⁷ Université Paris-Saclay, Institut Gustave Roussy, Inserm, Radiothérapie Moléculaire et Innovation Thérapeutique, 94800, Villejuif, France

⁸ Laboratoire d'Imagerie Translationnelle en Oncologie, U1288 Inserm, Institut Curie, PSL, Université Paris Saclay, France

⁹ OncoRay—National Centre for Radiation Research in Oncology, Technische Universität Dresden, Dresden, Germany

* contributed during the employment at 2 and 4, now with 3

** contributed during the employment at 4, now with 5

*** now with 4

Contact: julia.bauer@med.uni-heidelberg.de

Running title: Imaging of PET radionuclides induced by proton beams

Keywords: offline PET, in-beam PET, particle therapy, in-vivo range verification, positron emitter production yield

Objective: Reliable radionuclide production-yield data are a prerequisite for positron-emission-tomography (PET) based in-vivo proton treatment verification. In this context, activation data acquired at two different treatment facilities with different imaging systems were analyzed to provide experimentally determined radionuclide yields in thick targets and were compared with each other to investigate the impact of the respective imaging technique.

Approach: Homogeneous thick targets made of three different materials (PMMA, gelatine, and graphite) were irradiated with mono-energetic proton pencil beams at two distinct energies. Material activation was measured (i) *in-beam* during and after beam delivery with a double-head prototype PET camera and (ii) *offline* shortly after beam delivery with a commercial full-ring PET/CT scanner. Integral as well as depth-resolved β^+ - emitter yields were calculated from the PET data for the dominant positron-emitting radionuclides ^{11}C , ^{15}O , ^{13}N and (*in-beam* only) ^{10}C . Moreover, the *in-beam* data were used to investigate the qualitative impact of the time scheme of common monitoring workflows on activity depth profiles and their quantitative impact on count rates and total activity.

Main results: Production yields measured with the *in-beam* camera were comparable to or higher compared to respective *offline* results, confirming a quantification bias introduced by scattered coincidences in the *in-beam* data analysis, as recently shown by an in-silico study. Depth profiles of radionuclide-specific yields obtained from the *double-head* camera showed qualitative differences to data acquired with the *full-ring* camera with a more convex profile shape. Considerable impact of the imaging timing scheme on the activity profile was observed for gelatine only with a range variation of up to 3.5 mm. Evaluation of the coincidence rate and the total number of observed events in the considered workflows confirmed a strongly decreasing rate in targets with a large oxygen fraction. Thus, time delays have to be compensated by considerably prolonged acquisition times in order to reach comparable coincidence statistics.

Significance: The observed quantitative and qualitative differences between the datasets underline the importance of a thorough system commissioning. Due to the lack of reliable cross-section data, in-house phantom measurements are still considered a gold standard for careful characterization of the system response and to ensure a reliable beam range verification.

1. Introduction

Ion beam therapy, in particular proton therapy, has nowadays become a recognized treatment modality complementing conventional external photon irradiation for the treatment of various cancer diseases. Due to the characteristic inverted depth dose profile of ions stopping in the irradiated tissue, the beam range needs to be well controlled in the patient for a safe delivery of highly conformal dose distributions to complex target volumes. Range uncertainties can be minimized either by using high-energetic transmission beams to accurately determine the particle stopping power of the tissue (Johnson, 2018), or by secondary radiation detection to reconstruct the range of the primary beam particles (Knopf and Lomax, 2013). Based on the latter concept, two different techniques for in-vivo beam range monitoring have so far been studied clinically: (i) prompt gamma imaging (PGI) (Richter et al., 2016; Xie et al., 2017) and (ii) positron-emission-tomography (PET) imaging (Ferrero et al., 2018; Min et al., 2013; Miyatake et al., 2010; Nischwitz et al., 2015; Nishio et al., 2010; Parodi et al., 2007b). While PGI operates during beam delivery, the PET-based approach can be implemented in different schemes (Shakirin et al., 2011): for in-beam monitoring the PET camera is installed directly at the treatment place and detects the decay of irradiation induced PET-radionuclides during patient irradiation; in-room installations operate a PET-scanner in the treatment room but not integrated in the beam nozzle; offline installations use a nearby PET-scanner located outside the treatment room. In-beam installations are superior concerning the signal rate, thus requiring shorter acquisition time. However, the high demands of geometric compatibility of the camera with the (often rotating) beam nozzle have so far only been met by the use of limited-angle double-head cameras, with considerably affected imaging quality compared to full-ring PET cameras (Crespo et al., 2006). Offline installations, on the other hand, usually operate commercial full-ring PET-scanners with state-of-the-art imaging performance. However, the signal rate is low and spatial information is blurred due to physical decay and biological wash-out of the patient activation during the time delay caused by patient transport and repositioning after beam delivery.

Besides imaging performance and signal quality aspects, a reliable range verification is furthermore challenged by uncertainties in the production cross-section data of relevant radionuclides (Seravalli et al., 2012). These data are required to predict the expected patient activation and thus to assess the observed beam range. An experimental determination of radionuclide-specific production yields is therefore an important contribution to increase the accuracy of the verification strategy.

Treatment centers implementing a PET-based verification concept usually conduct various pre-clinical activity measurements in phantoms for system commissioning (Bauer et al., 2013a; Bisogni et al., 2016; Horst et al., 2019; Meißner et al., 2019; Parodi et al., 2007a; Shao et al., 2014; Zhu et al., 2011). An inter-facility comparison of the obtained production yield data, however, has so far been

1 impeded by differences in the respective irradiation setup. To overcome this issue, an experimental
2 campaign was conducted at the Heidelberg Ion-Beam Therapy Center (HIT, Heidelberg, Germany)
3 under irradiation conditions comparable to a previously conducted phantom activation
4 measurement performed at GSI (Helmholtzzentrum für Schwerionenforschung, Darmstadt,
5 Germany). At both facilities activation data were acquired in different homogeneous block-shaped
6 thick targets with different PET imaging concepts: an in-beam PET camera installed at GSI (Enghardt
7 et al., 2004) and an offline PET/CT scanner installed at HIT (Bauer et al., 2013b). With these data sets
8 we were able to:

- 9 (i) independently assess radionuclide production yield data obtained at different facilities
- 10 (ii) study the qualitative and quantitative impact of the respective imaging technique

11 Both aspects are considered relevant for the commissioning of future PET-monitoring installations
12 and workflow implementations relying on activity predictions for beam range evaluation. While
13 activation data obtained at the two irradiation facilities HIT and GSI for ions heavier than protons
14 have already been published separately (Bauer et al., 2019; Fiedler, 2008; Priegnitz et al., 2012;
15 Sommerer et al., 2009), this study reports the resulting integral and depth-resolved radionuclide-
16 specific production yields induced by therapeutic proton beams and, for the first time, an inter-
17 facility comparison of these datasets.

19 2. Materials and Methods

21 2.1. Phantom description

22 Table 1 provides an overview of relevant PET radionuclides (^{11}C , ^{10}C , ^{13}N , ^{15}O) produced in therapeutic
23 human tissue irradiation, the dominant production channels, and their respective half-life. Phantoms
24 made of gelatine (GEL) and graphite (GR) were used to study the respective production channels
25 separately, while the used PMMA (polymethyl methacrylate, $\text{C}_5\text{H}_8\text{O}_2$) phantoms are considered a
26 good surrogate for soft human tissue. Similar rectangular shaped homogeneous phantoms were used
27 at both centers, GSI and HIT. They were block-shaped with dimensions between 15 cm to 35 cm in
28 length and (5-10) cm in width and height. Geometrical details of the phantoms as well as the
29 respective photon attenuation coefficients required for the GSI image reconstruction pipeline are
30 reported in Table 2.

31 Gelatine phantoms at GSI were put on an acrylic glass plate of (2-5) mm thickness, respectively, for
32 irradiation and imaging to cope with their unstable structure. The plate was considered for activity
33 calculation. The gelatine phantoms at HIT were prepared in a dedicated PMMA box of 1 cm wall

thickness and kept in there to ensure its stability during transportation to and repositioning at the offline PET/CT scanner after irradiation.

Table 1: Production channels for the main radionuclides produced in graphite, PMMA and gelatine, corresponding production threshold energies E_{thr} and half-lives $T_{1/2}$ (Beebe-Wang et al., 2013).

Radionuclide	Production channel	E_{thr} [MeV]	$T_{1/2}$ [s]
^{11}C	$^{12}\text{C}(p,pn)^{11}\text{C}$	20.61	1,223
^{11}C	$^{16}\text{O}(p,3p3n)^{11}\text{C}$	27.50	1,223
^{15}O	$^{16}\text{O}(p,pn)^{15}\text{O}$	16.79	122
^{13}N	$^{16}\text{O}(p,2p2n)^{13}\text{N}$	5.66	568
^{10}C	$^{12}\text{C}(p,p2n)^{10}\text{C}$	34.5	19

Table 2: Overview of the homogeneous phantoms used at GSI (upper section) and HIT (lower section). Reported are the geometrical size (width W , height H and length L) and, for the GSI phantoms, the material-specific photon attenuation coefficients μ (at a photon energy of 511 keV) used for activity reconstruction, and the experiment-specific correction factors C used for the quantitative production yield calculation together with their uncertainty σ . Attenuation coefficients of the HIT phantoms were automatically determined during image reconstruction from a CT scan for the polymethyl methacrylate (PMMA) and the gelatine (GEL) phantom, and manually adjusted ('+corr', cf section 2.5) for the graphite phantom (GR).

Experiment ID	$W \times H \times L$ [cm]	μ [1/cm]	$C \cdot 10^3 / \sigma$
GR-GSI-E1	9 x 9 x 15	0.156	4.6 / 4%
GEL-GSI-E1-1	9 x 9 x 20	0.096	7.0 / 3%
GEL-GSI-E1-2	9 x 9 x 20	0.096	6.8 / 3%
PMMA-GSI-E1	9 x 9 x 20	0.111	6.6 / 3%
GR-GSI-E2	9 x 9 x 31	0.142	4.1 / 5%
GEL-GSI-E2	9 x 9 x 30	0.096	5.8 / 3%
PMMA-GSI-E2	9 x 9 x 30	0.111	5.4 / 4%
GR-HIT-E1/-E2	7 x 7 x 20	CT-based+corr	-
GEL-HIT-E1/-E2	10 x 5 x 35	CT-based	-
PMMA-HIT-E1/-E2	9 x 9 x 30	CT-based	-

2.2. Proton irradiation setup

At GSI, the phantoms were positioned at the experimental treatment place on the treatment table with the long side aligned to the beam direction for irradiation. Mono-energetic pencil-like proton beams were extracted from the synchrotron at a lower energy of 125.4 MeV/u (E1) in 100 spills and a higher energy of 176.95 MeV/u (E2) in 120 spills. The resulting irradiation times (t_{irr}) are reported in Table 3 together with the number of delivered protons (N_p) and the calculated intensity of the beam (I). The time structure of the beam delivery was reasonably stable with mean spill and pause times of $\bar{t}_s = 1.3$ s and $\bar{t}_p = 3.0$ s, respectively, with variations below 11%. At the time of experiments, a 4 mm thick protective PMMA plate, which corresponds to a range shift of 4.6 mm water-equivalent path length (WEPL), had accidentally been forgotten in the GSI beam line during all phantom

irradiation experiments. Thus, the energy of the primaries reaching the phantoms was slightly reduced and the production yields were expected to be lower. On the other hand, the WEPL of the GSI beam nozzle is about 1mm smaller as compared to the HIT beam nozzle. The impact of these two factors on the expected production yield was estimated with a Monte Carlo simulation using the FLUKA-based framework of (Bauer et al., 2014).

Phantom irradiation at HIT was performed under comparable experimental conditions (Bauer et al., 2013a). The phantoms were irradiated in the experimental room and positioned with the long side aligned to the beam direction such that the geometrical center of the phantoms coincided with the isocenter of the experimental room. Mono-energetic pencil-like proton beams of two energies, 125.67 MeV/u (E1) and 176.75 MeV/u (E2), respectively, were delivered over a time period of about 3 minutes by the synchrotron. Since the energy difference of about 0.2 MeV/u between the GSI and the HIT experiments is considered negligible, the labels E_1 and E_2 are in the following used for both experimental series. Beam delivery at HIT was stable within negligible deviations: the spill length and pause times between the spills were on average $\bar{t}_s = 1.7$ s and $\bar{t}_p = 4.3$ s with variations of below 5% and 2%, respectively (Bauer et al., 2013a).

Table 3: Parameters of different phantom experiments: experiment ID, proton beam energy E , total number of delivered particles N_p , average beam intensity I , and duration of irradiation t_{irr} , as well as the time delay Δt between the end of irradiation and start of PET data acquisition for the HIT experiments and (*) the time delay used to mimic an “offline” scenario for the GSI experiments (cf. section 2.4.5), and the PET acquisition frame time t_{decay} .

Exp ID	E [MeV/u]	N_p [$\cdot 10^{11}$]	I [10^8 /s]	t_{irr} [min:s]	Δt [min:s]	t_{decay} [min]
GR-GSI-E1	125.4	2.50	16.3	10:00	02:40*	30
GEL-GSI-E1-1	125.4	1.05	6.6	10:00	03:03*	35
GEL-GSI-E1-2	125.4	0.99	4.9	08:20	03:40*	25
PMMA-GSI-E1	125.4	1.16	8.5	10:00	01:54*	30
GR-GSI-E2	176.95	1.30	13.2	08:22	02:40*	30
GEL-GSI-E2	176.95	1.41	76.6	08:24	02:03*	30
PMMA-GSI-E2	176.95	1.46	13.8	08:19	02:04*	30
GR-HIT-E1	125.67	1.12	32.0	03:20	04:02	30
GEL-HIT-E1	125.67	1.12	32.0	04:35	04:35	30
PMMA-HIT-E1	125.67	1.12	32.0	03:43	03:43	30
GR-HIT-E2	176.75	1.12	32.0	04:11	04:11	30
GEL-HIT-E2	176.75	1.12	32.0	03:42	03:42	30
PMMA-HIT-E2	176.75	1.12	32.0	03:48	03:48	30

2.3. PET data acquisition

Imaging concepts as well as workflow implementation were fundamentally different at the two facilities, as sketched in Figure 1. At GSI, a double-head PET camera system was directly installed at the treatment place. This setup allowed for PET data acquisition during irradiation (in-beam) and after beam delivery (offline) without the need for transportation or repositioning of the phantoms. Each camera head consisted of 4 x 8 position-sensitive scintillation block detectors made of bismuth germanate (BGO). The centre of the PET camera field-of-view (FOV), where the imaging performance is best, coincided with the isocentre of the treatment room. Therefore, the phantoms were positioned such that the area of the expected distal activity fall-off, thus the activity maximum, was located at the isocenter (cf. Figure 1 left). Correct positioning was checked by means of a point sourced mounted on a reference point on the phantom prior to the irradiation (Parodi et al., 2002). A synchronization of the PET data acquisition with the beam control system, mapping the time course of irradiation and the acquired PET data, was used to suppress prompt gamma background radiation. PET data acquired in-spill could thus be discarded later on for the analysis. In order to study the activity decay, the in-beam PET acquisition was prolonged for 25 to 35 minutes after beam delivery.

At HIT, the irradiated phantoms were immediately transported to a nearby installed commercial full-ring PET/CT scanner (Siemens Biograph mCT[®]) based on lutetium oxyorthosilicate (LSO)-crystal technology. PET measurements began with a typical time delay Δt of 3 to 4 minutes after beam delivery (cf. Table 3), needed for transport and repositioning of the phantom. At the PET/CT, the phantoms were positioned with the long side transversal to the scanner axis, centered in the FOV using the internal laser alignment system and the indexed patient support table. On the measurement day, the PET/CT scanner was calibrated according to the standard procedure of the daily quality assurance to ensure the quantification reproducibility at clinically required level. The examination protocol was set up such that the 30 min PET acquisition directly followed the initial topogram scan, while the CT scan necessary to derive the attenuation and scatter correction of the PET data was performed at the end of the examination. In order to eliminate the scanner applied correction for the decay of the radionuclide as required for diagnostic PET applications, the administered radionuclide was set to ²²Na with a half-life of about 2.6 years, considering the β^+ branching ratio of this radionuclide of 90% later on in the data analysis.

Time structure information relevant for the PET imaging of both experimental series is reported in Table 3. At both facilities, PET data were acquired in “list mode”, imposing a time stamp every millisecond (every 10 ms) on the coincidences record acquired at HIT (GSI), which allows for a retrospective reconstruction of counts and PET images in arbitrary time frames.

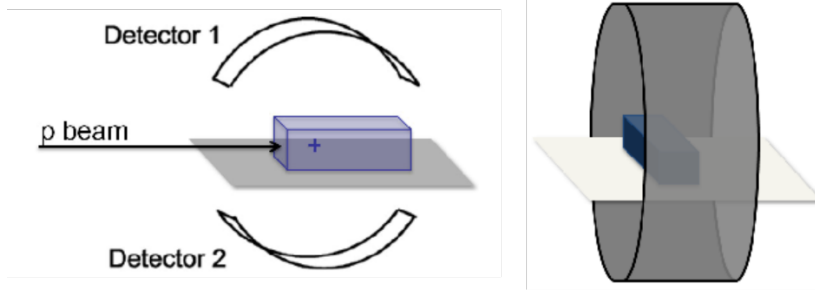


Figure 1: Schematic drawing of the irradiation place and the in-beam PET measurement at GSI (left) and of the offline PET/CT scanner at HIT (right).

2.4. Data analysis of the GSI experiment

The GSI data analysis pipeline has in general been adopted from (Parodi et al., 2002). In order to determine the absolute production yield, the signal counts detected by the GSI PET camera have to be reconstructed in several subsequent steps. First, the activity decay measured after beam delivery is used to extrapolate the radionuclide-specific activity present at the end of the irradiation. From this point, it is possible to calculate the radionuclide-specific production rate per second and eventually the radionuclide production yield per primary incoming proton considering the relevant irradiation parameters. Since the GSI analysis chain is based on raw PET camera data, the obtained production yields are relative numbers. Additional corrections for detector efficiency, annihilation photon attenuation and geometrical acceptance due to the solid-angle camera field-of-view need to be considered to obtain absolute production yields.

2.4.1. Fit of activity decay

The radionuclide-specific contribution to the total activity present at the end of irradiation, $A_{0,i}$ ($i = {}^{10}\text{C}, {}^{15}\text{O}, {}^{13}\text{N}$ and ${}^{11}\text{C}$) can in general be determined by a nonlinear fit of the measured activity decay according to the law of radioactive decay:

$$A(t) = \sum_i A_{0,i} \cdot e^{(-\lambda_i \cdot t)} \quad (1)$$

with λ_i being the half-life of the respective radionuclide. A customized software coded in IDL (Interactive Data Language, Version 5.3) was used for the regression analysis.

Various consistency checks were performed to ensure the reliability of the fitting procedure: (i) regression with fixed or variable radionuclide half-lives, (ii) regression using the full activity decay or only a selected time interval optimized for reasonable radionuclide abundances giving the best

results for the fitted half-lives compared to the expected half-lives of the radionuclides, and (iii) applying the regression code used for the HIT data (cf. section 2.5) as cross-check. The intermediate results of these checks have been documented in (Hildebrandt, 2012), the final activity values resulting from the investigation of Hildebrandt are summarized in Table 4 together with the deviations found for the cross-check. These A_0 values were used as input for the subsequent analysis pipeline (cf. equation 2).

Table 4: GSI radionuclide-resolved activity values at the end of irradiation A_0 [$10^3/s$] as obtained from the time analysis of the decay data for all considered phantoms and deviations Δ in [%] to the results from the fit cross-check.

Exp ID	$A_0 \cdot [10^3/s]$			
	$^{10}\text{C} / \Delta[\%]$	$^{15}\text{O} / \Delta[\%]$	$^{13}\text{N} / \Delta[\%]$	$^{11}\text{C} / \Delta[\%]$
GR-GSI-E1	4.2 / 11	-	-	29.0 / <1
GR-GSI-E2	5.4 / 5	-	-	21.6 / <1
GEL-GSI-E1-1	-	37.0 / <1	1.9 / <1	2.7 / <1
GEL-GSI-E1-2	-	42.0 / <1	2.3 / 4	2.5 / 2
GEL-GSI-E2	-	71.3 / 2	4.1 / 13	4.8 / 7
PMMA-GSI-E1	2.3 / 11	14.4 / <1	0.7 / 1	10.8 / <1
PMMA-GSI-E2	3.5 / 19	27.0 / <1	1.2 / 4	19.1 / <1

2.4.2. Determination of relative production rates

The radionuclide-specific production rate P_i can be calculated from the activity present at the end of irradiation, $A_{0,i}$ (cf. Table 4), considering the recorded time structure of the pulsed synchrotron irradiation, with beam delivery during spills (t_s) and pauses in between (t_p). Two different approaches have been pursued to calculate relative production rates P_i , analogous to (Parodi et al., 2002). The first approach assumes of a constant spill structure over N_s spills, thus using mean spill (T_s) and mean pause times (T_p), and a constant beam intensity. The production rate P_i is then given by:

$$P_i = \frac{A_{0,i}}{(1 - e^{-\lambda_i \cdot T_s}) \cdot \sum_{j=1}^{N_s} e^{-\lambda_i \cdot (j-1) \cdot (T_s + T_p)}} \quad (2)$$

A second approach considers the activity A_j measured at the end of each spill and the relative contribution f_i of the i^{th} radionuclide to the estimated amount of activity \mathcal{A}_j produced in the j^{th} spill, thus allowing energy and time variations during irradiation to be considered. The production rate is then calculated as follows:

$$P_{i,j} = \frac{\mathcal{A}_j \cdot f_i}{1 - e^{-\lambda_i \cdot t_{s,j}}} \quad (3)$$

with

$$\mathcal{A}_j = A_j - \sum_i \sum_{k=1}^{j-1} P_{i,k} \cdot (1 - e^{-\lambda_i t_{s,k}}) \cdot e^{-\lambda_i \sum_{m=1}^{j-1} (t_{s,m+k} + t_{p,m+k-1})} \quad (4)$$

1 The factor f_i is a weighting factor for the contribution of each considered radionuclide of the activity
 2 in each spill and is calculated as follows:

$$f_i = \frac{P_i \cdot (1 - e^{-\lambda_i t_s})}{\sum_i (P_i \cdot (1 - e^{-\lambda_i t_s}))} \quad (5)$$

3

4 2.4.3. Quantitative estimation of integral production yields

5 For an absolute estimation of the GSI production yields, spatial dependent correction factors for the
 6 material-specific γ -ray attenuation, the solid angle coverage relative to the centre of the camera as
 7 well as the measured absolute detector efficiency in the centre of the scanner have to be applied.
 8 Voxel-wise attenuation factor maps were created by a dedicated IDL routine for each phantom. The
 9 space around the phantom was filled with air with an attenuation factor of $\mu_{air} = 1.03677 \cdot$
 10 $10^{-4} [1/cm]$. Inside the phantom the respective attenuation factor of GR, GEL or PMMA was applied
 11 (cf. Table 2). For the border voxels containing both phantom material and air, the resulting
 12 attenuation factor had to be calculated proportionally. Based on these maps, attenuation corrections
 13 for each line of response (LOR) of the scanner were generated, using a home-made software (Pönisch
 14 et al., 2004). These attenuation corrections are inputs to the filtered backprojection image
 15 reconstruction in order to generate an attenuation corrected PET image, which can then be further
 16 corrected for the limited solid angle detection. An additional PET image was reconstructed using
 17 filtered backprojection without the attenuation and the solid angle correction. Based on these
 18 images the combined correction factor C was calculated similar to (Parodi et al., 2002) as:

$$C = \frac{\int A_u d\vec{r}}{\int A d\vec{r}} \cdot \alpha \quad (6)$$

19 where A_u is the uncorrected activity distribution, A the one corrected for attenuation and solid angle
 20 geometrical acceptance, both integrated over the whole volume where non-negligible activity is
 21 present. The detector efficiency α at the centre of the scanner was measured with a ^{22}Na point
 22 source of known activity as the ratio of true measured events over the number of radioactive decays,
 23 yielding $\alpha=0.022$.

24 The number of radionuclides produced by a single primary beam particle can then be calculated from
 25 the production rate P_i and the number of beam particles delivered per second I_s :

$$N_i = \frac{P_i}{C \cdot I_s} \quad (7)$$

2.4.4. Determination of depth dependent production rates

In order to obtain production rates that are spatially resolved along the irradiation axis (depth profiles), the phantoms were subdivided into several slices perpendicular to the beam direction, so-called regions of interest (ROIs). The full analysis pipeline (2.4.1.-2.4.3.) was then applied to these ROIs separately. The thickness of the ROIs varied in size between 3 mm to 50 mm. The choice represents a tradeoff between limited signal statistics and the desired resolution, which should be higher in regions of particular interest like the edge before the distal fall-off of the production yield. Therefore, smaller ROIs were chosen in this part, contrary to coarser ones in the entrance channel.

2.4.5. Time-dependent analysis of the reconstructed PET images

In-beam data acquired at GSI during and after irradiation were reconstructed in several time frames for hypothetical in-room and offline measurement scenarios. For the in-beam scenario only those data acquired during the irradiation pauses between the spills were considered, hence the time frame for this reconstruction began after the first and ended with the last spill. For the in-room scenario the activity was reconstructed for a time frame of 3 minutes, starting 2 minutes after the irradiation. The offline scenario started at the timepoint where the ratios of the dominating β^+ -emitting radionuclides matched the ratios of the offline HIT experiments and comprised the full PET acquisition time (cf. Table 3). It turned out that this timepoint was 2 to 3 minutes after completion of beam delivery, whereas realistic delays for patient examinations are in the order of 6 to 12 minutes. Therefore, an additional late-offline scenario being closer to clinical practice was defined starting 10 minutes after the end of irradiation.

2.5. Data analysis of the HIT experiment

Offline activity measurements at HIT were performed with a commercial PET/CT scanner, providing integrated state-of-the-art image reconstruction functionality. The analysis pipeline was therefore more straightforward, as reported in (Bauer et al., 2019, 2013a). For image reconstruction, the full 30 min acquisition time was split into 30 single frames of one minute each. Volumetric activity images were obtained from the scanner software, using an iterative ordered subset expectation maximization reconstruction (3 iterations, 24 subsets, voxel size: $2.036 \times 2.036 \times 3.0 \text{ mm}^3$) which automatically considered random, scatter and attenuation correction for activity quantification, and applying a Gaussian image smoothing with an isotropic kernel of $\sigma = 5 \text{ mm}$. These parameters were chosen according to the results of (Bauer et al., 2013a), who tested extensively the impact of several image reconstruction parameters on the measured activity signal. The CT-based attenuation factors

determined by the scanner software were checked and found to agree well with the literature value for PMMA and gelatine. The graphite attenuation factor, however, mismatched and was therefore corrected by a manipulation of the CT as reported in (Bauer et al., 2013a). Resulting PET images were exported from the scanner database for further external processing.

Data regression for the determination of the radionuclide-specific activity contribution $A_{0,i}^{PET}$ present at the beginning of each PET measurement was conducted with an in-house fitting program based on python and implemented in MeVisLab® (MeVisLab Medical Solutions AG and Fraunhofer MEVIS) analogous to equation (1):

$$A(\vec{r}, t) = \sum_i A_{0,i}^{PET}(\vec{r}) \cdot e^{-\lambda_i \cdot t} \quad (8)$$

with the spatially resolved $A_{0,i}^{PET}(\vec{r})$ as regression parameter, thus using tabulated half-lives for the considered radionuclides $i = {}^{11}\text{C}, {}^{13}\text{N}, {}^{15}\text{O}$. The activity values were already quantitative at this point, including corrections for random coincidences, photon scattering and attenuation as well as measured detection efficiency. Consideration of the experiment-specific transfer time Δt allowed to calculate the activity contribution of radionuclide species i at the end of beam delivery:

$$A_{0,i}(\vec{r}, t) = A_{0,i}^{PET}(\vec{r}) \cdot (e^{-\lambda_i \cdot \Delta t})^{-1} \quad (9)$$

Absolute production yield maps were determined according to the GSI analysis pipeline part 2.4.2 using the spatially resolved $A_{0,i}$ images as input and voxel-level application of equation (7) with $C=1$. Integral production yields were obtained from an integration of the full image, depth resolved distributions from lateral integration of the activity for each voxel position in depth.

3. Results

3.1. Comparison of integral production yields

Integral production yields obtained from the PET activation measurements at HIT and GSI together with their uncertainties are reported in Table 5. Uncertainties on the HIT total yields are conservatively estimated as 10% for all radionuclides, considering uncertainties in the calibration of the beam monitor system, thus the number of delivered primaries, the choice of reconstruction time points, the accuracy of the manual time delay measurement and fit uncertainties (Bauer et al., 2019, 2013a).

Uncertainties on the GSI yields were explicitly estimated for each radionuclide and experiment by quadratically summarizing uncertainties from different regression methods (cf. Table 4), variations in the time structure of the irradiation and the number of delivered protons (2%-40%, depending on

the experiment), and accuracy of the derived correction factor C (cf. Table 2). Regarding the impact of the different beam nozzle WEPLs and the PMMA plate in the GSI beam line, the conducted MC simulations predicted a small systematic effect of ~3% higher yields expected at HIT.

Ratios were calculated as the GSI yields divided by the corresponding HIT yields, with uncertainties obtained from maximum error estimate. GSI yields were found to be similar to or higher as the HIT yields, with the strength of the deviation depending on the phantom material.

Table 5: Production yields \widehat{N}_i with $i = {}^{11}\text{C}, {}^{15}\text{O}, {}^{10}\text{C}, {}^{13}\text{N}$ obtained from GSI and HIT data and their respective sum \widehat{N}_{all} .

Ratios are calculated as GSI yields divided by HIT yields. All yields are provided in units of $[\cdot 10^{-6} \text{ per primary proton}]$.

Uncertainties σ on the yields are provided in percentage, uncertainties on the ratio in absolute numbers.

Experiment ID	$\widehat{N}_{11\text{C}} / \sigma$	$\widehat{N}_{15\text{O}} / \sigma$	$\widehat{N}_{10\text{C}} / \sigma$	$\widehat{N}_{13\text{N}} / \sigma$	$\widehat{N}_{all} / \sigma$
GR-GSI-E1	42,057 / 39%	-	1,590 / 41%	-	43,647 / 41%
GR-HIT-E1	36,435 / 10%	-	-	-	36,435 / 10%
Ratio	1.1 ± 0.46	-	-	-	1.2 ± 0.51
GR-GSI-E2	94,215 / 12%	-	5,248 / 13%	-	99,463 / 13%
GR-HIT-E2	57,540 / 10%	-	-	-	57,540 / 10%
Ratio	1.6 ± 0.16	-	-	-	1.7 ± 0.17
GEL-GSI-E1-1	6,332 / 8%	24,894 / 8%	-	2,581 / 8%	33,807 / 8%
GEL-GSI-E1-2	7,490 / 27%	31,374 / 27%	-	3,706 / 27%	42,570 / 28%
GEL-HIT-E1	6,860 / 10%	23,733 / 10%	-	2,412 / 10%	33,005 / 10%
Ratio-1	0.9 ± 0.11	1.0 ± 0.13	-	1.0 ± 0.13	1.0 ± 0.13
Ratio-2	1.0 ± 0.32	1.3 ± 0.38	-	1.5 ± 0.44	1.2 ± 0.37
GEL-GSI-E2	10,879 / 8%	39,482 / 5%	-	4,963 / 14%	55,324 / 16%
GEL-HIT-E2	10,313 / 10%	38,369 / 10%	-	3,868 / 10%	52,550 / 10%
Ratio	1.0 ± 0.11	1.0 ± 0.11	-	1.2 ± 0.13	1.0 ± 0.11
PMMA-GSI-E1	28,196 / 10%	10,929 / 10%	1,607 / 15%	1,085 / 10%	41,817 / 15%
PMMA-HIT-E1	22,513 / 10%	8,101 / 10%	-	1,451 / 10%	32,065 / 10%
Ratio	1.2 ± 0.17	1.3 ± 0.19	-	0.7 ± 0.1	1.3 ± 0.18
PMMA-GSI-E2	52,637 / 12%	18,620 / 12%	2,090 / 22%	1,835 / 12%	75,182 / 23%
PMMA-HIT-E2	36,085 / 10%	14,030 / 10%	-	1,564 / 10%	51,679 / 10%
Ratio	1.4 ± 0.22	1.3 ± 0.20	-	1.1 ± 0.18	1.4 ± 0.22

3.2. Comparison of depth-resolved production yields

Depth-resolved production yields, stratified by radionuclide type are reported in Figure 2 (${}^{11}\text{C}$, ${}^{15}\text{O}$) and Figure 3 (${}^{13}\text{N}$, ${}^{10}\text{C}$), with each radionuclide-specific figure containing all relevant experiments, respectively. The ${}^{15}\text{O}$ depth curves from different targets are normalized to the respective ${}^{16}\text{O}$ density in gelatine and PMMA and should thus ideally be at the same level. Such normalization could not be performed for ${}^{11}\text{C}$ and ${}^{13}\text{N}$ due to multiple production channels (from ${}^{12}\text{C}$ and ${}^{16}\text{O}$) in the case of ${}^{11}\text{C}$ or large signal fluctuations in the case of ${}^{13}\text{N}$. The curves for ${}^{11}\text{C}$ in graphite are by definition expected to be ideally at the same level.

In general, a reasonably good quantitative agreement was observed between the GSI and HIT data within uncertainties, in particular regarding the aforementioned unconsidered scatter background. Larger deviations for ^{13}N as compared to the other radionuclide yields are mainly due to the scarce abundance of that radionuclide, resulting in higher uncertainties in the production yield quantification. The observed shorter activity range of the GSI data compared to the HIT profiles corresponds to the expected range shift due to the PMMA slab of 4.6 mm water-equivalent path length, which was accidentally present in the beamline during irradiation at GSI.

Despite the quantitative compatibility, the production yield curves of the GSI samples show a slightly convex shape whereas the HIT production yields rise continuously towards the maximum. A check of the impact of the PET image reconstruction, using filtered back-projection as well as an iterative reconstruction method for the GSI data, confirmed that the observed discrepancies are not due to the different reconstruction methods. The feature stems most likely from an insufficient correction of the limited-angle imaging artefacts of the double-head PET camera used at GSI.

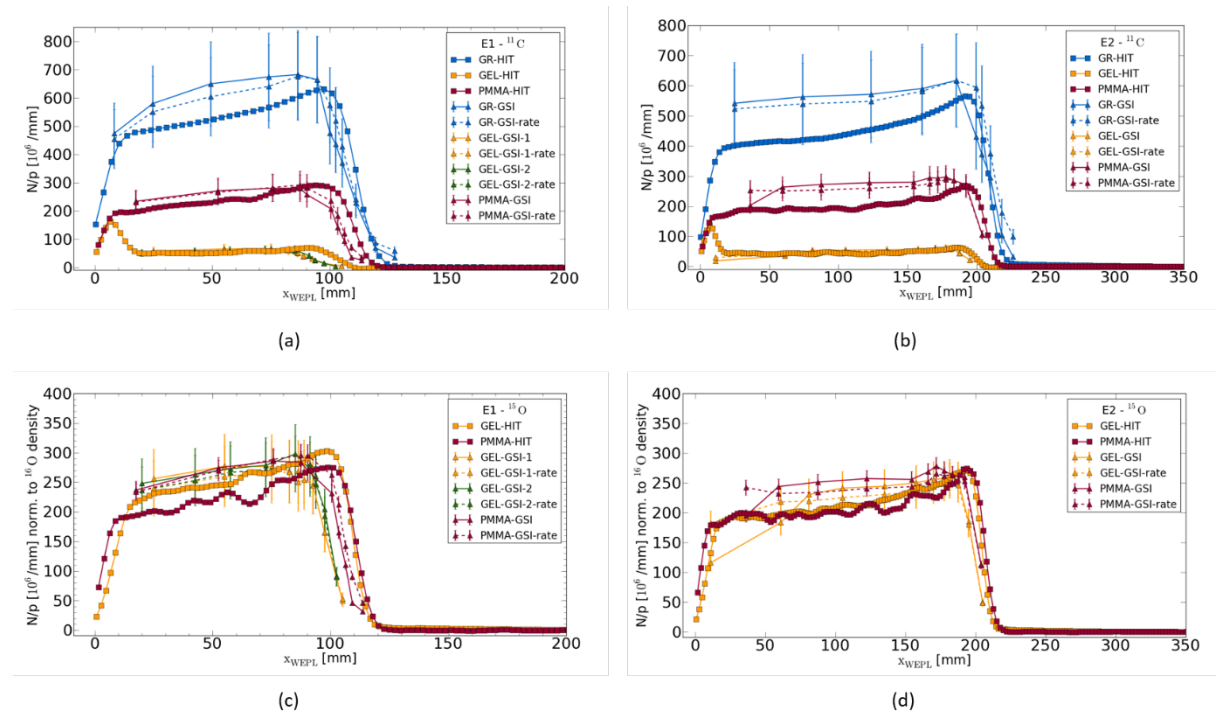


Figure 2: Production yields measured after phantom irradiation at GSI (triangles) and HIT (squares) in gelatine (GEL), graphite (GR) and polymethyl methacrylate (PMMA). Top: ^{11}C production yield measured in GEL, GR and PMMA after irradiation with E_1 (a) and E_2 (b). Bottom: ^{15}O production yield measured in GEL and PMMA after irradiation with E_1 (c) and E_2 (d). The x axis shows the phantom depth and is displayed in water equivalent path length in [mm], the y axis reports the number of radionuclides produced per proton per [mm] penetration depth. For the ^{15}O production yield the y axis is normalized to the ^{16}O density of the respective phantom material.

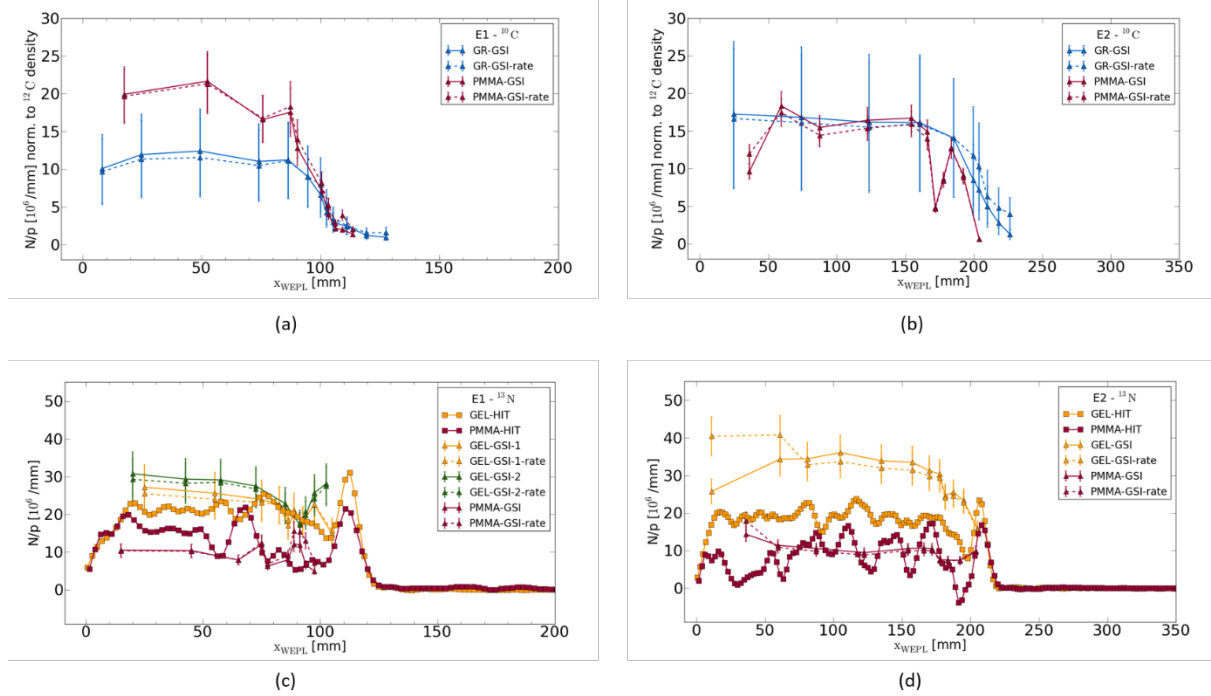


Figure 3: Production yields measured after phantom irradiation at GSI (triangles) and HIT (squares) in gelatine (GEL), graphite (GR) and polymethyl methacrylate (PMMA). Top: ^{10}C production yield measured at GSI in GR and PMMA after irradiation with E_1 (a) and E_2 (b). Bottom: ^{13}N production yield measured in GEL and PMMA after irradiation with E_1 (c) and E_2 (d). The x axis shows the phantom depth and is displayed in water equivalent path length in [mm], the y axis reports the number of radionuclides produced per proton per [mm] penetration depth. For the ^{10}C production yield the y axis is normalized to the ^{12}C density of the respective phantom material.

3.3. Comparison of PET imaging scenarios

Figure 4 shows activity depth profiles reconstructed for the low-energy GSI experiments in the different monitoring scenarios in-beam, in-room, offline and late-offline. The profiles were normalized in the proximal entrance region to the same activity level. Since these concepts start measuring at different times during or after the irradiation the radionuclide fractions contributing to the resulting activity curves are different and determine the shape of the activity depth profile in the respective material.

Activity profiles for the graphite and the PMMA phantoms shown in Figure 4 (a) and (b) are rather stable over time due to the predominance of ^{11}C with a half-life of about 20 min. Differences observed in graphite are due to the ^{10}C contribution in the in-beam scenario, which is not present in the later post-irradiation imaging frames. With a rather flat cross section and a higher production threshold energy for ^{10}C of 34.5 MeV (compared to 20.61 MeV for ^{11}C , cf. Table 1) the in-beam activity depth-profile shows a more convex plateau and a slightly shorter range compared to the other scenarios. Similar effects were observed in PMMA, though less pronounced since the

measured signal is composed of several radionuclides. As expected, activity depth-profiles measured in gelatine showed the largest differences over the four imaging frames. While the in-beam signal is largely dominated by the decay of ^{15}O , its fraction decreases with time and the shape of the depth-profile is more and more determined by the signal from the ^{11}C decay. Since the curve shapes of the energy-dependent production cross-sections are similar for ^{15}O and ^{11}C , the shape of the activity profile remains stable over time. Considerable changes were observed, however, in the activity range, defined as the position of the 50% activity fall-off. The largest difference found between in-beam and late-offline was 3.5 mm in case of the GEL-GSI-E1-1 experiment (cf. Figure 4 (c)). Differences in activity range are due to different production threshold energies, as reported in Table 1. In gelatine, ^{11}C and ^{15}O radionuclides are generated in proton-induced fragmentation reactions of the target oxygen nuclei only. The threshold energies are 27.50 MeV and 16.79 MeV, respectively, thus explaining the observed range decrease over time.

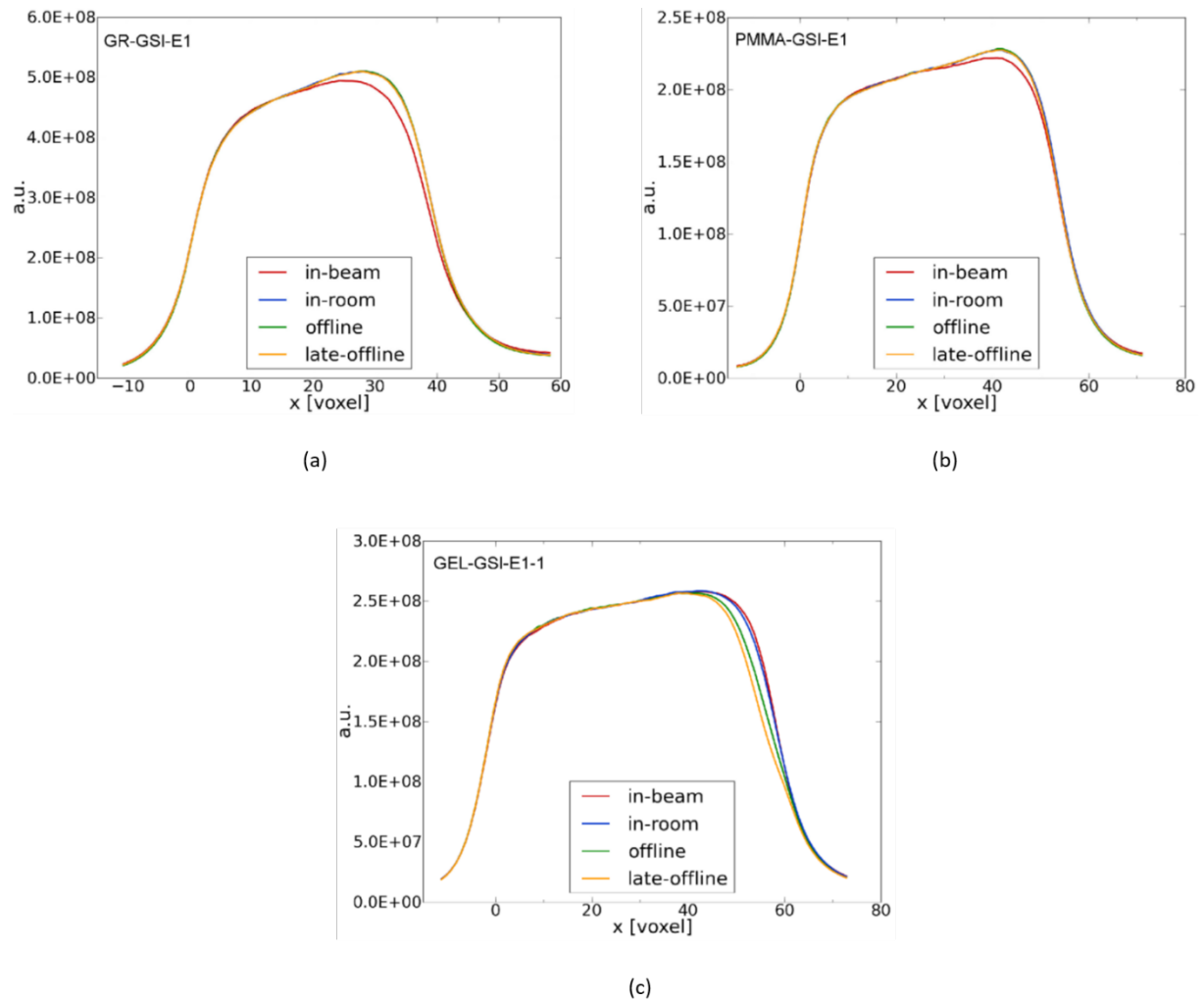


Figure 4: GSI activity profiles as a function of depth measured in the graphite (a), PMMA (b) and gelatine (c) phantoms after mono-energetic pencil beam irradiation with $E_1=125.4$ MeV/u for the different hypothetical in-room, offline and late-offline scenarios with the following time structure (acquisition start / acquisition time): in-beam: (end of first spill, duration of beam

delivery), in-room: (2 min after beam delivery, 3 min), offline: (Δt after beam delivery when radionuclide ratio matches HIT experiments (cf. Table 3), 30min- Δt), late offline (10 min after beam delivery, 20 min).

In addition to variations in the profile shape, the chosen imaging scenario also impacts the total number of measured coincidences, N_c , which are reported in Table 6. This number depends on the duration of the measurement as well as on the half-lives of the radionuclides. For the graphite phantoms, for example, the number of late-offline coincidences is more than doubled compared to the in-beam scenario. The reason is the long acquisition time of the late-offline scenario and the long half-life of ^{11}C whose decay can still be detected after several minutes. In the gelatine phantoms the signal is dominated by ^{15}O decays with a half-life of about two minutes. Thus, the late-offline coincidence number is smaller than the coincidence number of the earlier in-beam and in-room scenarios.

For PMMA, late-offline and in-beam coincidence numbers are almost the same because the longer-living radionuclide ^{11}C is still detected in the longer lasting late-offline scenario, while short-living radionuclides like ^{10}C and ^{15}O are detected in the shorter lasting in-beam scenario.

Table 46 also reports the coincidence rates R_c [1/s] that were obtained on average over the respective monitoring scenarios. Generally, the average count rate decreases with time delay to measurement start. For the graphite phantoms, with a high number of long-living ^{11}C radionuclides produced, the difference in coincidence rates between in-beam and late-offline is small with a ratio of about 1.5. For the gelatine phantoms, on the contrary, the number of in-beam coincidences is about a factor of 10 to 17 times higher than in the late-offline scenario. This is due to a transition in the radionuclide predominance with time.

Early scenarios are dominated by the short-lived ^{15}O , with a high production yield, while the long-lived ^{11}C radionuclides, generated in ^{16}O target fragmentation reactions with a much scarcer production yield, dominate the late-offline scenario. For the PMMA phantoms the in-beam coincidence rate is about three times higher than the late-offline coincidence rate. This ratio reflects the mixed radionuclide yield obtained in this composed material. The ratio is therefore larger than for graphite, due to the presence of ^{15}O and also due to the smaller amount of the longer-living ^{11}C radionuclides in PMMA compared with graphite. Compared with the ratio in gelatine, on the other hand, the PMMA coincidence ratio is smaller because the abundance of ^{11}C radionuclides in gelatine is negligible.

Table 6: Total number of coincidences N_c and corresponding coincidence rates R_c [1/s] averaged over the respective acquisition time for the hypothetical in-room, offline and late-offline scenarios of all GSI experiments with the following time structure (acquisition start / acquisition time): in-beam: (with first spill, duration of beam delivery), in-room: (2 min after

beam delivery, 3 min), offline: (Δt after beam delivery when radionuclide ratio matches HIT experiments (cf. Table 3), 30 min- Δt), late offline (10 min after beam delivery, 20 min).

Exp ID	in-beam		in-room		offline		late-offline	
	N_c	R_c [1/s]	N_c	R_c [1/s]	N_c	R_c [1/s]	N_c	R_c [1/s]
GR-GSI-E1	9.8	22.6	4.7	26.2	29.9	17.1	19.4	14.8
GR-GSI-E2	6.5	17.0	2.7	15.0	21.8	12.9	14.0	11.1
GEL-GSI-E1-1	13.5	30.7	2.8	15.7	6.5	3.2	3.0	1.8
GEL-GSI-E1-2	11.9	32.7	3.2	17.6	4.0	3.5	2.6	3.0
GEL-GSI-E2	21.2	57.0	5.2	29.1	13.6	7.9	4.8	3.9
PMMA-GSI-E1	7.6	17.8	2.7	14.9	12.8	7.7	7.0	5.9
PMMA-GSI-E2	12.4	34.2	4.9	27.0	25.1	12.2	14.9	9.4

Discussion

A comprehensive data analysis has been performed to calculate radionuclide-specific production yields from in-beam PET data acquired at GSI. While methodological uncertainties in the calculation of relative quantities could be well controlled, the determination of reliable quantitative numbers remained difficult due to the research character of the infrastructure. The prototype PET camera installed at GSI is a limited-angle double-head system with well-known limitations regarding a reliable activity quantification (Crespo et al., 2006). Moreover, the beam delivered by the GSI research accelerator suffered from intensity fluctuations and empty spills for some experiments. This resulted in additional uncertainties regarding the determination of the actual irradiation time and number of delivered protons and thus of the eventual production rate. Overall, a quantification of the production yield was possible, as integral number and spatially resolved along the beam path, but with non-negligible uncertainties.

In a second step, these results were compared to radionuclide-specific production yield data determined at HIT, thus providing the first inter-facility comparison of long-lived proton-induced PET radionuclide yields in terms of both integral figures and depth-resolved distributions. GSI production yields were found to be similar to or higher than HIT yields, with the strength of the deviation depending on the phantom material. This phenomenon has also been observed in a recent production yield comparison conducted for similar phantom activation experiments using heavy ion beams (Bauer et al., 2019). One presumable reason for the overestimation of production yields determined at GSI is an incorrect handling of scattered coincidences. A Monte-Carlo based in-silico study has recently been performed by Baumgartl (Baumgartl, 2014) investigating the fraction of scattered coincidences (SC) in the here presented GSI proton experiments, as well as the overall

1 signal quantification bias due to SC after full PET image reconstruction. The author determined
2 scatter fractions of 38.6%, 45.7% and 26.3% in the GSI PMMA, graphite and gelatine phantoms,
3 respectively, for the different GSI experiments. This SC fraction leads to an eventual scatter bias of
4 the reconstructed signal of 7.7% (PMMA), 10.6% (graphite) and 6.6% (gelatin). More details on the
5 study are provided in the supplementary material. These signal quantification uncertainties are
6 largely consistent with the discrepancies observed between the HIT and GSI production yields. They
7 have not been accounted for in the present study, since the analysis pipeline was chosen to be as
8 consistent as possible to previous works. However, the found scatter bias should be considered in
9 potential future quantitative analyses of in-beam activation data acquired with the GSI PET camera.

10 The overall agreement between GSI and HIT data within uncertainties, in particular for PMMA and
11 gelatine, is insofar remarkable as the imaging system used for activity detection and the analysis
12 pipeline to determine production yields were fundamentally different from each other. GSI data
13 were acquired with a prototype double-head PET camera and yields were calculated from raw count
14 rates. HIT data, on the other hand, were acquired with a commercial full-ring PET scanner providing
15 clinical accuracy and yields were calculated from spatially reconstructed activity maps.

16 Despite the applied signal corrections for the limited-angle FOV, qualitative differences remained in
17 the depth-profile shape between HIT and GSI data, which could not be attributed to the chosen
18 reconstruction algorithm.

19 Regarding the monitoring scenario, offline systems provide highest imaging performance since
20 standard full-ring scanners can be utilized. However, the time delay of this workflow causes a
21 considerable signal reduction and loss of spatial tissue activation information, resulting in a low
22 clinical verification power. As we also demonstrated by comparing different hypothetical monitoring
23 scenarios, the signal loss due to the physical activity decay during the time delay can be
24 compensated by considerably prolonged acquisition time frames. However, in the clinical workflow
25 this often conflicts with patient comfort and patient throughput, and the issue of signal distortion
26 due to activity washout from the target volume remains unsolved. Thus, the current trend in clinical
27 PET-imaging installations for beam range verification is towards in-beam solutions (Bisogni et al.,
28 2016). These systems typically require an open-ring geometry to ensure compatibility with the
29 rotating beam delivery and robotic patient positioning system, thus have to accept a limited-angle
30 signal detection with a compromised imaging quality. The quantitative and qualitative discrepancies
31 in production yields and depth-profile shape, respectively, that we observed between HIT and GSI
32 data underline the necessity of a thorough calibration of the imaging system. In particular
33 geometrical image corrections for prototype devices have to be carefully considered to avoid
34 artificial qualitative signal distortions that might impact clinical beam range verification.

1 In general, verification workflows involving a comparison of the measured signal to a pre-calculated
2 expectation have a high demand on the calculation accuracy. It has been demonstrated in the past
3 that available Monte Carlo based calculation engines used for PET-based beam range verification
4 have to be at least benchmarked against and usually further tuned with reliable experimental data
5 (Bauer et al., 2013a; Parodi et al., 2002; Pinto et al., 2020; Seravalli et al., 2012). The availability of
6 reliable production yield data, not only integral but also depth-resolved, is therefore of high interest
7 for this specific application. However, our findings also demonstrate that production yields calculated
8 from quantitative activation data can be very sensitive to the imaging and data analysis workflow.
9 Thus, a determination of energy dependent production cross sections from such thick target
10 experiments can be problematic. The conduction of activation measurements for the tuning of the
11 in-vivo activity signal prediction therefore remains an essential component of system commissioning
12 for the integration of PET-based beam range verification in the clinic.

14 **Conclusion**

15 We provide experimentally determined production yield data in thick homogenous phantoms
16 acquired at two treatment facilities with different PET imaging installations and processed by
17 different analysis pipelines. The observed quantitative and qualitative differences between the
18 resulting datasets underline the importance of a thorough system commissioning prior to clinical use
19 to ensure a reliable beam range verification. In-house phantom activation measurements can thus be
20 considered a gold-standard for careful system response characterization preceding the integration of
21 the PET-based range verification workflow into clinical routine.

23 **Acknowledgements**

24 The authors would like to thank Dieter Schardt and Peter Heeg for their assistance with the
25 experiments at GSI, and Stephan Brons for his support of experiments at HIT. JB acknowledges
26 funding from the German Federal Ministry for Research and Education (grant agreement numbers
27 01IB08002F, 01IB13001G, 13GW0436A) and by the University of Heidelberg (Olympia-Morata
28 fellowship 2021/2022).

30 **Conflicts of Interest**

31 The authors have no conflicts of interest to declare.

References

- Bauer, J., Sommerer, F., Mairani, A., Unholtz, D., Farook, R., Handrack, J., Frey, K., Marcelos, T., Tessonnier, T., Ecker, S., Ackermann, B., Ellerbrock, M., Debus, J., Parodi, K., 2014. Integration and evaluation of automated Monte Carlo simulations in the clinical practice of scanned proton and carbon ion beam therapy. *Phys. Med. Biol.* 59, 4635–4659. <https://doi.org/10.1088/0031-9155/59/16/4635>
- Bauer, J., Tessonnier, T., Debus, J., Parodi, K., 2019. Offline imaging of positron emitters induced by therapeutic helium, carbon and oxygen ion beams with a full-ring PET/CT scanner: experiments in reference targets. *Phys. Med. Biol.* <https://doi.org/10.1088/1361-6560/ab48b4>
- Bauer, J., Unholtz, D., Kurz, C., Parodi, K., 2013a. An experimental approach to improve the Monte Carlo modelling of offline PET/CT-imaging of positron emitters induced by scanned proton beams. *Phys. Med. Biol.* 58, 5193–5213. <https://doi.org/10.1088/0031-9155/58/15/5193>
- Bauer, J., Unholtz, D., Sommerer, F., Kurz, C., Haberer, T., Herfarth, K., Welzel, T., Combs, S.E., Debus, J., Parodi, K., 2013b. Implementation and initial clinical experience of offline PET/CT-based verification of scanned carbon ion treatment. *Radiother. Oncol.* 107, 218–226. <https://doi.org/10.1016/j.radonc.2013.02.018>
- Baumgartl, M., 2014. PET-based Hadrontherapy Monitoring: Monte Carlo Simulations for Nuclear Interaction Studies in Phantoms. LMU Munich and Université Paris Sud.
- Beebe-Wang, J., Peggs, S., Smith, L., 2013. Dependence of the Production Yields of Positron Emitters in Proton Therapy on the Cross Section Data Variations.
- Bisogni, M.G., Attili, A., Battistoni, G., Belcari, N., Camarlinghi, N., Cerello, P., Coli, S., Del Guerra, A., Ferrari, A., Ferrero, V., Fiorina, E., Giraudo, G., Kostara, E., Morrocchi, M., Pennazio, F., Peroni, C., Piliero, M.A., Pirrone, G., Rivetti, A., Rolo, M.D., Rosso, V., Sala, P., Sportelli, G., Wheadon, R., 2016. INSIDE in-beam positron emission tomography system for particle range monitoring in hadrontherapy. *J. Med. Imaging* 4, 011005. <https://doi.org/10.1117/1.JMI.4.1.011005>
- Crespo, P., Shakirin, G., Enghardt, W., 2006. On the detector arrangement for in-beam PET for hadron therapy monitoring. *Phys. Med. Biol.* 51, 2143–2163. <https://doi.org/10.1088/0031-9155/51/9/002>
- Enghardt, W., Parodi, K., Crespo, P., Fiedler, F., Pawelke, J., Pönisch, F., 2004. Dose quantification from in-beam positron emission tomography. *Radiother. Oncol. J. Eur. Soc. Ther. Radiol. Oncol.* 73 Suppl 2, S96-98. [https://doi.org/10.1016/s0167-8140\(04\)80024-0](https://doi.org/10.1016/s0167-8140(04)80024-0)
- Ferrero, V., Fiorina, E., Morrocchi, M., Pennazio, F., Baroni, G., Battistoni, G., Belcari, N., Camarlinghi, N., Ciocca, M., Del Guerra, A., Donetti, M., Giordanengo, S., Giraudo, G., Patera, V., Peroni, C., Rivetti, A., Rolo, M.D. da R., Rossi, S., Rosso, V., Sportelli, G., Tampellini, S., Valvo, F., Wheadon, R., Cerello, P., Bisogni, M.G., 2018. Online proton therapy monitoring: clinical test of a Silicon-photodetector-based in-beam PET. *Sci. Rep.* 8, 4100. <https://doi.org/10.1038/s41598-018-22325-6>
- Fiedler, F., 2008. Anwendung des in-beam PET Therapiemonitorings auf Präzisionsbestrahlungen mit Helium-Ionen. <https://doi.org/10.34657/1732>
- Hildebrandt, M., 2012. Experimental investigation of beta plus emitter yields induced by proton beams in different materials for improved modeling of in-vivo PET verification. Department of Physics and Astronomy, University of Heidelberg.
- Horst, F., Adi, W., Aricò, G., Brinkmann, K.-T., Durante, M., Reidel, C.-A., Rovituso, M., Weber, U., Zaunick, H.-G., Zink, K., Schuy, C., 2019. Measurement of PET isotope production cross sections for protons and carbon ions on carbon and oxygen targets for applications in particle therapy range verification. *Phys. Med. Biol.* 64, 205012. <https://doi.org/10.1088/1361-6560/ab4511>
- Jan, S., Benoit, D., Becheva, E., Carlier, T., Cassol, F., Descourt, P., Frisson, T., Grevillot, L., Guigues, L., Maigne, L., Morel, C., Perrot, Y., Rehfeld, N., Sarrut, D., Schaart, D.R., Stute, S., Pietrzyk, U., Visvikis, D., Zahra, N., Buvat, I., 2011. GATE V6: a major enhancement of the GATE simulation

platform enabling modelling of CT and radiotherapy. *Phys. Med. Biol.* 56, 881–901.
<https://doi.org/10.1088/0031-9155/56/4/001>

Jan, S., Santin, G., Strul, D., Staelens, S., Assié, K., Autret, D., Avner, S., Barbier, R., Bardiès, M., Bloomfield, P.M., Brasse, D., Breton, V., Bruyndonckx, P., Buvat, I., Chatziioannou, A.F., Choi, Y., Chung, Y.H., Comtat, C., Donnarieix, D., Ferrer, L., Glick, S.J., Groiselle, C.J., Guez, D., Honore, P.F., Kerhoas-Cavata, S., Kirov, A.S., Kohli, V., Koole, M., Krieguer, M., van der Laan, D.J., Lamare, F., Largeron, G., Lartzien, C., Lazaro, D., Maas, M.C., Maigne, L., Mayet, F., Melot, F., Merheb, C., Pennacchio, E., Perez, J., Pietrzyk, U., Rannou, F.R., Rey, M., Schaart, D.R., Schmidtlein, C.R., Simon, L., Song, T.Y., Vieira, J.M., Visvikis, D., Van de Walle, R., Wieërs, E., Morel, C., 2004. GATE: a simulation toolkit for PET and SPECT. *Phys. Med. Biol.* 49, 4543–4561. <https://doi.org/10.1088/0031-9155/49/19/007>

Johnson, R.P., 2018. Review of medical radiography and tomography with proton beams. *Rep. Prog. Phys. Phys. Soc. G. B.* 81, 016701. <https://doi.org/10.1088/1361-6633/aa8b1d>

Knopf, A.-C., Lomax, A., 2013. *In vivo* proton range verification: a review. *Phys. Med. Biol.* 58, R131–R160. <https://doi.org/10.1088/0031-9155/58/15/R131>

Meißner, H., Fuchs, H., Hirtl, A., Reschl, C., Stock, M., 2019. Towards offline PET monitoring of proton therapy at MedAustron. *Z. Med. Phys.* 29, 59–65.
<https://doi.org/10.1016/j.zemedi.2018.05.003>

Min, C.H., Zhu, X., Winey, B.A., Grogg, K., Testa, M., El Fakhri, G., Bortfeld, T.R., Paganetti, H., Shih, H.A., 2013. Clinical application of in-room positron emission tomography for in vivo treatment monitoring in proton radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.* 86, 183–189. <https://doi.org/10.1016/j.ijrobp.2012.12.010>

Miyatake, A., Nishio, T., Ogino, T., Saijo, N., Esumi, H., Uesaka, M., 2010. Measurement and verification of positron emitter nuclei generated at each treatment site by target nuclear fragment reactions in proton therapy. *Med. Phys.* 37, 4445–4455.
<https://doi.org/10.1118/1.3462559>

Nischwitz, S.P., Bauer, J., Welzel, T., Rief, H., Jäkel, O., Haberer, T., Frey, K., Debus, J., Parodi, K., Combs, S.E., Rieken, S., 2015. Clinical implementation and range evaluation of in vivo PET dosimetry for particle irradiation in patients with primary glioma. *Radiother. Oncol.* 115, 179–185. <https://doi.org/10.1016/j.radonc.2015.03.022>

Nishio, T., Miyatake, A., Ogino, T., Nakagawa, K., Saijo, N., Esumi, H., 2010. The development and clinical use of a beam ON-LINE PET system mounted on a rotating gantry port in proton therapy. *Int. J. Radiat. Oncol. Biol. Phys.* 76, 277–286.
<https://doi.org/10.1016/j.ijrobp.2009.05.065>

Parodi, K., Enghardt, W., Haberer, T., 2002. In-beam PET measurements of β^+ radioactivity induced by proton beams. *Phys. Med. Biol.* 47, 21–36. <https://doi.org/10.1088/0031-9155/47/1/302>

Parodi, K., Paganetti, H., Cascio, E., Flanz, J.B., Bonab, A.A., Alpert, N.M., Lohmann, K., Bortfeld, T., 2007a. PET/CT imaging for treatment verification after proton therapy: a study with plastic phantoms and metallic implants. *Med. Phys.* 34, 419–435.
<https://doi.org/10.1118/1.2401042>

Parodi, K., Paganetti, H., Shih, H.A., Michaud, S., Loeffler, J.S., DeLaney, T.F., Liebsch, N.J., Munzenrider, J.E., Fischman, A.J., Knopf, A., Bortfeld, T., 2007b. Patient study of in vivo verification of beam delivery and range, using positron emission tomography and computed tomography imaging after proton therapy. *Int. J. Radiat. Oncol. Biol. Phys.* 68, 920–934.
<https://doi.org/10.1016/j.ijrobp.2007.01.063>

Parodi, K., Polf, J.C., 2018. *In vivo* range verification in particle therapy. *Med. Phys.* 45.
<https://doi.org/10.1002/mp.12960>

Pinto, M., Kröniger, K., Bauer, J., Nilsson, R., Traneus, E., Parodi, K., 2020. A filtering approach for PET and PG predictions in a proton treatment planning system. *Phys. Med. Biol.* 65, 095014.
<https://doi.org/10.1088/1361-6560/ab8146>

- 1 Pönisch, F., Parodi, K., Hasch, B.G., Enghardt, W., 2004. The modelling of positron emitter production
2 and PET imaging during carbon ion therapy. *Phys. Med. Biol.* 49, 5217–5232.
3 <https://doi.org/10.1088/0031-9155/49/23/002>
- 4 Priegnitz, M., Fiedler, F., Kunath, D., Laube, K., Enghardt, W., 2012. An Experiment-Based Approach
5 for Predicting Positron Emitter Distributions Produced During Therapeutic Ion Irradiation.
6 *IEEE Trans. Nucl. Sci.* 59, 77–87. <https://doi.org/10.1109/TNS.2011.2172629>
- 7 Richter, C., Pausch, G., Barczyk, S., Priegnitz, M., Keitz, I., Thiele, J., Smeets, J., Stappen, F.V.,
8 Bombelli, L., Fiorini, C., Hotoiu, L., Perali, I., Prieels, D., Enghardt, W., Baumann, M., 2016.
9 First clinical application of a prompt gamma based in vivo proton range verification system.
10 *Radiother. Oncol.* 118, 232–237. <https://doi.org/10.1016/j.radonc.2016.01.004>
- 11 Seravalli, E., Robert, C., Bauer, J., Stichelbaut, F., Kurz, C., Smeets, J., Van Ngoc Ty, C., Schaart, D.R.,
12 Buvat, I., Parodi, K., Verhaegen, F., 2012. Monte Carlo calculations of positron emitter yields
13 in proton radiotherapy. *Phys. Med. Biol.* 57, 1659–1673. [https://doi.org/10.1088/0031-](https://doi.org/10.1088/0031-9155/57/6/1659)
14 [9155/57/6/1659](https://doi.org/10.1088/0031-9155/57/6/1659)
- 15 Shakirin, G., Braess, H., Fiedler, F., Kunath, D., Laube, K., Parodi, K., Priegnitz, M., Enghardt, W., 2011.
16 Implementation and workflow for PET monitoring of therapeutic ion irradiation: a
17 comparison of in-beam, in-room, and off-line techniques. *Phys. Med. Biol.* 56, 1281–1298.
18 <https://doi.org/10.1088/0031-9155/56/5/004>
- 19 Shao, Y., Sun, X., Lou, K., Zhu, X.R., Mirkovic, D., Pönisch, F., Grosshans, D., 2014. In-beam PET
20 imaging for on-line adaptive proton therapy: an initial phantom study. *Phys. Med. Biol.* 59,
21 3373–3388. <https://doi.org/10.1088/0031-9155/59/13/3373>
- 22 Sommerer, F., Cerutti, F., Parodi, K., Ferrari, A., Enghardt, W., Aiginger, H., 2009. In-beam PET
23 monitoring of mono-energetic ^{16}O and ^{12}C beams: experiments and FLUKA simulations for
24 homogeneous targets. *Phys. Med. Biol.* 54, 3979–3996. [https://doi.org/10.1088/0031-](https://doi.org/10.1088/0031-9155/54/13/003)
25 [9155/54/13/003](https://doi.org/10.1088/0031-9155/54/13/003)
- 26 Xie, Y., Bentefour, E.H., Janssens, G., Smeets, J., Vander Stappen, F., Hotoiu, L., Yin, L., Dolney, D.,
27 Avery, S., O’Grady, F., Prieels, D., McDonough, J., Solberg, T.D., Lustig, R.A., Lin, A., Teo, B.-
28 K.K., 2017. Prompt Gamma Imaging for In Vivo Range Verification of Pencil Beam Scanning
29 Proton Therapy. *Int. J. Radiat. Oncol.* 99, 210–218.
30 <https://doi.org/10.1016/j.ijrobp.2017.04.027>
- 31 Zhu, X., España, S., Daartz, J., Liebsch, N., Ouyang, J., Paganetti, H., Bortfeld, T.R., El Fakhri, G., 2011.
32 Monitoring proton radiation therapy with in-room PET imaging. *Phys. Med. Biol.* 56, 4041–
33 4057. <https://doi.org/10.1088/0031-9155/56/13/019>

37 Supplement

38 Determination of scatter correction bias on production yields

39 Sophisticated simulation studies are required to realistically estimate the scatter fraction, which were
40 not available at the time of previously published GSI data analyses (Fiedler, 2008; Parodi et al., 2002;
41 Sommerer et al., 2009). Such a dedicated Monte-Carlo (MC) study has been conducted by
42 (Baumgartl, 2014). The author reproduced the full experimental GSI set-up using the GATE MC
43 package (V6.2, based on Geant4 V9.5) (Jan et al., 2011, 2004). The simulated annihilation signal was
44 processed by the same data analysis and image reconstruction pipeline as the measured data. The
45 simulation framework considered also the macroscopic time structure of the pulsed beam delivery

for each experiment, which allowed to correlate the time dependent annihilation signal to the corresponding spills (i.e. beam-on times). Combining the time correlation with the information of the signal origin it was possible to investigate the impact of scatter coincidences for particular phantoms and irradiation setups.

The simulated scatter fraction determined after image reconstruction, A_{sf} , was defined as:

$$A_{sf} = \frac{A_{sc}}{A_{tot}} \quad (10)$$

where A_{sc} is the number of scattered coincidences and A_{tot} the sum of scattered and true coincidences (random coincidences were neglected). The found scatter fractions after image reconstruction for a PMMA (9x9x20cm³), graphite (9x9x15cm³) and gelatine (9x7.3x20cm³) phantom were 38.6%, 45.7% and 26.3%, respectively, assuming the default energy window of the GSI PET camera ($E=[250-850]$ keV).

In order to estimate the scatter bias on the signal quantification, $A_{S,bias}$, two different PET images were reconstructed from the simulated β^+ emitter maps. First, an image was reconstructed according to the experimental data analysis workflow: the photons penetrated the respective phantom material and were thus corrected for attenuation, both true and scattered photons by the same factor. The total amount of activity in the resulting reconstructed PET image is denoted as A_{Σ} . In a second step, the image was reconstructed in absence of the phantom, thus the annihilation photons were propagated in air neglecting the attenuation correction. The total amount of activity in this second image, containing only unscattered events, is denoted as A_{NoM} .

The activity bias due to scattered coincidences was then determined by comparing the amount of activity in the two reconstructed images:

$$A_{S,bias} = \frac{A_{\Sigma} - A_{NoM}}{A_{NoM}} \quad (11)$$

The scatter bias found for the aforementioned observed scatter fractions A_{sf} were 7.7%, 10.6% and 6.6% in the PMMA, graphite and gelatine phantom, respectively.