Physica Medica 145 (2026) 105772

Contents lists available at ScienceDirect

Physica Medica

journal homepage: www.elsevier.com/locate/ejmp

A phantom study to achieve comparable ^{124}I PET recovery and address scanner variability for dosimetry purposes in (re)differentiated thyroid cancer

Maaike Dotinga [a,b], Lioe-Fee de Geus-Oei [a,c,d], Petra Dibbets-Schneider [a], Mette K. Stam [a], Martin Pool [e], Lenka M. Pereira Arias-Bouda [a], Ellen Kapiteijn [f], Dennis Vriens [g,1], Floris H.P. van Velden [a,1,*]

[a] Department of Radiology, Section of Nuclear Medicine, Leiden University Medical Center, Leiden, the Netherlands
[b] Department of Nuclear Medicine, Netherlands Cancer Institute, Amsterdam, the Netherlands
[c] Biomedical Photonic Imaging, University of Twente, Enschede, the Netherlands
[d] Department of Radiation Science and Technology, Delft University of Technology, Delft, the Netherlands
[e] Department of Clinical Pharmacology and Toxicology, Leiden University Medical Center, Leiden, the Netherlands
[f] Department of Medical Oncology, Leiden University Medical Center, Leiden, the Netherlands
[g] Department of Medical Imaging, Radboud University Medical Center, Nijmegen, the Netherlands

ARTICLE INFO

Keywords:
^{124}I
Quantification
Phantom study
Dosimetry
Recovery correction
Thyroid cancer
PET/CT

ABSTRACT

Background: Accurate quantification of iodine uptake is essential for performing pre-treatment dosimetry of ^{131}I therapy after redifferentiation of radioiodine-refractory thyroid cancer. Standardized procedures for ^{124}I PET/CT-based dosimetry are currently lacking. We aim to evaluate the relation between ^{18}F and ^{124}I imaging for two different PET/CT scanners, the effect of bias correction on recovery across scanners and investigate the impact of clinically-encountered background (BG)-to-lesion ratios on recovery correction.
Methods: Cylindrical and NEMA body phantoms were scanned following vendor-recommended ^{124}I acquisition using clinically representative activity concentrations (5.6–5.7 kBq/mL in cylindrical phantom, 45.1–59.2 kBq/mL in NEMA spheres) and BG-to-sphere ratios (~1:125 to 1:infinity) using two digital PET/CT scanners (Philips Vereos and GE Healthcare Omni). Additionally, BG-to-sphere ratio ~1:10 was acquired to compare ^{124}I to the EARL ^{18}F standards 1 (EARL1). Calibration accuracy and recovery coefficients (RC_{max}, RC_{mean}) were compared between scanners with and without bias correction.
Results: ^{124}I recovery was ~20% higher for the Omni compared to the Vereos, showing calibration accuracies of 1.12–1.15 vs. 0.94, and RC_{mean} reaching 0.86 vs. 0.69. After bias correction, RC_{mean} was comparable between scanners (<1%) but below the lower limits of EARL1. A single fit for recovery correction ($R^2 = 0.97$) was obtained for different BG-to-sphere ratios for both scanners as RC_{mean} was comparable ($p > 0.4$).
Conclusion: Vendor-recommended ^{124}I acquisition and reconstruction leads to differences in quantification but can be compensated using a bias correction. Recovery correction is minimally affected by different BG-to-lesion ratios, suggesting that one RC curve is sufficient, simplifying ^{124}I calibration procedures in studies requiring ^{124}I quantification.

1. Introduction

In the last decade, several studies have been investigating the potential of tyrosine kinase inhibitors (TKIs) or other drugs to restore functional uptake of (radio)iodine in patients with dedifferentiated, radioiodine-refractory thyroid cancer (RAI-R DTC) to re-enable treatment with ^{131}I [1–10]. In these studies, quantification of RAI-uptake characteristics prior and after treatment with a redifferentiation agent

* Corresponding author at: Leiden University Medical Center, Dept. of Radiology, section of Nuclear Medicine, PO box 9600, Postzone C2-P, 2300 RC, Leiden, The Netherlands.
E-mail address: f.h.p.van_velden@lumc.nl (F.H.P. van Velden).
[1] Both authors contributed equally.

https://doi.org/10.1016/j.ejmp.2026.105772
Received 27 April 2025; Received in revised form 14 January 2026; Accepted 16 March 2026
Available online 23 March 2026

is essential to select patients in whom sufficient increase in RAI-uptake and −retention in tumors is observed to warrant effective ^{131}I therapy. Moreover, accurate determination of the activity concentration (AC) within a lesion or organ-at-risk over multiple time points allows for performing dosimetry. This provides further insight in the kinetics and absorbed doses to further improve patient selection and personalize treatment [11].

Dosimetry with ^{124}I positron emission tomography (PET) is preferred over ^{123}I and ^{131}I single photon emission computed tomography (SPECT) owing to its higher spatial resolution allowing smaller lesions to be quantified and its relatively long half-life (4.2 days) permitting delayed imaging [12–14]. Therefore, it is currently used in a majority of redifferentiation studies for RAI-R DTC. Despite its potential, the limited availability of ^{124}I has impeded its wide adoption. In addition, performing accurate quantification of ^{124}I PET/CT in these patients can be challenging for several reasons.

Technical challenges in detecting ^{124}I arise from its complex decay scheme comprising a low positron branching ratio of 23%, and a relative wide range of positron energies and gamma emissions. More than half of the gamma rays have an energy of 603 keV, which can be detected as true annihilation photon coincidences (511 keV), so-called prompt gammas, leading to higher random coincidence rates. Depending on the energy window selected, the prompt gamma coincidence fraction can vary from 2 to 33% [15]. Due to the high effective dose to the patient (strongly dependent on thyroid (remnant) uptake; 0.095 and 15 mSv/MBq for an uptake of 0% and 35%, respectively [16]), administered activities are generally limited compared to shorter-lived radionuclides such as ^{68}Ga and ^{18}F (typically 0.5–1 MBq/kg vs. 3 MBq/kg for contemporary PET/CT [16–18]). Overall, this results in relative low count statistics and degraded image contrast that have to be accounted for in order to accurately quantify ^{124}I uptake. In addition, ^{124}I's relatively large positron range further decreases spatial resolution and induces uncertainty in detecting the precise location of the positron emission in addition to the partial volume effect, especially in smaller lesions [12,19–21]. Using a standard clinical protocol (administered activity of 37 MBq and acquisition duration of 4 min. per bed position), well-differentiated millimetric lesions with sufficient iodine uptake can be identified by ^{124}I PET/CT [22].

In clinical practice, well-differentiated thyroid cancer lesions often exhibit high and lengthy ^{124}I uptake due to high expression of the sodium-iodine transporter and organification of the iodine. There is negligible to low uptake in most adjacent tissues due to specific RAI uptake in limited tissues of the body (predominantly lacrimal and salivary glands, thymus and gastro-intestinal tissue), and rapid renal and hepatobiliary elimination of free RAI. As RAI uptake and retention is adversely affected in RAI-R DTC, lower contrast between lesions and background can be encountered but BG-to-lesion ratios are typically above 1:10 [21]. Jentzen et al. observed median and mean 24 h-uptake of ^{124}I-accumulating metastatic lesions of 22 and 77 kBq/mL (range: 2.3–367 kBq/mL), respectively, after administration of 20–40 MBq ^{124}I [17]. Accurate quantification of ^{124}I PET/CT is further challenged by the observation that lesions are often still relatively small when a redifferentiation agent such as a TKI is started and morphological regression can be observed within a few weeks due to its anti-tumor effect. For TKI lenvatinib, tumor size is initially rapidly declining followed by a slower, continuous shrinkage. A median tumor size reduction of 25% was already observed at 8 weeks, and early tumor shrinkage was retrospectively confirmed at only two weeks of treatment.[23,24].

Detection of smaller lesions and lesions with a relative low ^{124}I uptake has been improved with the introduction of digital silicon photomultiplier-based PET/CT systems due to higher spatial and coincidence time resolution compared to conventional systems [25]. However, standardized quantification procedures are currently lacking. For dosimetry purposes, the effect of lesion size and BG-to-lesion ratios on recovery coefficients (RCs) should be characterized and corrected for in a reproducible way for different PET/CT systems [26].

Harmonized PET/CT calibration and quantification can already be established for ^{18}F, ^{89}Zr and ^{68}Ga by following the accreditation programs of EANM Research Ltd (EARL). When using recommended acquisition settings by the vendor and approved reconstruction parameters for EARL ^{18}F standards 1 accreditation (EARL1), ^{89}Zr RCs are directly related to the RCs obtained with ^{18}F when a bias correction is applied [27,28]. Similar to ^{124}I, ^{89}Zr is characterized by a physical half-life of several days, lower positron abundance of 23% and relative high positron energy and range in comparison to ^{18}F (Table 1). We aim to assess the relation between ^{124}I and ^{18}F RCs for two digital PET/CT systems from different vendors (Philips Vereos vs. GE Healthcare Omni), which differ in vendor-recommended energy windows (450–613 vs. 435–580 keV), type of crystals (lutetium-yttrium oxyothosilicate vs. bismuth germanate), sensitivity (5.2 vs. 47.6 cps/kBq) and ^{124}I branching ratio (0.227 vs. 0.229) [29,30]. For each scanner, we compare ^{124}I measurements with EARL1 and investigate whether bias correction is needed. The main aim of this study is to be able to apply recovery correction for ^{124}I, independent of the BG-to-lesion ratios typically encountered in clinical practice. Additionally, ACs and BG-to-lesion ratios are assessed from RAI-R DTC patient data to evaluate how these metrics are reflected in real patient data and whether they align with the phantom data findings.

2. Methods

2.1. Phantom preparation

A total of five phantom studies were conducted within a time frame of one year on two consecutive installed clinical PET/CT scanners within our department. Studies were performed using a cylindrical uniformity phantom of 6.7 L and the National Electrical Manufacturers Association (NEMA) international electrotechnical commission (IEC) body phantom. The latter phantom contains six fillable spheres with inner diameters of 10, 13, 17, 22, 28 and 37 mm, a lung insert and a background compartment of 9.7 L. Aqueous solutions with activities of [^{18}F]FDG (GE Healthcare, Leiderdorp, the Netherlands) and [^{124}I]NaI (3D Imaging LLC, Little Rock, AR, USA) were measured using a single local calibrated dose calibrator (VIK-202, Comecer Netherlands, Joure, the Netherlands) prior to filling the phantoms.

An overview of obtained ACs per isotope and scanner for the cylindrical phantom and NEMA body phantom is found in Tables 2 and 3, respectively. The small differences in prepared activity concentrations between scanners were due to the expected variation in drawing up and preparing the phantom. To achieve the different BG-to-lesion ratios, the AC in the spheres was kept constant while the AC in the background compartment was increased. ^{18}F ACs were prepared following EARL1 procedures [27,28]. Based on Jentzen et al., an ^{124}I AC of approximately 22–77 kBq/mL within the spheres was considered clinically representative of ^{124}I-avid lesions and used in this study [17]. A BG-to-sphere ratio of 1:10 was used to compare quantitative performance between ^{18}F and ^{124}I. In addition, BG-to-sphere ratios, simulating corresponding BG-to-lesion ratios, of 1:infinity, 1:500 and 1:125 were prepared for ^{124}I to evaluate recovery in the typical hotspot imaging as seen in RAI-avid lesions.

Table 1
Physical characteristics of ^{18}F, ^{124}I and ^{89}Zr [31].

	^{18}F	^{124}I	^{89}Zr
Half-life (h)	1.83	100.2	78.4
Positron branching ratio	97%	23%	23%
Mean positron energy (MeV)	0.25	0.83	0.40
Mean positron range in water (mm)	0.62	3.48	1.23
Single γ energy	–	602 keV (61%) 723 keV (10%) 1691 keV (11%)	909 keV (99.9%) 1657 keV (0.1%) 1713 keV (0.8%)

Table 2
Obtained ACs in the cylindrical phantom for ^{18}F and ^{124}I per scanner. AC: activity concentration.

		AC (kBq/mL)
Vereos	^{18}F	7.3
	^{124}I	5.6
Omni	^{18}F	7.8
	^{124}I	5.7

Table 3
Obtained ACs in the spheres and background compartment of the NEMA body phantom for ^{18}F and ^{124}I, specified per scanner. AC: activity concentration.

		Target BG-to-lesion ratio	Real BG-to-lesion ratio	AC (kBq/mL) spheres	AC (kBq/mL) BG
Vereos	^{18}F	1:10	1:11	19.1	1.8
	^{124}I	1:infinity	1:infinity	59.2	–
		1:500	1:493	59.2	0.12
		1:125	1:126	59.2	0.47
		1:10	1:11	74.1	6.9
Omni	^{18}F	1:10	1:10	18.1	1.8
	^{124}I	1:infinity	1:infinity	45.1	–
		1:500	1:644	45.1	0.07
		1:125	1:98	45.1	0.46
		1:10	1:10	45.1	4.52

2.2. Acquisition and reconstruction

Scans were acquired with two consecutive EARL-accredited clinical digital PET/CT systems: (A) Vereos (Philips Healthcare, Eindhoven, the Netherlands) and (B) Omni Legend 32 (GE Healthcare, Milwaukee, WI, USA). We adhered to EARL1 procedures. An acquisition duration of 5 min. per bed position was used for the cylindrical phantom for both ^{18}F and ^{124}I. For ^{18}F, EARL1 accreditation was acquired with two bed positions and an acquisition duration of 6 and 5 min. per bed position for the Vereos and Omni, respectively. For the NEMA phantom filled with ^{124}I, acquisitions with two bed positions of 4 min. per bed position were performed for the BG-to-sphere ratios 1:infinity, ~1:500 and ~1:10. Acquisition duration was decreased for BG-to-sphere ratio ~1:125 to account for the increase in total activity within the phantom, to the extent that count statistics remained comparable. All ^{124}I acquisitions were repeated for a total of 4 times, thereby repositioning the phantom prior to each acquisition. In addition, a single acquisition with 5-fold increase in duration was performed for the BG-to-sphere ratio 1:infinity in order to obtain higher count statistics and enable critical assessment of the chosen acquisition duration of 4 min. per bed position. For all NEMA scans, the spheres were located in the slices in the overlapping position between the two acquired bed positions. Prior to each acquisition, a low-dose CT scan for attenuation correction purposes was performed (Supplementary Table 1). For ^{124}I, vendor-specific settings for the acquisition were used that comprised above-mentioned energy windows for both PET/CT scanners.

Both ^{18}F and ^{124}I data were reconstructed using the approved parameters for EARL1 accreditation, similarly as described for EARL ^{89}Zr accreditation [27,28]. For the Vereos, we applied a 3D blob-based ordered-subsets expectation maximization (OSEM) algorithm with 3 iterations and 15 subsets and additional smoothing using a Gaussian filter with a full width at half maximum (FWHM) of 5 mm was used. For the Omni, data were reconstructed using Vue Point HD with 4 iterations and 12 subsets, and a Gaussian with a FWHM of 7 mm. All data were normalized and corrected for decay, scatter, randoms, attenuation and dead time and above-mentioned vendor-specific positron branching ratios. Prompt gamma correction for ^{124}I data was not available for the Vereos and could therefore not be applied. Note that the reconstruction of the Vereos incorporates hardware-based TOF information, while the Omni only has deep learning-based TOF enhancement that is currently only available for [^{18}F]FDG. Therefore, only for the Vereos we included TOF information into the reconstruction. Further note that, we did not incorporate any resolution recovery based on the point-spread function (PSF; Omni) and/or PSF deconvolution (Vereos) into the reconstructions, because 1) our focus on EARL1 and 2) these are optimized for ^{18}F and are, therefore, not applicable for ^{124}I due to its greater positron range and prompt gamma emissions, likely resulting in quantification errors.

2.3. Patient data

To quantify clinical BG-to-lesion ratios, the baseline scans of a total of nine RAI-R DTC patients who enrolled in the ongoing RESET trial (NTC04858876) were evaluated in this study [32,33]. The clinical study was approved by the institutional review board and written informed consent was obtained from all patients prior to their inclusion in the study. Prior to oral administration of 37 MBq ± 10% [^{124}I]NaI, patients adhered to an iodine-restricted diet (daily intake <100 μg) for seven days and received two 0.9 mg recombinant human TSH injections (Genzyme Ireland Ltd, Waterford, Ireland) 24 and 48 h prior to ingestion. PET/CT scans were acquired 24 and 96 h after ^{124}I administration. Four patients were scanned using the Vereos and the remaining subsequent five patients using the Omni PET/CT scanner at baseline. Patients were scanned with an acquisition duration of 4 min. per bed position and the scans were reconstructed using the aforementioned approved parameters for EARL1 accreditation.

2.4. Data analysis and statistics

Data were analyzed using Pmod Biomedical Image Quantification Software (PMOD v4.206, PMOD Technologies LLC, Zurich, Switzerland). Scans of the cylindrical phantom were used to assess the PET/CT calibration accuracy for ^{18}F and ^{124}I (i.e. the ratio between the reconstructed activity and the true activity as measured with the dose calibrator). The mean AC was determined using a spherical VOI of ~300 mL placed within the center of the cylindrical phantom. In addition, mean line profiles were determined over three slices across the phantom at ¼, ½ and ¾ length to confirm adequate scatter correction and the absence of artefacts.

Scans of the NEMA body phantom were used to assess RCs for the exact sphere size, including RC_{max} and RC_{mean}, and the coefficient of variation (CoV) as a measure of noise. The RC was defined as the measured AC in the sphere divided by the true AC within that sphere. RC_{max} was based on the AC of the voxel with the highest uptake and RC_{mean} on the mean AC value of an isocontour at 50% of the maximum voxel value adapted for background activity, aligning with the methodology used by Kaalep et al. [28]. Within the background compartment, 12 cylindrical VOIs (radius of 15 mm, length of 150 mm) were positioned after which the mean AC was determined to be able to determine background-adapted isocontour values, assess the calibration accuracy and CoV. The CoV was calculated by dividing the standard deviation by the mean of the pixel values within the cylindrical VOIs. Calibration accuracy was defined as the reconstructed AC divided by the true AC. For BG-to-lesion ratio 1:infinity, 1:500 and 1:125, RCs and CoVs were averaged over all four scans. The 1:infinity ratio was excluded from CoV analysis as it does not provide meaningful noise measurements due to the absence of background activity. For ^{124}I, the relation between RC_{mean} and sphere size was fitted using the following formula 1:

$$RC = \alpha \bullet \left(1 + \left(\frac{\gamma}{spherediameter}\right)^{\beta}\right)^{-1} \quad (1)$$

with α, β and γ as fit parameters.

Calibration accuracies and RCs were compared to EARL1 limits, i.e.

EARL1's acceptability criteria of the calibration accuracy (within the range of 0.90–1.10) and the known RC limits for each sphere size [34]. For ^{124}I, a bias correction was applied to correct for the difference in calibration accuracy between the PET/CT scanners. RC curves were rescaled using the accuracy derived from the NEMA body phantom's background compartment, similarly to Kaalep et al. [27,28].

For the patient data, baseline PET/CT images acquired at 24 h post-administration were analyzed by segmenting the ^{124}I-avid lesions using an isocontour at 50% of the maximum voxel value adapted for background activity. Background activity was determined using a spherical VOI with a diameter of 15 mm in adjacent (healthy) tissue. CoV in liver and lung tissue were determined using five spheres with a diameter of 15 mm that were distributed throughout the whole organ. Lesion volume was determined using three orthogonal measurements on low-dose CT and assuming an ellipsoid geometry.

The distributions of ^{124}I RCs between both scanners were compared statistically using a Mann-Whitney *U* test. RCs between different contrasts were compared using a Kruskal Wallis test. Goodness of fit was assessed using the coefficient of determination (R^2). Statistical analyses were performed using IBM SPSS Statistics 25.0 (IBM Corp., Armonk, New York, United Stated of America). The level of significance was set to 0.05.

3. Results

3.1. ^{18}F vs. ^{124}I.

Calibration accuracy of mean ^{18}F and ^{124}I ACs ranged from 0.94–1.15. Local cross-calibration was within 10% deviation of the dose calibrator except for ^{124}I using the Omni PET/CT scanner. Accuracy of ^{124}I was approximately 20% higher for the Omni in comparison to the Vereos (Table 4), with a tendency toward underestimation in the Vereos and overestimation in the Omni. In both scanners, line profiles of the cylindrical phantom for both ^{18}F and ^{124}I did not show a substantial concave upward or downward shape over the whole phantom length (Supplemental Fig. 1), indicating uniformity and properly applied scatter correction. The RCs for different sphere sizes for both ^{18}F and ^{124}I were plotted against the defined limits of EARL1 RC acceptability criteria (Fig. 1A and B). RCs were within these limits, except for ^{124}I RCs assessed using the Vereos, which were located around the lower limit of the EARL ^{18}F standards 1. Mean ^{124}I RCs were on average 9.6% (range: 3.3–14.3%) and 3.4% (range: 0.3–8.7%) lower than mean ^{18}F RCs for the Vereos and Omni PET/CT scanner, respectively. In comparison to the Omni PET/CT scanner, mean RCs were 12.6% (range: 7.8–17.9%) and 18.2% (range: 14.7–23.1%) lower for the Vereos for ^{18}F and ^{124}I, respectively.

After bias correction was applied, ^{124}I RCs were comparable between both scanners, with a mean deviation of 0.1% (range: −5.9–4.3%), and located around the lower limit of the EARL1 RC acceptability criteria for both scanners (Fig. 1C and D). Mean ^{124}I RCs were on average 3.8% (range: −2.9–8.9%) and 16.0% (range: 10.7–19.0%) lower than mean ^{18}F RCs for the Vereos and Omni, respectively.

Table 4
Calibration accuracy of ^{18}F and ^{124}I measurements in the cylindrical phantom and NEMA body phantom, and CoV in the background compartment of the latter phantom. Accuracy was measured for the Vereos and Omni PET/CT scanner.

Scanner	Accuracy (cylindrical phantom)		Accuracy (NEMA body phantom)	
	^{18}F	^{124}I	^{18}F	^{124}I
Vereos	0.94	0.94	0.96	0.94
Omni	1.06	1.12	0.98	1.15
Difference (%)	12.8	19.1	2.1	22.3

3.2. ^{124}I RCs for different BG-to-lesion ratios

For BG-to-lesion ratio 1:infinity, a 5-fold increase in acquisition duration did not lead to substantial improved RC, showing a mean deviation of <1% in both scanners (Supplemental Fig. 2). RCs for the different BG-to-lesion ratios were plotted for each scanner (Fig. 2). Maximum and mean RCs were significantly higher (mean increase of 19%) for the Omni in comparison to the Vereos for each sphere diameter ($p < 0.001$) without bias correction. In the BG-to-lesion ratios 1:125 and 1:500, CoV ranged between 56–66%. For each scanner, no statistical difference was found for maximum and mean RCs for the different BG-to-lesion ratios ($p > 0.4$) and a fit was made using the BG-adapted mean RCs (including ratios 1:125, 1:500 and 1:infinity) for recovery correction purposes per scanner. For the Vereos, the final fit deviated 0.1 ± 3.1% (range: −7.0–7.1%, $R^2 = 0.99$) from measured mean RCs. Deviations were < 5% for all sphere sizes except 13 and 17 mm. For the GE Omni, the final fit deviated 0.02 ± 2.2% (range: −5.0–3.3%, $R^2 = 0.99$) from measured mean RCs. Deviations were < 5% for all sphere sizes except for the smallest sphere (10 mm). After bias correction, one single fit could be obtained for both scanners including ratios 1:125, 1:500 and 1:infinity, which resulted in a mean deviation of −0.2 ± 8.0% (range: −17.0–17.7%, $R^2 = 0.97$). A tendency of overestimation and underestimation was observed for the Vereos and Omni, respectively, with a deviation range of −2.0–17.0% and −17.7–-0.2%. Both scanners showed the largest deviation in the two smallest spheres.

3.3. ^{124}I uptake in RAI-R DTC patients

A total of 5 of 9 patients showed limited ^{124}I-avid lesions at baseline (Fig. 3), with a median AC of 1.1 kBq/mL (range: 0.5–17.9 kBq/mL), corresponding to a median uptake of 0.0034%/mL (range: 0.0017–0.058%/mL). In one of these patients, ^{124}I-avid lesions could not be evaluated due to the small size (diameter < 10 mm). Mean BG-to-lesion ratio was 1:35.2 (range: 1:5.1 – 1:120.9). Mean CoV in healthy liver and lung tissue were 39% (range: 32–48%) and 41% (range: 34–56%), respectively. Parameters per scanner are found in Table 5.

4. Discussion

In this study, we wanted to obtain the relation in quantification performance between ^{18}F and ^{124}I PET imaging and assess ^{124}I recovery for different clinically relevant BG-to-lesion ratios for two different digital PET systems. This study shows a mean difference of 18% in ^{124}I recovery between two digital PET/CT scanners following vendor-recommended acquisition parameters and EARL1 reconstruction parameters. A bias correction similarly used as in the EARL ^{89}Zr accreditation program [27,28], substantially decreases the mean difference between scanners to <1% and then shows inferior recovery in comparison to ^{18}F for both scanners. Despite bias correction, our data suggest that, unlike for ^{89}Zr, ^{124}I RCs do not fall within the limits of acceptability of EARL1 accreditation. ^{124}I RCs for BG-to-sphere ratios 1:125 to 1:infinity are comparable, suggesting a single fit can be obtained to apply recovery correction for a range of BG-to-lesion ratios that are usually encountered in clinical practice.

Prior studies already showed that ^{124}I recovery is inferior to ^{18}F. Jentzen et al. found a difference in recovery between ^{18}F and ^{124}I of 12–14% [26], comparable to the results for the Vereos in which we observe a deviation up to 14.3%. Soderlund et al. found mean ^{124}I RCs up to 0.27–0.70 for the different sphere sizes using OSEM reconstruction [35], also in line with the results of the Vereos (RC mean: 0.31–0.70). Gregory et al. found an underestimation of approximately 8% and 15% in ACs of ^{18}F and ^{124}I, respectively [36]. In addition to the BG-to-sphere ratios investigated in this study, comparable RCs between BG-to-sphere ratios 1:10 to 1:20 were found by Preylowski et al. and are in concordance with the small observed spill-in effect in ratios above 1:10 by Jentzen et al. [37,38].

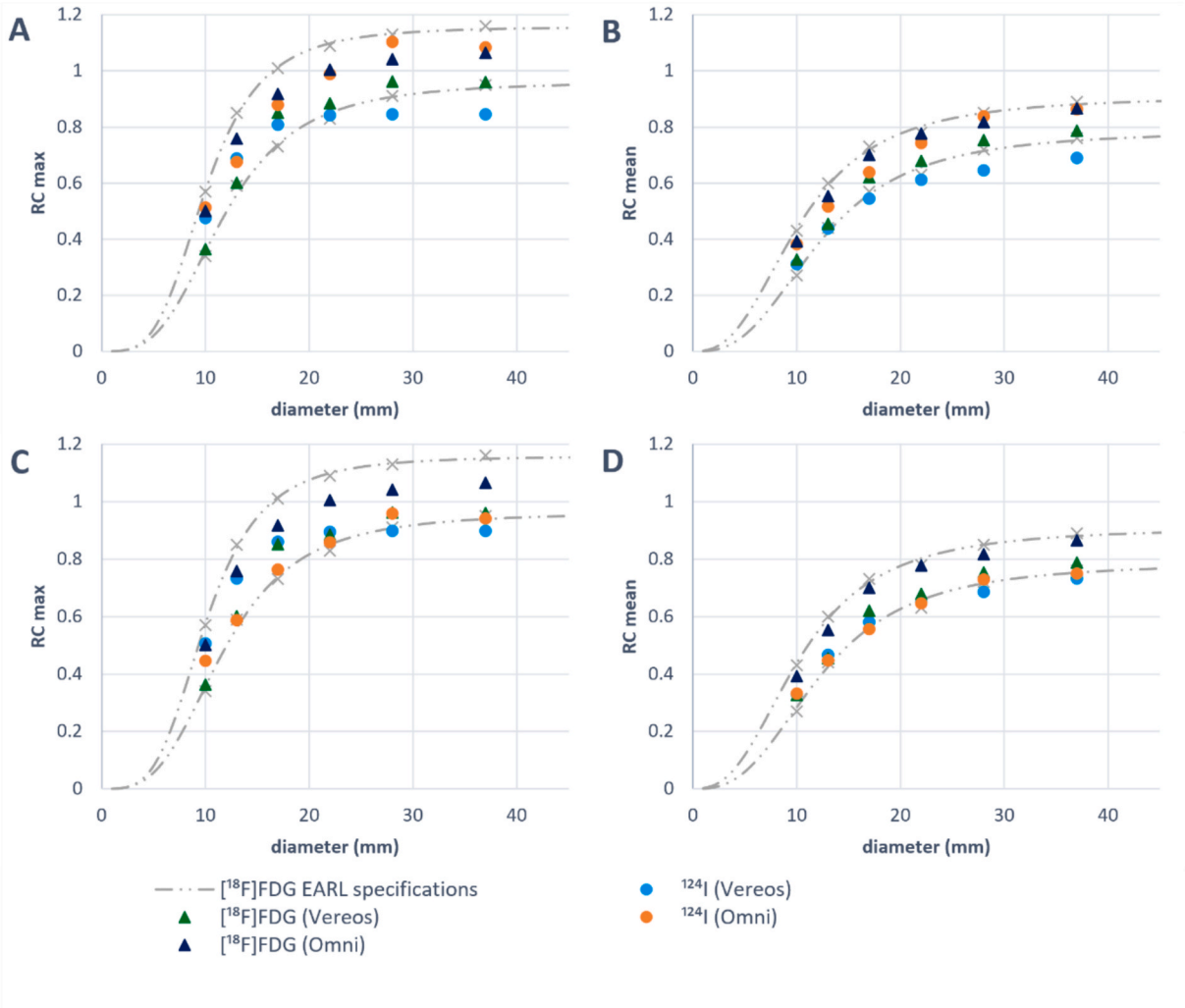

Fig. 1. Maximum and mean RCs of [^{18}F]FDG and ^{124}I for different sphere sizes measured using the NEMA body phantom for the Vereos and Omni using a background-to-lesion ratio of 1:10. RCs are presented without bias correction (A–B) and after bias correction was applied (C–D). RC: recovery coefficient.

The similar results in recovery after bias correction suggest that the initial differences were likely due to scanner-specific variations in detecting ^{124}I. To further enhance quantification performance, recommendations regarding energy windows might be needed. The prompt gammas with an energy of 602.7 keV were well within the energy window of the Vereos (450–613 keV) while this was less the case for the Omni PET/CT scanner (435–580 keV); additionally, the Omni applied prompt gamma correction. Gregory et al. showed that optimizing energy windows and applying single scatter simulation subtraction scatter correction can lead to a decrease of the underestimation up to 15.4% [36]. The observed CoV of >50% in the BG-to-lesion ratios 1:125 and 1:500 and CoV of approximately 40% in the patient data might indicate the need for post-processing corrections. Both time-of-flight (TOF) and positron range correction have shown to decrease background variability and reduce noise levels [39,40]. Without the use of prompt gamma correction, ^{124}I contrast is likely to be overestimated as prompt gamma correction has a main effect on background activity and only a small effect on the recovery of the sphere activity [37,39,41]. The use of resolution modeling based on the measured PSF shows improved hot sphere contrast and recovery, and can further increase reproducibility of RCs between different BG-to-lesion ratios [39]. Besides, it was not part of the EARL1-compliant reconstructions used for this study. The use of the different aforementioned techniques is limited by the lack of validation in patient data and availability on the scanner, further hampering comparison of results between different PET/CT systems across centers.

While beneficial for harmonization and standardization, a main drawback of the presented approach for ^{124}I recovery correction is its probable suboptimal reconstruction and subsequent ^{124}I quantification accuracy. Omitting the Gaussian smoothing filter, optimizing reconstruction parameters such as iterations and subsets, employing resolution recovery based on PSF (taking e.g. the positron range of ^{124}I into account), TOF correction and nuclide-specific scatter correction could improve quantification [36–41]. It remains to be determined whether optimizing reconstruction and implementation of the techniques mentioned can achieve compliance with ^{18}F EARL1 standards and/or, looking further ahead, enable alignment with updated ^{18}F EARL standards (EARL2), which supports recent acquisition and reconstruction technologies. In the context of precision dosimetry, a scanner-specific recovery correction framework might be preferred allowing RCs to exceed the current EARL standards. Depending on study design or clinical application, the current approach may still offer valuable consistency across scanners and sites, particularly in multi-center studies.

In this study, VOIs were defined using a background-adapted 50% isocontour. A limitation of this segmentation approach is that it does not necessarily reproduce the true inner sphere volume or true lesion size, particularly due to partial volume effects and spatial resolution [42]. We did not focus on optimizing volume estimation by PET via thresholding, as lesion volumes were derived from low-dose CT measurements within the dosimetry pipeline (or known spherical volumes from the NEMA phantom). The current approach remains feasible for dosimetry,

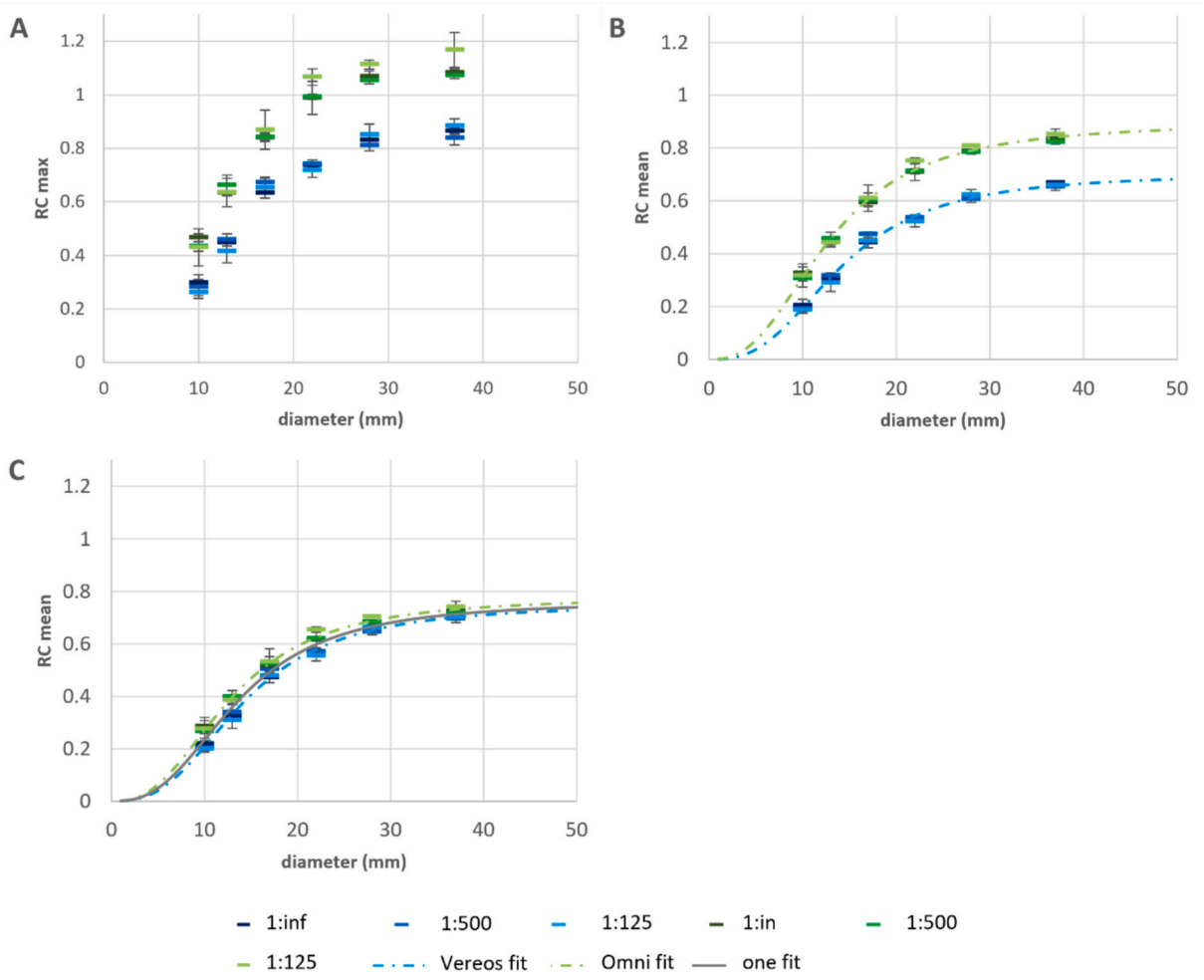

Fig. 2. Maximum and mean recovery coefficients of ^{124}I for different sphere sizes and background-to-lesion ratios measured using the NEMA phantom with different background-to-lesion ratios for the Vereos (blue) and Omni (green) PET/CT scanner. RCs are plotted for each scanner without bias correction (A-B) and after bias correction was applied (C). Mean RCs were fitted for recovery correction purposes. inf: infinity, RC: recovery coefficient. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

provided that lesion volume is accurately estimated, as this impacts the corresponding S-factor and subsequent dosimetric results [43]. We acknowledge that in this sense, the use of orthogonal measurements on a low-dose CT is not perfect either, as exact measurement is deteriorated due to limited resolution and motion artefacts (mainly around the diaphragm). We tried to limit these effects by measuring all dimensions/volumes in triple (based on all 3 low dose CTs available (within 1 week) for each patient.

Further limitations of this study include the limited number of scanners and the single-center approach, limited range of ^{124}I ACs and calibration of the dose calibrator without the use of a NIST traceable calibration source. Due to ^{124}I's high-energy emissions, accurate measurement of ^{124}I can be challenging [44]. Therefore, we ensured consistency in dose calibrator measurements by verifying the dose calibrator's accuracy for both vial and syringe geometry prior to the first ^{124}I experiment of each scanner using a source including a certificate provided by the supplier (3D Imaging LLC, Little Rock, AR, USA). However, we cannot entirely rule out potential errors in the calibration, which may introduce a systematic bias. In addition, the bias correction to rescale RC curves was based on the methodology used for ^{89}Zr [27,28], which relies on the accuracy of the NEMA body phantom's background rather than the cylindrical phantom. The latter may offer a more direct evaluation of scanner performance, being less prone to sampling errors, corrections and spill-in/spill-out effects, and remains a viable option for future work.

Although the acquisition duration for ^{18}F and ^{124}I scans were not identical, and ACs for the ^{124}I NEMA measurements could not be completely matched, we observe consistent trends in ^{124}I recovery across a range of BG-to-lesion ratios and different acquisition durations (Supplemental Fig. 2). This suggests that a similar trend can be expected when acquisition duration and ACs are matched, and that achieving higher count statistics would not lead to a substantial improvement in recovery. In the RESET trial, we observe lower contrast between lesions and BG than in the phantom study. The ratios in the phantom experiment were based on prior work by Jentzen et al. involving iodine-avid DTC lesions [17], while our population consisted of patients with RAI-R DTC, which likely accounts for the mismatch. Investigating ratios that are also encountered in redifferentiation studies as observed in the presented patient data would lead to more insight in quantification performance of ^{124}I and further validate recovery correction using the single fit approach. Despite these limitations, we expect that the results and trends will remain comparable if adjustments are made regarding ACs and acquisition times.

This study exemplifies a clinically representative situation that shows vendor-recommended ^{124}I acquisition leads to different quantification results and are currently not compliant with EARL1. Applying bias correction helps to compensate for scanner-specific differences after which comparable ^{124}I recovery is observed. In addition, we show that recovery correction can be achieved by a single fit of mean ^{124}I RCs in different BG-to-lesion ratios, potentially simplifying calibration

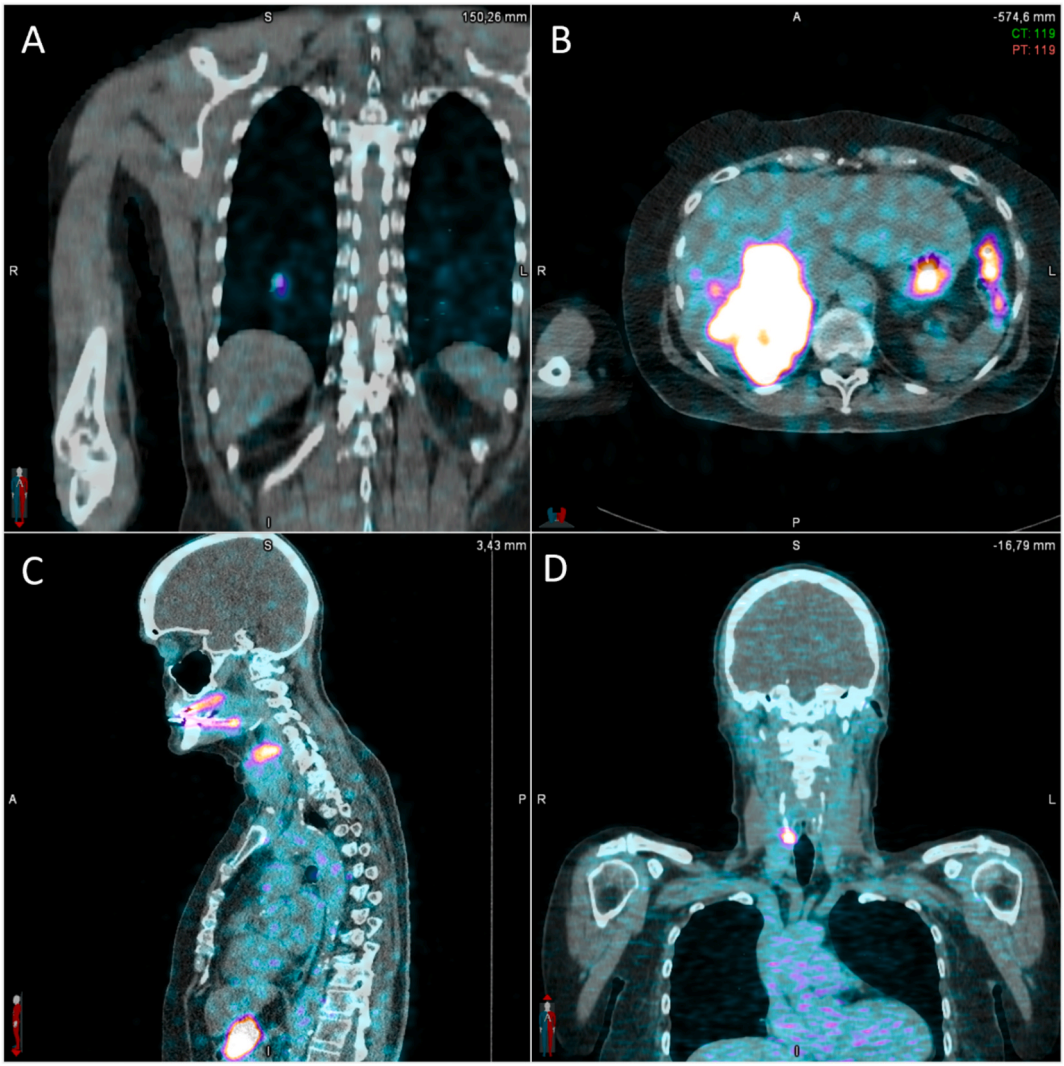

Fig. 3. Baseline ^{124}I PET/CT scans performed 24 h after administration of 37 MBq ^{124}I of four different RAI-R DTC patients with ^{124}I-avid lesions. The window level was set to 0–1.5 kBq/mL (equaling a body-weighted SUV of approximately 0–4 g/mL) for each patient. (A) a pulmonary macrometastasis in the right lower lobe, (B) a large liver metastasis in the right liver lobe, (C) a local recurrence in the left thyroid bed, (D) a local recurrence in the right thyroid bed.

Table 5
Details regarding patient's length and weight, and CoV of healthy liver and lung tissue per scanner. Values are displayed as median (range).

	Vereos	Omni
n	4	5
Length (cm)	171 (163–175)	165 (151–193)
Weight (kg)	67 (55–78)	89 (68–100)
CoV liver (%)	0.39 (0.34–0.48)	0.38 (0.32–0.45)
CoV lung (%)	0.41 (0.34–0.56)	0.42 (0.38–0.45)

procedures for ^{124}I PET quantification. The results of this study further indicate that RC correction in lesions >13 mm can be applied using a single fit across multiple scanners when bias correction is used. This approach results in maximum deviations <10%. For smaller lesions, scanner-specific RC corrections are more accurate. The use of TOF and PSF might further increase the applicability of RC correction in smaller lesions <10 mm [45].

In redifferentiation studies, comparable quantification results between different scanners are important as the limited patient population might require a multi-center approach and an absorbed dose threshold is often used to assess eligibility of ^{131}I therapy. Wierts et al. presented a simplified calibration procedure of ^{124}I PET quantification for multi-center studies, assuming high focal uptake and negligible background, and using a segmentation approach with oversized volume-of-interest (VOIs) [46]. Though the approach of oversized VOIs has been validated for varying ^{131}I background activity distributions [47], it is yet to be determined whether this approach will also suffice the situations encountered in redifferentiation studies to correct for the partial volume effect, especially in lesions which are nearby structures with physiological uptake (e.g. lesions in the neck near the salivary glands).

5. Conclusion

This study shows that bias correction helps compensate scanner-specific differences in detecting ^{124}I. Following vendor-recommended ^{124}I acquisition, EARL1 reconstruction and bias correction, RC curves were closely aligned for two digital PET/CT scanners. However, EARL1 acceptability criteria were not met. Recovery correction is minimally limited by BG-to-lesion ratio, allowing a single fit to be determined and applied across clinically encountered ratios. The current approach could potentially be useful for multicenter studies requiring ^{124}I quantification to ensure comparable results across different scanners.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used chatGPT in order to improve readability and language for the first draft of the manuscript. After using this service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Informed consent

The RESET trial has been approved by the Medical Ethics Committee Leiden-Den Haag-Delft (ref. no. P20.096). The study was conducted according to the principles of the Declaration of Helsinki (10th version, Fortaleza 2013) and in concordance with the Dutch Medical Research Involving Human Subjects Act and other applicable guidelines, regulations, and acts. Written informed consent was obtained from all individual participants included in the study. Trial registration: NIH ClinicalTrials.gov, NTC04858867.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

PD and FV have received speaker honoraria from GE HealthCare (fees received by institution). EK has consultancy/advisory relationships with Immunocore, Delcath and Lilly, and received research grants not related to this paper from Bristol Myers Squibb, Delcath, Novartis and Pierre-Fabre (fees received by institution).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejmp.2026.105772.

Data availability

Data linked to the RESET trial will be disseminated after completion in peer-reviewed journals and presented at conferences. Data generated and analyzed within the RESET trial shall only be shared upon request with researchers who provide a methodologically sound research proposal, at the discretion of the principal investigators. Only de-identified participant data from the final research dataset used in the published manuscripts can be shared.

References

[1] Tchekmedyian V, et al. Enhancing radioiodine incorporation in BRAF-mutant, radioiodine-refractory thyroid cancers with vemurafenib and the anti-ErbB3 monoclonal antibody CDX-3379: results of a pilot clinical trial. Thyroid 2022;32(3):273–82.
[2] Weber M, et al. Enhancing radioiodine incorporation into radioiodine-refractory thyroid cancer with MAPK inhibition (ERRITI): a single-center prospective two-arm study. Clin Cancer Res 2022;28(19):4194–202.
[3] Ho A, et al. Selumetinib-enhanced radioiodine uptake in advanced thyroid cancer. N Engl J Med 2013;368(7):623–32.
[4] Dunn L, et al. Vemurafenib redifferentiation of BRAF mutant, RAI-refractory thyroid cancer. J Clin Endocrinol Metab 2019;104(5):1417–28.
[5] Iravani A, et al. Mitogen-activated protein kinase pathway inhibition for redifferentiation of radioiodine refractory differentiated thryoid cancer: an evolving protocol. Thyroid Cancer Nodules 2019;29(11):1634–45.
[6] Hoftijzer H, et al. Beneficial effects of sorafenib on tumor progression, but not on radioiodine uptake, in patients with differentiated thyroid carcinoma. Eur J Endocrinol 2009;161(6):923–31.
[7] Jaber T, et al. Targeted therapy in advanced thyroid cancer to resensitize tumors to radioactive radioiodine. J Clin Endocrinol Metab 2018;103(10):3698–705.
[8] van Houten P, et al. Digoxin treatment does not reinduce radioiodine uptake in radioiodine refractory non-medullary thyroid carcinoma. Eur Thyroid J 2024;13(4).
[9] Rothenberg S, et al. Redifferentation of iodine-refractory BRAF V600E-mutant metastatic papillary thyroid cancer with dabrafenib. Clin Cancer Res 2014;21(5):1028–35.
[10] Leboulleux S, et al. A phase II redifferentiation trial with dabrafenib-trametinib and 131I in metastatic radioactive iodine refractory BRAF p.V600E-mutated differentiated thyroid cancer. Clin Cancer Res 2023;29(13):2401–9.
[11] Dotinga M, et al. Managing radioiodine refractory thyroid cancer: the role of dosimetry and redifferentiation on subsequent I-131 therapy. Q J Nucl Med Mol Imaging 2020;64:250–64.
[12] Weber M, et al. The role of I-124 PET/CT lesion dosimetry in differentiated thyroid cancer. Q J Nucl Med Mol Imaging 2019;63(3):235–52.
[13] Kuker R, Sztejnberg M, Gulec S. I-124 imaging and dosimetry. Mol Imaging Radionucl Ther 2017;26:66–73.
[14] Aashiq M, et al. Radioiodine-refractory thryoid cancer: molecular basis of redifferentation therapies, management and novel therapies. Cancers (Basel) 2019;11(9):E1382.
[15] Yu AR, et al. Comparison of imaging characteristics of (124)I PET for determination of optimal energy window on the siemens Inveon PET. Biomed Res Int 2016;2016:3067123.
[16] Tan-Phan, T.T.H. 124I PET/CT in Thyroid Cancer. 2016; Available from: https://richtlijnendatabase.nl/gerelateerde_documenten/f/17259/124I%20PETCT%20in%20Thyroid%20Cancer.pdf (accessed on 17-8-2025).
[17] Jentzen W, et al. Optimized 124I PET dosimetry protocol for radioiodine therapy of differentiated thyroid cancer. J Nucl Med 2008;49(6):1017–23.
[18] Boellaard R, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. Eur J Nucl Med Mol Imaging 2015;42(2):328–54.
[19] Conti M, Eriksson L. Physics of pure and non-pure positron emitters for PET: a review and a discussion. Eur J Nucl Med Mol Imaging Phys 2016;3(1):8.
[20] Jentzen W, Freudenberg L, Bockisch A. Quantitative imaging of (124)I with PET/CT in pretherapy lesion dosimetry. Effects impairing image quantification and their corrections. Q J Nucl Med Mol Imaging 2011;55(1):21–43.
[21] Lubberink M, Herzog H. Quantitative imaging of 124I and 86Y with PET. Eur J Nucl Med Mol Imaging 2011;38 Suppl 1(Suppl 1):S10–8.
[22] Kersting D, et al. Silicon-photomultiplier-based PET/CT reduces the minimum detectable activity of iodine-124. Sci Rep 2021;11(1):17477.
[23] Robinson B, et al. Characterization of tumor size changes over time from the phase 3 study of lenvatinib in thyroid cancer. J Clin Endocrinol Metab 2016;101(11):4103–9.
[24] Masaki C, et al. Lenvatinib induces early tumor shrinkage in patients with advanced thyroid carcinoma. Endocr J 2017;64(8):819–26.
[25] Kersting D, et al. Comparing lesion detection efficacy and image quality across different PET system generations to optimize the iodine-124 PET protocol for recurrent thyroid cancer. EJNMMI Phys 2021;8(1):14.
[26] Jentzen W, et al. Iodine-124 PET dosimetry in differentiated thyroid cancer: recovery coefficient in 2D and 3D modes for PET(/CT) systems. Eur J Nucl Med Mol Imaging 2008;17(3):R161–72.
[27] EANM. EARL PET/CT Accreditation User Manual. Version 4.2. 2023; Available from: https://earl.eanm.org/wp-content/uploads/2023/05/EARL_Manual_4.2.pdf (accessed on 25-3-2025).
[28] Kaalep A, et al. Feasibility of PET/CT system performance harmonisation for quantitative multicentre (89)Zr studies. EJNMMI Phys 2018;5(1):26.
[29] Rausch I, et al. Performance evaluation of the Vereos PET/CT system according to the NEMA NU2-2012 standard. J Nucl Med 2019;60(4):561–7.
[30] Kennedy JA, et al. An extended bore length solid-state digital-BGO PET/CT system: design, preliminary experience, and performance characteristics. Eur J Nucl Med Mol Imaging 2024;51(4):954–64.
[31] Lee YS, et al. Spatial resolution and image qualities of Zr-89 on siemens biograph TruePoint PET/CT. Cancer Biother Radiopharm 2015;30(1):27–32.
[32] Dotinga M, et al. Reinducing radioiodine-sensitivity in radioiodine-refractory thyroid cancer using lenvatinib (RESET): Study protocol for a single-center, open label phase II trial. Diagnostics (Basel) 2022;12(12).
[33] Dotinga M., et al. No reinduction of clinically relevant radioiodine uptake after lenvatinib treatment in radioidine-refractory differentiated thyroid cancer. Eur J Nucl Med Mol Imaging. 2025 Dec 16. doi: 10.1007/s00259-025-07662-9. Epub ahead of print. PMID: 41398087.
[34] Boellaard R, et al. EARL procedure for assessing PET/CT system specific patient FDG activity preparations for quantitative FDG PET/CT studies. 2013; Available from: https://eanm-earl-wordpress.esh.netkey.at/wp-content/uploads/2021/04/EARL-procedure-for-optimizing-FDG-activity-for-quantitative-FDG-PET-studies_version_1_1.pdf (accessed on 25-03-2025).
[35] Soderlund AT, et al. Beyond 18F-FDG: characterization of PET/CT and PET/MR scanners for a comprehensive set of positron emitters of growing application–18F, 11C, 89Zr, 124I, 68Ga, and 90Y. J Nucl Med 2015;56(8):1285–91.
[36] Gregory RA, et al. Optimization and assessment of quantitative 124I imaging on a Philips Gemini dual GS PET/CT system. Eur J Nucl Med Mol Imaging 2009;36(7):1037–48.
[37] Preylowski V, et al. Is the image quality of I-124-PET impaired by an automatic correction of prompt gammas? PLoS One 2013;8(8):e71729.
[38] Jentzen W. Experimental investigation of factors affecting the absolute recovery coefficients in iodine-124 PET lesion imaging. Phys Med Biol 2010;55(8):2365–98.
[39] Braad PE, et al. PET imaging with the non-pure positron emitters: (55)Co, (86)Y and (124)I. Phys Med Biol 2015;60(9):3479–97.

[40] Kertész H, et al. Positron range in combination with point-spread-function correction: an evaluation of different implementations for [124I]-PET imaging. EJNMMI Phys 2022;9(1):56.
[41] Kertész H, et al. Implementation of a spatially-variant and tissue-dependent positron range correction for PET/CT imaging. Front Physiol 2022;13:818463.
[42] Jentzen W, et al. Segmentation of PET volumes by iterative image thresholding. J Nucl Med 2007;48(1):108–14.
[43] Gear JI, et al. EANM practical guidance on uncertainty analysis for molecular radiotherapy absorbed dose calculations. Eur J Nucl Med Mol Imaging 2018;45 (13):2456–74.
[44] Saldarriaga Vargas C, et al. An international multi-center investigation on the accuracy of radionuclide calibrators in nuclear medicine theragnostics. EJNMMI Phys 2020;7(1):69.
[45] Kersting D, et al. Quantification performance of silicon photomultiplier-based PET for small (18)F-, (68)Ga- and (124)I-avid lesions in the context of radionuclide therapy planning. Phys Med 2023;114:103149.
[46] Kist JW, et al. Calibration of PET/CT scanners for multicenter studies on differentiated thyroid cancer with 124I. EJNMMI Res 2016;6(1):39.
[47] Carnegie-Peake L, et al. Quantification and dosimetry of small volumes including associated uncertainty estimation. EJNMMI Phys 2022;9(1):86.