



Original paper

Optimization of CT pulmonary angiography for pulmonary embolism using task-based image quality assessment and diagnostic reference levels: A multicentric study

Anaïs Viry^{a,*}, Veronika Vitzthum^a, Pascal Monnin^a, Julie Bize^a, David Rotzinger^b, Damien Racine^a

^a Institute of Radiation Physics, Lausanne University Hospital and University of Lausanne, Rue du Grand-Pré 1, 1007 Lausanne, Switzerland

^b Department of Diagnostic and Interventional Radiology, Lausanne University Hospital and University of Lausanne, Rue du Bugnon 46, 1011, Lausanne, Switzerland

ARTICLE INFO

Keywords:

Diagnostic reference levels
Image quality
Angiography CT
Optimization
Pulmonary embolism
Model observers

ABSTRACT

Purpose: To establish size-specific diagnostic reference levels (DRLs) for pulmonary embolism (PE) based on patient CT examinations performed on 74 CT devices. To assess task-based image quality (IQ) for each device and to investigate the variability of dose and IQ across different CTs. To propose a dose/IQ optimization.

Methods: 1051 CT pulmonary angiography dose data were collected. DRLs were calculated as the 75th percentile of CT dose index (CTDI) for two patient categories based on the thoracic perimeters. IQ was assessed with two thoracic phantom sizes using local acquisition parameters and three other dose levels. The area under the ROC curve (AUC) of a 2 mm low perfused vessel was assessed with a non-prewhitening with eye-filter model observer. The optimal IQ-dose point was mathematically assessed from the relationship between IQ and dose.

Results: The DRLs of CTDI_{vol} were 6.4 mGy and 10 mGy for the two patient categories. 75th percentiles of phantom CTDI_{vol} were 6.3 mGy and 10 mGy for the two phantom sizes with inter-quartile AUC values of 0.047 and 0.066, respectively. After the optimization, 75th percentiles of phantom CTDI_{vol} decreased to 5.9 mGy and 7.55 mGy and the interquartile AUC values were reduced to 0.025 and 0.057 for the two phantom sizes.

Conclusion: DRLs for PE were proposed as a function of patient thoracic perimeters. This study highlights the variability in terms of dose and IQ. An optimization process can be started individually and lead to a harmonization of practice throughout multiple CT sites.

1. Introduction

Computed tomography (CT) pulmonary angiography (CTPA) is currently the standard of care to quickly and accurately diagnose pulmonary embolism (PE) [1]. The number of CTPA examinations has considerably increased in the last fifteen years [2,3]. In Switzerland, a recent survey showed that despite the increased number of CT examinations, the effective dose has decreased due to efforts to optimize clinical protocols [2].

One tool for optimizing protocols is the use of diagnostic reference levels (DRLs) which was proposed by the International Commission on

Radiological Protection (ICRP) [4]. ICRP recommends that CT users optimize their protocols when the 50th percentile of the local dose indicator distribution systematically exceeds the corresponding national DRL [5]. ICRP also specifies that DRLs should be designed for specific clinical indications and not for anatomical regions, since different protocols and exposure levels could be required. Several efforts have been made to follow this recommendation and establish clinical indication-based DRLs [6,7]. Secondly, national DRLs are determined by calculating the 75th percentile of the dose distribution for a sample of patients close to a standard size [8]. It could be also helpful to optimize CT parameters for patients with different body builds, as proposed on

Abbreviations: DRL, Diagnostic reference levels; PE, Pulmonary embolism; CTPA, Computed tomography pulmonary angiography; IQ, Image quality; AUC, Area under the curve; HU, Hounsfield units; ICRP, International Commission on Radiological Protection; CTDI, Computed tomography dose index; DLP, dose length product; NPWE, Non-Prewhitening observer with eye-filter; TTF, Target transfer function; NPS, Noise Power Spectrum; IQR, Interquartile range; ROI, Region of interest; CM, Contrast medium.

* Corresponding author at: Rue du grand-pré 1, 1007 Lausanne, Switzerland.

E-mail address: anaïs.viry@chuv.ch (A. Viry).

<https://doi.org/10.1016/j.ejmp.2024.103365>

Received 20 November 2023; Received in revised form 12 February 2024; Accepted 19 April 2024

Available online 25 April 2024

1120-1797/© 2024 Published by Elsevier Ltd on behalf of Associazione Italiana di Fisica Medica e Sanitaria.

modern CT systems with automatic tube current modulation and tube voltage selection. In that case, it seems to be important to create size-specific DRLs[9], especially for pulmonary CTs. Several classifications of patients size have been used for this purpose, based on patient weight, BMI or water equivalent diameter [10,11]. Moreover, it is important that DRLs should not be assumed as dose limits for patients, especially for large-sized patients. Thirdly, DRLs are not directly linked to clinical image quality (IQ) criteria, which is still a topic for improvement [12].

Assessing IQ and patient exposure is crucial to ensure that a dose reduction still provides adequate IQ for an accurate diagnosis[13,14]. To our knowledge, there has been no previous analysis that jointly examines IQ and DRLs in the context of CTPA, specifically to ensure appropriate patient dose reduction.

For this purpose, qualitative IQ analysis on patient images by radiologists remains the gold standard for assessing diagnostic performance of CT protocols. However, it is a relatively time-consuming process that requires many observers. Moreover, traditional metrics such as contrast-to-noise ratio or standard deviation are not directly linked with diagnostic performances [15–18] and are not longer available using iterative algorithms [19,20]. Furthermore, it is not possible to conduct an experimental dynamic optimization process on patients to reduce step by step dose and compare IQ. Therefore, phantom studies are often used as a standard for optimizing both dose and IQ in CT [21,22] using a task-based paradigm [23]. Nevertheless, two questions remain: how representative are these phantom studies in comparison to patient cohort studies and what is the potential optimization that can be achieved.

Hence, the purpose of this work is to first establish size-specific regional DRLs using an innovative patients' classification for CTPA examinations conducted for suspected PE and performed on 74 CT devices. Secondly, phantom task-based IQ was assessed for each device and the variability of dose and IQ across different CTs was analysed. The agreement between phantom and patient doses was also studied. Finally, an optimization process was proposed on phantoms to ensure that the IQ is sufficient for an accurate diagnosis.

2. Materials and methods

2.1. Diagnostic reference levels

The regional DRLs for CTPA examinations performed for PE suspicion were investigated. Data was collected from 70 radiological institutes, which operated a total of 74 different CT devices. The dataset included 1051 examinations of adult patients (age older than 16). The four major manufacturers were represented, with a majority of GE Healthcare CT (47 %) (see Table 1). The patient size for each scan exam was defined as the thoracic perimeter calculated from the axial section at the level of the tracheal carina. The antero-posterior length (AP) was measured from the edge of the pectoral skin to the back, while the lateral length (Lat) was measured from the left edge to the right edge of the patient. The equivalent thoracic perimeter P was then determined considering an ellipsoidal thoracic shape using formula (1):

$$P = \pi \sqrt{0.5(AP^2 + Lat^2)} \quad (1)$$

The median of the equivalent thoracic perimeter distribution was used to separate the patients into the two categories medium and large-sized patients. The ICRP recommends determining the DRL as the 3rd quartile of the distribution of median doses for each institute [5]. Due to the limited number of scans collected from each radiological institute, with an average of 14 exams per CT, DRLs were calculated for the two groups as the third quartile (75th percentile) of the dose indicator distribution for all the radiology institutes. Regional DRLs were also calculated without patient size stratification and compared with national DRLs published in 2018. [8]. Since 87 % of the scans were performed using only an arterial phase, regional DRL (CTDI_{vol} and DLP) were calculated

Table 1

Number of CT systems involved in the study sorted by manufacturer and model.

| Manufacturer | Model | Number of CTs |
|---------------|-------------------------|---------------|
| Canon | Activion 16 | 1 |
| | Aquillion 64 | 3 |
| | Aquillion Prime | 5 |
| GE Healthcare | Brightspeed | 1 |
| | Lightspeed | 4 |
| | Optima CT 660 | 8 |
| | Revolution Apex | 2 |
| | Revolution EVO | 8 |
| | Revolution Frontier | 5 |
| | Revolution HD | 7 |
| | Brilliance 64 | 3 |
| Philips | iCT 256 | 2 |
| | iCT 6000 | 1 |
| | Incisive 128 | 3 |
| | Ingenuity 128 | 6 |
| | Ingenuity 32 | 1 |
| | Ingenuity 5000 | 1 |
| | iQon | 6 |
| | Definition AS | 1 |
| Siemens | Somatom Definition Edge | 1 |
| | Somatom Go up | 4 |
| | Somatom Force | 1 |
| | Total | 74 |

using only this phase and the cumulative DLP were calculated using all the phases.

2.2. Phantom acquisition

An anthropomorphic thoracic phantom (QRM, A PTW COMPANY) was used to assess the IQ of CTPA examinations for the search of PE. The phantom mimics different types of tissues, including lung, soft tissue, and vertebrae. To simulate larger patients, an additional ring with a thickness of 2.5 cm made of adipose-equivalent tissues was added. The perimeter of the two phantom sizes were 95 cm and 111 cm, respectively.

Three homemade PMMA modules were successively inserted into the thoracic phantom to measure contrast-dependent spatial resolution, noise and contrast (Fig. 1). The “contrast module” contains six 10 mm holes filled with iodinated contrast media solutions ranging from 0 to 6 mg/mL (0.0, 0.5, 1.0, 2.0, 4.0, and 6.0 mg/mL). The “spatial resolution module” is a PMMA rod (100 HU at 100 kVp) with a 5 cm-diameter central hole filled with an iodinated contrast media solution at 6.9 mg/mL (250 HU at 100 kVp). The transition of 150 HU between PMMA and iodine creates a sharp edge used to calculate the contrast-dependent spatial resolution. These solutions simulate the enhancement of different structures during CTPA. The “noise module” is homogeneous and was used to calculate the noise of the CT images.

Each phantom size was scanned with local settings (acquisition and reconstruction parameters) of the CTPA protocol on the 74 CT devices (Table 1). Tube voltage varied between CTPA protocols from 80 to 120 kVp. Tube current modulation has always been used in one or three directions, depending on CT technology. Images were reconstructed using the soft kernel designed for vessel visualization with a slice thickness between 1.25 and 2.5 mm. Iterative or deep-learning algorithms have been used to reconstruct the images. The displayed field of view was arbitrary set at 370 mm and 420 mm for the medium and large phantom, respectively. For each phantom size, the acquisitions were repeated two times to obtain enough data for the calculation of IQ metrics. The displayed CTDI_{vol} obtained using the local acquisition settings is referred to as the “local phantom dose”. The accuracy of the displayed CTDI is a part of the CT annual quality control, carried out by the manufacturer. This measurement is validated by the medical physicist and checked by the Swiss safety authority for approbation. The standard IEC 61223–3–5 recommends a maximum acceptable difference between the displayed and measured CTDI_w of 20 %. This tolerance is

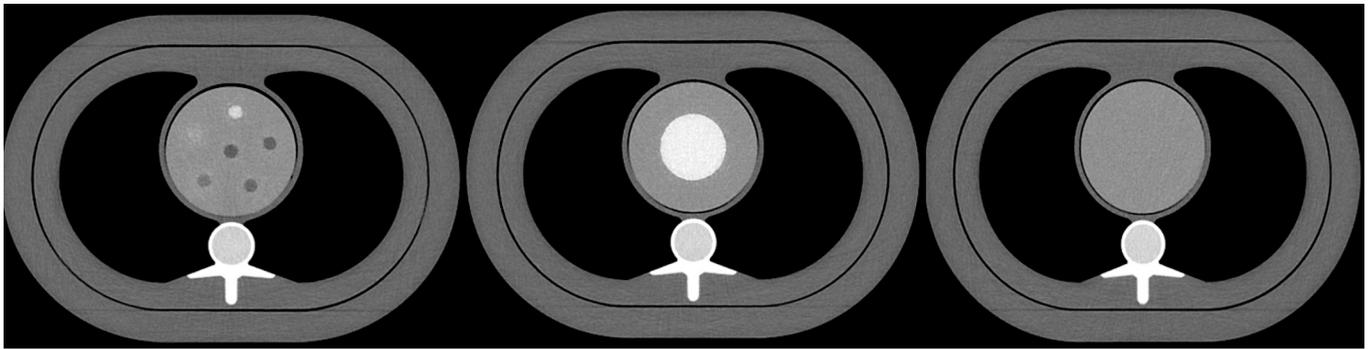


Fig. 1. CT image of the medium-sized thoracic phantom with (Left) the “contrast module”, (Middle) the “spatial resolution module” and (Right) the “noise module”.

respected for all CT devices involved in our study. The displayed CTDI value was used for the calculation of the size-specific DRLs as well as for the dose calculation in the phantom study.

For both phantom sizes, the distribution of local phantom doses across the CT systems was described using the 50th and the 75th percentile. The agreement between phantom and patients was evaluated by comparing the 50th and 75th percentile of the local phantom dose and patient dose distributions. For each category (medium and large), a non-parametric Mann-Whitney U test was performed.

In addition to the local settings phantoms were also scanned at three different dose levels around the local phantom dose by adjusting the settings of the automatic exposure control system in order to assess a potential of optimization for each CT. All other settings as tube voltage, pitch, rotation scan, collimation and reconstruction parameters were kept constant.

2.3. Image quality metrics

We investigated a clinically relevant task in CTPA, that is the contrast between a well-opacified macroscopic vessel with a 2 mm-diameter and a perfusion defect in the case of PE. This task aims to match the clinical question addressed by the imaging test: the absence or exclusion of PE. The presence, anatomic localization, clot burden, occlusive or non-occlusive nature of the thrombus will help determine treatment, follow-up, and allow for risk stratification [24]. One of the most challenging task performed by radiologists when reading CTPAs is to detect macroscopic circulating lumens with or without a perfusion defect. Subsegmental vessels have diameters ranging from 6 mm centrally to approximately 2 mm at the periphery. Smaller vessels (arterioles) are usually not detectable on CTPA [25].

To evaluate IQ of this task, a non-prewhitening model observer with eye filter (NPWE) [26] was used, based on the computation of contrast, spatial resolution and noise, calculated using an in-house software developed in IgorPro 9.

2.3.1. Contrast

On each CT and for each dose level the contrast was calculated as the mean difference over 50 slices between the signal measured in the two contrast media solutions at 2 and 4 mg/mL in the “contrast module”. The contrast media solution at 4 mg/mL represents a well opacified vessel. The concentration at 2 mg/mL represents a defect of perfusion.

2.3.2. Contrast-dependent spatial resolution

In the “spatial resolution module”, the edge between the iodine rod and the surrounding PMMA was used to measure the target transfer function (TTF), which is commonly accepted to assess contrast-dependent spatial resolution when dealing with iterative reconstruction algorithms [23]. Regions of interest (ROI) of 65x65 mm² centered in the iodine rod were extracted from 75 CT slices of the module in each acquisition. The circular iodine-PMMA edge was used to calculate

angular edge spread functions (ESF) by steps of 10 degrees over an angular aperture of 20 degrees. The angular TTFs were computed by applying the Fourier transform to the angular ESFs. The resulting mean TTF was obtained using a radial mean normalized to 1.0 at the zero-frequency.

2.3.3. Noise power spectrum (NPS)

We calculated the NPS using the methodology outlined in the report 87 from the International Commission on Radiation Units and Measurements [27]. The NPS is the state-of-the-art methodology used to evaluate both the magnitude and the texture of the noise in an image. It was calculated by using 400 ROIs of 60x60 mm² centered in the “homogeneous” module in 200 CT slices.

2.3.4. Non-prewhitening model observer with eye filter

The detection between a 2 mm well opacified macroscopic vessel and a perfusion defect is based on the calculation of the detectability index given by a NPWE model (Equation (2)). Contrast, spatial resolution and noise power spectrum were assessed using the various modules. The task was simulated using a 2 mm cylinder. In order to take into account the ability of a human observer to visualize some spatial frequencies, the virtual transfer function described by Samei and related to a CT experiment was used [25].

$$d'_{NPWE} = \sqrt{2\pi} [\Delta HU] \frac{\int_0^{f_{ny}} S^2(f) TTF^2(f) VTF^2(f) f df}{\sqrt{\int_0^{f_{ny}} S^2(f) TTF^2(f) NPS(f) VTF^4(f) f df}} \quad (2)$$

where $|\Delta HU|$ is the absolute value of the contrast measured using the “contrast module” between the two iodine concentrations at 4 mg/mL (representing well-opacified structures) and 2 mg/mL (representing hypo-perfused structures), f the radial spatial frequency, f_{ny} the Nyquist frequency and TTF and NPS were previously described. S represents the circular section of cylindrical vessel with a diameter of 2 mm [26] and VTF is the visual transfer function of the human eye [28]. To compute the detectability index, interpretation configuration was set with a zoom factor of 1.5, a 500 mm viewing distance and a field of view of 370 or 420 mm, depending on the phantom sizes. Unfortunately, model observers still lack recognized methodology to accurately evaluate the uncertainties, so only the mean value has been calculated. At the end, a monotonic function was used to link the detectability with the area under the receiver operating characteristics curve (AUC) [29]. AUC values typically ranged from 0.5 similar to a hazard detection, to 1 similar to a perfect detection. An AUC value above 0.8 is generally considered as good and a value above 0.9 is considered as excellent. AUC was used as the figure of merit to assess the detectability of macroscopic vessels.

2.4. Optimisation process: Trade-off between image quality and radiation dose

The medians of local phantom dose and corresponding AUC values enabled the classification of CTs into four quadrants based on their dose-IQ results:

- Q1: CT systems with a local phantom dose lower and an IQ higher than their respective medians
- Q2: CT systems with a local phantom dose higher and an IQ higher than their respective medians.
- Q3: CT systems with a local phantom dose lower and an IQ lower than the medians.
- Q4: CT systems with a local phantom dose higher and an IQ lower than their respective medians.

To propose an optimization process, the four AUC values obtained for each CT were fitted using a linear relationship between the logarithm of the dose and the logarithm of the detectability index. The theoretical optimum between dose and IQ can be mathematically described as the point where the curve has a maximum curvature, the “elbow” of the curve. At this point, a dose reduction would lead to an important loss of IQ and a dose increase would not lead to a major improvement of IQ. Interquartile range of the dose and AUC distributions was also calculated for each phantom size before and after optimization to quantify the impact of this proposed optimization.

3. Results

3.1. Diagnostic reference levels and phantom doses

The median of the equivalent thoracic perimeter distribution was 102 cm. The category “medium-sized patients” included 530 patients with a perimeter ranging from 75 cm to 102 cm (minimum and maximum values) and the category “large patients” included 557 patients with a perimeter ranging from 102 to 152 cm. The median values were 95 and 111 cm for the two categories, respectively.

The 50th and 75th percentile of the CTDI_{vol} and DLP for CTPA examinations are presented in Table 2 for the medium and large-sized patients. As expected, the dose values for medium-sized patients are lower than those for large-sized patients. 75th percentile of CTDI_{vol} and DLP values increased by approximately 60 % when comparing medium-sized patients to large-sized. The 75th percentile of CTDI_{vol} and DLP for all patient sizes for the arterial phase were 8.5 mGy and 320 mGy.cm.

The distribution of phantom and patient CTDI_{vol} values for the different CT systems is shown in Fig. 2 for the medium and the large size category. Due to the variability between CT devices and CTPA protocols, CTDI_{vol} values for the medium phantom size ranged from 1.9 to 13.1 mGy, with a median value of 4.3 mGy and 75th percentile of 6.3 mGy. For the large phantom size, CTDI_{vol} values ranged from 4.3 to 21.6 mGy

Table 2
75th percentile and 50th percentile of CTDI_{vol}, DLP for the arterial phase and cumulative DLP obtained for the two categories of patients.

| | | All patient sizes | Medium-sized patients | Large-sized patients |
|-----------------------|---------------------------|-------------------|-----------------------|----------------------|
| CT angiography | Number of CTPA analyzed | 1051 | 471 | 580 |
| 75th percentile (DRL) | CTDI _{vol} (mGy) | 8.5 | 6.4 | 10 |
| | DLP (mGy.cm) | 320 | 240 | 385 |
| | Cumulative DLP (mGy.cm) | 360 | 270 | 412 |
| 50th percentile | CTDI _{vol} (mGy) | 6.1 | 4.7 | 7.4 |
| | DLP (mGy.cm) | 230 | 181 | 278 |
| | Cumulative DLP (mGy.cm) | 250 | 202 | 304 |

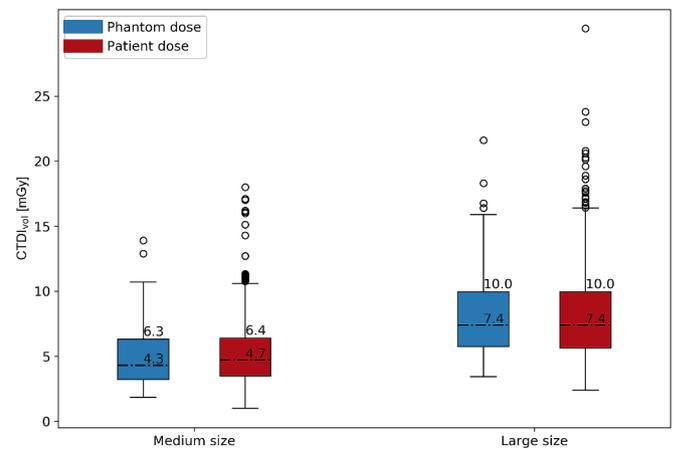


Fig. 2. Boxplots of CTDI_{vol} values obtained on each CT using local settings for CTPA protocol for the two patient categories and phantom sizes. Median and 75th percentile values were indicated in each boxplot.

with a median value of 7.4 mGy and a 75th percentile of 10.0 mGy. In comparison, the CTDI_{vol} values for medium-sized patients ranged from 1.0 mGy to 18.0 mGy, with a median value of 4.7 mGy and a 75th percentile of 6.4 mGy. For large-sized patients, CTDI_{vol} values ranged from 2.4 to 30.2 with a median value of 7.4 mGy and a 75th percentile of 10.0 mGy. The median perimeters of the two phantom sizes correspond exactly to the median perimeter of the two patient categories. Typical values (median and 75th percentile) for phantom and patient doses were statistically comparable ($p = 0.101$ for medium size and $p = 0.338$ for large size using a Mann-Whitney test).

3.2. Image quality and phantom doses

The relationship between IQ and radiation dose level was plotted for the 74 CT devices in Fig. 3 for the two phantom sizes. AUC values ranged from 0.821 to 0.999 with a median value of 0.980 and an interquartile range (IQR) of 0.047 for the medium-sized phantom. For the large phantom size, AUC values ranged from 0.766 to 0.998 with a median value of 0.963 and an IQR of 0.066. For the medium-sized phantom, 13, 23, 24 and 14 CT devices had their dose-IQ relationship located in the four quadrants Q1, Q2, Q3 and, Q4 respectively. For the large phantom size 12, 24, 25 and 13 CTs had their dose-IQ relationship in the quadrants Q1, Q2, Q3 and Q4, respectively. It is interesting to note that 20 CT devices had their dose-IQ relationship located in different quadrants for the medium and the large phantom size. As the automatic tube current modulation is activated during the acquisition, the CTDI_{vol} changed with the phantom size. In that case median CTDI_{vol} increased by 72 % whereas median AUC decreased by 1.7 % when comparing medium-sized to large-sized phantom.

3.3. Optimization process: a trade-off between image quality and radiation exposure

IQ values (AUC) were plotted as a function of CTDI_{vol} for the 74 CT systems to assess the variation in IQ at different dose levels. The two specific scanners shown as examples in Fig. 4 were selected based on their extreme IQ values, i.e. the CTs with the lowest IQ (referred as CT1) and the highest IQ (referred as CT2). A similar analysis was performed for the other 72 CTs (Detailed results for the other CT systems are not shown). For CT1, the AUC was 0.82 for a CTDI_{vol} of 2.6 mGy for the medium-sized phantom and 0.77 for a CTDI_{vol} of 3.8 mGy for the large-sized phantom. After optimization, the theoretical optimum increased the AUC by 15 % and 14 % (AUC = 0.95 and AUC = 0.88), and the CTDI_{vol} by 61 % and 21 % (CTDI_{vol} = 6.8 mGy and CTDI_{vol} = 8.4 mGy) for the medium and large phantom sizes, respectively. For CT2, the

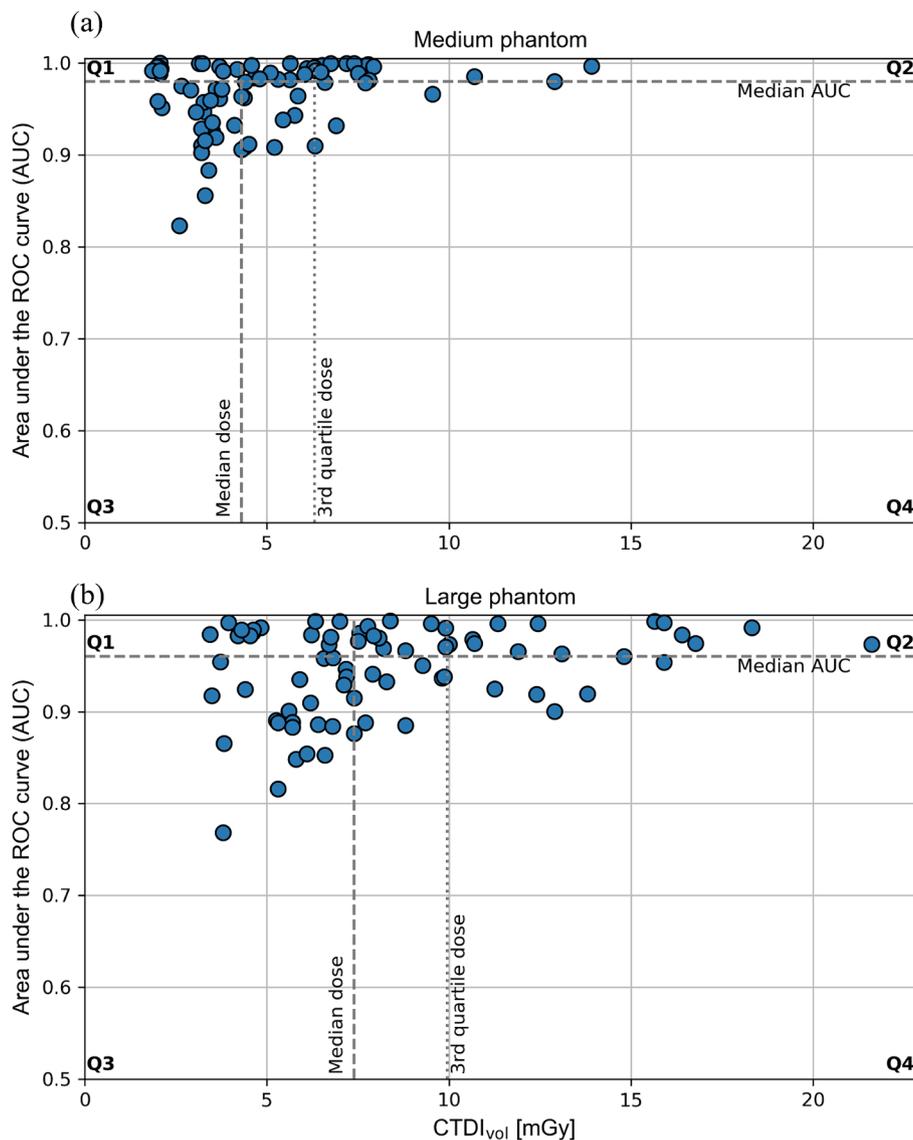


Fig. 3. Relationship between dose and task-based IQ assessed on the two phantom sizes (a) medium size and (b) large size for the various CT systems. Each blue dot represents the result obtained for a CT in a radiological institution using the local settings of CTPA protocol for suspected PE. Four quadrants were defined using the median of $CTDI_{vol}$ and the median of AUC values.

proposed optimal pair decreased the $CTDI_{vol}$ by 48 % and 53 % ($CTDI_{vol} = 2.9$ mGy and $CTDI_{vol} = 5.3$ mGy) for the medium-sized and large-sized phantom. As the AUC is already close to 1, the proposed shift of dose decreased IQ only by 0.02 % and 2 % (AUC = 0.99 and AUC = 0.97) for the medium-sized and large-sized phantoms, respectively.

The variability of IQ and dose level if the theoretical optimum was applied on each CT was plotted for the two phantom sizes in Fig. 5. The medium-sized phantom, $CTDI_{vol}$ would range from 1.6 to 7.4 mGy. In comparison with Fig. 3, the median dose value would increase from 4.3 to 4.9 mGy whereas the IQR would reduce from 3.1 mGy to 2.15 mGy. AUC values would range from 0.918 to 0.999, with a median value of 0.979 and the IQR would decrease from 0.047 to 0.025. For the large phantom size, the range of $CTDI_{vol}$ values would range from 3.3 to 8.8 mGy. The median dose value would decrease from 7.4 to 7.0 mGy and the IQR from 4.2 to 1.85 mGy. AUC values would range from 0.846 to 0.994 with a median value of 0.940 and the IQR would decrease from 0.066 to 0.057.

For the medium-sized phantom, the number of CT devices which had their dose-IQ relationship located in the quadrants Q1, Q2, Q3 and Q4 were 35, 1, 2 and 36, respectively. For the large-sized phantom 33, 3, 4

and 34 CT devices were located in the quadrants Q1, Q2, Q3 and Q4, respectively. Hence, the optimization process could homogenize the dose and IQ between the CT protocols and emptied the quadrants Q2 and Q3. The number of CT systems which a dose reduction could be proposed was 33 for the medium-sized phantom and 48 for the large-sized phantom. The mean percentage of dose reduction was 26 % and 30 % for the medium and large-sized phantom, respectively.

4. Discussion

Regional DRLs for PE suspicion were proposed in this study for two categories of patient size. The Swiss DRLs of $CTDI_{vol}$ and DLP published in 2018 (8.0 mGy and 300 mGy.cm, respectively) [8] are consistent with our values 8.5 mGy and 320 mGy.cm, calculated as the 75th percentile of $CTDI_{vol}$ and DLP for the arterial phase for all patients without respect of patient sizes. These results ($CTDI_{vol}$, DLP per phase and total DLP) are also consistent with clinical DRLs based on data obtained from nineteen european hospitals, published by Damalakis et al, which were 9 mGy for the $CTDI_{vol}$, 307 mGy.cm for the DLP per phase and 364 mGy.cm for the total DLP [30]. However, this study only considers patients with a

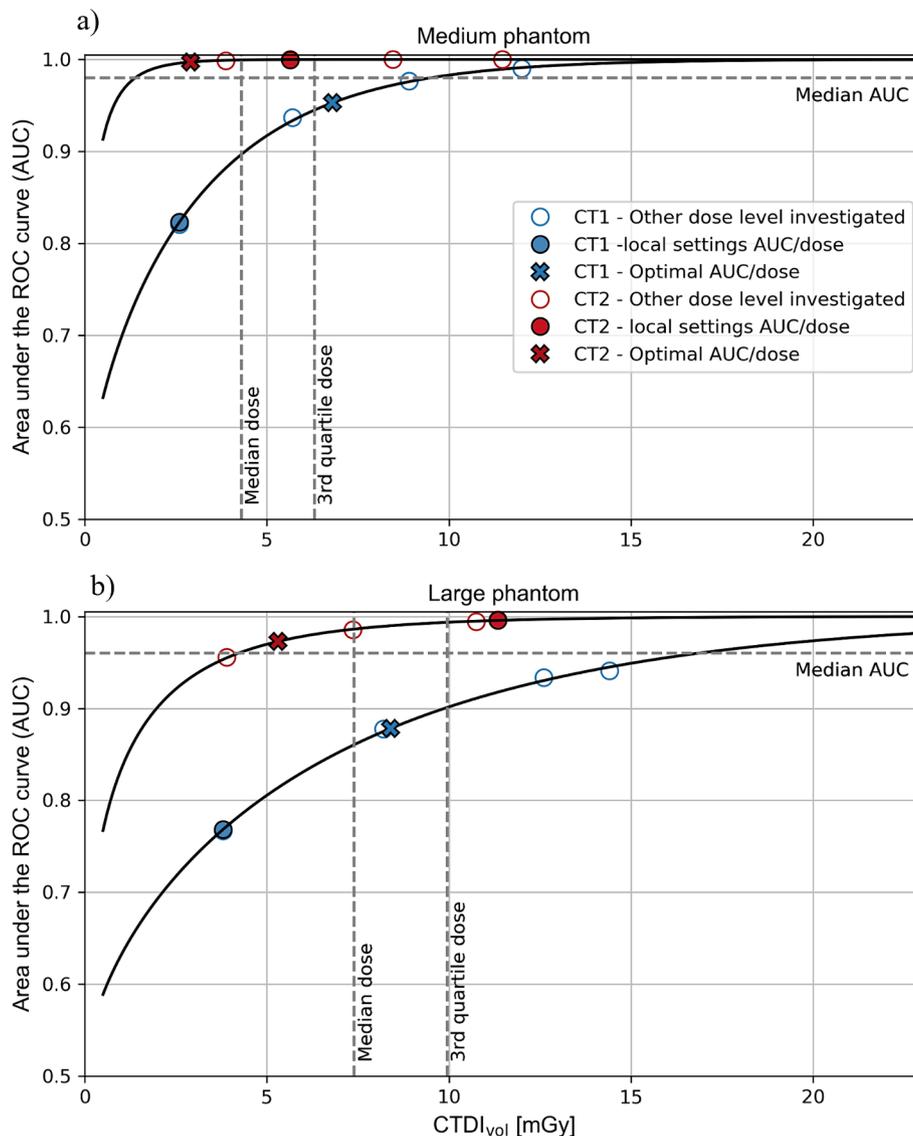


Fig. 4. IQ (expressed in terms of AUC) of two CT devices as a function of the radiation dose level (CTDI_{vol}) for (a) medium phantom size and (b) large phantom size. The filled points represent the pair (dose/AUC) obtained with local acquisition settings. The empty circles represent the supplementary dose levels investigated. The two crosses represent the optimal relationship (dose/AUC) obtained at the maximum curvature of the fitted curve.

weight interval of 70 ± 15 kg. Compared to the recent international DRL study published by Bos et al [31], our dose values are considerably lower than those related to the US practice, which were 14.9 for the CTDI_{vol} and 594 mGy.cm for the total DLP. However, our values were slightly lower than their DRLs calculated using data from four western european countries, which were 5.5 mGy for the CTDI_{vol} and 206 mGy.cm for the DLP. However, this study includes only large hospitals using dose management software. Our data comes from various radiological institutes such as hospitals, clinics, and small medical imaging centers, which can introduce a heterogeneity of practice and a large variety of CT generations. Recently, CT protocols adapted to the patient's size have been increasingly used, dealing with both tube voltage selection and tube current modulation techniques. Precise adjustment of parameter settings requires the use of size-specific DRLs. Hence, our classification based on the thoracic perimeter provides a reference for optimizing CT protocols related to various patient morphologies. Several publications have already proposed size-specific dose reference levels (DRLs) as a function of water equivalent diameter or body mass indexes (BMI) for CTPA protocols. [10,32,33]. However, water equivalent diameter can be challenging to access as it typically requires the use of dose management

software, and this information is generally only available after the examination. BMI is not really representative of the attenuation of a given anatomical localization, such as thoracic region. The thoracic perimeter can be easily measured on patients before the examination. The comparison between the planned dose level and our size-specific DRLs can provide a useful reference before the acquisition.

Generally, DRLs are not linked with IQ criteria, which should be considered a priority when optimizing imaging protocols. To overcome this limitation it is possible to use NPWE model observer to assess task-based IQ in CT [34,26,35] because this kind of model has shown good agreement with human observers [36]. Before the optimization process, all CT systems performed quite well for this task, with AUC values upper to 0.821 for the medium phantom size and 0.766 for the large phantom size, however the IQR was relatively high. This highlights substantial variation in IQ among CT systems, resulting in a difference in diagnostic information contained in clinical images. A previous similar study for abdominal protocols demonstrated similar variation in IQ [37]. The variability should be due to the differences between CT devices, acquisition, and reconstruction parameters for CTPA protocols.

The agreement between phantom and patient attenuation is rarely

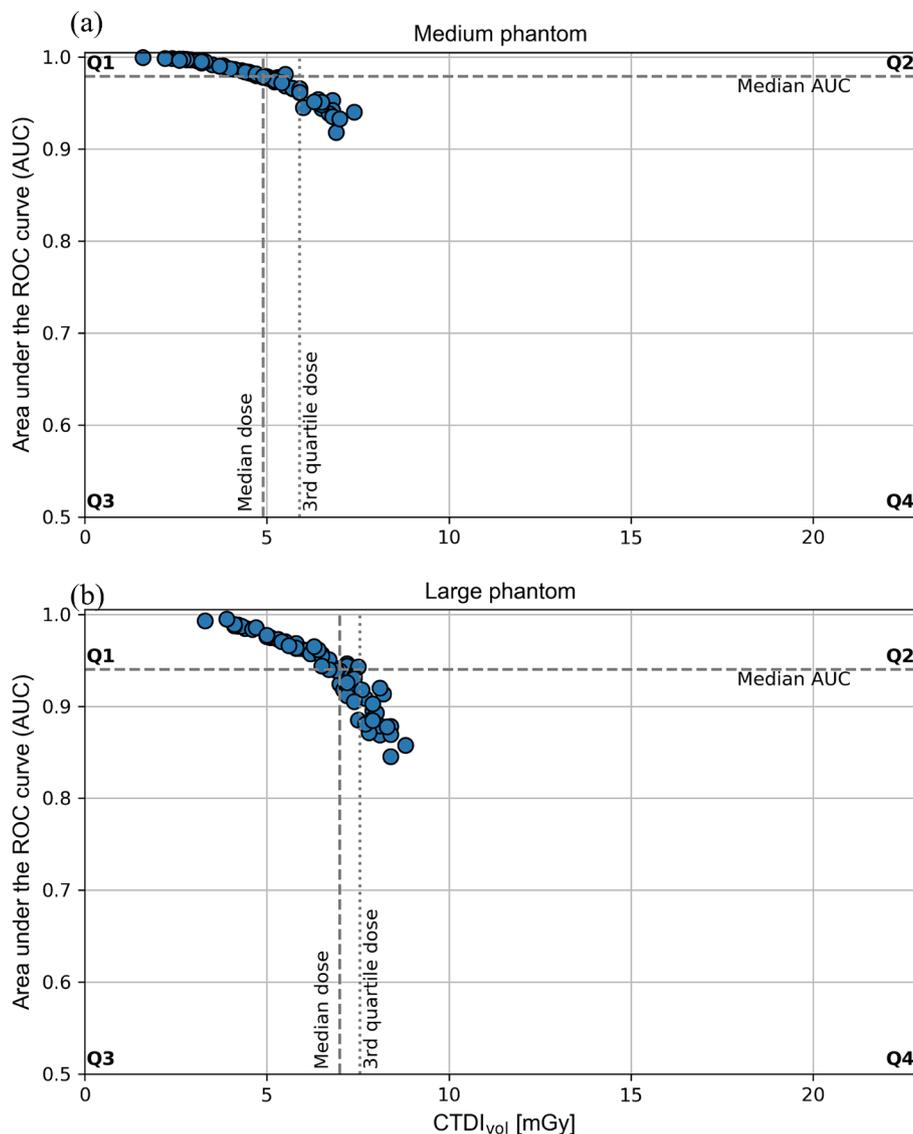


Fig. 5. Relationship between dose and task-based IQ assessed on the two phantom (a) medium size and (b) large size for the various CT systems after the proposed optimization process, where each blue dot represents the optimal relationship (dose/AUC) calculated for each CT.

assessed. Our study demonstrated a successful correlation between patient and phantom doses for the two categories. This phantom appeared to be a valuable tool for evaluating the radiation dose and IQ trade-off in CTPA protocols, even if using more realistic phantom can be a better approach with new algorithms based on artificial intelligence [38,39]. Indeed detecting enhanced vessels in a uniform background is a simple task compared to the complexity of a radiological diagnosis, which requires detecting small filling defects surrounded by CM in anatomical background, often complicated by flow artifacts. This shows the first limitation of our model which is based on static phantom studies where IQ is determined by a task-based model. It does not take into account the flow of CM through the blood and patient movement or breathing. Nevertheless with this approach a fast control of the dose-IQ relation is possible for multiple CTs and gives the possibility of starting an optimization process.

As proposed by ICRP, if the local median dose for standard patients exceeds the national DRL, dose reduction has to be considered. However, if the local median dose is below the 50th percentile value of the national DRL, IQ should be considered the priority for further optimization [40]. In this study, the optimization process was based on the fitted curve between IQ and radiation dose values for each CT device.

The optimal pair of dose and IQ was determined using the maximum curvature of the fitted curve. With this optimization process, the third quartile of radiation dose levels and interquartile values of dose and IQ could be reduced. This means that the proposed methodology could lead to a global harmonization of IQ and dose level across different CT systems. The number of CT devices which had their dose-IQ relationship in the quadrants Q2 and Q3 has drastically decreased. This means that a CT device initially located in the second quadrant (Q2) can mostly be optimized by decreasing the dose level. A system in the third quadrant (Q3) can also be optimized by increasing the dose level in order to improve image quality. Hence, the optimized relationship between dose and IQ for each CT are largely located in the quadrants Q1 and Q4. A CT device initially located in the quadrant Q1 respects the balance between IQ and dose level and an optimization might not be necessary. Optimizing a CT scanner in the fourth quadrant (Q4) may be challenging, since the proposed optimization process does not guarantee that the dose level stays below the DRL. The proposed dose levels were only obtained by modifying the automatic tube current modulation parameters. Also, IQ is dependent on other acquisition or reconstruction parameters such as tube voltage, slice thickness, iterative reconstruction algorithm or kernels. When optimizing a CT scanner, it is essential to

verify if other parameters have been appropriately adjusted to address the clinical question, as proposed by Hendricks and al [41]. In some cases, i.e. with older CT generations, it might be necessary to operate at a dose level slightly higher than the DRL value to deliver a diagnostically useful IQ. This case demonstrates that it is essential to assess IQ when applying DRL in order to ensure an appropriate diagnosis. Further, the optimization process also needs to be evaluated as a function of the patient size when the CT dose-IQ relationships are in different quadrants for the medium and the large phantom sizes. Almost thirty percent of the CT devices were initially in this case. Optimizing one phantom size might lead to a non-optimal dose-IQ relationship for the other phantom size if only one protocol is used for all patient sizes. Hence, either a compromise needs to be found to optimize the IQ-dose relationship for both phantom sizes using a single protocol, or at best two separate protocols need to be established for the two patient categories. In the latter case, the discrimination between patient sizes can be done using the thoracic perimeter as explained above.

This study has several limitations. First, the assessment of uncertainties for IQ on each CT was not evaluated since no reference standard exists to assess the uncertainties of AUC values using a NPWE model observer. The uncertainty on dose levels (CTDI and DLP) was estimated at 20 % due to the fact that this accuracy is required in quality controls. Mathematically, the optimization process chose one optimal pair between IQ and dose. Taking into account the uncertainties, a zone can be defined where the optimal dose-IQ pair is located. This means that around the optimum, an individual optimization process can still give the priority to the dose level or the IQ, depending on patient radiosensitivity and criticality of clinical task. This provides a starting point for the optimization with a theoretical and optimal dose value. It was shown in clinical studies that even high dose reductions can lead to similar and still acceptable image quality [42,43]. Nevertheless care should be taken in reducing the dose level too drastically as it was shown that reduction in dose worsens the readers' subjective assessment of diagnostic confidence and image quality [44]. Each facility should be involved in the optimization process and care needs to be taken to apply clinically the values found in phantom studies. Particularly, a lower patient dose reduction than the theoretical value should never be applied.

5. Conclusion

Specific adult DRLs for suspected PE have been proposed as a function of patient thoracic perimeters. Their application for optimizing CTPA protocols requires the assessment of IQ using anthropomorphic phantoms representative of patient attenuation. The dose and IQ variation between CT systems was highlighted. With this information the number of CTs that can be optimized might be identified and an optimization process is proposed for each CT in order to maximize the balance between IQ and radiation exposure. This provides a starting point for the optimization with an optimal dose value and can lead to global harmonization of IQ for a wide variety of CT systems while still applying the DRL concept.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors gratefully acknowledge the support from the participating radiological institutes in data collection and phantom measurements.

References

- [1] Moore AJE, Wachsmann J, Chamrathy MR, Panjikaran L, Tanabe Y, Rajiah P. Imaging of acute pulmonary embolism: an update. *Cardiovasc Diagn Ther* 2018;8:225–43. <https://doi.org/10.21037/cdt.2017.12.01>.
- [2] Viry A, Bize J, Trueb PR, Ott B, Racine D, Verdun FR, et al. Annual exposure of the swiss population from medical imaging in 2018. *Radiat Prot Dosim* 2021;195:289–95. <https://doi.org/10.1093/rpd/ncab012>.
- [3] Brenner DJ, Hall EJ. Computed tomography — an increasing source of radiation exposure. *N Engl J Med* 2007;357:2277–84. <https://doi.org/10.1056/NEJMra072149>.
- [4] Icrp I. Publication 73: radiological protection and safety in medicine. *Ann ICRP* 1996;26.
- [5] Vaňo E, Miller DL, Martin CJ, Rehani MM, Kang K, Rosenstein M, et al. ICRP publication 135: diagnostic reference levels in medical imaging. *Ann ICRP* 2017;46:1–144. <https://doi.org/10.1177/0146645317717209>.
- [6] Paulo G, Damilakis J, Tsapaki V, Schegerer AA, Repussard J, Jaschke W, et al. Diagnostic reference levels based on clinical indications in computed tomography: a literature review. *Insights Imaging* 2020;11:96. <https://doi.org/10.1186/s13244-020-00899-y>.
- [7] Tsapaki V, Damilakis J, Paulo G, Schegerer AA, Repussard J, Jaschke W, et al. CT diagnostic reference levels based on clinical indications: results of a large-scale European survey. *Eur Radiol* 2021;31:4459–69. <https://doi.org/10.1007/s00330-020-07652-5>.
- [8] Aberle C, Ryckx N, Treier R, Schindera S. Update of national diagnostic reference levels for adult CT in Switzerland and assessment of radiation dose reduction since 2010. *Eur Radiol* 2020;30:1690–700. <https://doi.org/10.1007/s00330-019-06485-1>.
- [9] Kanal KM, Butler PF, Sengupta D, Bhargavan-Chatfield M, Coombs LP, Morin RLUS. Diagnostic reference levels and achievable doses for 10 adult CT examinations. *Radiology* 2017;284:120–33. <https://doi.org/10.1148/radiol.2017161911>.
- [10] Brat H, Zanca F, Montandon S, Racine D, Rizk B, Meicher E, et al. Local clinical diagnostic reference levels for chest and abdomen CT examinations in adults as a function of body mass index and clinical indication: a prospective multicenter study. *Eur Radiol* 2019;29:6794–804. <https://doi.org/10.1007/s00330-019-06257-x>.
- [11] International Electrotechnical Commission (IEC). IEC 62985 : Methods for calculating size specific dose estimates (SSDE) for computed tomography 2019.
- [12] Rehani MM. Limitations of diagnostic reference level (DRL) and introduction of acceptable quality dose (AQD). *BJR* 2015;88:20140344. <https://doi.org/10.1259/bjr.20140344>.
- [13] Viry A, Aberle C, Lima T, Treier R, Schindera ST, Verdun FR, et al. Assessment of task-based image quality for abdominal CT protocols linked with national diagnostic reference levels. *Eur Radiol* 2022;32:1227–37. <https://doi.org/10.1007/s00330-021-08185-1>.
- [14] Samei E, Järvinen H, Kortesiani M, Simantirakis G, Goh C, Wallace A, et al. Medical imaging dose optimisation from ground up: expert opinion of an international summit. *J Radiol Prot* 2018;38:967–89. <https://doi.org/10.1088/1361-6498/aac575>.
- [15] Burgess AE. The rose model, revisited. *J Opt Soc Am A* 1999;16:633. <https://doi.org/10.1364/JOSAA.16.000633>.
- [16] Solomon J, Marin D, Roy Choudhury K, Patel B, Samei E. Effect of radiation dose reduction and reconstruction algorithm on image noise, contrast, resolution, and detectability of subtle hypoattenuating liver lesions at multidetector CT: filtered back projection versus a commercial model-based iterative reconstruction algorithm. *Radiology* 2017;284:777–87. <https://doi.org/10.1148/radiol.2017161736>.
- [17] Schindera ST, Odedra D, Raza SA, Kim TK, Jang H-J, Szucs-Farkas Z, et al. Iterative reconstruction algorithm for CT: can radiation dose be decreased while low-contrast detectability is preserved? *Radiology* 2013;269:511–8. <https://doi.org/10.1148/radiol.13122349>.
- [18] Benz DC, Benetos G, Rampidis G, von Felten E, Bakula A, Sustar A, et al. Validation of deep-learning image reconstruction for coronary computed tomography angiography: impact on noise, image quality and diagnostic accuracy. *J Cardiovasc Comput Tomogr* 2020;14:444–51. <https://doi.org/10.1016/j.jcct.2020.01.002>.
- [19] Mileto A, Guimaraes LS, McCollough CH, Fletcher JG, Yu L. State of the art in abdominal CT: the limits of iterative reconstruction algorithms. *Radiology* 2019;293:491–503. <https://doi.org/10.1148/radiol.2019191422>.
- [20] Racine D, Becce F, Viry A, Monnin P, Thomsen B, Verdun FR, et al. Task-based characterization of a deep learning image reconstruction and comparison with filtered back-projection and a partial model-based iterative reconstruction in abdominal CT: a phantom study. *Phys Med* 2020;76:28–37. <https://doi.org/10.1016/j.ejmp.2020.06.004>.
- [21] Barrett HH, Myers KJ. *Foundations of image science*. Hoboken, NJ: Wiley-Interscience; 2004.
- [22] Rehani M. *Dose reduction in CT while maintaining diagnostic confidence: a feasibility/demonstration study*. Vienna, Austria: International Atomic Energy Agency; 2009.
- [23] Samei E, Bakalyar D, Boedeker KL, Brady S, Fan J, Leng S, et al. Performance evaluation of computed tomography systems: summary of AAPM Task Group 233. *Med Phys* 2019;46. <https://doi.org/10.1002/mp.13763>.
- [24] Rotzinger DC, Knebel J-F, Jouannic A-M, Adler G, Qanadli SD. CT pulmonary angiography for risk stratification of patients with nonmassive acute pulmonary embolism. *radiology: Cardiothoracic Imaging* 2020;2:e190188. <https://doi.org/10.1148/ryct.2020190188>.

- [25] Chen G, Dong H, Wang X, Zheng Y, Xu M, Zhao Z. Measurement of the diameter of segmental and sub-segmental pulmonary arteries of the right inferior lobe by multislice spiral CT angiography. *Acta Anatomica Sinica* 2014;45:248–52.
- [26] Ott JG, Becce F, Monnin P, Schmidt S, Bochud FO, Verdun FR. Update on the non-prewhitening model observer in computed tomography for the assessment of the adaptive statistical and model-based iterative reconstruction algorithms. *Phys Med Biol* 2014;59:4047–64. <https://doi.org/10.1088/0031-9155/59/4/4047>.
- [27] Boone JM, Brink JA, Edyvean S, Huda W, Leitz W, McCollough CH, et al. ICRU report No. 87: radiation dose and image-quality assessment in computed tomography. *J ICRU* 2012;12:1. <https://doi.org/10.1093/jicru/ndt007>.
- [28] Solomon JB, Samei E. Correlation between human detection accuracy and observer model-based image quality metrics in computed tomography. *J Med Imag* 2016;3:1. <https://doi.org/10.1117/1.JMI.3.3.035506>.
- [29] Verdun FR, Racine D, Ott JG, Tapiovaara MJ, Toroi P, Bochud FO, et al. Image quality in CT: from physical measurements to model observers. *Phys Med* 2015;31:823–43. <https://doi.org/10.1016/j.ejmp.2015.08.007>.
- [30] European Commission. Directorate General for Energy. European study on clinical diagnostic reference levels for X-ray medical imaging: EUCLID. LU: Publications Office; 2021.
- [31] Bos D, Yu S, Luong J, Chu P, Wang Y, Einstein AJ, et al. Diagnostic reference levels and median doses for common clinical indications of CT: findings from an international registry. *Eur Radiol* 2022;32:1971–82. <https://doi.org/10.1007/s00330-021-08266-1>.
- [32] Klosterkemper Y, Thomas C, Bethge OT, Appel E, Aissa J, Boeven J, et al. Implementation of institutional size-specific diagnostic reference levels for CT angiography. *Acad Radiol* 2019;26:1661–7. <https://doi.org/10.1016/j.acra.2019.01.019>.
- [33] Boere H, Eijssvoogel NG, Sailer AM, Wildberger JE, de Haan MW, Das M, et al. Implementation of size-dependent local diagnostic reference levels for CT angiography. *Am J Roentgenol* 2018;210:W226–33. <https://doi.org/10.2214/AJR.17.18566>.
- [34] Rotzinger DC, Racine D, Beigelman-Aubry C, Alfudhili KM, Keller N, Monnin P, et al. Task-based model observer assessment of a partial model-based iterative reconstruction algorithm in thoracic oncologic multidetector CT. *Sci Rep* 2018;8:17734. <https://doi.org/10.1038/s41598-018-36045-4>.
- [35] Greffier J, Frandon J, Si-Mohamed S, Dabli D, Hamard A, Belaouni A, et al. Comparison of two deep learning image reconstruction algorithms in chest CT images: a task-based image quality assessment on phantom data. *Diagn Interv Imaging* 2022;103:21–30. <https://doi.org/10.1016/j.diii.2021.08.001>.
- [36] Gang GJ, Stayman JW, Zbijewski W, Siewerdsen JH. Task-based detectability in CT image reconstruction by filtered backprojection and penalized likelihood estimation: task-based detectability in FBP and PL. *Med Phys* 2014;41:081902. <https://doi.org/10.1118/1.4883816>.
- [37] Racine D, Ryckx N, Ba A, Becce F, Viry A, Verdun FR, et al. Task-based quantification of image quality using a model observer in abdominal CT: a multicentre study. *Eur Radiol* 2018;28:5203–10. <https://doi.org/10.1007/s00330-018-5518-8>.
- [38] Conzelmann J, Genske U, Emig A, Scheel M, Hamm B, Jahnke P. Comparison of low-contrast detectability between uniform and anatomically realistic phantoms—influences on CT image quality assessment. *Eur Radiol* 2022;32:1267–75. <https://doi.org/10.1007/s00330-021-08248-3>.
- [39] Ba A, Abbey CK, Racine D, Viry A, Verdun FR, Schmidt S, et al. Channelized Hotelling observer correlation with human observers for low-contrast detection in liver CT images. *J Med Imaging (Bellingham)* 2019;6:025501. <https://doi.org/10.1117/1.JMI.6.2.025501>.
- [40] Ogino H, Vañó E. *Diagnostic reference levels in medical imaging*. Thousand Oaks, Calif: SAGE; 2017.
- [41] Hendriks BMF, Eijssvoogel NG, Kok M, Martens B, Wildberger JE, Das M. Optimizing pulmonary embolism computed tomography in the age of individualized medicine: a prospective clinical study. *Invest Radiol* 2018;53:306–12. <https://doi.org/10.1097/RLI.0000000000000443>.
- [42] Litmanovich D, Boiselle PM, Bankier AA, Kataoka ML, Pinykh O, Raptopoulos V. Dose reduction in computed tomographic angiography of pregnant patients with suspected acute pulmonary embolism. *J Comput Assist Tomogr* 2009;33:961–6. <https://doi.org/10.1097/RCT.0b013e318198cd18>.
- [43] Haubold J, Zensen S, Hosch R, Schaarschmidt BM, Bos D, Schmidt B, et al. Individualized scan protocols for CT angiography: an animal study for contrast media or radiation dose optimization. *Eur Radiol Exp* 2023;7:24. <https://doi.org/10.1186/s41747-023-00332-1>.
- [44] MacKenzie JD, Nazario-Larrieu J, Cai T, Ledbetter MS, Duran-Mendicuti MA, Judy PF, et al. Reduced-dose CT: effect on reader evaluation in detection of pulmonary embolism. *Am J Roentgenol* 2007;189:1371–9. <https://doi.org/10.2214/AJR.07.2686>.