Objective comparison of high-contrast spatial resolution and low-contrast detectability for various clinical protocols on multiple CT scanners

Damien Racine* and Anaïs Viry*
Institute of Radiation Physics, Lausanne University Hospital, Rue du Grand-Pré 1, 1007 Lausanne, Switzerland

Fabio Becce and Sabine Schmidt
Department of Diagnostic and Interventional Radiology, Lausanne University Hospital, Rue du Bagnon 46, 1011 Lausanne, Switzerland

Alexandre Ba and François O. Bochud
Institute of Radiation Physics, Lausanne University Hospital, Rue du Grand-Pré 1, 1007 Lausanne, Switzerland

Sue Edyvean
Medical Dosimetry Group, Centre for Radiation Chemicals and Environmental Hazards, Public Health England, Didcot OX11, UK

Alexander Schegerer
Department of Medical and Occupational Radiation Protection, Federal Office for Radiation Protection (BfS), Ingolstädt, Landstrasse 1, 85764 Neuhberg, Germany

Francis R. Verdun
Institute of Radiation Physics, Lausanne University Hospital, Rue du Grand-Pré 1, 1007 Lausanne, Switzerland

(Received 30 September 2016; revised 3 February 2017; accepted for publication 7 March 2017; published 13 September 2017)

Purpose: We sought to compare objectively computed tomography (CT) scanner performance for three clinically relevant protocols using a task-based image quality assessment method in order to assess the potential for radiation dose reduction.

Methods: Four CT scanners released between 2003 and 2007 by different manufacturers were compared with four CT scanners released between 2012 and 2014 by the same manufacturers using ideal linear model observers (MO): prewhitening (PW) MO and channelized Hotelling (CHO) MO with Laguerre-Gauss channels for high-contrast spatial resolution and low-contrast detectability (LCD) performance, respectively. High-contrast spatial resolution was assessed using a custom-made phantom that enabled the computation of the target transfer function (TTF) and noise power spectrum (NPS). Low-contrast detectability was assessed using a commercially available anthropomorphic abdominal phantom providing equivalent diameters of 24, 29.6, and 34.6 cm. Three protocols were reviewed: a head (trauma) and an abdominal (urinary stones) protocol were applied to assess high-contrast spatial resolution performance; and another abdominal (focal liver lesions) protocol was applied for LCD. The liver protocol was tested using fixed and modulated tube currents. The PW MO was proposed for assessing high-contrast detectability performance of the various CT scanners.

Results: Compared with older generation CT scanners, three newer systems displayed significant improvements in high-contrast detectability over that of their predecessors. A fourth, newer system had lower performance. The CHO MO was appropriate for assessing LCD performance and revealed that an excellent level of image quality could be obtained with newer scanners at significantly lower dose levels.

Conclusions: This study shows that MO can objectively benchmark CT scanners using a task-based image quality method, thus helping to estimate the potential for further dose reductions offered by the latest systems. Such an approach may be useful for adequately and quantitatively comparing clinically relevant image quality among various scanners. © 2017 The Authors. Medical Physics published by Wiley periodicals, Inc. on behalf of American Association of Physicists in Medicine. [https://doi.org/10.1002/mp.12224]

Key words: computed tomography, high-contrast spatial resolution, image quality, low-contrast detectability, model observers

1. INTRODUCTION

In most Western countries, the radiation exposure of the population due to computed tomography (CT) examinations has increased steadily for 20 yr.¹ A survey performed in 2006 in the United States showed that the average effective dose due to CT reached 1.5 mSv per capita, per year.² The last surveys performed in Switzerland in 2008 and 2013 showed a similar trend, with the average dose per capita from CT increasing from 0.8 mSv to 1.0 mSv within this 5-year period.³,⁴
In this context, the radiation protection requirements in diagnostic radiology (justification of the examination and optimization of the imaging protocol) need to be reinforced. The justification aspect is beyond the scope of this article. The optimization of a CT examination is achieved when image quality enables the clinical question to be answered while keeping patient radiation dose as low as reasonably achievable (ALARA). This goal is, however, difficult to apply in practice. The quality of a CT examination depends on a wide range of parameters such as acquisition time, temporal resolution, and energy resolution when dealing with kV optimization or spectral CT imaging, and other factors. Thus, the actual determination of the clinical performance of a CT scanner is quite complex, and the clinical question needs to be clarified to enable a standard for image quality level to be set. Task-oriented image quality criteria can then be used as surrogates for the assessment of actual clinical image quality. They will necessarily be simple in comparison to the clinical situations, but will make it possible, for example, to predict the ability to detect simple structures of high and low contrast within homogeneous backgrounds. This represents the most basic task that can be considered a surrogate for measuring clinical image quality.

To add complexity, making the task more realistic, one could then not only consider the detection but also the determination of the correct position of the detected structure. Then, the performance with which these tasks are performed could be assessed over more realistic structured backgrounds that mimic the actual anatomy. To go a step further, one could also check if the sizes and contrasts measured on the images correspond to the actual values. This strategy is still far from actual clinical image quality assessment, but may aid in the optimization of clinical protocols.

The aim of this study was to propose a way to objectively compare CT scanner performance using the simplest task-based image quality assessment: detection. This method was used in particular to evaluate the impact of technological developments on the potential for radiation dose reduction. In addition, we also wanted to investigate if major differences in performance existed among different manufacturers in the limited image quality criteria chosen. We compared the outcomes of four CT scanners released by the four major manufacturers from 2003 to 2007 with the outcomes of newer systems introduced from 2012 to 2014, using ideal model observers (MO) on three clinically relevant protocols.

2. MATERIALS AND METHODS

2.A. CT scanners and clinical protocols

This study was conducted using eight different CT scanners: two per major manufacturer including models released between 2003 and 2007 (referred as “older”), and models released between 2012 and 2014 (referred as “newer”). These eight CT scanners were, listed as “older”/”newer”: Lightspeed VCT/Revolution CT (GE Healthcare, Milwaukee, WI, USA), Brillance 40/Ingenuity Core 128 (Philips Medical Systems, Best, the Netherlands), Somatom Sensation 64/Force (Siemens Healthcare, Forchheim, Germany), and Aquilion 16/Aquilion Prime (Toshiba Medical Systems, Tokyo, Japan). Basic characteristics and the image reconstruction methods used for this study are summarized in Table I.

For all CT systems, the displayed weighted computed tomographic dose index (CTDIw, data were verified by measuring the normalized weighted computed tomographic dose index (nCTDIw) using a 32-cm diameter CTDI test object and a 10-cm long CT pencil ionization chamber connected to an electrometer (model 1035-10.3 CTDI chamber and MDH model 1015 electrometer, Radcal, Monrovia, CA, USA), calibrated in RQR9 and RQA9 beams according to IEC 61267 and traceable to the Swiss Federal Office of Metrology. The volume CT dose index (CTDIvol) is defined as the CTDIw divided by the helical pitch factor; the values used in this study were taken directly from the displayed ones.

The image acquisition protocols were used to compare the performance of the CT units were proposed by a panel of four senior radiologists working in three different University Hospitals in Switzerland. Among a large number of clinically relevant protocols we focused on three: two requiring a relatively high level of spatial resolution for the detection of high-contrast structures in the head and abdomen, and one requiring a high level of low-contrast resolution in the abdominal region. When dealing with the assessment of low-contrast resolution performance in the abdominal region, two approaches were chosen: one using fixed dose levels (5, 10, and 15 mGy), with the 15 mGy dose level corresponding to the Swiss abdominal CT Diagnostic Reference Level (DRL) using one phantom size; and the other using the tube current modulation option (with the local settings used for the clinical indication of the acquisitions) using three phantom sizes. The details of the acquisition parameters used for each protocol are given in Table I. For technical reasons, the acquisition parameters used were not exactly the same.

2.B. Phantoms and data acquisitions

2.B.1. High-contrast performance

The assessment of high-contrast performance for head and abdominal protocols was made using a dedicated custom-made phantom containing cylindrical rods of different contrast materials (Teflon® or polytetrafluoroethylene [PTFE], polyethylene, and polymethylmethacrylate [PMMA]) (Fig. 1). The edge of this internal cylinder, at different z-axis positions, is used as the interface for the high-contrast numerical evaluation. The external diameter of this phantom is 250 mm; the high-contrast internal cylinder diameter is 100 mm.

2.B.2. Low-contrast performance

A modified anthropomorphic abdominal phantom (QRM 401, QRM, Moehrendorf, Germany) (Fig. 2) was used to investigate the low-contrast resolution performance of the CT units. It is made of calibrated tissue-equivalent material.
The body of the phantom (equivalent diameter of 24 cm) contains muscle, liver, spleen, and bone (vertebrae) tissue equivalents. A module can be inserted into the phantom body that includes spheres of different sizes: 8, 6, 5, 4, and 3 mm; each size having a contrast 20 HU relative to the background at 120 kV. For practical reasons, only three spheres (5, 6, and 8 mm) were used. Two additional annuli (increasing the phantom’s effective diameter to 29.6 cm and 34.6 cm, respectively) were added to simulate a range of body habitus (from an approximate patient weight of 50 kg for the equivalent diameter of 24 cm to 75 kg and 100 kg for the equivalent diameter of 29.6 cm and 34.6 cm).

Ten successive scans of the phantom fixed in place were performed to obtain 40 regions of interest (ROIs) with the spheres, and 100 ROIs without any target. This phantom was scanned using two protocols. First, the small phantom was...
scanned at three dose levels to assess the baseline values of the CT scanners. Then, datasets were again acquired, this time with the two additional rings installed, to investigate the effect of body habitus on low-contrast detectability (LCD).  

2. Image quality metrics  

2.1. High-contrast performance  

Spatial resolution: The parameter usually used to assess spatial resolution when dealing with CT images is the Modulation Transfer Function (MTF). However, iterative reconstruction (IR) algorithms are known to be highly nonlinear and therefore might introduce a dependency of the image contrast over the spatial resolution. Boone16 and Richard et al.17 proposed target transfer function (TTF) metrics to overcome this problem by characterizing spatial resolution taking into account contrast properties. MTF and TTF are similar metrics, except that TTF may be applied on different contrast levels. In this study we took a similar approach. Using the rod phantom, TTF was calculated for each rod from the radial mean of the edge spread function (ESF) profiles. The ESF’s raw data were fitted and analytically differentiated to provide line spread functions (LSFs). Finally, performing a Fourier Transform on the LSFs gave the TTFs, which were normalized to 1 at the zero frequency. More details about the methods can be found in Ott et al.14  

Noise power spectrum: The rod phantom also allows the assessment of the noise power spectrum (NPS). ROIs of 100 x 100 pixels, which were located in the center of a homogeneous region from 10 images, were used to calculate the NPS. The 2D NPS was computed using the following equation:

\[
NPS_{2D}(f_x,f_y) = \frac{\Delta_x \Delta_y}{L_x L_y N_{ROI}} \sum_{i=1}^{N_{ROI}} |FT_{2D}(ROI_i(x,y) - \bar{ROI}_i)|^2 
\]  

(1)

Where \(\Delta_x, \Delta_y\) are the pixel sizes in the x and y dimensions, \(L_x, L_y\) are ROI sizes for the two directions (\(L_x = L_y = 100\) pixels), \(N_{ROI}\) is the number of ROIs (\(N_{ROI} = 10\)), and \(\bar{ROI}_i\) is the mean pixel value for the ith ROI. The 2D NPS was then radially averaged to provide the 1D NPS_{1D} according to the methodology presented in ICRU 87.18  

Prewhitening model observer: To perform a task-based image quality assessment of high-contrast structures, the detectability index (\(d'\)) of different diameters structures having a nominal contrast of 1080 HU at 120 kVp (PTFE/water), 120 HU (PMMA/water), and -80 HU (polyethylene/water) was computed using the prewhitening mathematical model observer, PW,19 and to reduce inconsistencies due to the use of iterative reconstructions, the MTF function, that should be used in that model, was replaced by the TTF function (see Eq. 2)14.

\[
d' = \sqrt{2\pi|\Delta HU|} \sqrt{\int_{f_N}^{f_{Ny}} \frac{S_2(f)TTF^2(f)f}{NPS_{1D}(f)f} df} 
\]

(2)

where \(f\) is the frequency, \(f_{Ny}\) is the Nyquist frequency of the image, |\(\Delta HU\)| is the absolute contrast difference between the signal and the background and \(S(f)\) is the Fourier transform of the input signal, \(S(f) = \frac{1}{2}J_1(2\pi f \cdot Rf)\) with \(J_1\), a Bessel function of the first kind. In our study, the rod phantom was used to provide the estimation of TTF and NPS, which are needed for PW MO, but not for the direct measurement of the small size disks’ high-contrast detectability. To overcome
this limitation, we simulated a virtual disk with a radius R varying from 0.5 mm to 2.5 mm.

It is of note that scatter reduces not only image contrast but also the amplitude of the TTF. It was decided to take into account the scatter effect by using the measured contrasts rather than the nominal ones. Thus, TTFs were fitted to avoid the effect of scatter (spatial resolution drop in the low frequency range) as presented in reference.14

The uncertainty of the PW outcome was assessed by varying randomly the contrast and TTF values in the range of their standard deviations, considering a Gaussian distribution, measured on 30 images. The NPS parameter was not considered due to the fact that its uncertainty is negligible compared to that of the contrast and TTF parameters.

2.C.2. Low-contrast performance

Channelized Hotelling observer: LCD was evaluated in the image domain using a channelized Hotelling observer (CHO), with Laguerre-Gauss (LG) channels. This model is an estimation of the Hotelling observer, which itself is the ideal linear MO. The use of LG channels is appropriate in this case because they are known to maximize task performance.20,21 The computation was made up to the tenth order of the LG polynomials and for one orientation only (due to the circular symmetry of the structure to be detected), resulting in a total of 10 channels. The pth channel (up) is obtained by multiplying the Laguerre polynomial at the pth order by a Gaussian function:

\[ u_p(r|a_u) = \frac{\sqrt{2}}{a_u} \exp\left(-\pi r^2/a_u^2\right) L_p\left(\frac{2\pi r^2}{a_u^2}\right) \]  

(3)

where \( L_p \) is a Laguerre polynomial, \( r \) is a two-dimensional spatial coordinate, and \( a_u \) is the width of the Gaussian function (taken to be \( a_u = 9 \) in the present study).

Laguerre polynomials are defined by:

\[ L_p(x) = \sum_{k=0}^{p} (-1)^k \binom{p}{k} x^k k! \]  

(4)

The image is passed through the 10 LG channels. The channel output is a scalar \( v_i \) obtained by the dot product between the channel \( u_p \) and the image \( g \):

\[ v_i = U^T g \]  

(5)

Where \( U \) represents the matrix of the channels, each column is one of the 10 channels:

\[ U = [u_1, u_2, \ldots, u_{10}] \]  

(6)

The CHO is then computed from the template \( w_{LG} \):

\[ w_{LG} = \left(K_{v_i/n}\right)^{-1} \left(\langle v_i \rangle - \langle v_n \rangle\right) \]  

(7)

where \( K_{v_i/n} \) is the covariance matrix calculated from 100 signal absent images as perceived through the channels (channelized images), \( \langle v_i \rangle \) represents the mean of 40 channelized signal images and \( \langle v_n \rangle \) the mean of 100 channelized absent signal images.

The decision variable \( \lambda_{LG} \) of the CHO model is obtained by combining the template \( w_{LG} \) and the channelized image \( v_i \):

\[ \lambda_{LG} = w_{LG}^T v_i \]  

(8)

In the end, the MO was tested with the same set of images as with the training set although this could overestimate its performance.22 For each category (lesion and phantom size as well as dose levels), a receiver operating characteristic (ROC) curve was calculated with 50 threshold levels.23 To summarize the information, an area under the ROC curve (AUC) was calculated using the trapezoidal method. The average and standard deviation of the model observers were estimated by performing a bootstrap method.24 In practice, 500 ROC experiments were performed for each category.

For each dose level and phantom size, we used an image quality metric called “AUCw” which combines the AUCs for lesions of different sizes. This metric is computed thus:

\[ AUC_w = \frac{\sum_{i \in \{8,6,5\}} AUC_{lesion(i)}}{\sum_{i \in \{8,6,5\}} AUC_{lesion(i),max}} \]  

(9)

where \( i \) represents the lesion sizes: 8, 6, or 5 mm, \( AUC_{lesion(i)} \) represents the outcome of model observer for each lesion size. With such a definition, \( AUC_{lesion(L,max)} \) corresponds to the value of this metric when the performance is maximal for each lesion size (\( AUC_{lesion(L,max)} = 1.0 \)).

3. RESULTS

To ensure the impartiality of this work, the results are reported in an anonymous manner consistently throughout the manuscript. A capital letter (A, B, C, and D) was assigned to each manufacturer and the lower case letter “a” and “b” was added for respectively “newer” and “older” CT units. Differences between the displayed and measured CTDIvol were within 15%, in conformity with Swiss legal requirements (limit of ± 20%).

3.A. Image quality for high-contrast structures

3.A.1. High-contrast detection for the head protocol

For the detection of high-contrast structures, a \( d’ \) was calculated for different contrast values for each CT using a head protocol. As expected, the \( d’ \) increased with the diameter and the nominal contrast of structures to be detected (Figs. 3(a)–3(c)).

Comparison of performance of new and old scanner models from each manufacturer: Figure 3(a) shows that for manufacturers A, C, and D, there was a noticeable improvement of the detectability when switching from the older CT scanners to newer ones while a slight reduction was observed for manufacturer B. The largest improvement was observed for manufacturer C (283% for lesions from 3–5 mm), whereas moderate improvements were found for manufacturers A and D (18% and 37%, respectively).
In Figs. 3(b) and 3(c), similar behavior was observed for manufacturers B, C, and D. For manufacturer A, no major difference appeared between the older and newer CT units.

Differences between manufacturers (newer CT models): For the three contrast levels tested with the newer CT models presented in Figs. 3(a)–3(c), the $d'$ reached the highest value...
for manufacturer C, and manufacturers A and D provided better results than manufacturer B.

**High-contrast detection for the abdomen protocol:** The same methodology was applied to assess the detectability of high-contrast structures for the abdominal protocol.

**Comparison of performance of new and old scanner models from each manufacturer:** In Fig. 4(a), for the highest contrast level, detectability improved when switching from older to newer CTs for manufacturers A, C, and D. A major improvement was noted for manufacturer C (86%) and a moderate improvement was noted for manufacturer A and D (23% and 40%, respectively). For manufacturer B, a trend similar to the one identified in the head protocol was observed (Fig. 3(a)).

In Figs. 4(b) and 4(c), similar behaviors were observed for manufacturers B, C, and D. For manufacturer A, smaller differences appeared between the older and newer CT units for materials PMMA and polyethylene than for PTFE.

**Differences among manufacturers (newer CT models):** For each contrast level and each structure’s diameter, the results of the comparison of newer CT models was very similar to the results for the head protocol. Manufacturer C reached the highest performance. Manufacturers A and D provided better results than manufacturer B.

3.B. **Image quality for low-contrast detectability**

3.B.1. **Abdomen low-contrast detection — CTDI\textsubscript{vol} variation**

Imaging the small abdomen phantom with a CTDI\textsubscript{vol} of 15 mGy (Fig. 5) showed no major differences among the various scanners. Reducing the CTDI\textsubscript{vol} to 10 mGy, the image quality metrics slightly decreased for all scanners (AUC\textsubscript{w} going from 1.0 to 0.985), with a larger reduction observed for scanner “Db” (AUC\textsubscript{w} going from 1.0 to 0.945). These variations are statistically significant as an uncertainty of 0.003 was set for these measurements ($P < 0.05$). At the lowest CTDI\textsubscript{vol}, we investigated (5 mGy), all newer scanners provided better results than the older ones except for scanner “Aa”.

To investigate the robustness of the method used, the measurements were repeated five times on the same scanner “Da” using the small abdomen phantom with a CTDI\textsubscript{vol} of 5 mGy (Fig. 5) under “positioning uncertainties,” and demonstrate that comparable results could be obtained when repositioning the phantom several times.

To investigate how the method would vary when characterizing various scanners of the same type, the methodology was applied on five different “Da” scanners (Fig. 5), and represented as “CT machine uncertainties.”

Comparable results were found with different machines of the same type.

3.B.2. **Abdomen low-contrast detection – Phantom size variation**

Using automatic tube current modulation and the small abdominal phantom, Fig. 6 shows that it is possible to reach a similar level of image quality for all scanners (differences within 5%). However, this high level of image quality is obtained at noticeable different CTDI\textsubscript{vol} values (almost 300%).

When using the medium abdominal phantom a significant drop in image quality is observed for three scanners (“Db”, “Bb”, and “Ba”). For the other scanners, comparable image quality is preserved but again within a large range of CTDI\textsubscript{vol} values (Fig. 7).

Finally, when using the largest anthropomorphic abdominal phantom, large differences in behaviors were observed (Fig. 8).

**Comparison of performance of new and old scanner models from each manufacturer:** For all manufacturers but one, major improvements were demonstrated with the newer models. It was possible to reach similar image quality levels at significantly lower CTDI\textsubscript{vol} levels. For manufacturer A, image quality level was slightly decreased but the patient exposure was reduced by 50%. For manufacturer C, a noticeable improvement in image quality was obtained at less than half the dose from the older model. For manufacturer D, a major improvement of image quality was obtained with a lower dose reduction than manufacturers A and C (30%). The only manufacturer where no major improvement was noted was manufacturer B, where similar image quality level was obtained at a slightly higher CTDI\textsubscript{vol} value.

**Differences between manufacturers (newer CT models):** When using the largest size of the phantom (simulating a patient of 100 kg), all newer CT scanners reached a high level of image quality (AUC\textsubscript{w} > 0.850). Nevertheless, this level of image quality was reached with CTDI\textsubscript{vol} differences within a range of 300%.

4. **DISCUSSION**

A full characterization of CT scanner units would require the assessment of a large number of parameters. Among these parameters one could mention: the acquisition time, the standard high- and low-contrast resolutions, the temporal resolution, and the energy resolution when dealing with kV optimization or dual energy imaging.

We chose to use simple task-based image quality assessment methodologies that do not include the whole range of potential performance of the scanners. We assessed the performance regarding image quality of high- and low-contrast structures.
using ideal MO for eight CT systems and three clinical protocols. This benchmark provided a large panel of image quality levels for older (2003–2007) and newer CTs (2011–2014).

The first aim of the study was to investigate if technological improvements over time could be shown using our limited set of image quality criteria. For the high-contrast...
detectability, scanners could be discriminated using the PW MO, and performance improvement was noted for manufacturers A, C, and D. The \( d' \) values were systematically very high, indicating that the detection of a structure > 2 mm in diameter with such a nominal contrast value was trivial. For better discrimination one could add complexity to the task, for example: the estimation of shape, size, and contrast. Concerning the low-contrast resolution, performance improvements were observed also for three (manufacturers A, C, and D) out of four manufacturers with a drastic dose reduction to reach similar high image quality levels.

The second aim of our study was to investigate if major differences in performance existed between newer CT scanners of various manufacturers. For the limited criteria chosen in this study, manufacturer C reached the highest performance with the chosen reconstruction kernels. However, our measurements have two limitations: the first one deals with the choice of the reconstruction kernel that could not be the same for all manufacturers.\(^ {25} \)

To investigate if some kernels used in this study would give advantage to a particular range of target sizes, we computed the \( d' \) values for structures 1–5 mm in diameter. A regular increase in \( d' \) with the object diameter was noted for all manufacturers. In addition, the clinical parameters used for the high-contrast resolution would not necessarily represent the maximum theoretical capability of the system.
second limitation deals with the Slice Sensitivity Profile, a parameter that was not considered in our methodology. Nevertheless, comparable reconstruction slice thicknesses (within the range 2.5–3.0 mm) were used, so no major influence of this parameter is expected in our results. Moreover, as the 2D NPS was not isotropic in the phantom, using symmetric channels could impact the performance of the model; but this impact would be minor.

Dealing with the protocol using fixed CTDI\textsubscript{vol} values, we proposed first to use a range of dose levels. The highest value, 15 mGy, corresponds to our DRL for a “standard” patient of 75 kg. At 15 mGy, no difference between CTs appeared when we used the small abdominal phantom. However, differences were detected at lower dose levels. Thus, such a phantom should be imaged over a lower dose range (e.g., 2–10 mGy), or replaced by a larger version as shown in Figs. 7 and 8 to produce useful results. Finally, with a larger phantom size, the use of a fixed tube current might be a limitation that introduces weaknesses of the image quality which are not relevant for clinical applications where tube current modulation is generally used.

The use of the protocol with tube current modulation on various phantom sizes is certainly more realistic in the framework of patient dose optimization. When comparing the products of one manufacturer, the results were generally straightforward. The major difficulty of such an approach is
to compare various manufacturers who propose different strategies to manage the balance between image quality and patient exposure. In this study we decided to use the local settings. The indication “search for focal liver lesions” requires particularly high image quality. Using the small phantom, a high level of image quality was reached by all CTs using a large range of CTDIvol values. For the larger phantom, a high and particularly high image quality. Using the small phantom, a high level of image quality could not be reached by certain units despite the use of a large range of CTDIvol values.

The outcome of this work clearly demonstrates the weaknesses of the DRL concept; indeed, a similar dose level cannot be reached on different scanners without impairing the LCD. To improve this situation, one could associate the DRL to an image quality criterion such as the LCD estimated in a standardized phantom.

5. CONCLUSION

This study shows that MOs can objectively benchmark CT scanners using a task-based image quality method. Such an approach may be useful for quantitatively comparing the clinically relevant image quality among various scanners to aid in the estimation of the potential dose reduction without missing the detection of critical lesions.

ACKNOWLEDGMENTS

This work was supported by the German Radiation Protection Agency (BfS – “UFO-Plan Vorhabens 3613S20007”). The authors thank Prof. Peter Vock and PD Dr Sebastian Schindera, the two other senior radiologists involved in this project for their help when dealing with the choice of the protocols and diagnostic tasks. The authors also thank MSc Pascal Monnin for his advices concerning the estimation of the TTF functions.

CONFLICT OF INTEREST

The authors have no relevant conflicts of interest to disclose.

*Author to whom correspondence should be addressed. Electronic mail: damien.racine@chuv.ch; Telephone: +41 21 314 92 49.

REFERENCES

1. Brenner DJ, Hall EJ. Computed tomography — An increasing source of radiation exposure. N Engl J Med. 2007;357:2277–2284.
2. Schauer DA, Linton OW. National council on radiation protection and measurements report shows substantial medical exposure increase. Radiology. 2009;253:293–296.
3. Samara ET, Aroua A, Bochud FO, et al. Exposure of the swiss population by medical x-rays: 2008 review. Health Phys. 2012;102:263–270.
4. Couture RL, Bize J, Champendal M, et al. Exposure of the swiss population by radiodiagnosics: 2013 review. Radiat Prot Dosimetry. 2015;169:221–224.
5. Barrett HH, Myers KJ, Hoeschen C, Kupinski MA, Little MP. Task-based measures of image quality and their relation to radiation dose and patient risk. Phys Med Biol. 2015;60:R60.
6. Vennart W. ICRU report 54: medical imaging — The assessment of image quality. Radiography. 1996;3:243–244.
7. Racine D, Ott JG, Ba A, Ryckx N, Bochud FO, Verdun FR. Objective task-based assessment of low-contrast detectability in iterative reconstruction. Radiat Prot Dosimetry. 2016;169:73–77.
8. Saitprasad G, Flibiben J, Peskin A, et al. Evaluation of low-contrast detectability of iterative reconstruction across multiple institutions, CT scanner manufacturers, and radiation exposure levels. Radiology. 2015;277:124–133.
9. Vaishnav JY, Jung WC, Popescu LM, Zeng R, Myers KJ. Objective assessment of image quality and dose reduction in CT iterative reconstruction. Med Phys. 2014;41:71904.
10. Verdun FR, Racine D, Ott JG, Tapiovara MJ, et al. Image quality in CT: from physical measurements to model observers. Phys Medica Eur J Med Phys. 2015;31:823–843.
11. Schindera ST, Odedra D, Raza SA, et al. Iterative reconstruction algorithm for CT: can radiation dose be decreased while low-contrast detectability is preserved? Radiology. 2013;269:511–518.
12. International Electrotechnical Committee. Medical diagnostic X-ray equipment - Radiation conditions for use in the determination of characteristics. 1994. Standard IEC # 61267. https://webstore.iec.ch/publication/21937. Published 2015. Accessed September 27, 2016.
13. Office fédéral de la santé publique OFSP. Notice R-06-06:Niveaux de référence diagnostiques en tomodensitométrie. 2010.
14. 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;59:4047–4064.
15. Racine D, Ba AH, Ott JG, Bochud FO, Verdun FR. Objective assessment of low contrast detectability in computed tomography with channelized hotelling observer. Phys Medica Eur J Med Phys. 2016;32:76–83.
16. Boone JM. Determination of the presampled MTF in computed tomography. Med Phys. 2001;28:356–360.
17. Richard S, Husarik DB, Yadava G, Murphy SN, Samei E. Towards task-based assessment of CT performance: system and object MTF across different reconstruction algorithms. Med Phys. 2012;39:4115–4122.
18. International Commission on Radiation Units and Measurements. ICRU report no. 87: radiation dose and image-quality assessment in computed tomography. J ICRU. 2012;12:1–149.
19. Barrett HH, Yao J, Rolland JP, Myer KJ. Model observers for assessment of image quality. Proc Natl Acad Sci U S A. 1993;90:9758–9765.
20. Park S, Barrett HH, Clarkson E, Kupinski MA, Myers KJ. Channelized-ideal observer using Laguerre-Gauss channels in detection tasks involving non-Gaussian distributed lumpy backgrounds and a Gaussian signal. J Opt Soc Am A Opt Image Sci Vis. 2007;24:B136–B150.
21. Gallas BD, Barrett HH. Validating the use of channels to estimate the ideal linear observer. J Opt Soc Am A. 2003;20:1725–1738.
22. Barrett HH, Myer KJ. Foundations of Image Science. Wiley-Interscience; 2004.
23. International Commission on Radiation Units and Measurements. Receiv- er operating characteristic analysis in medical imaging. In: ICRU Report n 79, Vol 79. Bethesda, MD: International Commission on Radiation Units and Measurements; 2008:79.
24. Efron B, Tibshirani RJ. An Introduction to the Bootstrap. New York: CRC Press; 1994.
25. Solomon JB, Christianson O, Samei E. Quantitative comparison of noise texture across CT scanners from different manufacturers. Med Phys. 2012;39:6048–6055.