Published in final edited form as:

Phys Med Biol. 2015 November 7; 60(21): 8381–8397. doi:10.1088/0031-9155/60/21/8381.

A cross-platform survey of CT image quality and dose from routine abdomen protocols and a method to systematically standardize image quality

Christopher P. Favazza, Ph.D.¹, Xinhui Duan, Ph.D.¹, Yi Zhang, Ph.D.¹, Lifeng Yu, Ph.D.¹, Shuai Leng, Ph.D.¹, James M. Kofler, Ph.D.¹, Michael R. Bruesewitz, RT(R)¹, and Cynthia H. McCollough, Ph.D.^{*,1}

¹Department of Radiology, Mayo Clinic, Rochester, MN

Abstract

Through this investigation we developed a methodology to evaluate and standardize CT image quality from routine abdomen protocols across different manufacturers and models. The influence of manufacturer-specific automated exposure control systems on image quality was directly assessed to standardize performance across a range of patient sizes. We evaluated 16 CT scanners across our health system, including Siemens, GE, and Toshiba models. Using each practice's routine abdomen protocol, we measured spatial resolution, image noise, and scanner radiation output (CTDI_{vol}). Axial and in-plane spatial resolutions were assessed through slice sensitivity profile (SSP) and modulation transfer function (MTF) measurements, respectively. Image noise and CTDI_{vol} values were obtained for three different phantom sizes. SSP measurements demonstrated a bimodal distribution in slice widths: an average of 6.2 ± 0.2 mm using GE's "Plus" mode reconstruction setting and 5.0 ± 0.1 mm for all other scanners. MTF curves were similar for all scanners. Average spatial frequencies at 50%, 10%, and 2% MTF values were 3.24 ± 0.37 , 6.20 ± 0.34 , and 7.84 ± 0.70 lp/cm, respectively. For all phantom sizes, image noise and CTDI_{vol} varied considerably: 6.5–13.3 HU (noise) and 4.8–13.3 mGy (CTDI_{vol}) for the smallest phantom; 9.1–18.4 HU and 9.3–28.8 mGy for the medium phantom; and 7.8–23.4 HU and 16.0–48.1 mGy for the largest phantom. Using these measurements and benchmark SSP, MTF, and image noise targets, CT image quality can be standardized across a range of patient sizes.

Keywords

CT; protocol standardization; image noise; automatic exposure control

1. Introduction

For large healthcare institutions, standardizing CT protocols across manufacturers and scanner models, particularly for a large, diverse fleet of scanners, is a non-trivial task. As CT vendors continue to offer new innovations, CT scanners become more complex and

*Corresponding author, 200 First Street SW, Rochester, MN 55905, Ph: (507) 284-6875, Fax: (507) 284-2405, mcollough.cynthia@mayo.edu.

Author Conflict of Interest: Cynthia McCollough Recipient of a research grant from Siemens Healthcare.

differences between scanner models increase (Bhalla and Arora 2013). Without a thorough understanding of the intricacies of each individual scanner, a simple review of scan protocol parameters often provides insufficient information to ensure commensurate image quality. For instance, implementations of automatic exposure control (AEC) systems, which are widely utilized on clinical scanners, vary with manufacturer and significantly impact radiation output and image quality.

Briefly, the basic function of an AEC system is to automatically adjust x-ray output to account for differences in x-ray attenuation found both within individual patients and among different patients (Gies *et al* 1999, Kalender *et al* 1999, Kalra *et al* 2004, 2005, Lee *et al* 2008, McCollough *et al* 2006). An AEC system provides two main benefits: One, it serves to provide similar image quality across different patient sizes, increasing the tube current for larger patients and reducing the tube current for smaller patients. Two, it can reduce dose to the patient by up to 40–50%, without a reduction in image quality (Gies *et al* 1999, Kalender *et al* 1999, Kalra *et al* 2004, 2005, McCollough *et al* 2006). Currently, there are several strategies being employed by AEC systems to determine tube current settings for a given patient or phantom. Through one method, the scanner attempts to maintain a constant noise value across all patient sizes (Lee *et al* 2008, Keat 2005, Favazza *et al*). Through a second approach, the scanner adjusts the tube current relative to a selected “reference” effective mAs value, which is defined for a fixed patient size. This approach allows for image noise to vary with patient size to account for different diagnostic demands (Lee *et al* 2008, Keat 2005, Favazza *et al*). A third strategy relies on a reference image for a given diagnostic task and adjusts the tube current to achieve a similar noise value as compared to the reference image (Lee *et al* 2008, Keat 2005, Favazza *et al*). Thus, standardizing acquisition parameters and image quality for a given protocol and single patient size does not guarantee equivalent image quality across all patient sizes and scanners.

Scanner-specific protocol review is considered a best practice (Cody *et al* 2013). Moreover, periodic review of CT protocols is a fundamental component of a quality assurance program and is a requirement of the American College of Radiology CT Accreditation program (American College of Radiology 2013). The primary objective of such a review program is to safeguard against overdosing patients while at the same time producing images that meet minimum quality metrics (e.g. image uniformity, contrast-to-noise ratio, spatial resolution, etc.). The American Association of Physicists in Medicine provides guidance on CT protocol review practices; however, these guidelines address review objectives only and do not establish a roadmap to standardize CT protocols across manufacturers and scanner models.

Recent developments in dose reporting have led to commercially available tools to assist with protocol management and monitoring (Couch and Couch 2012); however, these tools are limited to information regarding scanner radiation output (i.e., a surrogate for patient dose). Central to these tracking tools is a method to assess radiation dose across all patient exams. There are no direct markers to establish, track, and assess image quality, which is necessary to ensure that dose is being effectively delivered by each specific scanner. Although others have previously reported methods to equate noise texture (Solomon *et al* 2012) and scanner output for a given AEC setting (McKenney *et al* 2014) across scanner

platforms, currently, there is no published systematic strategy to measure and standardize image quality across CT manufacturers and models.

Through this investigation, we have developed a systematic assessment methodology to evaluate image quality across different CT platforms. The aim of this approach is to establish “first order” image quality-based protocol standardization throughout our healthcare system that is applicable to a wide range of patient sizes. To this end, we measured spatial resolution, both axial and in-plane, and image noise on **14** different clinical scanner models (16 total different scanners). For image noise measurements, three different phantom sizes were investigated to directly probe the impact of AEC system differences on image noise and radiation output for different patient sizes. Lastly, we propose a systematic procedure to standardize CT protocols based on these three image quality measurements.

2. Methods

A total of 16 different diagnostic CT scanners located at 11 different sites within our healthcare system were investigated. These 16 scanners represent 14 different scanner models within our health system, which includes the following: GE BrightSpeed 16, GE BrightSpeed Elite 16, GE Discovery 750 HD, GE LightSpeed 16, GE LightSpeed 64 VCT (GE Healthcare, Waukesha, Wisconsin), Siemens Emotion 16, Siemens Sensation 16, Siemens Sensation 40, Siemens Sensation 64 and Siemens Somatom Definition FLASH (Siemens Healthcare, Forchheim, Germany), Toshiba Aquilion 16, Toshiba Aquilion 32, Toshiba Aquilion 64 and Toshiba Aquilion Premium (Toshiba America Medical Systems, Tustin, California). Note that for 2 scanner models, two separate scanners at different locations were evaluated.

To assess image quality, three measurements were performed: [1] axial spatial resolution using the slice sensitivity profile (SSP), [2] in-plane spatial resolution using the modulation transfer function (MTF), and [3] image noise using the standard deviation of pixel values. Additionally, CTDI_{vol} values reported at the scanner console were recorded to assess scanner radiation output (CTDI_{vol}) as phantom size changed.

2.1. Axial Spatial Resolution—SSP

To assess the axial resolution, the SSP was measured for each scanner using the site’s routine clinical abdomen protocol, with the exception that the mAs value was fixed (i.e. AEC was disengaged). A gold foil phantom (1 mm diameter; 25 μ m thick; Figure 1a) was placed in the scanner such that the circular face of the gold foil was parallel with the axial plane of the scanner and subsequently imaged. The raw data were reconstructed into 5 mm thick images at 0.1 mm intervals. The average value of a 3 $\times$ 3 pixel² region of interest (ROI) in the centre of the gold foil image was calculated and recorded for each image, as depicted in Figure 1b. These average values as a function of z-axis position comprised the raw data of the SSP curve. This resulting curve, i.e. average pixel value vs. slice location, was normalized to the peak average pixel value after subtracting the average background pixel value. Image thickness was calculated as the full-width at half maximum (FWHM) value of the normalized curve. The area under the curve was also calculated for each SSP to determine and compare the relative amount of signal used to create each image.

2.2. In-plane spatial resolution—MTF

To assess the high contrast, in-plane spatial resolution, the MTF was calculated for each scanner using the site's routine abdomen clinical protocol, with the exception that the mAs value was fixed (i.e. AEC was disengaged). A specially designed phantom was used for the MTF measurements, as shown in Figure 2a. Specifically, a 0.125 mm diameter tantalum wire comprised the primary component of the phantom and was positioned perpendicular to the axial plane of the scanner. Tantalum offers high contrast relative to the surrounding air while generating HU values within the standard range of CT numbers for all scanners. This characteristic enables acquisition of images without necessitating the use of an expanded CT number feature, which is not available on all scanner models. Following acquisition of the wire phantom images, MTFs were calculated through the following process: (1) reconstruct a cross-sectional image from the centre of the scan using the smallest available field of view (5 cm for Siemens and Toshiba, 9.6 cm for GE), shown in Figure 2b, (2) re-size the image using basic linear interpolation and zero-pad the image, (3) apply fast Fourier transform (FFT) to the resized image and normalize pixel values of FFT image from 0 to 1, shown in Figure 1c, and (4) average pixel values over 2π radians at each radial value, which yields the MTF curve.

2.3. Image Noise and Dose

Three different semi-anthropomorphic abdomen phantoms (Model 007TE, CIRS, VA) made of tissue equivalent materials were imaged as if they were actual patients using each site's routine clinical abdomen protocol. The phantoms represented the following patient size profiles: (1) 15 year old or small adult (23.9 cm $\times$ 18.5 cm), (2) medium adult (32.4 cm $\times$ 25.0 cm), and (3) large adult (38.7 cm $\times$ 31.0 cm), shown in Figure 3a. Following image acquisition, scan parameters were recorded. Notable parameters, including AEC settings, are listed in Table 1.

Image noise for each phantom and acquisition was calculated as the standard deviation of the pixel values within four manually placed ROIs, as depicted in Figure 3b. Each circular ROI had a diameter of 15 mm and was placed inside one of the central rod inserts of the phantom, which facilitated ROI placement consistency among all acquisitions. An average standard deviation value across all four ROIs was determined from noise measurements obtained from the central seven cross-sectional images of the phantom.

All console CTDI_{vol} values were recorded. These values are calculated (for all but two Toshiba scanners, see below) from the time-averaged tube current over the scan duration. Due to the relatively narrow extent of the abdomen phantoms along the z-axis, the average tube current varied along the z-axis, despite the uniform size of the phantom, due to edge effects. Average tube current values near the edges of the acquisition (which are near the physical edges of the phantoms) in some cases varied significantly from the values near the middle of the acquisition, which is where the tube current value accurately responds to the true size of the phantom. Consequently, the console CTDI_{vol} can be smaller or larger than the value in the centre of the phantom, depending on the edge tube current values. To obtain CTDI_{vol} values that reflect the scanner output in the centre of the phantom, where noise was measured, each console CTDI_{vol} value was multiplied by the average tube current

$(\overline{mA}_{central})$ generated during the production the 7 central images of the acquisition (i.e. images that were used for noise measurements) and divided by the average tube current over the entire acquisition $(\overline{mA}_{full})$. This calculation is shown in the following equation:

$$CTDI_{vol}^{true} = CTDI_{vol}^{console} \times \frac{\overline{mA}_{central}}{\overline{mA}_{full}}. \quad (1)$$

Notably, for the two Toshiba Aquilion 64 scanners, the console $CTDI_{vol}$ values were calculated differently than prescribed by IEC Standard 60601-2-44 Edition 3. Briefly, the scanner software used the peak mA value reached during the entire scan to calculate the $CTDI_{vol}$ value, as opposed to the using the average mA throughout the scan to determine $CTDI_{vol}$, which is what the IEC standard prescribes and is done by other scanners. This procedure generated artificially high values. To determine comparable $CTDI_{vol}$ values, we performed the following rescaling. First, we obtained a console $CTDI_{vol}$ value from a calibration scan under the condition of a fixed mA setting (i.e. the AEC system was deactivated). Second, we calculated the average effective mAs $(\overline{mAs}_{eff})$ output during the scan of anthropomorphic phantom with the following equation:

$$\overline{mAs}_{eff} = \frac{\overline{mA} \times t_{rot}}{pitch} \quad (2)$$

where $\overline{mA}$ is the average tube current during the scan, t_{rot} is the scanner rotation time, and $pitch$ is the helical pitch setting during the acquisition. Lastly, we calculated the $CTDI_{vol}$ value for each anthropomorphic phantom using the following equation:

$$CTDI_{vol} = CTDI_{vol}^{fixed\ mAs} \times \frac{\overline{mAs}_{eff}}{mAs_{fixed}} \quad (3)$$

where $CTDI_{vol}^{fixed\ mAs}$ is the value recorded under the condition of a fixed mAs value, mAs_{fixed} is the fixed mAs value during the calibration scan, and $(\overline{mAs}_{eff})$ the value is the described in equation 2.

3. Results

3.1. Axial resolution—SSP

For all scanners, the clinical imaging protocol called for 5 mm thick image slices. The SSPs for the investigated scanners and protocols demonstrated a bimodal distribution of FWHM values, and the shapes of these profiles were manufacturer dependent. These results are shown in Figure 4 and Table 2. Specifically, images reconstructed using GE's "Plus" mode setting possessed slice thicknesses that were on average 24% thicker than images reconstructed using GE's "Full" mode setting and images acquired with Toshiba and Siemens scanners. An average 6.2 ± 0.2 mm slice thickness was measured on GE scanners using the "Plus" mode setting for reconstruction; whereas an average slice thickness of 5.0 ± 0.1 mm was measured on all other scanners. In addition to differences in FWHM values, there was significant variation in the areas under the curves of the SSPs. For images

reconstructed using “Plus” mode, the average area under the curve was 5.91 (AU); whereas the average value from all other SSP measurements was 4.87 (AU).

3.2. In-plane spatial resolution—MTF

The calculated MTF curves are shown in Figure 5 along with MTF curves generated from other common convolution kernels to provide reference for the measured inter-scanner MTF variation. The MTF curves across all investigated scanners were similar. The average spatial frequencies at 50%, 10%, and 2% MTF values were 3.24 ± 0.37 , 6.20 ± 0.34 , and 7.84 ± 0.70 lp/cm, respectively. These data are reported in Table 3.

3.3. Image Noise and Dose

For all phantom sizes, the image noise varied significantly. The measured image noise values had the following ranges: 6.5–13.3 HU for the 15 year old phantom; 9.1–18.4 HU for the medium adult phantom; and 7.8–23.4 HU for the large adult phantom. These results are shown in Figure 6(a–c) along with target noise values and ranges for each phantom size. The target noise values were determined as the average noise value for each of phantom sizes produced by three specific scanners (Siemens: Sensation 16, Sensation 64, and Definition Flash) located at our primary campus that were selected to be reference scanners. The targeted noise range was set as $\pm 10\%$ of the target noise value.

As with the SSPs, both the image noise and change in image noise as function of phantom size also varied with manufacturer. These results are shown in the box plot depicted in Figure 7a. This plot shows that Siemens scanners yielded the highest average noise values. Images acquired from Toshiba scanners had the lowest average noise values for all three phantom sizes.

Also shown in this plot is the average change in image noise as a function of phantom size. For each scanner, the best linear fit of the noise data versus phantom size was calculated. Scanner-specific linear fits for each manufacturer were then averaged together and are shown as the solid line in the plot (Figure 7a). The shading around each fit-line corresponds to the standard deviation of the manufacturer-specific slopes determined by the linear fits. GE scanners demonstrated the smallest change in average noise values as a function of phantom size, which increased at a rate of 0.27 HU/cm, and was closely followed by the Toshiba scanners which yielded an average change of 0.31 Hu/cm. The performance of the Siemens scanners was substantially different. Image noise increased at a rate of 0.67 Hu/cm for these scanners.

CTDI_{vol} values varied considerably across all phantom sizes. CTDI_{vol} values ranged from 4.8–13.3 mGy for the 15 year phantom, 9.3–28.8 mGy for the medium phantom, and 16.0–48.1 mGy for the large phantom. These results are shown in Figure 6(d–f). Similar to the results for the image noise measurements, CTDI_{vol} values were correlated with manufacturer. These results are shown in the box plot depicted in Figure 7b. Toshiba scanners produced images using the largest dose to phantoms and demonstrated the greatest change in CTDI_{vol} values as a function of phantom size. Siemens scanners yielded the lowest CTDI_{vol} values and implemented the smallest change in CTDI_{vol} as a function of patient size. Solid linear fit-lines in Figure 7b were calculated just like those in Figure 7a,

and the shading around these curves indicates the standard deviation in the slopes of these linear fits.

4. Discussion

Standardizing a fleet of scanners requires a set of performance benchmarks to target. Here we have adopted the performance metrics of three scanners (Siemens: Sensation 16, Sensation 64, and Definition Flash) to serve as such benchmarks. These scanners are operated at our main site and the protocols for these systems have been carefully developed to deliver diagnostic quality images at the lowest doses achievable. After establishing benchmark values, the image quality of the other scanners should be standardized by adjusting acquisition parameters such that each target metric is achieved in the following order: (1) axial resolution—SSP, (2) in-plane spatial resolution—MTF, and (3) image noise. This procedure is illustrated in Figure 8. **In addition to these three steps, tube potential and collimation should be adjusted to best match benchmark tube potential and collimation selections.**

Through our first evaluation step, axial resolution was assessed through measurements of the scanners' SSPs. With the exception of images acquired using GE's "Plus" mode, variations in measured SSPs were minimal. For images acquired with "Plus" mode, both the FWHM and the area under the curve were, respectively, on average 24% and 22% greater than corresponding values measured on all other scanners. These results indicate that images produced using "Plus" mode incorporate more total projection data, including more data from adjacent slices, which has the effect of increasing blur along the z-axis and reducing noise in the image. Thus, as a first step in standardizing protocols (Figure 8), it is important to consider the SSP and achieve the target slice profile prior to matching image noise, as the image noise value is dependent on this attribute. Aside from changing from "Plus" to "Full" mode, there are few options to adjust the SSP; however, knowledge of the scanner's SSP is important when adjusting protocols based on image noise. If SSPs of different scanners substantially vary, it may be appropriate to utilize different slice thickness values.

The next step in our scanner evaluation procedure involves assessing in-plane resolution through MTF measurement. Note, the MTF measurements presented here employ a unique, home-made phantom; however, there are numerous other strategies that rely on more commonly found phantoms, such as the ACR CT accreditation phantom (Friedman *et al* 2013, Richard *et al* 2012, Takenaga *et al* 2014). In our study, the MTF measurements yielded similar response curves across all of the scanners, which indicate consistent in-plane spatial resolution. Image reconstruction methods constitute the primary influence on in-plane spatial resolution. Thus, implicit in these results is that the convolution kernels utilized to reconstruct the images performed similarly. It is well understood that the selection of the convolution kernel will greatly impact image noise. Therefore, adjustments to image reconstruction should be completed prior to adjusting radiation output to reach a specific image noise target. Accordingly, matching in-plane spatial resolution (i.e. MTF) by selecting the appropriate convolution kernel is the second step in standardizing image quality across different scanners (Figure 8). Previous investigations regarding matching convolution kernels of different vendors through experimental measurements can help guide adjustment

of this protocol parameter (Solomon *et al* 2012). Alternatively, multiple acquisitions using different convolution kernels can be executed to find the kernel that best matches the desired in-plane resolution.

The last step of our evaluation procedure entails measuring image noise. These noise measurements across a range of phantom sizes provide many insights into scanner performance, which can be used to establish cross-platform consistency. Our image noise measurements yielded several findings. First, absolute noise values showed correlation between manufacturers (Figure 7). This correlation is likely due to two main factors. First, the majority of scanners investigated are located at smaller, community medical practices, and such practices less frequently adjust acquisition parameters from the manufacturers' default values. Additionally, three of the five investigated Siemens scanners are located at our main site, and utilize protocols that have been developed on site to produce images of a standard quality level. The one Siemens scanner (Emotion 16) that produced image noise values that significantly deviated from those of the other Siemens scanners (Figure 6b) is located at a smaller satellite practice.

The image noise measurements also demonstrate how different scanners' AEC systems function in regards to phantom size. As shown in Figure 7, image noise is more strongly dependent on phantom size for Siemens scanners as compared to Toshiba or GE scanners. Siemens' AEC systems apply weaker tube current modulation as compared to other manufacturers, which yields different image noise values for different patient sizes to better meet clinical demands. Typically, smaller patients require higher image quality (i.e. lower noise) than larger patients. Consequently, maintaining a constant image noise level across all patient sizes could result in overexposure of larger patients and/or underexposure of smaller patients (i.e. non-diagnostic quality images) (McCollough *et al* 2006, Wilting *et al* 2001).

Although Toshiba and GE AEC systems are theoretically designed to yield constant image noise across patient sizes, our measurements show a slight increase in image noise as the phantom size increases. One method to avoid the aforementioned overexposure or underexposure of patients that can result from constant noise based AEC systems is to set tube current limits. Setting a maximum tube current limit can prevent overexposure of large patients by prohibiting the tube current from reaching values necessary to maintain constant image noise. Consequently, the image noise in this scenario will increase relative to the target value. Conversely, a minimum tube current limit prohibits the scanner from reducing the tube current to a level low enough to achieve the preset noise values for smaller patients. Thus, minimum tube current limits serve to lower image noise relative to the target value, resulting in higher image quality.

With few exceptions, most of the investigated GE scanners reached tube current limitations, which explains the varied image noise values as a function of phantom size. For the 15 year old phantom all but one of the scanners reached the user prescribed minimum tube current limit (Figure 6c). The one scanner that did not reach its tube current minimum was operated with a size-specific Noise Index value, i.e. a lower Noise Index value was set for small patients as compared to larger patients (see Table 1). For the large phantom, all but one of the scanners reached the user prescribed maximum tube current limit. The scanner that did

not reach the maximum tube current limit yielded a noise value that was consistent with the noise value produced in the medium phantom images.

Measurements on the Toshiba scanners also demonstrated a slight dependence of image noise on phantom size, despite an AEC system operating under a constant noise paradigm. In several instances, tube current minimum and maximum limits were reached; however, there are multiple instances in which the tube current was freely modulated and the noise value still changed with respect to the phantom size. We have not found any other parameters that contributed to the measured size dependency on the image noise. As noted, Toshiba AEC systems are purportedly based on a constant noise paradigm. However, it is possible that the manufacturer actually adjusts the tube current such that a constant noise value is not maintained; rather the tube current is modulated to a lesser degree resulting in less noisy images for small patients and noisier for larger patients. Others have shown also a similar dependence of image noise on phantom size for Toshiba scanners (Keat 2005).

As mentioned above, typically higher image quality (lower noise) is required for small patients, while relatively noisier images are acceptable for larger patients. Accordingly, our target noise values reflect these clinical constraints. AEC systems that are predicated on a constant noise paradigm do not natively accommodate these constraints, and thus acquisition protocols must be adjusted to account for these variations. There are two basic strategies to generate size dependent image noise. As discussed above, one method entails imposing tube current limits, which causes image noise to increase for larger patients and decrease for smaller patients. Through an alternative strategy to produce size-specific image quality, different AEC parameters are set for different patient sizes, essentially creating a size-dependent technique chart. Both strategies have unique advantages. Tube current limitations enable a fixed protocol for a wide spectrum of patients, which makes the acquisition procedure more automated and consistent. Conversely, size-specific AEC settings require the user to select the protocol, perhaps based on weight or BMI, which may be sub-optimal discrimination metrics. Hence, this procedure is more susceptible to variability in outcomes. However, size-specific AEC settings preserve tube current modulation during the scan, which has been shown to drastically reduce patient dose (Gies *et al* 1999, Kalender *et al* 1999, Kalra *et al* 2004, 2005, McCollough *et al* 2006). On the other hand, tube current limits may result with the scanner operating with a fixed tube current value (i.e. the minimum or maximum limit); thereby eliminating tube current modulation and inefficiently administering radiation dose. Further, the size-specific AEC strategy allows for the establishment of multiple size classes; whereas, tube current limits allow for three patient-size categories: (1) minimum fixed tube current, (2) modulated tube current, and (3) maximum fixed tube current. Due to its dose efficiency and protocol flexibility, the size-specific AEC settings strategy is a more advantageous method to standardize image quality.

Baseline measurements of image noise can guide protocol adjustments to achieve a target image noise value, which is the final step of standardizing scanner performance (Figure 8). Further, estimates of AEC settings can be determined based on the relation between the target value and the initial measurements. For scanners that utilize a constant noise-based AEC system, the “Noise Index” value (GE) and the “SD” parameter (Toshiba) can be assumed to vary proportionally with image noise. For scanners that use an effective mAs

value (Siemens) or a reference image (Philips), estimated adjustments to the mAs values (either the effective mAs AEC value, or the mAs value used to acquire the reference image)

may be determined with the relation of mAs and image noise (σ), $\sigma = \frac{1}{\sqrt{mAs}}$. Additionally, Siemens scanners provide a second parameter, which can be adjusted to better achieve a noise target across patient sizes. In these systems, the strength of tube current modulation as function of patient size can be adjusted with the following setting options: “weak,” “average,” or “strong.”

We chose image noise as a principal metric to standardize image quality, as opposed to a dose surrogate (e.g. CTDI_{vol} or size-specific dose estimate) because it is more representative of the objective of the imaging task. The goal of clinical imaging is to generate images of the patient from which diagnosis can be reliably rendered. Thus, it is paramount that all CT protocols achieve this goal, and standardizing image quality based on patient dose or a surrogate dose metric, does not ensure this goal will be met. Particularly when managing a diverse fleet of scanners, there is likely a range of dose efficiency among the fleet. Consequently, standardizing protocols by standardizing dose in such a scenario will lead to the production of varied image quality among scanners.

Presented here is a strategy to provide basic image quality standardization across different scanners. As a primary tenet of this strategy, the magnitude of image noise is a chief measure of image quality. This simplistic metric possesses several limitations in regards to its use as a standardization tool. Images with the same noise magnitude may possess different noise textures and consequently vastly different image quality (Boedeker *et al* 2007, Solomon *et al* 2012). Although image noise texture is not specifically evaluated in this standardization strategy, it is indirectly assessed. Noise texture is predominantly affected by the convolution kernel used during reconstruction. Through the MTF measurements, performance of the convolution kernels were indirectly evaluated, which largely obviates the need for additional noise texture measurements.

Additionally, image contrast, which contributes to the conspicuity of an object, is excluded from the proposed standardization framework. Despite using the same tube potential, contrast differences may exist between scanner models due to differences in beam filtration, detector sensitivities, and reconstruction parameters. Hence, image contrast could be used to adjust the target noise value in the standardization procedure, or more simply, a contrast-to-noise ratio could be used as the primary image quality metric. This was not done in our work for two primary reasons. First, contrast measurements from a high contrast material (i.e. bone mimicking material) performed on images acquired at the same tube voltage showed relative differences within 10% of a target value (defined by the same three scanners used to define the target noise) for all but 2 of the 38 scanner-phantom combinations (data not shown). Thus, contrast-based adjustments to the target noise value for each scanner and patient size would have a minimal impact, yet would complicate the standardization procedure. Second, contrast differences can vary both with source material and location in a phantom. Thus, the appropriateness of using contrast-to-noise ratio as an image quality metric for standardization depends on the structure of interest and the diagnostic task. For example, for contrast-enhanced vascular exams, use of the contrast-to-noise ratio of the

iodinated blood relative to the background soft tissue is appropriate; for contrast-enhanced routine abdominopelvic exams, a noise constrained iodine contrast-to-noise ratio is appropriate (Winklehner *et al* 2011, Yu *et al* 2010); and for non-contrast abdominopelvic exams, use of a soft tissue based contrast-to-noise ratio is more appropriate. Therefore, there is no universal contrast-to-noise ratio that can be used as an appropriate image quality metric. As mentioned above, for scanners using the same tube potential setting, contrast differences were minimal. Consequently, matching the noise level is a good first order approximation to matching the CNR, regardless of the contrast amount. Thus, image noise was used as the metric by which to standardize image quality among exams from different scanners, when spatial resolution was matched.

Another limitation regarding the use of noise as an image quality metric pertains to its applicability across different reconstruction methods. Image noise is a suitable performance metric when images are reconstructed through filtered back projection; however, it may insufficiently describe image quality for images produced through iterative reconstruction algorithms. Investigations have demonstrated that images produced through iterative reconstruction may suffer from reduced low contrast resolution despite possessing low image noise (McCollough *et al* 2015). Thus for images reconstructed with iterative methods, the magnitude of image noise may be an inappropriate standardization metric. Currently, many scanners offer iterative-based image reconstruction. In fact, three of the investigated scanners presented here employed iterative reconstruction techniques (see Table 1). However, we have developed this standardization approach for a straightforward acquisition protocol for our existing scanner fleet. Application of this method to more complex acquisition protocols and inclusion of more advanced CT scanners (e.g. dual energy) may require more sophisticated assessment metrics, such as observer models.

5. Conclusions

We have established a general CT scanner evaluation procedure to compare image quality. Included in our evaluation procedure are measurements of SSP, MTF, and image noise. The scanners across our health system demonstrated varied performance, particularly in regards to image noise. These differences in image noise are greatly impacted by differences in AEC implementations. Using results from our scanner survey, we have formulated a general strategy to standardize CT image quality across different manufacturers and models. Through this methodology, first a set of benchmark values must be established. Then, SSP should be standardized via selection of the slice thickness or adjustment of the reconstruction mode on GE scanners (“Full” vs. “Plus”). Next, MTF values should be standardized through appropriate selection of the convolution kernel. Lastly, image noise values across a range of phantom sizes should be standardized through modification of AEC settings and the application of size-specific AEC parameters (i.e., Noise Index and SD) for scanners based on a constant noise paradigm.

Acknowledgements

Funding: The project described was supported by Grant number EB 017095 from the National Institute of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Health.

References

- American College of Radiology. CT Accreditation Program Requirements. 2013. Online: <http://www.acr.org/~media/ACR/Documents/Accreditation/CT/Requirements.pdf>
- Bhalla, AS.; Arora, AD. Diagnostic Radiology: Recent Advances and Applied Physics in Imaging. Gupta, AK.; Chowdhury, V.; Khandelwal, N., editors. New Delhi: JP Medical Ltd; 2013.
- Boedeker KL, Cooper VN, McNitt-Gray MF. Application of the noise power spectrum in modern diagnostic MDCT: part I. Measurement of noise power spectra and noise equivalent quanta. *Phys. Med. Biol.* 2007; 52:4027–4046. [PubMed: 17664593]
- Cody DD, Fisher TS, Gress DA, Rick Robert Layman J, McNitt-Gray MF, Robert J, Pizzutiello J, Lynne A, Fairbent AS. AAPM Medical Physics Practice Guideline 1.a: CT Protocol Management and Review Practice Guideline. *J. Appl. Clin. Med. Phys.* 2013; 14
- Couch G, Couch J. Dose estimation service system configured to support multiple computerized medical imaging scan providers. 2012
- Favazza CP, Yu L, Leng S, Kofler JM, McCollough CH. Automatic exposure control systems designed to maintain constant image noise: effects on computed tomography dose and noise relative to clinically accepted technique charts. *J. Comput. Assist. Tomogr.* 39:437–442. [PubMed: 25938214]
- Friedman SN, Fung GSK, Siewerdsen JH, Tsui BMW. A simple approach to measure computed tomography (CT) modulation transfer function (MTF) and noise-power spectrum (NPS) using the American College of Radiology (ACR) accreditation phantom. *Med. Phys.* 2013; 40:051907. [PubMed: 23635277]
- Gies M, Kalender WA, Wolf H, Suess C, Madsen MT. Dose reduction in CT by anatomically adapted tube current modulation. I. Simulation studies. *Med. Phys.* 1999; 26:2235. [PubMed: 10587204]
- Kalender WA, Wolf H, Suess C. Dose reduction in CT by anatomically adapted tube current modulation. II. Phantom measurements. *Med. Phys.* 1999; 26:2248. [PubMed: 10587205]
- Kalra MK, Maher MM, Toth TL, Schmidt B, Westerman BL, Morgan HT, Saini S. Techniques and applications of automatic tube current modulation for CT. *Radiology.* 2004; 233:649–657. [PubMed: 15498896]
- Kalra MK, Naz N, Rizzo SMR, Blake MA. Computed tomography radiation dose optimization: scanning protocols and clinical applications of automatic exposure control. *Curr. Probl. Diagn. Radiol.* 2005; 34:171–181. [PubMed: 16129235]
- Keat, N. Report 05016: CT scanner automatic exposure control systems. London: 2005.
- Lee CH, Goo JM, Ye HJ, Ye S-J, Park CM, Chun EJ, Im J-G. Radiation dose modulation techniques in the multidetector CT era: from basics to practice. *Radiographics.* 2008; 28:1451–1459. [PubMed: 18794318]
- McCollough CH, Bruesewitz MR, Kofler JM. CT dose reduction and dose management tools: overview of available options. *Radiographics.* 2006; 26:503–512. [PubMed: 16549613]
- McCollough CH, Yu L, Kofler JM, Leng S, Zhang Y, Li Z, Carter RE. Degradation of CT Low-Contrast Spatial Resolution Due to the Use of Iterative Reconstruction and Reduced Dose Levels. *Radiology.* 2015:142047.
- McKenney SE, Seibert JA, Lamba R, Boone JM. Methods for CT automatic exposure control protocol translation between scanner platforms. *J. Am. Coll. Radiol.* 2014; 11:285–291. [PubMed: 24589404]
- 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. [PubMed: 22830744]
- Solomon JB, Christianson O, Samei E. Quantitative comparison of noise texture across CT scanners from different manufacturers. *Med. Phys.* 2012; 39:6048–6055. [PubMed: 23039643]
- Takenaga T, Katsuragawa S, Goto M, Hatemura M, Uchiyama Y, Shiraishi J. Modulation transfer function measurement of CT images by use of a circular edge method with a logistic curve-fitting technique. *Radiol. Phys. Technol.* 2014; 8:53–59. [PubMed: 25142743]

- Wiltink JE, Zwartkruis A, van Leeuwen MS, Timmer J, Kamphuis AG, Feldberg M. A rational approach to dose reduction in CT: individualized scan protocols. *Eur. Radiol.* 2001; 11:2627–2632. [PubMed: 11734970]
- Winklehner A, Goetti R, Baumuehler S, Karlo C, Schmidt B, Raupach R, Flohr T, Frauenfelder T, Alkadhi H. Automated attenuation-based tube potential selection for thoracoabdominal computed tomography angiography: improved dose effectiveness. *Invest. Radiol.* 2011; 46:767–773. [PubMed: 21730872]
- Yu L, Li H, Fletcher JG, McCollough CH. Automatic selection of tube potential for radiation dose reduction in CT: A general strategy. *Med. Phys.* 2010; 37:234. [PubMed: 20175486]

Figure 1.

(a) Photograph of the slice sensitivity profile phantom with an arrow pointing to the gold foil object. (b) Sample cross-sectional images depicting the slice sensitivity profile measurement procedure. A 3×3 pixel² region of interest (ROI) was algorithmically centered inside the gold foil image (center slice). The average pixel value was determined inside that ROI location for all image slices.

Figure 2.

(a) Photograph of the modulation transfer function measurement phantom with an arrow pointing to the tantalum wire. (b) Sample cross-sectional image of the wire after re-sizing the image and interpolating pixels. (c) Sample fast Fourier transform image of the real-space wire image shown in (b).

Figure 3.

(a) Photograph of the anthropomorphic abdomen phantoms positioned in a CT scanner. (b) Sample cross-sectional image of the medium abdomen phantom used for noise calculations. The 15 mm diameter regions of interest outlined by the red circles depict the regions from which image noise was measured.

Figure 4. Plot of four representative slice sensitivity profiles obtained from Siemens (red curve), Toshiba (green curve), and GE scanners, both “Full” (solid blue) and “Plus” (dashed blue) reconstruction modes on GE scanners.

Figure 5.

Plot of the MTF curves measured on all of the investigated scanners (each manufacturer plotted with a different colour: GE—blue, Siemens—red, and Toshiba—green), along with MTF curves measured using a range of Siemens convolution kernels (dashed black lines), where B10 is the smoothest body kernel and B50 is a very sharp body kernel.

Figure 6.

(a–c) Average noise values for all Toshiba (a), Siemens (b) and GE (c) scanners. Average and standard deviation values were calculated from image noise measurements obtained from the central 11 slices for each scanner and phantom combination. Also shown are target noise values (black stars) determined as the average noise values from 3 well calibrated scanners (Siemens Sensation 64, Sensation 16, Definition Flash), which are operated at the main campus of our institution. The target range of image noise values (dashed black lines) represent $\pm 10\%$ of the target noise values. Red crosses indicate instances when the scanner reached a tube current limit. (d–f) CTDI_{vol} values for all Toshiba (d), Siemens (e), and GE (f) scanners.

Figure 7.

(a) Manufacturer specific box plots of image noise as a function of phantom size (GE in blue, Siemens in red, and Toshiba in green). Data from Toshiba and Siemens scanners were artificially offset on the x-axis for display purposes. Black stars and dashed lines represent target noise values and ranges, respectively, for each phantom size. Black dots represent average values obtained from each manufacturer specific set of scanners. Solid lines represent the manufacturer-specific average of the best linear fits of image noise versus phantom size, which were determined for each individual scanner. The shading surrounding the solid line represents the standard deviation of manufacturer-specific linear fits. (b) Manufacturer-specific CTDI_{vol} box plots, similar to the plots in (a).

Figure 8.

A flow chart of the general strategy to standardize image quality performance across a fleet of CT scanners comprised of different manufacturers and models.

Table of acquisition parameters for each scanner investigated. The second column shows the primary AEC settings for all scanners. SD is the primary AEC parameter on Toshiba scanners, which represents the standard deviation of CT numbers (i.e. noise). QRM, which stands for Quality Reference mAs, is the primary AEC parameter for Siemens scanners and represents the effective mAs for a reference patient size. NI, which stands for Noise Index, is the primary AEC parameter for GE scanners and represents the target noise value in the acquired image. Iterative based reconstruction is noted by AIDR 3D for the Toshiba scanner and SS30 (ASIR) setting for the GE scanners in the convolution kernel column.

Table 1

Scanner	SD/QRM/NI (15 yr Med Lg)	Effective mAs			kVp (15 yr Med Lg)	Convolution Kernel	Slice Thick. (mm)	
		15 yr	Med	Lg				
Toshiba	Aquilion 16	12.5	53	91	267	120	FC13	5
	Aquilion 32	12.5	95	150	364	120	FC08	5
	Aquilion Premium	12.5	107	114	236	120	FC13(AIDR 3D)	5
	Aquilion 64	12.5	49	129	232	120	FC13	5
	Aquilion 64	12.5	76	119	381	120	FC13	5
Siemens	Emotion 16	150	71	110	368	130	B31s	5
	Sensation 16	240	70	116	187	120	B40f	5
	Sensation 40	210	100	165	286	120	B41f	5
	Sensation 64	350 350 240	70	154	357	100 100 120	B40f	5
	Definition FLASH	240	85	240	242	120	B40f	5
GE	Discovery 750 HD	15.9	261	346	366	120	STD (SS30)	5 (Plus)
	LightSpeed 64 VCT	11.5	74	150	263	120	STD (SS30)	5 (Plus)
	LightSpeed 64 VCT	10.4 12.5 14.5	101	218	302	80 100 120	STD	5 (Full)
	BrightSpeed Elite 16	11.5	72	251	319	120	STD	5 (Plus)
	LightSpeed 16	11.5	73	254	322	120	STD	5 (Plus)
	BrightSpeed 16	11.5	72	182	319	120	STD	5 (Plus)

Table 2

Statistics of slice sensitivity profile measurements—full width at half maximum and area under the curve. Included are the minimum and maximum measured values along with average and standard deviations derived from three groups of scanners: [1] all scanners (*total*), [2] Siemens + Toshiba + GE—“Full” mode reconstruction (*non-plus mode*), and [3] GE—“Plus” mode reconstruction (*plus mode*).

SSP Statistics		
	FWHM (mm)	Slice Area (AU)
average (<i>total</i>)	5.4	5.22
std. dev. (<i>total</i>)	0.6	0.53
average (<i>non-plus mode</i>)	5.0	4.87
std. dev. (<i>non-plus mode</i>)	0.1	0.16
average (<i>plus mode</i>)	6.2	5.91
std. dev. (<i>plus mode</i>)	0.2	0.11
minimum	4.7	4.57
maximum	6.3	6.01

Table 3

Statistics of spatial frequencies that yielded modulation transfer function values of 50%, 10%, and 2%. Also included spatial frequencies that yielded modulation transfer function values of 50%, 10%, and 2% for three different Siemens convolution kernels.

MTF Statistics			
	50% (lines/cm)	10% (lines/cm)	2% (lines/cm)
average	3.54	6.20	7.84
std. dev.	0.37	0.34	0.70
minimum	2.84	5.81	7.11
maximum	4.10	7.19	10.13
B10	2.4	4.8	5.6
B30	3.2	5.6	7.2
B50	6.4	8.8	9.6