Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Acad Radiol. Author manuscript; available in PMC 2019 April 24.

Published in final edited form as:

Acad Radiol. 2017 July ; 24(7): 876–890. doi:10.1016/j.acra.2016.12.017.

Estimation of Observer Performance for Reduced Radiation Dose Levels in CT: Eliminating Reduced Dose Levels that are Too Low Is the First Step

Joel G. Fletcher, MD[1,*], **Lifeng Yu, PhD**[1], **Jeff L. Fidler, MD**[1], **David L. Levin, MD, PhD**[1], **David R. DeLone, MD**[1], **David M. Hough, MB, ChB**[1], **Naoki Takahashi, MD**[1], **Sudhakar K. Venkatesh, MD**[1], **Anne-Marie G. Sykes, MD**[1], **Darin White, MD**[1], **Rebecca M. Lindell, MD**[1], **Amy L. Kotsenas, MD**[1], **Norbert G. Campeau, MD**[1], **Vance T. Lehman, MD**[1], **Adam C. Bartley, MS**[2], **Shuai Leng, PhD**[1], **David R. Holmes III, PhD**[3], **Alicia Y. Toledano, PhD**[4], **Rickey E. Carter, PhD**[2], and **Cynthia H. McCollough, PhD**[1]

[1]Department of Radiology, Mayo Clinic, Rochester, MN, Mayo Clinic, 200 First Street SW, Rochester, MN 55905

[2]Department of Health Sciences Research, Mayo Clinic, 200 First Street SW, Rochester, MN 55905

[3]Department of Physiology and Biomedical Engineering, Mayo Clinic, 200 First Street SW, Rochester, MN 55905

[4]Biostatistics Consulting, LLC, 10606 Wheatley Street, Kensington, MD 20895

Abstract

Rationale and Objectives: To estimate observer performance for a range of dose levels for common CT exams (detection of liver metastases or pulmonary nodules, and cause of neurologic deficit) in order to prioritize non-inferior dose levels for further analysis.

Materials and Methods: Using CT data from 131 exams (44 abdominal CT, 44 chest CT, 43 head CT), CT images corresponding to 4 – 100% of routine clinical dose were reconstructed with filtered back projection (FBP) or iterative reconstruction (IR). Radiologists evaluated CT images, marking specified targets, providing confidence scores, and grading image quality. Non-inferiority was assessed using reference standards, reader agreement rules and jacknife alternative free-response receiver operating characteristic (JAFROC) figure of merit (FOM). Reader agreement required that a majority of readers at lower dose identify target lesions seen by the majority of readers at routine dose.

Results: Reader agreement identified dose levels lower than 50% and 4% to have inadequate performance for detection of hepatic metastases and pulmonary nodules respectively, but could not

[*]**Corresponding Author:** 200 First Street SW, Rochester, MN 55905, USA, Phone: (507) 284-2511, Fax: (507) 266-3661, fletcher.joel@mayo.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

exclude any low dose levels for head CT. Estimated differences in JAFROC FOM between routine and lower dose configurations found that only the lowest dose configurations tested (i.e., 30%, 4% and 10% of routine dose levels for abdominal, chest and head CT, respectively) did not meet criteria for non-inferiority. At lower doses, subjective image quality declined before observer performance. IR was only beneficial when FBP did not result in non-inferior performance.

Conclusion: Opportunity exists for substantial radiation dose reduction using existing CT technology for common diagnostic tasks.

Keywords

CT; radiation dose; radiation dose reduction; iterative reconstruction; observer performance

Introduction

Computed tomography (CT) imaging is widely used in medical practice to quickly and accurately diagnose a large number of medical conditions and guide appropriate medical or surgical management. CT is commonly available and readily interpreted by radiologists and referring clinicians, and can provide images with high spatial and temporal resolution that cover large regions of the body, with few contraindications to imaging.

The most notable concern for CT imaging is that it requires exposure to ionizing radiation, and there is both provider and patient concern over the risk of radiation-induced malignancy, with this risk being low but uncertain for each patient, and a subject of ongoing debate (1, 2). The medical justification for performing CT imaging is that anticipated benefit exceeds anticipated radiation exposure risk. The optimal dose of radiation that would accomplish the diagnostic task but not compromise diagnostic performance should be used (3). However, this “optimal” dose is not known for even the most common CT exams, resulting in a wide spectrum of radiation doses being used, even for common diagnostic tasks (4). Alternatively, some radiology practices reduce radiation dose so much that observer performance is compromised, negating the beneficial impact of CT imaging. The lack of available evidence on what dose levels are required to achieve specific results is due to many factors, e.g., evolving technology, inability or impracticality to obtaining multiple dose levels simultaneously or consecutively, fear that substantially lowering of dose may result in compromised diagnostic benefit, and difficulty in study design and expense (5). To this end the National Institute of Biomedical Imaging and Bioengineering (NIBIB) has held a series of expert conferences and sponsored grants to lay out a pathway to systematically lower radiation dose from CT imaging (6, 7).

Several recent technical and clinical advances have expanded opportunities to lower radiation dose: automatic exposure control (8) and automated tube potential selection adapt radiation dose to patient size and diagnostic task (9); iterative reconstruction substantially improves the image quality of lower dose CT images (10), but with an unclear impact performance for many diagnostic tasks (11–15); large scale screening studies have demonstrated that low radiation doses can be used to detect early lung cancer or colorectal cancers and polyps (16, 17).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

We have begun a multiphase, systematic evaluation to determine the relationship of radiation dose to observer performance for three of the most common diagnostic tasks in our imaging practice: the detection of hepatic metastases with contrast-enhanced abdominal CT, the detection of pulmonary nodules with unenhanced chest CT, and the detection of causes of acute neurologic deficit with unenhanced head CT. We aim to achieve this goal using the ability to archive the CT projection data in patients with proven pathology and reconstruct CT images from the projection data that accurately simulates lower dose images (18). This goal can be achieved in a multiphase study design, wherein the first step uses a small number of readers to winnow down a wide range of radiation dose levels to a smaller number of dose levels that can be tested with a larger number of cases, eventually culminating in a large multi-reader, multi-case study.

The first phase of our investigation employs a smaller number of subspecialized radiologist readers and a smaller number of cases to identify lower doses that can be expected to result in inferior reader performance. This selection prioritizes of a smaller number of critical dose levels and reconstruction methods for subsequent comprehensive examination.

Materials and Methods

This multi-reader study was approved by our Institutional Review Board and was HIPAA compliant. For this study, we archived CT projection data for contrast-enhanced abdominal (n=789, from 9/25/2012 to 12/17/2013), non-contrast chest (n=118, from 8/28/2013 to 12/17/2013), and non-contrast head CT (n=924, from 8/19/2013 to 12/17/2013). For each of these examined regions, a subspecialized radiologist (“diagnostic task leader”) oversaw the collection of CT projection data, and evaluated all potential patient cases for their ability to meet inclusion criteria. Appropriate normal controls were also selected for each of these regions.

Inclusion criteria included: (1) appropriate CT exam type performed on a 128- slice multi-detector CT platform (Siemens FLASH or Siemens AS+; Siemens Healthcare, Malvern, PA); (2) patient consent to retrospective evaluation of medical records and imaging data for research purposes; and (3) successful archival of CT projection data. For contrast-enhanced abdominal and unenhanced head CT exams, diagnostic task leader review of CT images and all medical records confirmed that imaging findings met reference standard criteria for positive or negative exams (Table 1). For unenhanced head CT, the potential causes of neurologic deficit to be included in the study included infarction (acute, subacute, or chronic), traumatic and non-traumatic intracranial hemorrhage, or mass.

Exclusion criteria included (1) failure to meet all inclusion criteria, (2) for abdominal CT exams, more than seven hepatic lesions (metastases or non-metastatic benign lesions), or hepatic lesions larger than 5 cm, and (3) for chest CT exams, diagnostic task leader assessment of images and all medical data indicating prior thoracic, surgery, active pneumonia or air-space disease, or an excessive number of pulmonary nodules (>10).

The overall study schema for this study is provided in Figure 1. At the end of the data accrual phase, each diagnostic task leader selected cases for inclusion, with the goal of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

selecting 50% of the cases so that they were positive. For positive cases meeting inclusion criteria, each diagnostic task leader selected the most visually challenging cases for inclusion. For abdominal CT and head CT, half of the normal cases were selected to have a benign liver lesion or leukoaraiosis, respectively.

CT acquisition and image reconstruction of varying dose levels

Patients were scanned according to our routine clinical CT acquisition protocol (Table 2). CT examinations were performed on a Definition FLASH or AS+ CT system from Siemens Healthcare (2 × 128 slices, 0.28 second rotation time), with the same x-ray tube and automatic exposure control system (Straton® tube, CareDose4D; Siemens Healthcare). Unenhanced chest and head CT exams were performed at 120 kVp. For contrast-enhanced abdominal CT exams, a vendor-supplied kV selection system (CarekV, Siemens Healthcare, Malvern, PA) was additionally used, and a validated noise insertion program was used to normalize the “baseline” or routine abdominal CT examination if the original quality reference mAs (an automatic exposure control setting) was above 200 (18). The mean size-specific dose estimate (SSDE) was calculated for each exam using dose information from exam images as previously described (5)

Lower-dose projection data were created by using the validated noise insertion program (5, 18). The altered CT projection data was loaded back onto the original CT scanner and reconstructed using the routine filtered back projection (FBP) or corresponding iterative reconstruction (IR) method, as detailed in Table 2. In this manner, 6 (abdominal and head CT) or 5 (chest CT, owing to increased interpretation time) dose levels for each diagnostic task using FBP and IR, yielding 10–12 configurations produced per original patient exam.

Image evaluation

For this phase of our study, a total of nine radiologists, three subspecialized radiologists for each diagnostic task, interpreted cases. Cases were anonymized and de-identified for both the dose and reconstruction method. A randomized reading schedule was created for each reader such that one of the dose / reconstruction configurations for each patient was selected per reading session. This required each reader to read multiple reading sessions in order to evaluate all configurations under study (the number of sessions equaling the number of configurations under study), with separate sessions separated by at least 7 days to reduce potential recall bias related to recognition of prior positive cases by the radiologists. An open-source, dedicated computer workstation was constructed to facilitate examination of lower dose configurations (19), and was adapted for each imaging task to record the location of reader and reference markings, as well as all other data input by the readers. Axial and coronal images were evaluated for abdominal CT exams, thin section and overlapping 10 mm MIP images were evaluated for chest CT exams, and only axial images were evaluated for unenhanced spiral head CT exams. Readers were blinded to all clinical data, as well as radiation dose level and image reconstruction method.

Prior to study initiation, the principal investigator met personally with each radiologist reader for hands-on training of the diagnostic workstation functionality, including methods used for annotating of lesions of interest, and instruction on appropriate methods for

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

recording reader diagnosis, confidence, and image quality. For each case, radiologists were directed to scroll through the dataset using multiple window and level settings. When a target lesion corresponding to the imaging task was identified, the radiologist was instructed to circumscribe the lesion and provide a diagnosis using a drop-down menu (Table 3), using a slider bar to indicate the level of confidence that the lesion was an hepatic metastasis, indeterminate pulmonary nodule, or particular cause of acute neurologic deficit (0 – 100 scale). For example, readers were instructed that even if they thought a lesion was benign, the confidence assigned was the confidence that the relevant target lesion was present. For example, a reader might identify a cystic liver lesion and diagnose it as a hepatic cyst, but then assign a confidence level of 10, indicating a low likelihood of 10% for the lesion to represent a cystic metastasis. After all lesions had been identified and selected, readers were given an image quality survey, which has been previously reported (19). The workstation permitted only focal lesions 3 mm or larger in size to be marked, and readers were informed of this limitation. However, readers were instructed that there was no minimum size to lesions that should be marked, except for chest CT, where readers were informed that they should mark only lesions 5 to 15 mm in size. After completing a case, the reader was not allowed to revisit it.

Following identification of target lesions of interest, each radiologist was asked to grade subjective image quality using modified European Quality criteria (5, 20), with overall diagnostic image quality rated using a 5 point scale (1=non-diagnostic due to excessive noise artifacts; 2=diagnosis questionable due to excessive noise/artifacts; moderate decrease in confidence; 3=diagnostic with moderate but acceptable noise/artifacts; 4=mild noise, no change in confidence; 5= routine diagnostic image quality).

Reference Standard Assignment and Matching of Reader and Reference Lesions

The subspecialized radiologist functioning as the diagnostic task leader directed the selection of cases and documentation of the reference standard for every lesion, as outlined in Table 1. For contrast-enhanced abdominal CT and unenhanced head CT, diagnostic task leaders compared the routine dose exam to all medical records, including prior and subsequent imaging. Every focal lesion within the target organ was circumscribed using a circular or spline region of interest (ROI), and the same diagnoses available to the readers were selected based on reference criteria. For every lesion within every patient, a confidence score of 100 was assigned when the lesion met reference criteria, and a score of less than 100 was assigned when reference criteria were not met. For the detection of indeterminate pulmonary nodules, the diagnostic task leader and another thoracic radiologist evaluated routine-dose images unblinded to the medical record (and prior and subsequent exams) and circled non-calcified pulmonary nodules of target size; the reference standard for chest CT exams was nodules circled by both of these reference readers.

Matching of reader detections and target reference lesions was performed automatically by the computer workstation software using an overlapping spheres method for hepatic metastases and pulmonary nodules (5). For head CT findings, the workstation was modified so that reader and reference markings were displayed simultaneously in different colors, with subsequent semi-automated matching of reference vs. reader markings (19) by the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

neuroradiology task leader. This manual matching algorithm allowed identification of reader markings as false-positive or irrelevant (if extracranial), and unpaired reference markings as false negative exams. Manual matching was used for this head CT as critical findings might occur at multiple levels within the imaged volume so that reader and reference markings may not overlap (e.g., intraventricular hemorrhage). At the conclusion of the study, we discovered that one neuroradiologist reader had circumscribed only acute findings on the computer workstation, and listed subacute and chronic findings and their location on hard copy sheets returned to the study coordinator. In this case, the head CT diagnostic task leader was instructed to match hard copy findings with reference markings manually. This matching sufficed for determination of reader agreement rules. For other calculations, the study team determined that the confidence level would be input as 100.

In addition to marking and documenting target lesions, diagnostic task leaders closely circumscribed all target lesions for automated calculation of imaging features included size, CT number difference compared to background, and contrast-to-noise. For this purpose, they also placed regions of interest in the adjacent liver or white matter (for abdominal and head CT).

Statistical Analysis

The studies were designed as a novel adaption of the Simon two-stage design commonly used for oncology clinical trials (21). Full details on the adapted design are found in the Appendix. The results presented here are for the first stage (planned interim analysis) for each of the three diagnostic tasks. The sample size for stage I of the study was estimated to be 44 cases based on an optimal two-stage design with the null success rate of 80%, an alternative success rate of 90%, and 90% power at alpha=0.05 (one-sided).

The screening process for imaging strategies was based on a binary reader agreement rules defining “successful” interpretation. If 37 of the 44 cases were read successfully (defined below) for a particular dose / reconstruction approach, it would be considered for continuation onto stage II, consisting of fewer imaging strategies but more cases and a similar number of readers. Additionally, standard multi-reader, multi-case analysis employing the JAFROC Figure of Merit (FOM) was also used. The latter took into account relative confidence between true positives and false positives localizations. For statistical analysis, benign lesions with diagnoses not relevant to the diagnostic task (e.g., benign liver lesions) were not considered when the assigned confidence score was zero.

For binary reader agreement rules, we constructed a list of “essential” lesions for each case that consisted of target reference lesions identified by the majority of readers (i.e., at least 2 of 3) when interpreting the routine dose exams (Appendix). Binary reader agreement rules rated a case as being “successfully” interpreted when the majority of readers identified every “essential” lesion within a positive exam, and when a majority of readers made no false positive localizations (i.e., false positive detections) for a negative exam.

To determine if lower radiation dose CT exams resulted in non-inferior observer performance, we calculated the estimated differences in the JAFROC FOM for the routine dose dataset compared to the lower dose configurations, along with 95% confidence

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

intervals, using −0.10 as our limit of non-inferiority. Non-inferiority was demonstrated when the lower limit of the 95% confidence interval of the difference between the routine dose and lower dose configuration was above −0.10 (Figure 2). The limit of non-inferiority was determined prior to the start of the study based on the investigators' judgment and the sample size.

To determine factors that may contribute to whether or not a lesion was detected by radiologists at the different dose levels, the Spearman correlation coefficient between the mean reader confidence for each lesion and objective imaging features (i.e., lesion size, CT number difference from adjacent normal parenchyma, image noise, contrast-to-noise ratio) was examined for each dose level, with p-values less than 0.05 considered significant. A confidence score of 0 was assigned if a reader failed to detect the lesion of interest. Contrast-to-noise ratio was defined as the CT number difference between the lesion and the adjacent organ parenchyma divided by the image noise of the lesion.

General statistical analyses were conducted using the SAS System version 9.4 (Cary, NC). Estimation of the JAFROC FOM was estimated using the RJafroc software version 4.2.1.

Results

There were 44 patients that underwent contrast-enhanced abdominal CT. Seventy-five hepatic proven hepatic metastases were present in 22 patients, with median lesion size of 14 mm (25^{th} – 75^{th} percentile: 10 – 18 mm), and a median absolute CT number difference from background liver of 18 HU (25^{th} – 75^{th} percentile: 8 – 35 HU). These metastases originated from a variety of primary tumors (colorectal, 9; neuroendocrine, 3; sarcoma, 3; pancreatic, 2; melanoma, 2; other, 3). There were 60 benign hepatic lesions (cyst, 26; hemangioma, 10; vascular perfusion defect, 12; post-op or post-radiofrequency ablation defect, 6, focal fat, 6) present in 15 patients.

Of 44 patients undergoing routine unenhanced chest CT for pulmonary nodule detection, 21 patients had 28 indeterminate pulmonary nodules with a median size of 6.8 mm (mean, 6 mm; 25^{th} – 75^{th} %tile: 5.3 – 8.5 mm), with three of these nodules rated as being semi-solid and the remaining being solid nodules.

Of 43 patients with suspected neurologic deficit undergoing unenhanced spiral head CT, 21 of 43 patients had 30 intracranial findings (16 infarction [acute, subacute, or chronic], 9 traumatic and non-traumatic intracranial hemorrhage, and 5 masses; non-exclusive listing), with a median CT number difference from adjacent normal white matter being 11 HU (25^{th} – 75^{th} percentile: 6 – 20 HU), with the median size of positive findings being 3.7 cm^2 (25^{th} – 75^{th} percentile: 1.4 – 10.6 cm^2) in maximal cross-sectional dimension on an axial slice. For abdominal, chest, and head CT exams, the mean CT volume dose index ($CTDI_{vol}$) for the routine dose exams in this study was 11.8 ± 4.1, 5.5 ± 1.7, and 38.2 ± 0 mGy respectively, and the mean size-specific dose estimate (SSDE) was of 14.3 ± 3.4, 7.2 ± 1.6, and 26.2 ± 1.2 mGy, respectively (22, 23).

The number of successful interpretations based reader agreement rules is shown in Table 4 for each dose-kernel configuration. The median time interval between reading sessions was

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

9 days (25th – 75th percentile: 7 – 15 days). For hepatic metastases detection, dose levels below an AEC setting of 100 QRM (nominal $CTDI_{vol}$ 6.8 mGy) failed to meet reader agreement rules unless IR was used, and both FBP and IR configurations failed below this dose level. For unenhanced chest CT, all dose levels and configurations met the reader agreement criteria (despite a greater than 10-fold difference in dose levels), except for the lowest dose level tested in chest CT exams reconstructed with FBP. Similarly, all lower dose configurations met reader agreement rules for unenhanced head CT (despite 10-fold reduction in dose). There were a few positive cases within each exam type in which fewer than 2 radiologist readers identified the target lesion at routine dose (liver - 2, chest - 2, head - 3), so negative findings at lower dose configurations were not interpreted as unsuccessful when these subtle lesions were not identified (i.e., in Table 4 these cases were categorized as “cases without essential lesions”). Unsuccessful interpretations in cases without essential lesions occurred when 2 or 3 radiologists created false positive markings. From Table 4 it can be seen that there were very few unsuccessful interpretations in negative head CT exams across all dose levels (0 or 1), and no trend towards increase in false positive markings as dose declined for unenhanced chest CT or contrast-enhanced abdominal CT. Based on reader agreement rules alone, dose reductions to 50%, 4% and 10% of currently utilized clinical doses met criteria (at least 37 successful interpretations) for further evaluation in the next phase of this study. Figure 3 provides a representative example of successfully interpreted lower dose configurations for each task.

Figure 4 shows the forest plots describing the estimated differences in the JAFROC figures of merit between the routine dose and each lower dose configuration. For each diagnostic task, all but 1 or 2 of the lowest dose configurations were non-inferior in observer performance given our pre-set criteria for non-inferiority, reflecting both a large potential for substantial dose reduction without compromising observer performance and the sample size.

Figure 5 shows the average diagnostic image quality for each diagnostic task for every dose configuration. For contrast-enhanced abdominal CT, IR significantly improved the image quality of lower dose CT images with AEC settings below 120 QRM. In contrast, in chest and head CT, IR only improved the image quality in the higher dose configurations, but did not improve the image quality of the very noisy, very low-dose exams.

For liver metastasis and pulmonary nodule detection, lesion size significantly correlated with reader confidence detection of target lesions for all doses and reconstruction methods (range of Spearman rank correlations: 0.46 – 0.81, $p<0.05$). For hepatic metastasis detection, neither absolute CT number difference or signal-to-noise ratio were predictive observer confidence, and there were no trends in detection for solid vs. semisolid pulmonary nodules in chest CT due to few semisolid nodules. No objective imaging criteria were correlated with reader confidence in interpreting head CT exams.

Discussion

We report an important first step in a systematic effort to define the lowest dose levels common diagnostic CT imaging tasks in three anatomic regions (liver, chest and brain) for which diagnostic performance is maintained. This effort was undertaken using a 128-slice

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

CT scanner from a single vendor in order to eliminate differences in dose efficiency and exposure between CT systems. The CT exam types examined in our study (contrast-enhanced abdominal CT, unenhanced chest CT and head CT) represent over one-third of CT examination types performed in our large clinical CT practice, and over 50% of the exams reported to have widely varying dose levels between and within institutions (4). The diagnostic tasks were also selected based on imaging features of the target lesions and the background organ of interest, with detection of hepatic metastases considered one of the most challenging low contrast detection tasks in diagnostic CT, unenhanced chest CT being considered a high contrast detection task, and unenhanced head CT being both--coupled with the challenge of beam-hardening artifacts created by the calvarium. Additionally, because IR has contrast-dependent spatial resolution (i.e., spatial resolution with IR is better for objects with higher contrast differences compared to their background) and because it is widely thought of as a method that facilitates radiation dose reduction (13, 24–27), we tested observer performance with and without IR. In order to estimate the lowest dose levels at which acceptable performance might be maintained, we tested a wide range of dose levels to include inferiorly performing dose levels, with dose reductions ranging from 70% to over 90% compared to current, clinically-utilized dose levels.

For the diagnostic tasks examined, there appears to be substantial opportunity for radiation dose reduction, with dose reductions of 70% or more demonstrating non-inferiority compared to current routine clinical dose levels. Specifically, our results indicate that substantial, but varying degrees of, dose reduction should be evaluated in the next phase of this study. The degree of dose reduction suitable for further evaluation was greater with the non-enhanced chest and head CT exams compared to the contrast enhanced abdominal exams.

Additionally, while early studies demonstrated that IR substantially improves image quality (10, 24), our results suggest that the positive impact of IR on observer performance only occurs near the threshold dose levels where observer performance appears to drop off. At doses above this threshold, FBP performs equally well, while at doses substantially lower than this threshold, performance suffers no matter which reconstruction technique is used. This finding is important because it implies that substantial dose reduction can be achieved on older CT systems not equipped with IR options, and will have broad application to heterogeneous CT fleets used by most departments of radiology.

Considering contrast-enhanced abdominal CT, dose levels corresponding to dose reductions of greater than 50% did not meet our preset criteria for success using reader agreement rules, and the lowest dose at 70% dose reduction did not demonstrate non-inferiority. We also found little benefit to IR in preserving performance at the lower dose levels despite the fact that it significantly improved subjective image quality. These results may be surprising in that our current clinical threshold was already based on observer performance studies validating one noise-reduction approach (5), but underscore the limitations of current methods of IR and CT noise reduction for low contrast tasks such as detection of liver lesions (12–14).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The National Lung Cancer Screening Trial (NLST) used a dose level corresponding to an effective dose of approximately 1.5 mSv (17, 28). This dose level is about two-thirds of the dose level used for the clinical "low-dose" chest CT exams in this study. While the NLST confirmed that low-dose chest CT resulted in a 20% reduction in mortality due to lung cancer (17), it is unclear if nodules seen at higher dose levels could be missed at the dose level employed, or if more dramatic dose reductions are possible without compromising observer performance. Both reader agreement rules for determination of "successfully interpreted" cases and the JAFROC FOM suggest that dose levels well below 1 mSv can perform non-inferiorly, with near identical performance using a simulated quality reference mAs of 5 mAs (corresponding to a $CTDI_{vol}$ of 0.3 mGy, and a dose reduction of over 80% compared to the NLST dose level). While dose reduction varies in chest CT depending on the diagnostic task under study (26, 29, 30), prior work has largely supported the use of lower doses using tube currents in the range of 20 – 60 mAs for the detection of pulmonary nodules (29–31). Importantly, however, this preliminary study had few semisolid nodules, which are more difficult to detect (30), and we did not test for other pathologic conditions other than pulmonary nodules.

Unenhanced head CT examinations were not acquired using automatic exposure control, which has been shown to reduce dose (32), but which also requires very precise centering of the head, which when not achieved, can results in degraded image quality. Rather, we employed a fixed mA technique, which was consistent at the start of the study with the 2012 recommendations by the American Association of Medical Physicists (AAPM) (33), which stated that AEC use was optional, as many sites did not elect to use AEC for head examinations. Most efforts to lower radiation dose in head CT examinations utilize subjective or objective measures of image quality rather than observer performance estimates (27, 32, 34, 35). Prior authors have found that iterative reconstruction is beneficial in improving image quality in unenhanced head CT (35). We found that despite the fact that many reference intracranial findings had subtle CT number differences with the surrounding parenchyma, that only the lowest dose levels performed non-inferiorly by JAFROC analysis, and that reader agreement rules did not eliminate any lower dose configurations. Our findings could have resulted from a combination of very subtle cases missed by readers at routine dose and very obvious cases detected by all readers at nearly all dose levels. We did not ask neuroradiologists to grade the age of infarcts due to the uncertainty in the characterizing the acuity of patient symptoms and imaging findings. Our findings could have resulted from a combination of very subtle cases missed by readers at routine dose and very obvious cases detected by all readers at nearly all dose levels. We did not ask neuroradiologists to grade the age of infarcts due to the uncertainty in the characterizing the acuity of patient symptoms and imaging findings. In subsequent and ongoing phases of this study, we will examine a larger number of cases, and additionally institute a visual conspicuity score for CT findings, to exclude CT findings that are likely readily visible at any dose level, in order to ensure that subtle CT findings can be detected at lower dose configurations. Unlike liver and lung lesions, lesion size did not contribute to whether lesions were identified with higher confidence, but this may be due to the small number of positive cases. We hypothesize that size may not be as an important factor at head CT exam, as an array of different findings seen on multiple slices often contribute to diagnosis (e.g.,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

findings at the gray-white junction might be present along with a dense MCA sign, or intraventricular blood would be seen in multiple ventricles). We have previously shown that dramatic reductions in head CT dose can be employed for clinical evaluation of ventricular size (36), and have employed this lower dose protocol in patients with suspected shunt malfunction across our practice, but have not extended this method to other clinical scenarios due to the fear that observer performance might be compromised.

We have learned several lessons in conducting this first step in a systematic evaluation of lower dose configurations for these CT tasks. First, substantial opportunity appears to exist for dramatic reductions in CT dose using existing CT technology across the diagnostic tasks studied, and this study has taken the first step in defining the range of dose levels that result in acceptable performance. Second, radiologist confidence in diagnostic image quality generally declines before objective measures of observer performance. For both chest CT and head CT, even subspecialized radiologists familiar with the look of lower dose CT images reported substantial reductions in image quality and confidence at dose levels where performance was maintained. Radiologists might be concerned that even though an abnormality was detected, confidence in evaluating the remainder of the CT examination as normal is substantially reduced on the noisier images. The lack of objective evidence in observer performance for lower dose CT, and the difficulty in obtaining this data, has been a major impediment to the systematic lowering of radiation dose: without objective data, radiologists have to rely on their subjective impressions of image quality, which are idiosyncratic and vary between radiologists (37)—and these subjective impressions often lead to widely disparate doses and acquisition methods across and even within institutions. Thirdly, we found that target lesions that are too obvious or too subtle are not helpful in identifying inferiorly performing dose levels. Fourthly, lesion size is an important variable affecting reader detection at lower radiation doses for abdominal CT and chest CT. We subsequently will select abdominal and chest CT exams with smaller lesions in future work, and will develop a subjective conspicuity score for the selection of positive head CT exams with more subtle lesions. Finally, the impact of this work on CT findings that are not under study (potentially unexpected but beneficial) is uncertain.

We acknowledge several limitations to our work, principally the small number of patients and readers. However, our approach permits examination of a larger number of lower dose configurations in fewer patients in the early phase, and will provide important preliminary data to enable appropriate selection of relevant dose levels in a larger and subsequent multi-reader, multi-case study employing fewer dose levels with many more cases. Subsequent steps in our process will be to examine nearly double the number of cases with longer intervals between interpretation sessions using the same readers in order to winnow down potential lower dose configurations to an acceptable number and to then complete a larger multi-reader, multi-case study using 10 or more readers for each task. Our study design required both archival of CT projection data and proof of target lesions. These requirements meant that exams were archived daily, but that consecutive cases were not included. This determination was made as we wanted to include a selection of visually challenging lesions. We simulated lower dose levels using a validated noise insertion method that takes into account the bowtie filter, automatic exposure control, and electronic noise of the CT system (18). This noise insertion method is highly accurate for noise insertion and reproduction of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

some artifacts, but is not perfect, particularly at very low doses in all cases; nevertheless, we have used it extensively to lower radiation dose in head and abdominal CT over a wide range of simulated doses, with our clinical practice rapidly accepting the diagnostic image quality of these lower dose CT images (5, 38). Additionally, we employed a memory extinction period of 7 days to reduce potential recall bias, but must acknowledge that this may have occurred due to the number of configurations examined. Lastly, for identification of indeterminate pulmonary nodules, we relied on identification of the same nodule by two reference readers using clinical dose settings. While this methodology can result in substantial variability in estimates of reader performance (39, 40), we are estimating the effect of dose on performance, so the relative performance of the readers across the dose levels (rather than the absolute estimate) is the variable of interest.

In summary, we have explored the potential to substantially lower radiation dose based on observer performance for some of the most common tasks in diagnostic CT imaging. This study represents an important first step in systematically defining the lowest radiation doses that preserve observer performance for common but narrowly defined specific diagnostic tasks in CT imaging. Because such a large variability exists between the doses needed to accomplish different diagnostic tasks, radiation dose should be reduced by adapting scan parameters to the diagnostic task performed by the radiologist rather than simply the body part being imaged. Further work with more readers and cases will be required to better define lower radiation dose levels that do not degrade observer performance. This is the first step in providing objective data to minimize the radiation exposure risks related to CT examinations, while maintaining diagnostic benefit.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

The project described was supported by Grant numbers U01 EB17185 and R01 EB17095 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. They appreciate the support of the Dr. Kent Thielen, chair of the Department of Radiology at Mayo Clinic Rochester, for his commitment to ensure that a large number of staff could participate in this endeavor over a long period of time. We thank Kris Nunez for her assistance in preparation of the manuscript and Maria Shiung, our study coordinator, for her assistance in managing data collection.

List of Abbreviations:

JAFROC	jackknife alternative free-response receiver operating characteristic
FOM	figure of merit
AEC	Automatic exposure control, which modulates x-ray tube current in the x-y and z planes to deliver a desired image quality or image noise
QRM	Quality reference mAs, the automatic exposure control setting of the CT system used in this study
SSDE	size-specific dose estimate (mGy)

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$CTDI_{vol}$	volume computed tomography dose index (mGy), a standard measure of the radiation output of a CT system during a scan as measured in an acrylic phantom
mAs	milli-ampere seconds
IR	iterative reconstruction
FBP	Filtered back projection

References

1. Brenner DJ, Hall EJ. Computed tomography--an increasing source of radiation exposure. The New England journal of medicine. 2007; 357(22):2277–84. [PubMed: 18046031]
2. McCollough CH, Bushberg JT, Fletcher JG, Eckel LJ. Answers to Common Questions About the Use and Safety of CT Scans. Mayo Clin Proc. 2015; 90(10):1380–92. [PubMed: 26434964]
3. McCollough CH, Guimaraes L, Fletcher JG. In defense of body CT. AJR Am J Roentgenol. 2009; 193(1):28–39. [PubMed: 19542392]
4. Smith-Bindman R, Lipson J, Marcus R, et al. Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer. Arch Intern Med. 2009; 169(22):2078–86. [PubMed: 20008690]
5. Fletcher JG, Yu L, Li Z, et al. Observer Performance in the Detection and Classification of Malignant Hepatic Nodules and Masses with CT Image-Space Denoising and Iterative Reconstruction. Radiology. 2015; 276(2):465–78. [PubMed: 26020436]
6. McCollough CH, Chen GH, Kalender W, et al. Achieving routine submillisievert CT scanning: report from the summit on management of radiation dose in CT. Radiology. 2012; 264(2):567–80. [PubMed: 22692035]
7. Boone JM, Hendee WR, McNitt-Gray MF, Seltzer SE. Radiation exposure from CT scans: how to close our knowledge gaps, monitor and safeguard exposure-- proceedings and recommendations of the Radiation Dose Summit, sponsored by NIBIB, February 24–25, 2011. Radiology. 2012; 265(2): 544–54. [PubMed: 22966066]
8. Mulkens TH, Bellinck P, Baeyaert M, et al. Use of an automatic exposure control mechanism for dose optimization in multi-detector row CT examinations: clinical evaluation. Radiology. 2005; 237(1):213–23. [PubMed: 16126917]
9. Yu L, Fletcher JG, Grant KL, et al. Automatic Selection of Tube Potential for Radiation Dose Reduction in Vascular and Contrast-Enhanced Abdominopelvic CT. AJR Am J Roentgenol. 2013; 201(2):W297–306. [PubMed: 23883244]
10. Prakash P, Kalra MK, Kambadakone AK, et al. Reducing abdominal CT radiation dose with adaptive statistical iterative reconstruction technique. Invest Radiol. 2010; 45(4):202–10. [PubMed: 20177389]
11. Gandhi NS, Baker ME, Goenka AH, et al. Diagnostic Accuracy of CT Enterography for Active Inflammatory Terminal Ileal Crohn Disease: Comparison of Full- Dose and Half-Dose Images Reconstructed with FBP and Half-Dose Images with SAFIRE. Radiology. 2016:151281.
12. Goenka AH, Herts BR, Dong F, et al. Image Noise, CNR, and Detectability of Low-Contrast, Low-Attenuation Liver Lesions in a Phantom: Effects of Radiation Exposure, Phantom Size, Integrated Circuit Detector, and Iterative Reconstruction. Radiology. 2016:151621.
13. Pickhardt PJ, Lubner MG, Kim DH, et al. Abdominal CT with Model-Based Iterative Reconstruction (MBIR): Initial Results of a Prospective Trial Comparing Ultra- Low Dose with Standard-Dose Imaging. AJR Am J Roentgenol. 2012; 199(6):1266–74. [PubMed: 23169718]
14. 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(2):511–8. [PubMed: 23788715]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

15. Goenka AH, Herts BR, Obuchowski NA, et al. Effect of reduced radiation exposure and iterative reconstruction on detection of low-contrast low-attenuation lesions in an anthropomorphic liver phantom: an 18-reader study. Radiology. 2014; 272(1):154–63. [PubMed: 24620913]
16. Johnson CD, Chen MH, Toledano AY, et al. Accuracy of CT colonography for detection of large adenomas and cancers. The New England journal of medicine. 2008; 359(12):1207–17. [PubMed: 18799557]
17. Aberle DR, Adams AM, Berg CD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. The New England journal of medicine. 2011; 365(5):395–409. [PubMed: 21714641]
18. Yu L, Shiung M, Jondal D, McCollough CH. Development and validation of a practical lower-dose-simulation tool for optimizing computed tomography scan protocols. J Comput Assist Tomogr. 2012; 36(4):477–87. [PubMed: 22805680]
19. Fletcher JG, Hara AK, Fidler JL, et al. Observer performance for adaptive, image-based denoising and filtered back projection compared to scanner-based iterative reconstruction for lower dose CT enterography. Abdom Imaging. 2015; 40(5):1050–9. [PubMed: 25725794]
20. Froemming AT, Kawashima A, Takahashi N, et al. Individualized kV Selection and Tube Current Reduction in Excretory Phase Computed Tomography Urography: Potential for Radiation Dose Reduction and the Contribution of Iterative Reconstruction to Image Quality. J Comput Assist Tomogr. 2013; 37(4):551–9. [PubMed: 23863531]
21. Simon R Optimal two-stage designs for phase II clinical trials. Control Clin Trials. 1989; 10(1):1–10. [PubMed: 2702835]
22. American Association of Physicists in Medicine. Size-Specific Dose Estimates (SSDE) in Pediatric and Adult Body CT Examinations (AAPM Report No. 204). Book Size-Specific Dose Estimates (SSDE) in Pediatric and Adult Body CT Examinations (AAPM Report No. 204). City: American Association of Physicists in Medicine; 2011.
23. American Association of Physicists in Medicine. Use of water-equivalent diameter for calculating patient size and size-specific dose estimates (SSDE) in CT (AAPM Report No. 220). Book Use of water-equivalent diameter for calculating patient size and size-specific dose estimates (SSDE) in CT (AAPM Report No. 220). City: American Association of Physicists in Medicine; 2014.
24. Sagara Y, Hara AK, Pavlicek W, Silva AC, Paden RG, Wu Q. Abdominal CT: comparison of low-dose CT with adaptive statistical iterative reconstruction and routine-dose CT with filtered back projection in 53 patients. AJR Am J Roentgenol. 2010; 195(3):713–9. [PubMed: 20729451]
25. Solomon J, Mileto A, Ramirez-Giraldo JC, Samei E. Diagnostic Performance of an Advanced Modeled Iterative Reconstruction Algorithm for Low-Contrast Detectability with a Third-Generation Dual-Source Multidetector CT Scanner: Potential for Radiation Dose Reduction in a Multireader Study. Radiology. 2015; 275(3):735–45. [PubMed: 25751228]
26. Pontana F, Billard AS, Duhamel A, et al. Effect of Iterative Reconstruction on the Detection of Systemic Sclerosis-related Interstitial Lung Disease: Clinical Experience in 55 Patients. Radiology. 2016; 279(1):297–305. [PubMed: 26583761]
27. Komlosi P, Zhang Y, Leiva-Salinas C, et al. Adaptive statistical iterative reconstruction reduces patient radiation dose in neuroradiology CT studies. Neuroradiology. 2014; 56(3):187–93. [PubMed: 24384672]
28. Aberle DR, Berg CD, Black WC, et al. The National Lung Screening Trial: overview and study design. Radiology. 2011; 258(1):243–53. [PubMed: 21045183]
29. Kubo T, Ohno Y, Kauczor HU, Hatabu H. Radiation dose reduction in chest CT-- review of available options. Eur J Radiol. 2014; 83(10):1953–61. [PubMed: 25066756]
30. Christe A, Charimo-Torrente J, Roychoudhury K, Vock P, Roos JE. Accuracy of low-dose computed tomography (CT) for detecting and characterizing the most common CT-patterns of pulmonary disease. Eur J Radiol. 2013; 82(3):e142–50. [PubMed: 23122673]
31. Funama Y, Awai K, Liu D, et al. Detection of nodules showing ground-glass opacity in the lungs at low-dose multidetector computed tomography: phantom and clinical study. J Comput Assist Tomogr. 2009; 33(1):49–53. [PubMed: 19188784]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

32. Smith AB, Dillon WP, Lau BC, et al. Radiation dose reduction strategy for CT protocols: successful implementation in neuroradiology section. Radiology. 2008; 247(2):499–506. [PubMed: 18372456]

33. American Association of Physicists in Medicine. Adult Routine Head CT Protocols (Version 2.0); . Book Adult Routine Head CT Protocols (Version 2.0); . City2016.

34. Wu TH, Hung SC, Sun JY, et al. How far can the radiation dose be lowered in head CT with iterative reconstruction? Analysis of imaging quality and diagnostic accuracy. European radiology. 2013; 23(9):2612–21. [PubMed: 23645331]

35. Korn A, Fenchel M, Bender B, et al. Iterative reconstruction in head CT: image quality of routine and low-dose protocols in comparison with standard filtered back-projection. AJNR Am J Neuroradiol. 2012; 33(2):218–24. [PubMed: 22033719]

36. Gabriel S, Eckel LJ, DeLone DR, et al. Pilot study of radiation dose reduction for pediatric head CT in evaluation of ventricular size. AJNR Am J Neuroradiol. 2014; 35(12):2237–42. [PubMed: 25082822]

37. Fletcher JG, Krueger WR, Hough DM, et al. Pilot study of detection, radiologist confidence and image quality with sinogram-affirmed iterative reconstruction at half-routine dose level. J Comput Assist Tomogr. 2013; 37(2):203–11. [PubMed: 23493209]

38. Gabriel S, Eckel LJ, DeLone DR, et al. Pilot study of radiation dose reduction for pediatric head CT in evaluation of ventricular size. AJNR Am J Neuroradiol. 2014; 35(12):2237–42. [PubMed: 25082822]

39. Rubin GD. Lung nodule and cancer detection in computed tomography screening. J Thorac Imaging. 2015; 30(2):130–8. [PubMed: 25658477]

40. Armato SG 3rd, Roberts RY, Kocherginsky M, et al. Assessment of radiologist performance in the detection of lung nodules: dependence on the definition of “truth”. Academic radiology. 2009; 16(1):28–38. [PubMed: 19064209]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

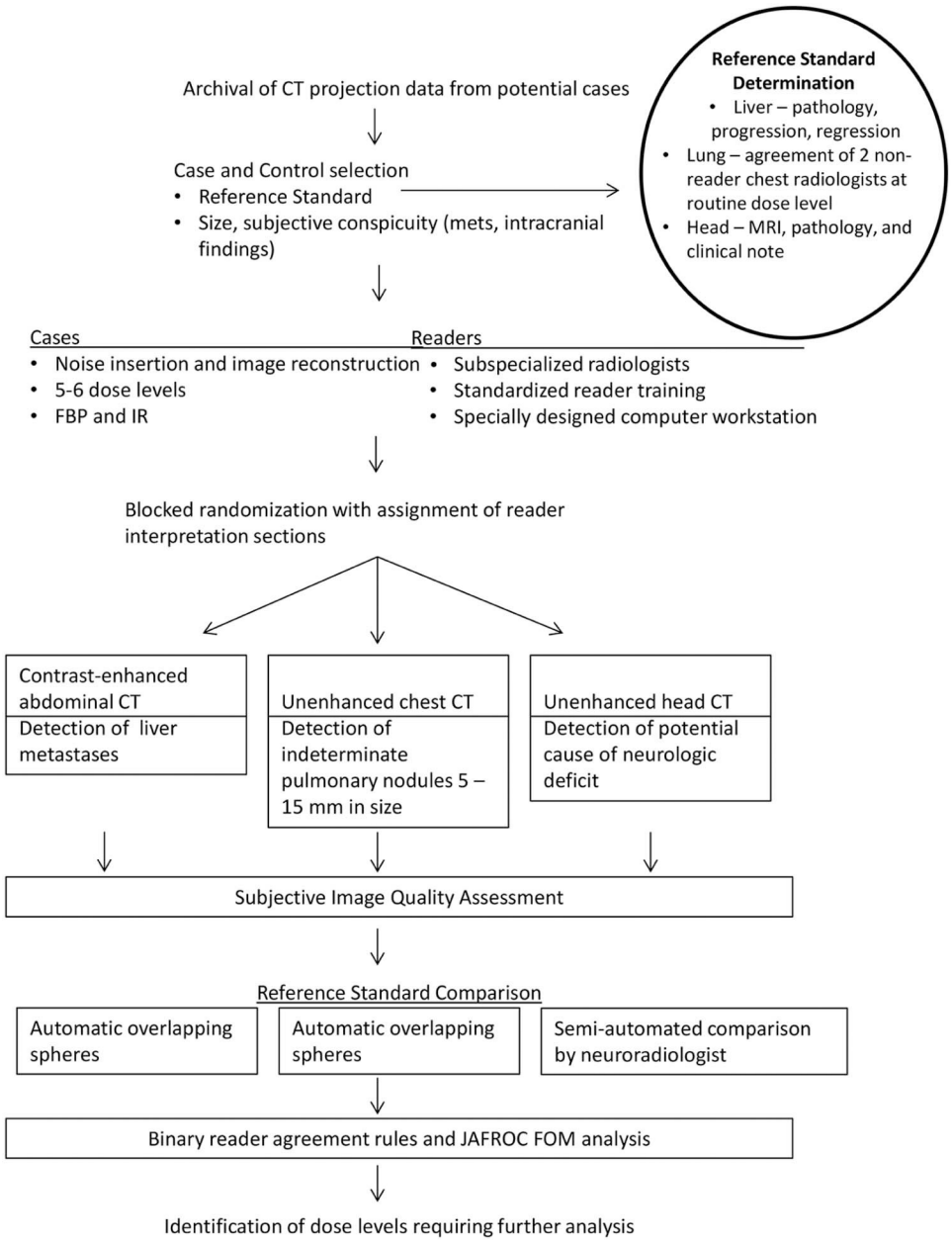

Figure 1.

Study schema. For each of three common diagnostic tasks in CT imaging (contrast-enhanced abdominal CT to detect liver metastases, unenhanced chest CT to detect indeterminate pulmonary nodules, and unenhanced head CT to detect cause of neurologic deficit), three subspecialized radiologists examined CT images corresponding to 5 or 6 dose levels reconstructed with both FBP and IR, using a specially designed computer workstation to mark target lesions of interest, with reader markings and confidence compared to an established reference standard. Reader agreement rules and a JAFROC figure of merit were used to compare observer performance at varying dose levels.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

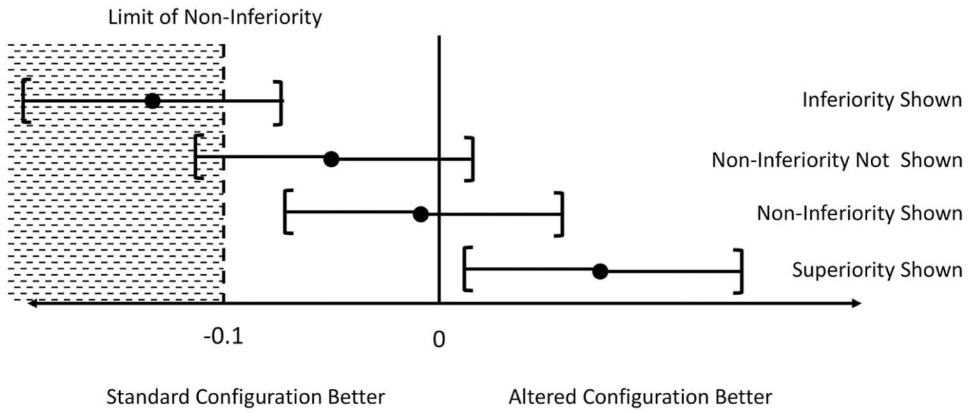

Figure 2.

Graphical demonstration of non-inferiority forest plot interpretation, with each dot representing the estimated difference in performance in the JAFROC figure of merit between the routine dose and lower dose configuration, with bars representing the 95% confidence intervals for this estimate. For lower doses that demonstrate non-inferiority compared to routine dose, the lower limit of the 95% confidence interval for the difference in performance (i.e., the left bar) will be above (or to the right of) the limit of non-inferiority.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

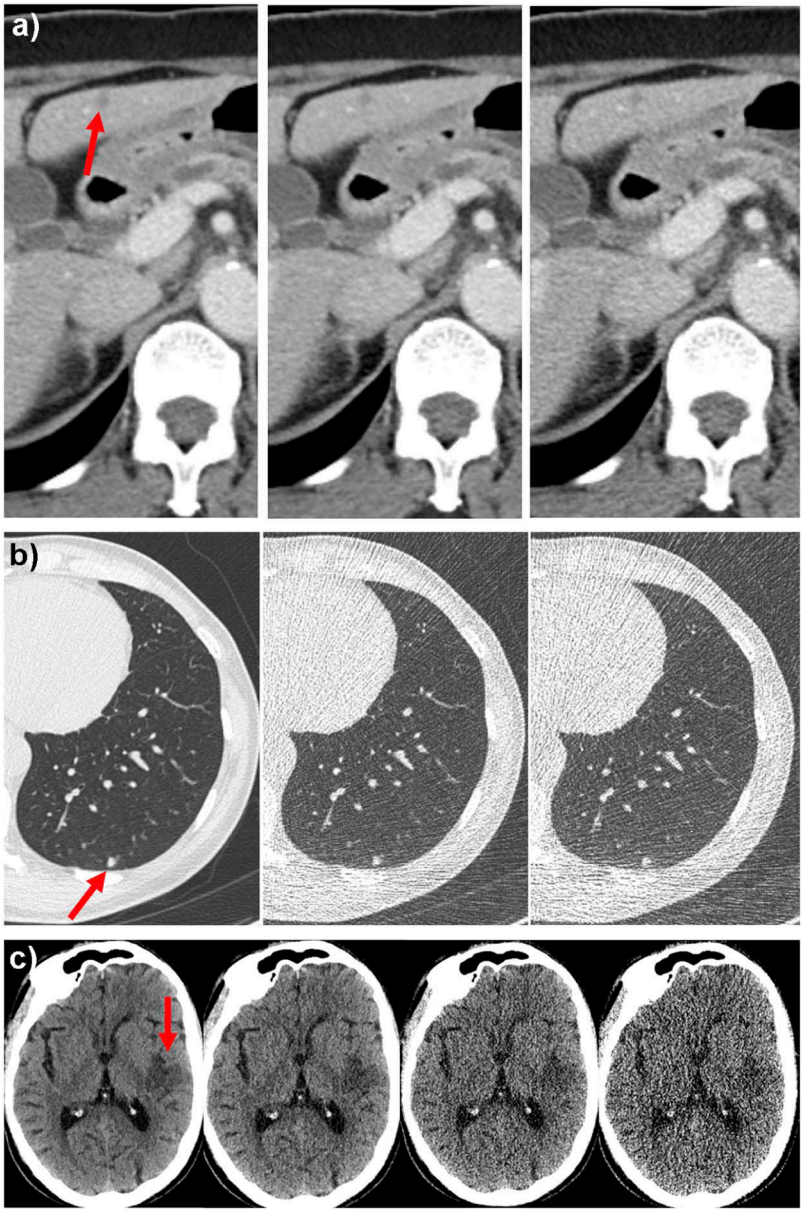

Figure 3.

Representative study images demonstrating observer performance at varying dose levels. (A) a 6 mm metastasis in segment 3 of the liver was identified by 2, 3 and 1 abdominal radiologists at routine dose with filtered back projection (left, 200 quality reference mAs), 50% routine dose with IR (middle, 100 quality reference mAs), and 50% routine dose with FBP (right, 100 quality reference mAs), respectively, at contrast-enhanced abdominal CT. (B) a 5 mm solid nodule was identified by 2, 3 and 3 thoracic radiologists at routine dose with FBP (left, 70 quality reference mAs), 7% routine dose with FBP (middle, 5 quality reference mAs), and 4% routine dose with IR (2.5 quality reference mAs), respectively, at unenhanced chest CT. (C) an acute infarct of the posterior left insula and adjacent parietal operculum and temporal lobe was identified by 3, 3, 3 and 1 neuroradiologists using routine dose with IR (250 effective mAs), 40% routine dose with IR (100 effective mAs), 20%

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

routine dose with IR (50 effective mAs), and 10% routine dose with IR (25 effective mAs), respectively, at unenhanced head CT.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

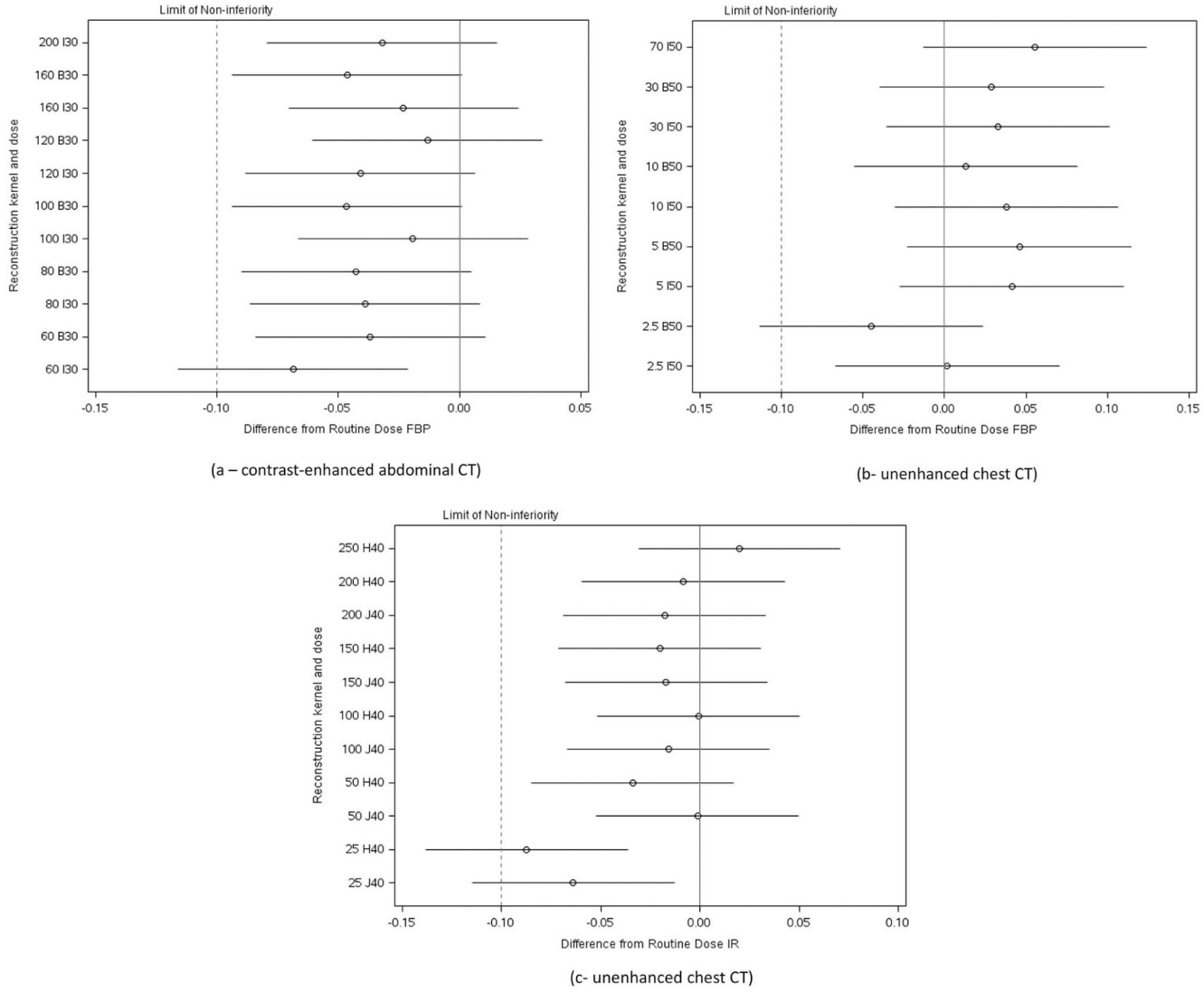

Figure 4.

Forest plots indicating estimated differences in observer performance of between routine dose and reconstruction method and lower dose configurations using JAFROC figures of merit. The dotted line represents the limit of non-inferiority selected for this study (i.e., −0.10). Point estimates are given by circles, with lines representing the 95% confidence intervals in the estimated differences. When the lower limit of the 95% confidence interval lies to the right of the limit of non-inferiority, the lower dose is shown to be non-inferior. Radiation doses are given corresponding to the automatic exposure control setting of the CT scanner (in quality reference mAs) for each lower dose configuration, as CT systems modulate tube current appropriately for patients of different sizes. Reconstruction kernels are given as named by the CT manufacturer using a letter and number, with “B” and “H” kernels used in FBP, and “I” and “J” kernels corresponding to IR kernels.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

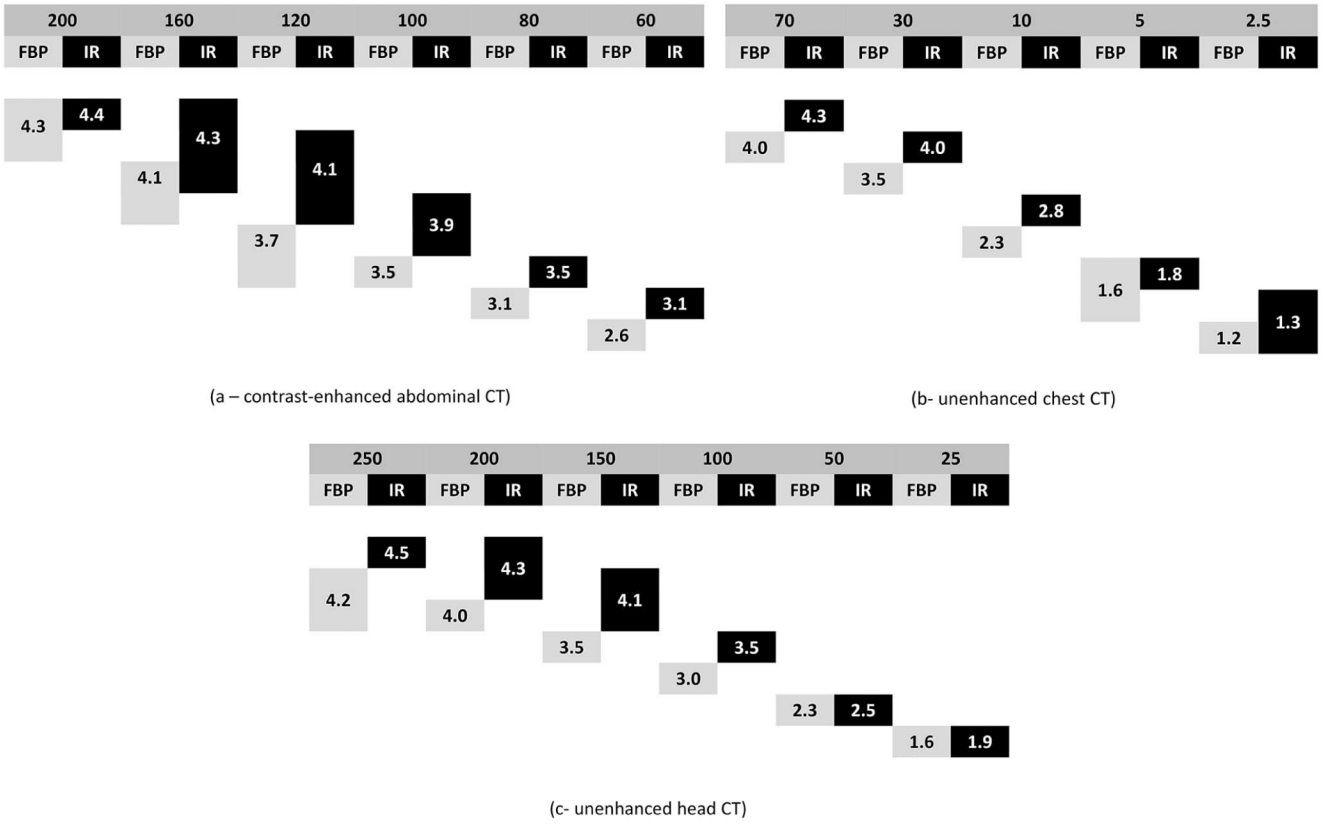

Figure 5.

Subjective measures of diagnostic image quality pooled across three subspecialized radiologists for the three common diagnostic tasks in CT imaging for the routine and lower dose configurations. The top line shows the radiation doses corresponding to the automatic exposure control setting or tube current, as appropriate, and the second line lists the reconstruction type (FBP or IR). Boxes below indicate the spread of diagnostic image quality scores. For each dose level, significant differences in diagnostic image quality exist when boxes do not overlap (i.e., share an adjacent side).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1.

Reference standard criteria for target lesions for each diagnostic task in this study

Cohort of cases	Reference Standard		
	Hepatic Metastasis	**Detection of Indeterminate Pulmonary Nodules**	**Cause of potential acute neurologic deficit**
Positive cases	*For each case* • proven malignancy • at least one hepatic lesion that meets criteria below *For each lesion* • hepatic metastasis by histopathology (surgical extirpation or biopsy) • progression (increase in size) over serial cross-sectional exams • regression (decrease in size) over serial cross-sectional exams while on medical treatment	Two chest radiologists, not participating as readers in this study, independently evaluate routine dose CT images, unblinded to prior and subsequent exams • Both chest radiologists circumscribe every indeterminate nodule 5–15 mm in size • Nodules circled by both radiologists are reference nodules	Head CT findings visible in retrospect by diagnostic task leader, and (one of the following)… • Neurologist or attending physician or neurosurgical note indicating clinical findings that correspond to CT findings, *or* • Subsequent confirmation on MRI, *or* • Subsequent progression on serial CT exams, *or* • Confirmation of CT findings at neurosurgery and/or histology
Negative cases without other findings	• Agreement between individual assessment by 2 GI radiologists that no liver lesions are present • A subsequent CT or MRI performed ≥ 6 months later also showing absence of liver lesions	Two non-reader chest radiologists verify absence of indeterminate pulmonary nodules comparing to prior and subsequent CT exams	2 non-reader neuro-radiologists to agree on absence of CT findings on routine dose images that would indicate acute neurologic deficit, as well as corresponding neurological assessment prior to discharge
Negative cases with benign findings	Characteristic imaging features [19] plus stability on separate CT or MR exam performed ≥ 5 months from index CT	Not applicable (granulomas are not marked or tracked)	Leukoaraiosis was confirmed and noted when present on subsequent MRI exam

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 2.

CT acquisition and reconstruction parameters

Acquisition Parameter	CT Diagnostic Tasks		
	Detection of liver metastases (Contrast-enhanced abdominal CT)	**Detection of indeterminate pulmonary nodules (Unenhanced chest CT)**	**Detection of potential causes of acute neurologic deficit (Unenhanced head CT)**
Tube energy (kVp)	Vendor-supplied automatic kV selection	120 kV	120 kV
Tube current setting of routine clinical exam	200 QRM *	70 QRM	250 eff. mAs **
Automatic exposure control	On	On	Off
Axial slice thickness/reconstruction interval (mm)	3/2	1.5/1	5/5
Other reconstruction volumes (slice thickness/ reconstruction interval [mm])	Coronal 3/2	Thick maximum intensity projection 10/2.5	N/A
Reconstruction Kernels (FBP/IR) ***	B30/I30(2)	B50/I50(2)	H40/J40(2)
Dose levels (as represented by tube current setting) examined by all readers (Top value is routine setting	200 QRM 160 QRM 120 QRM 100 QRM 80 QRM 60 QRM	70 QRM 30 QRM 10 QRM 5 QRM 2.5 QRM -	250 eff. mAs 200 eff. mAs 150 eff. mAs 100 eff. mAs 50 eff. mAs 25 eff. mAs
Dose levels examined by nominal CTDIvol (mGy)	13.5 10.8 8.1 6.8 5.4 4.1	4.7 2.0 0.7 0.3 0.2	38.3 30.6 23.0 15.3 7.7 3.8
Total number of dose/kernel configurations	12	10	12

* QRM = quality reference effective mAs, the setting for the CT system automatic exposure control, that delivers equivalent image quality at a given setting across patients of different sizes

** eff. mAs = effective milli-ampere second

*** FBP = filtered backprojection; IR = iterative reconstruction

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3.

Potential diagnoses provided the subspecialized radiologists for circumscribed lesions using a drop-down menu. Readers specified location and diagnosis by circumscribing each lesion with region-of-interest tools.

Hepatic Metastases	**Chest CT**	**Potential causes of acute neurologic deficit***
• Metastasis • Benign Cyst • Hemangioma • Adenoma • Vascular - Shunt/fistula • Vascular - Perfusion defect • Benign neoplasm (e.g., FNH) • Post op/Post RFA defect • Focal fat/focal fatty sparing	• Solid nodule • Sub-solid nodule • Calcified (benign) nodule	• Infarction - – acute, subacute, chronic, or indeterminate age • Traumatic hemorrhage • Non-traumatic hemorrhage • Mass

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 4.

Number of successfully interpreted cases for each diagnostic task - at every dose-reconstruction configuration - using binary reader agreement rules. For positive cases, a majority of readers had to identify all lesions seen by a majority of readers using the routine dose level; for negative cases, the majority of readers made no non-lesion localizations (i.e., no false positive markings). Numbers in red indicate failure to meet preset thresholds for successful dose-reconstruction configurations. The number of cases with essential lesions reflects cases with target lesions identified by 2 or more readers at routine dose level.

Contrast-enhanced abdominal CT for Hepatic Metastasis							
Dose Level by AEC setting (QRM)	**Reconstruction Kernel (FBP/IR)**	**# of readers with successful interpretation**				**Successful Interpretations All Cases**	**Successful Interpretations Cases with (n=21): Cases without (n=23) essential Lesions**
		0/3	**1/3**	**2/3**	**3/3**		
200	B30	0	3	9	32	41	20:21
	I30	0	3	9	32	41	19:22
160	B30	0	4	12	28	40	18:22
	I30	0	4	10	30	40	18:22
120	B30	2	4	5	33	38	16:22
	I30	1	4	8	31	39	18:21
100	B30	**3**	**6**	**7**	**28**	**35**	17:18
	I30	2	4	9	29	38	17:21
80	B30	**3**	**6**	**10**	**25**	**35**	14:21
	I30	**4**	**4**	**8**	**28**	**36**	14:22
60	B30	**4**	**4**	**10**	**26**	**36**	14:22
	I30	**7**	**4**	**8**	**25**	**33**	14:19
Unenhanced chest CT for Pulmonary Nodule detection							
Dose Level by AEC setting (QRM))	**Reconstruction Kernel (FBP/IR)**	**# of readers with successful interpretation**				**Successful Interpretations All Cases**	**Successful Interpretations Cases with (n=19): Cases without (n=25) essential Lesions**
		0/3	**1/3**	**2/3**	**3/3**		
70	B50	1	3	12	28	40	19:21
	I50	1	1	10	32	42	19:23
30	B50	0	3	9	32	41	17:24
	I50	1	3	7	33	40	17:23
10	B50	0	4	8	32	40	18:22
	I50	1	2	5	36	41	17:24
5	B50	1	2	7	34	41	18:23
	I50	1	1	6	36	42	19:23
2.5	B50	**2**	**6**	**4**	**32**	**36**	13:23
	I50	2	3	8	31	39	16:23
Unenhanced spiral head CT to detect Intracranial Findings Causing Neurologic Deficit							

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Contrast-enhanced abdominal CT for Hepatic Metastasis							
Dose Level by AEC setting (QRM)	**Reconstruction Kernel (FBP/IR)**	**# of readers with successful interpretation**				**Successful Interpretations All Cases**	**Successful Interpretations Cases with (n=21): Cases without (n=23) essential Lesions**
		0/3	**1/3**	**2/3**	**3/3**		
Dose Level by effective mAs	**Reconstruction Kernel (FBP/IR)**	**# of readers with successful interpretation**				**Successful Interpretations All Cases**	**Successful Interpretations Cases with (n=18): Cases without (n=25) essential Lesions**
		0/3	**1/3**	**2/3**	**3/3**		
250	H40(FBP)	0	2	2	39	41	16:25
	J40 (IR)	0	1	7	35	42	17:25
200	H40(FBP)	0	2	4	37	41	17:24
	J40 (IR)	0	2	7	34	41	17:24
150	H40(FBP)	0	1	6	36	42	17:25
	J40 (IR)	1	0	4	38	42	17:25
100	H40(FBP)	0	0	10	33	43	18:25
	J40 (IR)	0	2	6	35	41	16:25
50	H40(FBP)	1	1	9	32	41	17:24
	J40 (IR)	0	1	6	36	42	17:25
25	H40(FBP)	2	3	8	30	38	14:24
	J40 (IR)	2	0	11	30	41	16:25

AEC = automatic exposure control setting, which adjusts the modulation of the CT system tube current; QRM = quality reference milliampere-seconds (mAs), the AEC setting that will produce the image quality at a similar effective mAs for a hypothetically normal-sized patient; FBP = filtered back projection; IR = iterative reconstruction