LUND UNIVERSITY

Optimization of image quality and radiation dose in neuroradiological computed tomography

Löve, Askell

2013

Link to publication

Citation for published version (APA):
Löve, A. (2013). *Optimization of image quality and radiation dose in neuroradiological computed tomography*. [Doctoral Thesis (compilation), Diagnostic Radiology, (Lund)]. Diagnostic Radiology, (Lund), Lund University.

Total number of authors:
1

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Optimization of image quality and radiation dose in neuroradiological computed tomography

Áskell Löve

LUND UNIVERSITY

AKADEMISK AVHANDLING
som för avläggande av doktorsexamen i medicinsk vetenskap vid
Medicinska fakulteten, Lunds universitet, kommer att offentligen
försvaras i demonstrationsrum 10, plan 4, BFC, SUS Lund,
fredagen den 20 september 2013, kl 13.00

Organization
LUND UNIVERSITY
Department of Diagnostic Radiology,
Clinical Sciences Lund, Faculty of Medicine.

Document name
DOCTORAL DISSERTATION

Date of issue
17 June 2013

Author(s)
Áskell Löve

Sponsoring organization

Title and subtitle

Optimization of image quality and radiation dose in neuroradiological computed tomography

Abstract

Background: The goal of clinical computed tomography (CT) is to produce images of diagnostic quality using the lowest possible radiation exposure. Degradation of image quality, with increased image noise and reduced spatial resolution, is a major limitation for radiation dose reduction in CT. This can be counteracted with new post-processing image filters and iterative reconstruction (IR) algorithms that improve image quality and allow for reduced radiation doses. Implementation of new methods in clinical routine requires prior validation in phantoms and clinical feasibility studies including comprehensive evaluation of image quality.

Aims: The main objectives of this thesis were to assess new methods for improvement of image quality in CT, explore the associated potential for radiation dose reduction, and to outline a comprehensive approach for evaluation of image quality.

Methods: Extensive phantom testing was performed and a total of 100 human subjects were included in the clinical studies. Image quality and diagnostic acceptability were assessed in brain CT acquired with 30% reduced radiation dose in combination with post-processing filter (Paper I) and IR (Paper II). The potential of IR for image quality improvement, without concomitant radiation dose reduction, was assessed in craniocervical CT angiography (CTA) (Paper III). The performance of six IR algorithms was evaluated in a brain CT phantom model (Paper IV), using different combinations of radiation dose levels and iterative image optimization levels. Throughout the studies, various approaches for subjective and objective evaluation of image quality were used and assessed.

Results: Post-processing image filtering (Paper I) and IR (Paper II) compensated partly or entirely for the loss of image quality caused by 30% reduced radiation dose in brain CT. In both studies, considerable inter-observer variation was seen. In Paper II a discrepancy was seen between results of objective and subjective evaluation of image quality and also between grading and ranking, indicating observer bias. Statistical IR improved image quality in craniocervical CTA (Paper III) with fairly good inter-observer agreement. Despite having different strengths and weaknesses, the six iterative reconstruction algorithms evaluated in Paper IV all improved image quality. Best over-all improvement was seen for one of the model-based IR algorithms, especially at lower radiation doses.

Conclusion: All evaluated methods improved image quality and showed potential for radiation dose reduction while maintaining diagnostic quality. Careful study design and comprehensive evaluation of image quality including objective and subjective evaluation steps may reduce observer bias and improve reliability of study results.

Key words
Computed tomography; image quality; radiation dose; neuroradiology; post-processing image filters; iterative reconstruction; filtered back-projection brain CT; craniocervical CT angiography; phantom.

Classification system and/or index terms (if any)

Supplementary bibliographical information

Language
English

ISSN and key title
1652-8220

ISBN
978-91-87449-52-9

Recipient´s notes

Number of pages
98

Price

Security classification

Distribution by (name and address): Áskell Löve, Department of Neuroradiology, Skåne University Hospital, Lund, Sweden.

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature: Áskell Löve Lund, 17 June 2013

Optimization of image quality and radiation dose in neuroradiological computed tomography

Áskell Löve

Department of Diagnostic Radiology,
Clinical Sciences Lund, Faculty of Medicine,
Lund University, Skåne University Hospital
2013

CLIMATE
COMPENSATED
PAPER

Thesis for the degree of Doctor of Philosophy in Medical Science
Department of Diagnostic Radiology, Clinical Sciences Lund, Faculty of Medicine,
Lund University and Skåne University Hospital Lund

Copyright ©2013 Áskell Löve

ISBN 978-91-87449-52-9
Printed in Sweden by Media-Tryck, Lund 2013

In CT, too low or too high image quality is equally inappropriate

– Áskell Löve

Contents

Abstract 7

Abbreviations used 9

Original papers 11

Introduction 13

Background 15
- Computed tomography 15
  - Technique 15
  - Radiation dose 16
  - Image reconstruction 17
  - Parameters of image quality 17
  - Image quality improvement 18
- Evaluation of image quality 19
  - Objective evaluation 19
  - Subjective evaluation 21
  - Observer bias 22
  - Evaluation scales 23
  - Research subjects 25

Aims 27
- Paper I 27
- Paper II 27
- Paper III 27
- Paper IV 28

Subjects and methods 29
Study subjects 29
CT technique 30
Image reconstruction 32
Image quality assessment 33
Statistical evaluation 35

Results 37
Paper I 37
Paper II 39
Paper III 41
Paper IV 43

Discussion 47
Image quality and radiation dose 47
Comprehensive evaluation of image quality 49
Clinical impact 51

Conclusions 53
Paper I 53
Paper II 53
Paper III 54
Paper IV 54

Acknowledgements 55

References 57

Appendix 63

Abstract

Background: The goal of clinical computed tomography (CT) is to produce images of diagnostic quality using the lowest possible radiation exposure. Degradation of image quality, with increased image noise and reduced spatial resolution, is a major limitation for radiation dose reduction in CT. This can be counteracted with new post-processing image filters and iterative reconstruction (IR) algorithms that improve image quality and allow for reduced radiation doses. Implementation of new methods in clinical routine requires prior validation in phantoms and clinical feasibility studies including comprehensive evaluation of diagnostic image quality.

Aims: The main objectives of this thesis were to assess new methods for improvement of image quality in CT, explore the associated potential for radiation dose reduction, and to outline a comprehensive approach for evaluation of image quality.

Methods: Extensive phantom testing was performed and a total of 100 human subjects were included in the clinical studies. Image quality and diagnostic acceptability were assessed in brain CT acquired with 30% reduced radiation dose in combination with post-processing filter (Paper I) and IR (Paper II). The potential of IR for image quality improvement, without concomitant radiation dose reduction, was assessed in craniocervical CT angiography (CTA) (Paper III). The performance of six IR algorithms was evaluated in a brain CT phantom model (Paper IV), using different combinations of radiation dose levels and iterative image optimization levels. Throughout the studies, various approaches for subjective and objective evaluation of image quality were used and assessed.

Results: Post-processing image filtering (Paper I) and IR (Paper II) compensated partly or entirely for the loss of image quality caused by 30% reduced radiation dose in brain CT. In both studies, considerable inter-observer variation was seen. In Paper II a discrepancy was seen between results of objective and subjective evaluation of image quality and also between grading and ranking, indicating observer bias. Statistical IR improved image quality in craniocervical CTA (Paper III) with fairly good inter-observer agreement. Despite having different strengths and weaknesses, the six iterative reconstruction algorithms evaluated in Paper IV all improved image

quality. Best over-all improvement was seen for one of the model-based IR algorithms, especially at lower radiation doses.

Conclusion: All evaluated methods improved image quality and showed potential for radiation dose reduction while maintaining diagnostic quality. Careful study design and comprehensive evaluation of image quality including objective and subjective evaluation steps may reduce observer bias and improve reliability of study results.

Abbreviations used

AAPM	American association of physicists in medicine
AIDR	Adaptive iterative dose reduction (Toshiba)
ALARA	As low as reasonably achievable
ASIR	Adaptive statistical iterative reconstruction (GE)
CNR	Contrast-to-noise ratio
CT	Computed tomography
CTA	Computed tomography angiography
CTDI	Computed tomography dose index
DLP	Dose length product
FBP	Filtered back-projection
FLD	Filtered low-dose (Paper I)
HU	Hounsfield unit (attenuation)
ID2	iDOSE^4 noise reduction level 2 (Paper II)
ID4	iDOSE^4 noise reduction level 4 (Paper II)
IMR	Iterative model reconstruction (Philips)
IR	Iterative reconstruction
IRIS	Iterative reconstruction in image space (Siemens)
kV	Kilovolt (tube voltage)

LD	Low dose (Paper I)
mA	Milliampere (tube current)
mAs	Milliampere second (tube charge, radiation quantity)
mGy	Milligray (absorbed radiation dose)
mSv	Millisievert (equivalent radiation dose)
MDCT	Multi-detector computed tomography
MPR	Multi-planar reconstruction
MTF	Modulation transfer function
ND	Normal dose (Papers I–III)
NPS	Noise-power spectrum
RD	Reduced dose (Papers II, III)
ROC	Receiver operating characteristics
ROI	Region of interest
SAFIRE	Sinogram affirmed iterative reconstruction (Siemens)
SD	Standard deviation
SNR	Signal-to-noise ratio
VGA	Visual grading analysis

Original papers

This thesis is based on the following papers, which will be referred to in the text by their Roman numerals. The complete papers are appended at the end of the theses. Unpublished paper IV is excluded from the e-published version of this thesis. Reprints were made with permission from the respective publishers.

I R. Siemund, **A. Löve**, D. van Westen, L. Stenberg, C. Petersen, I. M. Björkman-Burtscher.

Radiation dose reduction in CT of the brain: Can advanced noise filtering compensate for loss of image quality?

Acta radiologica 2012 May 1;53(4):468-72.

II **A. Löve**, R. Siemund, P. Höglund, D. van Westen, L. Stenberg, C. Petersen, I. M. Björkman-Burtscher.

Hybrid iterative reconstruction algorithm in brain CT – A radiation dose reduction and image quality assessment study

Acta radiologica (accepted 2013 May).

III **A. Löve**, R. Siemund, P. Höglund, B. Ramgren, P. Undrén, I. M. Björkman-Burtscher.

Hybrid iterative reconstruction algorithm improves image quality in craniocervical CT angiography

American journal of roentgenology (accepted 2013 March).

IV **A. Löve**, M-L Olsson, R. Siemund, F. Stålhammar, I. M. Björkman-Burtscher, M. Söderberg.

Six iterative reconstruction algorithms in brain CT – A phantom study on image quality at different radiation doses

In manuscript.

Introduction

Computed tomography (CT) is an important diagnostic tool in neuroradiology. It uses ionizing radiation to produce images of the subject for diagnostic purposes. Continuously increasing utilization of clinical CT imaging has raised concerns over the risks associated with increased population radiation exposure.[1-4]

The main limitation for radiation dose reduction in CT is the resulting degradation of image quality due to increased noise and reduced spatial resolution.[5, 6] Therefore, maintenance of diagnostic image quality while reducing radiation exposure is a major technical challenge. In recent years, various technological advances have been introduced to address the problem. In the beginning, emphasis was mainly on hardware improvement, but in the last few years focus has shifted towards advanced software solutions.

Post-processing image filters (Paper I) and iterative reconstruction (IR) algorithms (Papers II–IV) are two software innovations directly targeting CT image quality. These operate at different stages in the image generation process, with the common objective of reducing noise while maintaining structural information. Both solutions have been shown to improve image quality and allow for radiation dose reduction in CT.[7-13]

In diagnostic radiology, implementation of new methods in clinical routine is usually preceded by validation in phantoms and clinical feasibility studies. Typically, such evaluation involves both objective and subjective assessment of diagnostic image quality. Subjective assessment is particularly challenging as it introduces various biases – systematic errors from the truth – that can affect study results.[14, 15] Knowledge about these biases is important as they can be counteracted with robust study design. Furthermore, the choice of evaluation scales and study groups deserves careful consideration.

This thesis explores the potential of various software innovations to improve image quality in brain CT and craniocervical CT angiography (CTA), and the potential for

radiation dose reduction, in both clinical data and in phantoms. In addition, various aspects of image quality assessment and study design are explored.

Background

Computed tomography

Computed tomography is widely used in diagnostic imaging. Brain CT and craniocervical CT angiography are fundamental examinations in diagnostic neuroradiology and important first-line diagnostic tools in many clinical situations, e.g. in acute stroke.[16, 17]

Technique

In CT, a rotating source of X-rays (tube) is used to acquire volumetric images of the subject. During the scan the tube continually rotates around the subject while a detector on the opposite side records the remaining beam intensity – how much the subject attenuates the X-rays – at different angles (projections). Attenuation is expressed in Hounsfield units (HU).

Multi-detector CT (MDCT) uses an array of detector rows, permitting simultaneous acquisition of data from multiple parallel slices, thus reducing examination time. Current mainstream CT systems allow simultaneous acquisition of 64 to 128 slices, with up to 512-slice systems being available. *Beam collimation* is a product of the number of detector rows and the detector row width. In a 64-slice system with a detector row width of 0.625 mm the collimation is 64 x 0.625 mm, resulting in a total coverage of 4 cm. A disadvantage of MDCT is slightly reduced image quality with increasing number of parallel slices due to cone beam artifacts and scattered radiation.[18]

CT scans can be run in different modes. In *axial (sequential) mode* the subject is fixed while the X-ray source and detector rotate around it in the x-y plane. In *spiral (helical) mode* the subject simultaneously moves along the z-axis, perpendicular to the plane of rotation of the X-ray source and detector. In axial mode, image quality is marginally better, but acquisition time is increased in scans longer than the detector width.[19] Speed and continuity of image data acquisition are the main benefits of spiral mode, a

benefit generally considered to outweigh the slight associated reduction in image quality.[19] To allow enough attenuation data to be gathered for image reconstruction, helical CT scans need to complete at least an extra half rotation before the start and beyond the end of the imaged volume.[20] Such *over-ranging* (or over-scanning) is an important source of increased radiation dose in spiral MDCT compared with axial CT, particularly at shorter scan lengths, and increases with collimation and pitch.[21, 22] In some recent wide detector scanner models this problem has been addressed with *dynamic z-axis collimation*.[23]

Radiation dose

The main disadvantage of CT is the use of ionizing radiation, with the associated risk of tissue damage and cancer induction. In CT, radiation dose depends on tube current (mA), slice scan time (s) and peak tube voltage (kV).[24] Image *noise* is inversely proportional to radiation dose. According to the Poisson distribution, increase in noise associated with radiation dose reduction equals $\sqrt{(mAs_{original}/mAs_{reduced})}$.[6] Therefore, 30% radiation dose reduction is expected to increase image noise by 20% ($\sqrt{(100/70)} = 1.20$), thus reducing image quality.

Image noise primarily affects the ability of the CT system to reproduce two adjacent objects with similar CT numbers as separate structures (i.e. *low-contrast resolution*).[5, 25] In brain CT, the attenuation difference between gray and white matter is small (~7 HU), requiring low noise levels and high radiation doses to adequately reproduce internal brain anatomy. In chest CT, however, noise is less of a problem since the attenuation difference between vessels or bronchial walls and the surrounding parenchyma is large (~850 HU), allowing for much higher noise levels (and thus lower radiation doses).[5]

In clinical CT imaging, radiation dose levels should be adjusted according to the anatomical region being examined and the diagnostic task. Basic understanding of CT dosimetry is a prerequisite for successful radiation dose optimization.

The fundamental dosimetric quantity in spiral CT is the *computed tomography dose index by volume* ($CTDI_{vol}$), expressed in milligrays (mGy). It represents the scanner radiation output, and is directly proportional to tube current in milliamperes (mA).[26, 27] The $CTDI_{vol}$ is calibrated with standardized 16 cm (head) or 32 cm (body) phantoms, and is not a direct measure of dose to the particular subject.[24, 28] *Patient dose* is the dose (mGy) absorbed by the particular subject and is dependent on $CTDI_{vol}$ and both subject volume and composition.[28] For subjects smaller than the

calibration phantom, patient dose is higher than indicated by the $CTDI_{vol}$, and vice versa. Total radiation exposure of a CT examination is defined by the *dose-length product* (DLP = $CTDI_{vol}$ x scan length in centimeters).[6] *Radiosensitivity* indicates the relative susceptibility of tissues to the harmful effects of ionizing radiation. *Effective dose equivalent*, the overall estimated risk associated with radiation exposure, can be estimated using conversion factors (E_{DLP}) reflecting radiosensitivity of the exposed tissues or body regions, and is expressed in millisievert (mSv = DLP x E_{DLP}).[29]

Image reconstruction

Image contrast is determined by differences in attenuation within the examined volume.[24] During the CT image reconstruction process, attenuation data from a large number of projections *(projection domain)* is mathematically processed to create an image of the examined volume *(image domain)*. Conventionally this is achieved through a fast mathematical procedure called *filtered back-projection* (FBP).[6]

Prior to reconstruction, the projection data is *filtered* to counteract blurring, and to achieve appropriate balance between detail resolution and noise for the diagnostic task.[6] *Back-projection* involves reversal of the attenuation measurement process where data from thousands of projections is used to reconstruct an image of the subject.

To ensure fast reconstruction, the FBP algorithm assumes an ideal system.[30] All measurements are treated equally and only processed once.[6] This is usually not an issue in CT examinations with standard radiation dose levels, where the signal is much higher than the noise. The limitations of FBP are revealed in low radiation dose acquisitions where image quality is compromised by disproportionally high levels of noise and image artifacts. In recent years, these limitations have been addressed with the introduction of iterative reconstruction (IR) algorithms – discussed further on.

Parameters of image quality

The goal of clinical CT imaging is to produce images of diagnostic quality using the lowest possible radiation exposure – often referred to as the *ALARA-principle* (as low as reasonably achievable).[5, 31, 32] This implies adequate reproduction of clinically important anatomical structures and pathological processes.

There are four fundamental determinants of image quality in CT systems:[24]

- *Low-contrast resolution* is the ability of the system to reproduce two adjacent objects with similar CT numbers as separate structures. Low contrast resolution is primarily increased with decreased image noise.[5, 24]

- *Spatial resolution* (high contrast resolution) is the ability of the system to resolve image detail. Spatial resolution is determined by object-to-detector distance, focal spot size, detector size, reconstruction matrix resolution, and slice thickness.[5]

- *Noise* represents the local statistical fluctuation in CT numbers of individual picture elements.[24] Noise decreases with increased radiation exposure (mAs, kV) and is also affected by slice thickness and reconstruction algorithm.[5] Noise power spectrum (NPS) is a measure of noise power as a function of spatial frequency (i.e. noise distribution),[33] and characterizes noise texture.

- *Image artifact* is defined as a systematic discrepancy between CT numbers in the reconstructed image and true attenuation coefficients of the examined volume. Artifacts can be based on *physics* (e.g. beam hardening, photon starvation, partial volume), the *patient* (e.g. metal and motion artifacts), or the *CT system* (e.g. ring and distortion artifacts).[34]

Image quality improvement

Image quality in CT is a result of interaction between many factors. The optimal balance between image quality and radiation dose is achieved by adjusting scan parameters and the image reconstruction process to fit diagnostic requirements, with respect to clinical indication and the anatomical region in question.

Radiation dose is a major determinant of image quality. Steadily increasing utilization of CT in clinical practice has raised concerns regarding the risks associated with radiation exposure.[1-4] In recent years various techniques have been developed that allow radiation dose to be reduced while maintaining diagnostic image quality. Two of these innovations directly target image quality:

- *Post-processing image filters* [35, 36] optimize images that have already been reconstructed with FBP. They use adaptive noise suppression and edge enhancement to reduce noise while maintaining structural information,

resulting in improved image quality.[13] As they operate in the *image domain* only, post-processing filters are unable to retrieve image information already lost during the image reconstruction process. This may be a limitation of the method, especially in combination with aggressive dose reduction, where photon starvation artifacts and high level of noise predominate. Post-processing filters have both been available as proprietary solutions incorporated into a CT systems (e.g. IRIS, Siemens Healthcare, Forchheim, Germany), and as separate vendor independent modules (e.g. SharpView CT, ContextVision AB, Linköping, Sweden).[13]

- *Iterative reconstruction algorithms* [37-41] are the latest advance in CT technology and come in two basic designs:[42, 43] *a) Statistical iterative optimization,* based on photon statistics, assuming ideal system, and *b) Model-based iterative optimization,* that additionally attempts to model the system and the acquisition process, including system optics. These algorithms perform iterative image optimization, not only in the *image domain*, but also in the *projection domain*. Optimization of projection data *prevents* noise and artifacts in the projection domain from propagating into the image domain where they might be more difficult to remove. As a result, the IR algorithms can better deal with low signal levels in modern low dose acquisitions, and will ultimately replace conventional FBP.[43]

Both post-processing image filters[9, 10, 13, 44-47] and IR algorithms[7, 8, 11, 12, 48-52] have been shown to improve image quality in CT, and allow for considerable radiation dose reduction while maintaining diagnostic image quality.

Evaluation of image quality

Optimization of CT image quality requires reliable methods for evaluation of the resulting images to ensure adequate diagnostic quality.[53] Although many key image quality parameters can be directly measured, appreciation of human readers is also important. Ideally, the evaluation process should therefore consist of objective and subjective evaluation steps.

Objective evaluation

A variety of objective image quality parameters can be measured in clinical CT images.

- *CT numbers* (HU) are ideally measured in a homogeneous area in a phantom or an organ, to prevent structural information from contributing to the measurements.[24] Although exact absolute CT numbers are infrequently used for diagnostic purposes, they can be important, such as for characterization of incidental adrenal masses.[54]

- *Noise* in CT is typically expressed as standard deviation (SD) of the CT numbers in a ROI, assuming normal distribution.[24] Using SD for direct comparison of noise levels between FBP and extensively manipulated IR may be inappropriate if noise distribution differs (Paper IV).[55] NPS provides information about spatial characteristics (texture) of noise and is assessed in homogeneous phantoms using advanced mathematics.[33] It is ideal for comparison of noise distributions in different reconstructions.[55]

- *Signal-to-noise ratio* can be expressed as signal (HU) divided by noise (SD) in the same homogenous ROI (SNR = HU / SD).[5] High SNR indicates that true information (signal) overpowers noise.

- *Low-contrast resolution* represents the ability of the system to reproduce two adjacent objects with similar CT numbers as separate structures.[5] It can be subjectively evaluated as the smallest discernible object in an image, with specific difference in contrast relative to the adjacent background.[56] Low contrast resolution is highly dependent on image noise,[5, 24] and can be objectively assessed by the *contrast-to-noise ratio* (CNR), where the contrast between two structures is expressed as a function of noise. Examples of formulas for calculation of CNR between two structures (A and B) in CT images are:

  a) $CNR = (HU_A - HU_B) / [(SD_A^2 + SD_B^2)/2]^{0.5})$, and

  b) $CNR = (HU_A - HU_B) / SD_B$.

  The key difference between these is that the first formula takes into account noise levels (SD) in both structures, while the other only includes noise in one of the structures. If noise levels in both structures are identical – as is common in evaluation of low-contrast resolution – the formulas will return identical results. If there are considerable differences in noise levels between the structures (e.g. large attenuation differences), the first formula might produce more accurate absolute CNR values. However, in studies on CT

image quality the result of interest is typically the relative difference in CNR between two methods being compared, and not absolute CNR values.

- *Spatial resolution* (high contrast resolution) represents the ability of the system to resolve image detail.[5] To eliminate interference from image noise, spatial resolution is usually determined by using test objects with large differences in CT numbers in a phantom. Spatial resolution can be assessed subjectively by visualization of line pairs of increasing density.[24] Objectively it can be assessed by the *modulation transfer function* (MTF) which involves complex calculations based on the degree of sharpness observed in the image of a specific test object, typically a high density bead or wire.[33, 56]

Subjective evaluation

The ultimate goal of clinical CT imaging is to reproduce anatomical and pathological information for diagnostic purposes – not to produce technically flawless or aesthetically appealing images. The diagnostic process culminates in human interpretation of images and to simulate this, a subjective evaluation step is important in studies on diagnostic image quality.

There are two principal approaches to subjective evaluation of diagnostic image quality:[57]

- *Diagnostic acceptability* is an indirect measure of diagnostic performance. In *visual grading analysis* (VGA), image criteria consisting of clinically important anatomical structures are chosen and visually graded using an ordinal scale. The underlying theory is that the ability to detect pathology correlates well with accurate reproduction of anatomy.[57] Strictly speaking, ordinal grading data from VGA cannot be converted to numerical data, limiting the statistical evaluation to non-parametric rank-invariant methods that can be difficult to comprehend. It may, however, be argued that ordinal data can be treated as normally distributed interval variables,[58] especially in large data sets, and this is widely accepted in medical literature. The main benefit of the VGA approach is that it is not dependent on specific pathological processes or diagnoses, thereby simplifying the study design.

- *Diagnostic accuracy* is a direct measurement of diagnostic performance. It involves detection of abnormal cases against a background of normal cases. Typically, observers are asked to state if the examination is normal or

abnormal and to define their confidence level. The results are presented with receiver operating characteristics (ROC) curves where the true-positive fraction is plotted against the false-positive fraction. Different methods can be compared by measuring the area under the curve. However, assigning interval values to the area under the curve is equally debatable here as in VGA since the original variables are ordinal data.[57, 59]

Although it can be argued that *diagnostic accuracy* is a better measure of diagnostic image quality, such studies are more difficult to carry out as they require a large number of subjects with a balanced mix of abnormal and normal cases where the true status of each case is known. Furthermore, patient recruitment with informed consent may be problematic in important applications, such as in acute ischemic stroke, due to the risk of treatment delay. Because of this and due to the rapid technical advances in CT imaging a fast and effective method for subjective evaluation of image quality is favorable, explaining the preference for *diagnostic acceptability* above diagnostic accuracy in many recent studies.

Observer bias

Subjective evaluation of image quality is dependent on human observers. It can be argued that subjective evaluation of image quality is less reliable than objective measurements due to the introduction of various sources of *observer bias*.[14, 15, 60, 61] Nevertheless, subjective assessment is difficult to omit in studies on diagnostic image quality. Knowledge about the most important sources of observer biases is therefore important, as many of these can be counteracted with careful study design, resulting in more reliable study results.

Different subtypes of *observer bias* are important in the evaluation of image quality:

- *Adaptation bias* occurs when observers accustomed to a certain image appearance and noise texture prefer these images over clinically equivalent images with slightly different appearance. This bias can cause reduced acceptability during the introduction of new methods or equipment that affects image appearance.

- *Recognition bias* becomes evident when observers who are intended to be blinded for an evaluation variable (e.g. a new method or equipment) nevertheless can, or believe that they can, identify which reconstruction they are looking at. Trained radiologists may be able to recognize even the

slightest deviations from usual image characteristics, making complete blinding of observers very difficult to achieve.

- *Confirmation bias* can arise when observers themselves are responsible for the study design, allowing them to choose study variables to suit their personal preference. Individually tailored study variables serve to maximize the chance for positive results and subsequent publication, but are not necessarily representative for radiologists in general.

Substantial observer variation highlights the fact that in order to produce representative results, studies on subjective image quality not only require a sufficient number of cases, but also an adequate number of observers.

Many recent studies on image quality and radiation dose include only two observers,[12, 46, 62-64] although it can be assumed that observer diversity, caused by disparate visual and cognitive abilities, will lead to individual differences in interpretation.[65] There is no simple answer to the question how many observers are appropriate, but in general the more observers the better. An article by Obuchowski provides some background information and suggestions based on different study phases.[66] Furthermore, to avoid confirmation bias, it is strongly suggested to keep the number of observers directly involved in the study design at a minimum.

Evaluation scales

Subjective evaluation of image quality involves grading using *ordinal scales* where the order of grades is defined, but the degree of difference between them is not. The type of ordinal scale used for subjective evaluation of image quality deserves some consideration as it can affect the study results (Paper II).

Ranking of image quality

Ranking of image quality involves a simple form of ordinal scale. Observers are asked to do side-by-side ranking of image quality from best to worst according to image quality criteria (Papers I, II). The ranking steps are only relatively defined and do not include a description of the expected image quality. To compensate for this, one of the ranked images can be used as an internal reference. The internal reference usually represents the image quality produced by the current method, allowing rank comparison with the new method.[57]

The main disadvantage of ranking scales is that observers are usually forced to apply different ranks to images, even in cases where there are no differences in image quality. Furthermore, the size of each ranking step is variable, both between steps and between individual assessments, making it impossible to quantify the difference in image quality between the assigned ranks. Yet another disadvantage of the ranking procedure is that side-by-side presentation unavoidably facilitates recognition bias that may affect the results (Paper II).

Grading of image quality

Grading of image quality involves a more sophisticated type of ordinal scales. Observers are typically asked to grade individual images according to image quality criteria (Papers II, III). Each grading step is defined using an absolute description (instead of only relative as in ranking), preferably complemented with an example image to improve consensus between observers. Grading steps can be defined to reflect an approximately linear improvement in image quality, allowing some degree of quantification of differences. Compared with ranking, the main benefits of grading are that the scale is described in absolute terms, allowing images with similar image quality to be assigned the same grade, and that images are reviewed separately (not side-by-side), reducing the risk for recognition and adaptation bias.

Grading scales

Grading scales should be tailored for the study purpose. Ideally, in clinical studies on subjective image quality the number and size of the steps should reflect clinically relevant differences. The number of steps in the grading scale is a balance between resolution (small steps) and reproducibility (ease of use). In smaller studies, detection of small differences in image quality might require scales with high resolution. However, scales with many small steps can be difficult to apply in a consistent manner.

It is common for studies on image quality to employ five-step scales. However, through clinical correlation such grading scales can be reduced to only four steps, making them easier to use without significant loss in resolution:

- Grade 1: *Non-diagnostic* image quality.
- Grade 2: *Sufficient* image quality, some diagnostic limitations.
- Grade 3: *Standard* image quality, no significant diagnostic limitations.

- Grade 4: *Excellent* image quality, no diagnostic limitations.

The goal of CT optimization is to achieve image quality without significant diagnostic limitations using the smallest possible radiation dose *(ALARA-principle)*, corresponding to grade 3 in the scale example above. Due to patient diversity and technical limitations this is impossible to achieve for all patients, and therefore more scale steps are required.

To reflect basic diagnostic requirements it is important that the scale allows clear discrimination between diagnostic (grades 2–4) and non-diagnostic (grade 1) examinations. For clinical purposes it is usually not important to know how poor the non-diagnostic examinations are, eliminating the need for further grades at the lower end of the scale. Likewise, grade 4 represents image quality better than what is clinically required, implying usage of unnecessarily high radiation dose that should be lowered. Therefore, further grades at the higher end of the scale are unnecessary.

Results based on such four-grade scales in CT are easy to interpret. First it should be ensured that no examinations are graded 1 (non-diagnostic). Then, as many of the examinations as possible should be graded 3 (standard). Finally, the ratio between grade 2 (sufficient) and grade 4 (excellent) can be used to decide upon an adjustment of the radiation dose. Overrepresentation of examinations with grade 4 image quality would for example indicate a too high radiation dose that should be lowered.

In CT, too low or too high image quality is equally inappropriate.

Research subjects

Evaluation of new algorithms begins in phantoms where repetitive adjustments can be performed under standardized conditions allowing for exact comparison of results. In phantom studies, objective image quality can be automatically evaluated using special software, eliminating human influence with the accompanying variability and bias. Phantom studies are usually followed by clinical studies before the techniques are introduced in clinical routine. Such studies usually involve evaluation of diagnostic acceptability or diagnostic accuracy in a rather small population (<100).

Studies on diagnostic image quality should preferably be dimensioned to be able to detect clinically significant differences. Detection of small intervention effects, such as the effect of different image reconstruction methods on image quality, against a background of relatively large inter-individual variation (i.e. anatomical differences)

requires an adequate number of cases and good matching between cases and controls.[67] Better matching allows for fewer cases to be used. Ideally this involves having identical case and control populations with each case serving as their own control. Using the same case and control population when evaluating different radiation doses can involve increased radiation exposure to the research subjects. The associated risk is dependent on radiation dose, radiosensitivity of the exposed tissues, and age of the subjects.[24] These risk factors can be adjusted accordingly to minimize the resulting risk, for example by only including elderly research subjects (Papers I, II).

Aims

The main objectives of this thesis were to assess new methods for improving image quality in CT, explore the associated potential for radiation dose reduction, and to outline a comprehensive approach for evaluation of image quality.

Specific objectives of the individual papers were as follows:

Paper I

- To evaluate the effect of post-processing image filtering (SharpView CT) on image quality in brain CT acquired with 30% reduced radiation dose.

Paper II

- To evaluate the effect of statistical iterative reconstruction (iDOSE^4) on image quality in brain CT acquired with 30% reduced radiation dose.
- To illustrate and discuss methodological pitfalls and strengths associated with the methods used for subjective evaluation of image quality.
- To improve on methodological limitations of paper I.

Paper III

- To evaluate the potential of statistical iterative reconstruction (iDOSE^4) for improvement of image quality in craniocervical CT angiography.
- To assess inter-observer and intra-observer performance.
- To further improve on methodological limitations of papers I and II.

Paper IV

- To evaluate image quality produced by six iterative reconstruction algorithms (AIDR, ASIR, iDOSE[4], IMR, SAFIRE, Veo) in a brain CT phantom model, using different radiation dose levels and iterative image optimization levels.

Subjects and methods

Study subjects

The thesis includes a total of 100 human subjects in two human studies (Papers I, III) and one combined human and phantom study (Paper II). The fourth study was a phantom study. Human subjects were recruited amongst patients referred to the examinations on clinical indications not related to the studies. As the two first studies (Papers I, II) involved an additional CT examination leading to increased radiation exposure, informed consent was required. This requirement was waived for the third study by the ethical committee. An overview of study subjects is presented below (Table 1).

	Paper I	Paper II	Paper III	Paper IV
Examined organ	Brain	Brain	Head/neck	–
Number of cases/controls [a]	30	40	30	–
Subject age range (years)	65–92	65–93	16–80	–
Median/Mean age (years)	79 / 78	76 / 77	63 / 58	–
Examinations per subject	2	2	1	–
Radiation protection approval	Yes	Yes	Yes	–
Research ethics approval	Yes	Yes	Yes	–
Informed consent	Yes	Yes	Waived	–
Phantom examinations	No	Yes	No	Yes

Table 1 – Overview of study subjects.

[a] In each of the clinical studies (Papers I–III), each patient served as his/her own control.

CT technique

One of the studies (Paper III) was a sheer image quality evaluation study, while the remaining three (Papers I, II, IV) combined evaluation of image quality with radiation dose reduction. The baseline CT protocols (Table 2) used for the clinical studies (Papers I–III) were based on the clinical routine protocols in use at our clinic at that time. For the phantom study (Paper IV), the reference radiation dose was adjusted to match the *European guidelines on quality criteria for CT*,[5] while all other examination parameters were chosen according to recommendations for adult brain CT published by the *American Association of Physicists in Medicine (AAPM)*.[68] In studies evaluating the effect of different radiation dose levels (Papers I, II, and IV), 30–90% dose reduction was accomplished by adjusting tube charge (mAs).

	Paper I	Paper II [a]	Paper III	Paper IV
CT system	Philips	Philips	Philips	GE, Philips, Siemens, Toshiba
Tube voltage (kV)	120	120	120	120
Tube current (mAs)	355 248	498 349	121 [b]	Variable [c]
$CTDI_{vol}$ (mGy) (% of baseline dose)	57 (100%) 40 (70%)	57 (100%) 40 (70%)	6.8 [b] (–)	120 [d] (100%) 84 (70%) 48 (40%) 12 (10%)
Slice collimation (mm)	16 x 0.625	64 x 0.625	64 x 0.625	Variable [c]
Pitch	0.683	0.578	0.98 [b]	Variable [c]
Rotation time (s)	0.75	0.75	0.50 [b]	Variable [c]
Automatic dose modulation	Disabled	Disabled	Enabled	Disabled
Intravenous contrast medium	No	No	Yes	No

Table 2 – Overview of CT examination parameters used in the individual papers.

[a] Provided parameters are valid for the clinical study (not the phantom experiments).
[b] Values represent means since automatic dose modulation was enabled.

[c] Values adapted from the recommendations for adult brain CT published by the *American Association of Physicists in Medicine (AAPM)*.

[d] The baseline dose of 120 mGy for the 20 cm image quality phantom was chosen to roughly equal 60 mGy in a standard 16 cm phantom.[19]

Image reconstruction

FBP reconstruction was used as baseline for all four studies (Table 3). In one of the studies (Paper I) post-processing filtering was used for improving image quality. IR algorithms were utilized for the three remaining studies (Papers II–IV), using up to three different IR levels. In all other aspects, preparation of images for subjective and objective evaluation followed clinical routine for brain CT or craniocervical CTA, respectively.

	Paper I	Paper II	Paper III	Paper IV
FBP reconstruction	Yes	Yes	Yes	Yes
Post-processing filter	SharpView CT	–	–	–
IR algorithms	–	iDOSE^4	iDOSE^4	AIDR, ASIR, iDOSE^4, IMR SAFIRE, Veo
IR levels (*n*)	–	2	1	1–3 [a]
Slice thickness	5 mm	5 mm	3 mm	5 mm
MPR-mode	Average	Average	MIP	Average

Table 3 – Overview of image generation processes for individual papers.

[a] The IR algorithms IMR and Veo were evaluated using one IR level while AIDR, ASIR, iDOSE^4 and SAFIRE were assessed using three different IR levels.

AIDR = Adaptive Iterative Dose Reduction 3D.
ASIR = Adaptive statistical iterative reconstruction
iDOSE^4 = (product name, not acronym)
IMR = Iterative model reconstruction
SAFIRE = Sinogram affirmed iterative reconstruction
Veo = (product name, not acronym)

Image quality assessment

A variety of evaluation methods were employed in the studies (Table 4). The phantom studies (Papers II, IV) emphasized on objective evaluation of image quality, while clinical studies (Papers I–III) focused more on subjective evaluations.

	Paper I	Paper II	Paper III	Paper IV
Objective methods				
CT numbers (HU)	•	•	•	•
Noise (SD)	•	•	•	•
Signal-to-noise ratio (SNR)			•	•
Contrast-to-noise ratio (CNR)		•	•	•
Noise-power spectrum (NPS)		•		•
Spatial resolution (MTF)		•		•
Subjective methods				
Number of observers	5	6	5	5
Number of quality criteria	3	3	8	2
Low-contrast resolution	•	•		•
Spatial resolution			•	
Ranking of quality criteria	•	•		
Grading of quality criteria		•	•	
Inter-observer agreement		•	•	
Intra-observer agreement			•	

Table 4 – Overview of subjective and objective image quality assessment methods used in the different studies.

In all four studies, multiple independent observers were employed for subjective evaluation of carefully chosen clinically important image quality criteria. Images were presented in randomized order under standardized viewing conditions using identical workstations and medical grade display monitors and standardized window level/width. In the clinical studies (Papers I–III), the radiologists were free to adjust window level/width and zoom level, simulating standard clinical reading process.

Clinical images were anonymized and information on dose level and reconstruction type was masked.

In the brain CT studies (Papers I, II) subjective evaluation focused on low-contrast resolution, employing quality criteria sensitive for this: *gray-white matter discrimination* and *delineation of basal ganglia structures*. The subjective evaluation in the contrast-enhanced craniocervical CTA study (Paper III) focused on spatial resolution through evaluation of *delineation of vessels* in multiple locations. In addition, the studies all included subjective evaluation of *general image quality*, reflecting the over-all image impression.

Two types of subjective evaluation scales – ranking and grading – were used in the clinical studies (Papers I–III) and in one of these the two methods were compared (Paper II). The number of ranking steps was three (Paper I) or four (Paper II), dependent on the number of images being compared. The grading scales used in Papers II and III both included four steps, adhering to the following basic design:

- Grade 1: *Poor (non-diagnostic)*.
- Grade 2: *Fair (sufficient)*.
- Grade 3: *Good (standard)*.
- Grade 4: *Excellent (better than required)*.

In the phantom study (Paper IV) the subjective evaluation of low-contrast resolution was done in consensus between five observers, using two criteria: *smallest discernible cylinder*, and *smallest sharply defined cylinder*.

Finally, observer agreement was evaluated in two clinical studies (Papers II, III).

Statistical evaluation

In all four studies descriptive statistics were used, both for evaluation of subjective and objective data (Table 5). The subjective ranking and grading procedures in Papers I–III produce ordinal data that preferably should be statistically evaluated with non-parametric rank-invariant methods such as two-tailed sign test. This is especially true for forced ranking procedures (Papers I, II), where the sizes of the ranking steps can vary widely. In grading procedures (Papers II, III), the combination of scales reflecting approximately linear improvement in image quality and a large number of observations, allows the resulting ordinal data to be treated as normally distributed interval variables. This permits the use of traditional statistical methods in the form of descriptive statistics and linear mixed-models. In addition, grading data were presented using cumulative distribution graphs where the ordinal nature of the data was preserved. In Paper IV each automated measurement was performed only once and the subjective evaluation done in consensus. Therefore, data could be adequately presented using descriptive statistics only.

	Paper I	Paper II	Paper III	Paper IV
Objective methods				
Descriptive statistics	•	•	•	•
Linear mixed-model		• [a]	• [a]	
Subjective methods				
Descriptive statistics	•	•	•	•
Two-tailed sign test	• [b]	• [c]		
Percent agreement		• [d]	• [d]	
Cumulative distribution		• [e]	• [e]	
Linear mixed-model		• [f]	• [f]	

Table 5 – Overview of statistical methods used in the different papers.

[a] Used in evaluation of differences in mean attenuation and noise.
[b] Used for statistical evaluation or ranking data (non-parametric rank-invariant).
[c] Here only used for supplementing evaluation of ranking data to illustrate potential consequences of observer variability in combination with only two observers.
[d] Used for assessment of complete observer agreement.
[e] Used in evaluation of grade distribution between algorithms within observer.
[f] Used for differences in mean grade, including *p*-values and confidence intervals.

Results

Paper I

The study evaluated the combination of 30% radiation dose reduction and post-processing image filtering (SharpView CT) in brain CT (Figure 1).

In pooled data (Table 6), image quality was significantly improved in filtered low dose (FLD) compared with low dose (LD), for all three subjective quality criteria. Objective noise was reduced with about 15%, and thereby in parity with normal dose (ND). Subjectively, ND was still considered superior to FLD, although this difference was not statistically significant.

Considerable inter-observer variation was noted between the 5 observers.

Image quality assessment	Normal dose (ND)	Filtered low lose (FLD)	Low dose (LD)
Objective methods			
Mean CT numbers (HU)	25.9	25.8	25.7
Mean noise, SD (HU)	**3.0**	**3.0**	3.4
Subjective methods			
Mean rank, gray-white matter discrimination [a]	**1.78**	1.82 [b]	2.40
Mean rank, delineation of basal ganglia [a]	**1.65**	1.85 [b]	2.50
Mean rank, general image quality [a]	**1.65**	1.98 [b]	2.37

Table 6 – Summary of results from Paper I. Compared with low dose (LD), image quality was improved in filtered low dose (FLD), although subjectively normal dose (ND) was still considered superior.

[a] Lower rank indicates better image quality.
[b] Significant difference compared with LD ($P < 0.05$).

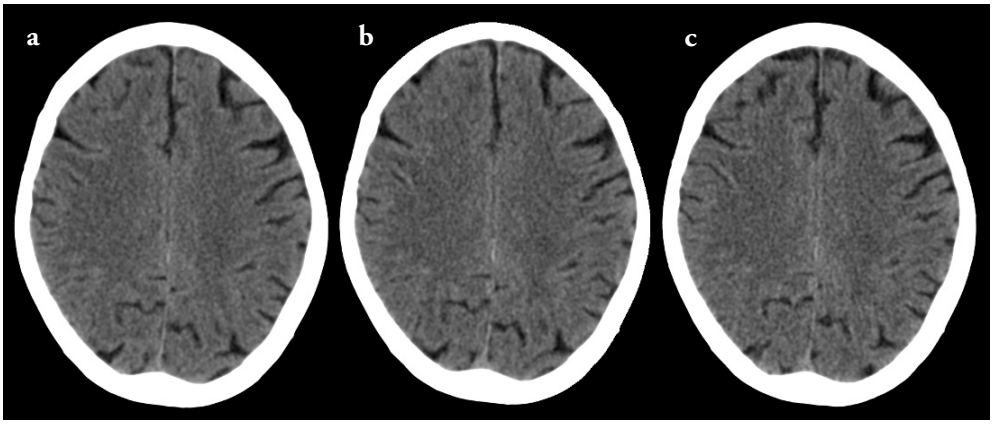

Figure 1 – Typical reconstruction results from a single patient at the level of the centrum semiovale showing: (a) normal dose (ND), (b) filtered low dose (FLD), and (c) low dose images. FD and FLD have very similar appearances, while slight increase in noise can be seen in LD. *CT images may not be adequately reproduced in print.*

Paper II

The study evaluated the combination of 30% radiation dose reduction and iterative reconstruction (iDOSE^4) in brain CT (Figure 2) and in a phantom.

Of the two iterative reconstruction levels, iDOSE^4 level 4 (ID4) was generally superior to iDOSE^4 level 2 (ID2) (Table 7). Objectively, image quality was significantly better in ID4 than in ND. Subjectively, this was reversed, with ND being both graded and ranked significantly better than ID4 (and ID2).

Inter-observer percent agreement for the 6 observers ranged from 50–70%.

Image quality assessment	Normal dose (ND)	iDOSE^4 level 2 (ID2)	iDOSE^4 level 4 (ID4)	Reduced dose (RD)
Objective methods				
CT numbers (HU)	25.6	25.5	25.5	25.5
Noise, SD (HU)	2.4	2.5 [a]	**2.2** [a]	2.8 [a]
CNR precentral gyrus	3.5	3.6	**3.9** [a]	3.2 [a]
CNR caudate nucleus	2.1	2.1	**2.4** [a]	1.9 [a]
CNR lentiform nucleus	2.2	2.4	**2.6** [a]	2.0 [a]
Subjective methods				
Mean grade, pooled criteria [b]	**3.22**	2.87 [a]	2.91 [a]	2.72 [a]
Mean rank, pooled criteria [c]	**1.6**	2.4 [a]	2.7 [a]	3.3 [a]

Table 7 – Summary of results of the clinical brain CT evaluation in Paper II. According to objective measures, image quality was best in iDOSE^4 level 4 (ID4) and iDOSE^4 level 2 (ID2). Subjectively, normal dose (ND) was considered superior to all others.

[a] Significant difference compared with ND ($P < 0.05$).
[b] Higher grade indicates better image quality.
[c] Lower rank indicates better image quality.

In the phantom evaluation, the IR algorithm reduced noise (SD) and improved SNR in reduced dose acquisitions, while maintaining spatial resolution (MTF) unchanged. Noise-power spectra were practically identical in IR and FBP.

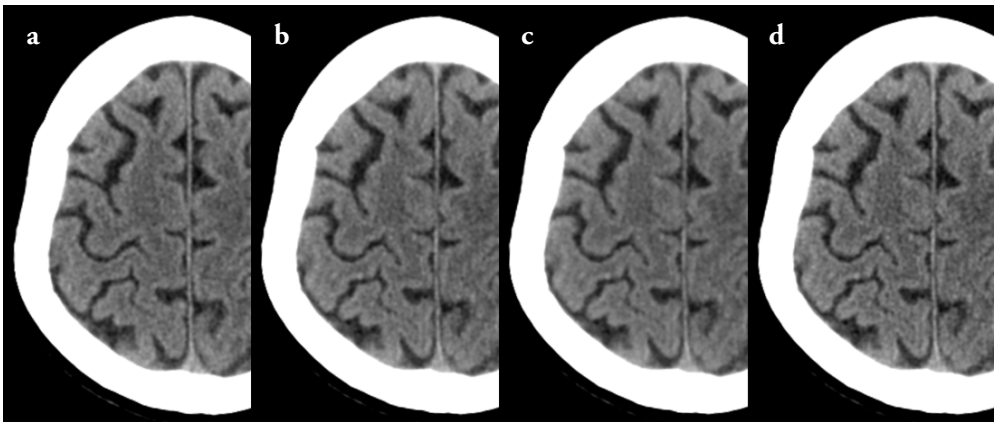

Figure 2 – Typical reconstruction results from a single patient at the level of centrum semiovale showing: (a) normal dose (ND), (b) iDOSE [4] level 2 (ID2), (c) iDOSE[4] level 4, and (d) reduced dose (RD) images. *CT images may not be adequately reproduced in print.*

Paper III

The study evaluated the potential of iterative reconstruction (iDOSE[4]) to improve image quality in craniocervical CT angiography (Figure 3).

As shown in Table 8, image quality was significantly better in IR than FBP for all criteria, both objective and subjective. Greatest subjective and objective improvement was seen at the level of the vertebral arteries. After excluding one patient that swallowed during the examination, the number of arterial segments receiving "poor" rating was reduced from 19 with FBP to 0 with IR.

Overall the intra-observer agreement was higher (mean 79%, range 69–88%) than the inter-observer agreement (mean 66%, range 59–75%).

Image quality assessment	FBP	IR
Objective methods		
CNR, vertebral arteries	2.8	**4.8** [a]
CNR, carotid bulbs	6.3	**6.9** [a]
CNR, basilar artery	5.7	**6.5** [a]
CNR, middle cerebral artery	5.7	**6.5** [a]
Subjective methods		
Mean grade, vertebral arteries [b]	2.45	**3.46** [a]
Mean grade, carotid bulbs [b]	3.40	**3.56** [a]
Mean grade, basilar artery [b]	3.41	**3.56** [a]
Mean grade, middle cerebral artery [b]	3.29	**3.45** [a]
Mean grade, general image quality [b]	3.17	**3.58** [a]

Table 8 – Summary of key results from Paper III. Image quality was significantly better in iterative reconstruction (IR) than filtered back-projection (FBP) for all criteria, both objective and subjective. Greatest subjective and objective improvement was seen at the level of the vertebral arteries.

[a] Significant difference compared with FBP ($P < 0.05$).
[b] Higher grade indicates better image quality.

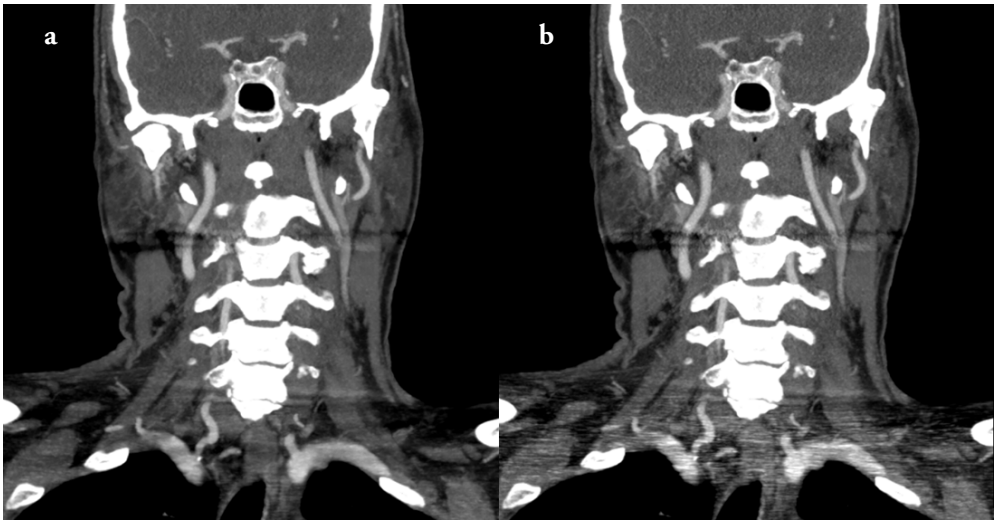

Figure 3 – Typical coronal reconstruction results from a single patient. The improvement in image quality in (a) iterative reconstruction, compared with (b) filtered back-projection, is best seen at the level of the shoulders. *CT images may not be adequately reproduced in print.*

Paper IV

This study evaluated image quality produced by six different IR algorithms, and the effects of radiation dose reduction in a brain CT phantom model.

As seen in Table 9, the greatest noise reduction was achieved with Philips IMR and GE IR3. All GE and Philips algorithms manage to reduce noise while maintaining spatial resolution (MTF). The model-based algorithms (Philips IMR, GE Veo) reduce noise progressively with reduced radiation dose.

	120 mGy			84 mGy			48 mGy			12 mGy		
	SD %	SNR %	MTF %	SD %	SNR %	MTF %	SD %	SNR %	MTF %	SD %	SNR %	MTF %
GE												
FBP	100	100	100	100	100	100	100	100	100	100	100	100
IR1	86	87	100	87	87	101	89	89	101	94	95	101
IR2	62	117	103	63	117	104	66	117	104	68	124	104
IR3	**42**	**170**	106	**43**	**167**	106	**46**	**160**	107	**46**	**175**	108
Veo	97	69	**136**	98	70	**141**	89	78	**149**	67	103	**134**
Philips												
FBP	100	100	100	100	100	100	100	100	100	100	100	100
IR1	97	104	100	94	107	100	91	110	102	88	113	100
IR2	79	127	101	77	129	100	80	126	102	77	130	100
IR3	**66**	**152**	97	63	158	100	67	149	102	64	157	102
IMR	69	141	**119**	**59**	**168**	**117**	**52**	**189**	**117**	**44**	**224**	**113**
Siemens												
FBP	100	100	**100**	100	100	**100**	100	100	**100**	100	100	**100**
IR1	85	117	88	85	118	87	87	116	93	83	120	90
IR2	71	140	87	71	142	89	74	137	92	67	150	87
IR3	**59**	**170**	88	**58**	**175**	89	**60**	**171**	91	**53**	**194**	86
Toshiba												
FBP	100	100	**100**	100	100	**100**	100	100	**100**	100	100	**100**
IR1	77	124	93	77	124	95	78	128	83	65	153	73
IR2	**66**	**139**	95	66	141	95	62	152	84	57	175	72
IR3	67	138	94	**61**	**146**	91	**58**	**162**	81	**54**	**188**	67

Table 9 – Results from objective measurements of noise (SD), signal-to-noise ratio (SNR) and $MFT_{10\%}$ (MTF) presented as *percentages relative to FBP*. Values in blue indicate improvement; values in red indicate deterioration; and bold style the best result for the respective CT system/dose.

CNR was generally improved with increasing radiation dose and higher IR-level (Figure 4). Same trend was seen with subjective visual evaluation of cylinders in the low-contrast module (Figure 5).

Visual evaluation of NPS curves revealed almost identical curve shapes (noise texture) between Philips FBP and IR, while they were very dissimilar for both model-based IR algorithms (Philips IMR, GE Veo) indicating different noise textures (Figure 6).

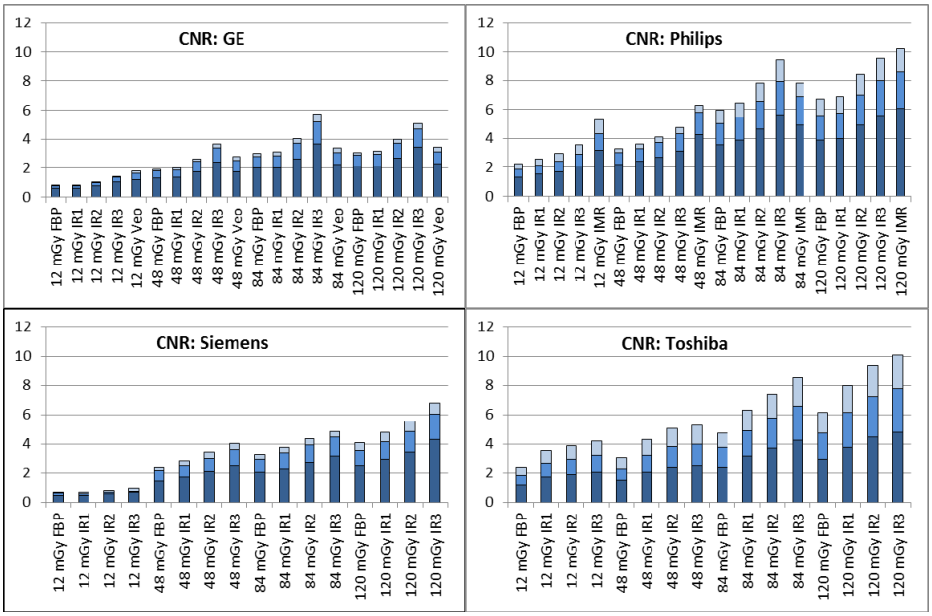

Figure 4 – Cumulative contrast-to-noise (CNR) ratios for all combinations of radiation doses and reconstruction algorithms from objective measurements at three nominal contrast levels: 1.0% (dark blue), 0.5% (blue), and 0.3% (light blue).

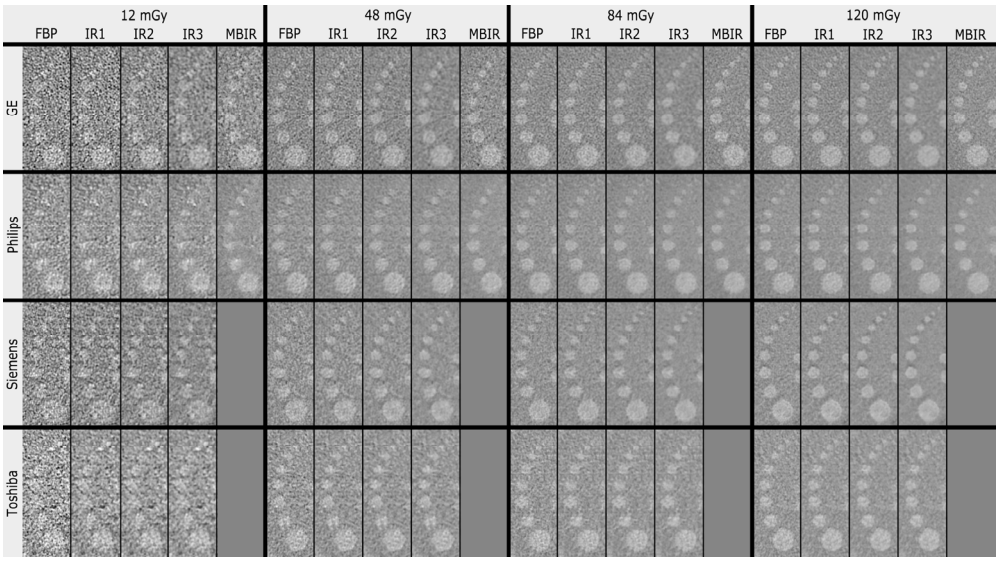

Figure 5 –All 72 combinations of radiation dose and reconstruction algorithm in the low-contrast phantom at 1.0% nominal contrast level. *CT images may not be adequately reproduced in print.*

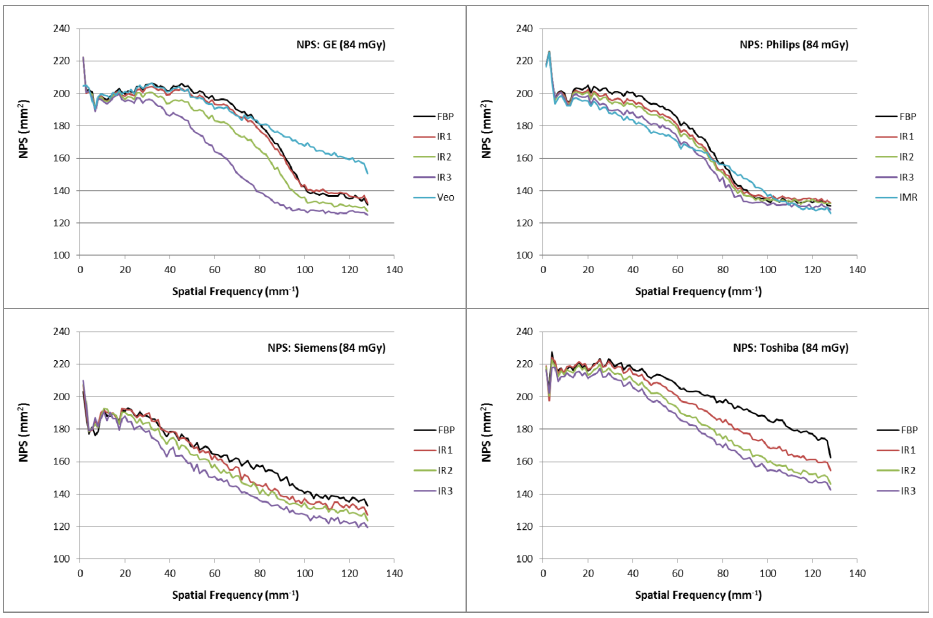

Figure 6 – Noise power spectra for all reconstruction algorithms at 84 mGy.

Discussion

The studies included in this thesis show that software solutions can be used to improve image quality and reduce radiation doses in neuroradiological CT examinations. Results of the clinical studies (Papers I–III) and the phantom study (Paper IV) all revealed significant improvement of image quality compared with conventional FBP reconstruction. With all other examination parameters kept constant, image quality using the new methods surpassed the reference method (Papers III, IV). Even with considerable radiation dose reduction, certain image quality parameters were improved by the new method compared with the standard dose reference method (Papers II, IV).

Image quality and radiation dose

The post-processing image filter SharpView CT (Paper I), is a stand-alone solution that performs image optimization in the image domain only.[36] The main benefit of the post-processing approach is the ability to serve multiple CT systems from different vendors. A theoretical limitation of the post-processing approach is the inability to optimize projection data.

Optimization of projection data is a key feature of IR algorithms, such as iDOSE[4], with the main purpose being to prevent noise and artifacts from propagating to the image domain where they might be more difficult to remove. As a result, image quality in IR can be expected to be superior to post-processing filtering.

Unfortunately, no publications have directly compared SharpView CT and the IR algorithms. In a white paper published by Siemens,[40] their statistical IR algorithm (SAFIRE) is stated to improve image quality compared with a post-processing image filter from the same vendor (IRIS). However, this has not been confirmed in the peer reviewed literature. The only publication including both SAFIRE and IRIS[69] does not explicitly compare the two, but the results do not indicate a clear benefit of the IR algorithm compared with the post-processing image filter.

Post-hoc comparison of results from Papers I and II is difficult to perform because of different evaluation approaches used in the studies. Despite both studies using the same internal reference (normal dose FBP), the results of the subjective ranking procedure cannot be directly compared as a different number of ranks was used in the studies: three in Paper I and four in Paper II. However, noise was measured in a similar way in both studies allowing comparison of relative noise reduction. With FBP as baseline, SharpView CT reduced noise by 11%, while iDOSE[4] reduced noise by 10.8% (ID2) and 21.4% (ID4). Thus, more noise reduction was achieved with the IR algorithm than the post-processing image filter.

Review of the literature shows that both post-processing image filters and IR algorithms improve image quality and show potential for radiation dose reduction. This is in line with the results presented in this thesis. SharpView CT has been shown to improve image quality, with a potential for 13–30% radiation dose reduction in brain CT[13, 44], 30–50% dose reduction in abdominal CT,[10, 70, 71] and 58% in CT of the paranasal sinuses.[9]

Likewise, iDOSE[4] has been shown to improve image quality in various applications,[72-74] (Papers II–IV), but so far the potential for radiation dose reduction in clinical examinations has only been explored in cardiac CTA (55–63% reduction)[8, 75] and temporal bone CT (50% reduction).[76] One study on lesion detection rates in a liver phantom model reports no benefit of iDOSE[4] compared with FBP.[77]

Similar results have been published for IR algorithms from other vendors. However, studies comparing IR algorithms from different vendors are lacking. Only two studies have compared IR algorithms from two vendors.[7, 11] All other published studies (and Papers II, III) have compared the IR algorithms with FBP in the same CT-system. Therefore the comparison of six iterative IR algorithms in four CT systems in Paper IV gives a unique insight into the strengths and weaknesses of the algorithms.

Although all IR algorithms evaluated in Paper IV improved image quality compared with FBP reconstruction, they showed different strengths and weaknesses. The overall best performing algorithm in the study was the prototype model-based IR algorithm IMR (Philips), which managed to greatly reduce noise and at the same time improve spatial resolution. Both evaluated model-based IR algorithms (GE Veo and Philips IMR) were most effective at lower radiation dose levels – a finding indicating that this new generation of IR might allow further radiation dose reduction in clinical protocols. This has already been confirmed for GE Veo in various applications.[11, 78-81]

Comprehensive evaluation of image quality

Diagnostic radiology is constantly evolving through introduction of new technologies and continuous optimizations of current methods. Every innovation or major adjustment necessitates validation before implementation in clinical routine. Ideally, this entails multicenter randomized case-control clinical trials that measure diagnostic clinical performance compared with the clinical reference standard – the best available method for establishing the true status of a subjects' condition.[82] However, such clinical trials are complicated and time consuming. Considering the fast rate of development in diagnostic radiology, simpler evaluation methods are needed.

In Papers I–III, major simplification was achieved by using *diagnostic acceptability* testing as it involves evaluation of important anatomical structures instead of specific disease processes.[57] In this approach, diagnostic performance is indirectly measured through evaluation of carefully selected clinically important image quality criteria. The lack of *diagnostic accuracy* testing can be considered a weakness of the studies. However, diagnostic accuracy is not the ultimate evaluation method and studies exploring *diagnostic/therapeutic impact* and *patient outcome* are considered superior.[53]

Most studies on subjective image quality and IR only include multiple research subjects, but only two observers. This is the case for all published studies on IR in brain CT.[12, 46, 62-64] Employing two observers is sufficient only if inter-observer agreement is high and if the observers can be assumed to be representative for observers in general. Results from Papers I–III show that inter-observer variability is a problem and that it is unreasonable to assume that any two observers are representative for observers in general. Therefore, as recommended in the literature,[66] more observers are needed. However, a weakness of Papers I–IV is that all observers were recruited from the same center. This might for instance have contributed to overestimation of observer agreement.

Increasing the number of observers is not the only way to improve reliability of study results. Observer agreement can be improved by careful study design. In Paper I, considerable observer variation was seen in the subjective ranking procedure. To further explore this, Paper II included both side-by-side ranking and individual grading procedures supported with a more extensive objective evaluation. From this we concluded that grading was more reliable for the assessment of image quality, especially when complemented with a thorough objective assessment. However, observer variation was still considerable in Paper II, indicating that further methodological improvement was possible. An improved version of the grading scale

was introduced in Paper III, including image examples for each grade. This may have contributed to the observed improvement in subjective grading consistency, resulting in reduced observer variation.

As explained above, the main problems associated with subjective evaluation of image quality are related to consistency and reproducibility. Since these are the main strengths of objective evaluation methods, the two methods complement each other well. Various methods for assessment of image quality have been employed throughout this thesis in an attempt to improve on the methodology. Based on the findings presented in the papers, the following outlines for *comprehensive evaluation of image quality* have been constructed:

- *Evaluation:* Assessment of image quality benefits from using both subjective and objective methods (Papers I–IV) as these have different strengths and weaknesses. An important weakness of subjective evaluation is that it is prone to various observer biases that might affect study results. A disadvantage of objective evaluation is that it does not evaluate human perception of images and observer variation. Therefore, a combined approach should result in more reliable study results.

- *Observers:* In subjective assessment of image quality, multiple observers are required to compensate for observer diversity as illustrated in Papers I–III. This is especially important when the observers themselves are responsible for the study design and therefore also subject to *confirmation bias*.

- *Evaluation scales:* Image quality is preferably evaluated with grading of clinically important quality criteria. Grading scales should be clinically correlated and the grading steps should reflect clinically important differences. Four-grade scales may provide all the necessary information in the context of optimization of image quality and radiation dose in CT (Papers I–III).

- *Research subjects:* The use of identical case and control groups eliminates confounding factors related to individual anatomical and pathological diversity. Therefore, differences in results between the two groups can with greater certainty be assumed to be caused by differences between the methods being evaluated. Due to the perfectly paired situation, significant results can be achieved using small study samples (Papers I–III).

Clinical impact

Positive study results should preferably lead to clinical implementation of the observed advances. Although the results of Paper I (SharpView CT) were positive, they were never implemented in clinical routine at our clinic since the iterative reconstruction algorithm evaluated in Paper II (iDOSE[4]) was launched close in time and also proved to be effective. However, the results of both Papers II and III have had direct clinical impact on the current routine brain CT and craniocervical CTA protocols at our clinic, leading to reduced radiation doses and improved image quality.

In the near future, continuing advances in CT technology can be expected to further improve image quality. If possible this should lead to further reduction in radiation dose, since *in CT, too low or too high image quality is equally inappropriate.*

Conclusions

The main objectives of this thesis were to assess new methods for improving image quality in CT, explore the associated potential for radiation dose reduction, and to outline a comprehensive approach for evaluation of image quality.

The following conclusions can be drawn from the individual papers:

Paper I

- Post-processing image filtering (SharpView CT) improves image quality in brain CT acquired with 30% reduced radiation dose.
- Noise levels were reduced and equivalent with normal dose. Subjective image quality was improved although normal dose still was considered non-significantly superior.
- Ranking of image quality is associated with considerable observer variation.

Paper II

- Statistical iterative reconstruction (iDOSE^4) improves image quality in both phantom and clinical brain CT acquired with 30% reduced radiation dose.
- Noise levels and contrast-to-noise ratios were significantly improved beyond normal dose examinations. Subjective image quality was also improved although normal dose still was considered significantly superior.
- Observer bias is a problem associated with subjective evaluation of image quality, although less so in grading than in side-by-side ranking.
- To produce representative results studies on subjective image quality require an adequate number of both research subjects and observers.

- Modest inter-observer agreement was seen, possibly due to the use of using written grading scale only, lacking accompanying example images.

Paper III

- Statistical iterative reconstruction (iDOSE[4]) improves image quality in craniocervical CT, especially at the thoracic inlet.
- Noise levels, signal-to-noise ratios, and contrast-to-noise ratios were all significantly improved compared with filtered back-projection. Subjective image quality was also improved at all levels.
- Perfect matching of cases and controls eliminated many confounding factors possibly affecting image quality in CTA.
- Improved inter-observer agreement was seen, probably reflecting a well-defined written grading scale in combination with example images.

Paper IV

- The iterative reconstruction algorithms have different strengths and weaknesses, but all improve image quality compared with filtered back-projection from the same vendor.
- Noise levels, signal-to-noise ratios, and contrast-to-noise ratios all improved, although to a variable degree. Three of the algorithms also improved spatial resolution. Subjective image quality was generally improved for all algorithms.
- The model-based algorithms improved image quality progressively with decreasing radiation doses, while the effect of statistical algorithms was more constant irrespective of radiation dose.
- Iterative reconstruction affected noise-power spectra, explaining differences in image texture, especially for the model-based algorithms.

Acknowledgements

Working on my dissertation has been an inspiring and fulfilling experience. It has given me the opportunity to work with many wonderful people that have contributed to the success of this project.

First, I want to thank my senior supervisor, Isabella Björkman-Burtscher, for believing in me and encouraging me when I needed it the most. Also I want to thank my assistance supervisor, Roger Siemund, who first introduced me to radiology research. Thank you for your relentless effort in getting this project done in a timely manner!

Next, I want to thank Peter Hochbergs, head of radiology, and professors Pia Maly-Sundgren and Stig Holtås for their interest in supporting my academic efforts and for giving me the opportunity to focus on this project.

Then, I want to thank Peter Höglund for his invaluable statistical guidance and advice, and for always being there for me when I got stuck.

Furthermore, I would like to thank my other co-authors: Danielle van Westen, Lars Stenberg, Cecilia Petersén, Per Undrén, Birgitta Ramgren, Marcus Söderberg, Marie-Louise Olsson and Fredrik Stålhammar – without your help I would never have succeeded.

Besides my co-authors, I also want to acknowledge my other colleagues at the department of neuroradiology: Åke Forssell, Mats Cronqvist, Olof Rudling, Martin Larmark, Håkan Sjunnesson, Jan Anders Ahnlide, Tomas Nilsson, Kasim Abul-Kasim, and Angelica Overgaard – thank you for covering for me at work when I was busy doing research!

My colleague and roommate, Mats Geijer, deserves a special mention. Working in parallel with him on his pursuit of becoming an associate professor, provided the competitive atmosphere that sometimes is necessary to get things done!

Also, I want to thank many radiographers, medical physicists, secretaries and administrators for much appreciated assistance: Mette Forseth, Eva Borglin, Karolina Larsson, Siv Tjärnberg, Helena Szaboova, Amir Vahabi, Alexandra Trudel, Marina Jamtheim, Sonny La, Lars Weber, Mats Nilsson, Åsa Hansson, Gunilla Angelin, Annika Samuelsson, Karin Jönsson, Anita Nilsson, Annika Törling-Ring, and Inger Bjarke ...with apologies if I forgot anyone – well done all of you!

Last, and above all, I want to thank my wife Elva Rut and my two wonderful daughters Sif (6) and Saga (4). Without their patience and support I would never have made it – I love you!

References

1. Hall EJ, Brenner DJ. Cancer risks from diagnostic radiology. *The British journal of radiology* 2008;81:362-378.
2. Berrington de Gonzalez A, Mahesh M, Kim KP, et al. Projected cancer risks from computed tomographic scans performed in the United States in 2007. *Archives of internal medicine* 2009;169:2071-2077.
3. Radiation exposures from medical applications. *The Federal Office for Radiation Protection*, Germany; 2011.
4. Hall P, Adami HO, Trichopoulos D, et al. Effect of low doses of ionising radiation in infancy on cognitive function in adulthood: Swedish population based cohort study. *BMJ* 2004;328:19.
5. Bongartz G GS, Jurik AG, et al. European Guidelines on Quality Criteria for Computed Tomography (EUR 16262 EN). *European commission*, Luxembourg; 1999.
6. Goldman LW. Principles of CT and CT technology. *Journal of nuclear medicine technology* 2007;35:115-128; quiz 129-130.
7. Desai GS, Thabet A, Elias AY, et al. Comparative assessment of three image reconstruction techniques for image quality and radiation dose in patients undergoing abdominopelvic multidetector CT examinations. *The British journal of radiology* 2013;86:20120161.
8. Hou Y, Liu X, Xv S, et al. Comparisons of image quality and radiation dose between iterative reconstruction and filtered back projection reconstruction algorithms in 256-MDCT coronary angiography. *American journal of roentgenology* 2012;199:588-594.
9. Kropil P, Cohnen M, Andersen K, et al. [Image quality in multidetector CT of paranasal sinuses: potential of dose reduction using an adaptive post-processing filter]. *RoFo : Fortschritte auf dem Gebiete der Rontgenstrahlen und der Nuklearmedizin* 2010;182:973-978.
10. Leander P, Soderberg M, Falt T, et al. Post-processing image filtration enabling dose reduction in standard abdominal CT. *Radiation protection dosimetry* 2010;139:180-185.
11. Mieville FA, Gudinchet F, Brunelle F, et al. Iterative reconstruction methods in two different MDCT scanners: physical metrics and 4-alternative forced-choice detectability experiments--a phantom approach. *Physica medica* 2013;29:99-110.

12. Ren Q, Dewan SK, Li M, et al. Comparison of adaptive statistical iterative and filtered back projection reconstruction techniques in brain CT. *European journal of radiology* 2012;81:2597-2601.

13. Siemund R, Love A, van Westen D, et al. Radiation dose reduction in CT of the brain: can advanced noise filtering compensate for loss of image quality? *Acta radiologica* 2012;53:468-472.

14. Sackett DL. Bias in analytic research. *Journal of chronic diseases* 1979;32:51-63.

15. Sica GT. Bias in research studies. *Radiology* 2006;238:780-789.

16. Love A, Siemund R, Andsberg G, et al. Comprehensive CT Evaluation in Acute Ischemic Stroke: Impact on Diagnosis and Treatment Decisions. *Stroke research and treatment* 2011;2011:726573.

17. Latchaw RE, Alberts MJ, Lev MH, et al. Recommendations for imaging of acute ischemic stroke: a scientific statement from the American Heart Association. *Stroke; a journal of cerebral circulation* 2009;40:3646-3678.

18. Endo M, Mori S, Tsunoo T, et al. Magnitude and effects of x-ray scatter in a 256-slice CT scanner. *Medical physics* 2006;33:3359-3368.

19. Flohr TG, Schaller S, Stierstorfer K, et al. Multi-detector row CT systems and image-reconstruction techniques. *Radiology* 2005;235:756-773.

20. Schilham A, van der Molen AJ, Prokop M, et al. Overranging at multisection CT: an underestimated source of excess radiation exposure. *Radiographics* 2010;30:1057-1067.

21. Tien CJ, Winslow JF, Hintenlang DE. A methodology for direct quantification of over-ranging length in helical computed tomography with real-time dosimetry. *Journal of applied clinical medical physics* 2011;12:3326.

22. Kroft LJ, Roelofs JJ, Geleijns J. Scan time and patient dose for thoracic imaging in neonates and small children using axial volumetric 320-detector row CT compared to helical 64-, 32-, and 16- detector row CT acquisitions. *Pediatric radiology* 2010;40:294-300.

23. Christner JA, Zavaletta VA, Eusemann CD, et al. Dose reduction in helical CT: dynamically adjustable z-axis X-ray beam collimation. *American journal of roentgenology* 2010;194:W49-55.

24. Goldman LW. Principles of CT: radiation dose and image quality. *Journal of nuclear medicine technology* 2007;35:213-225; quiz 226-218.

25. Judy PF, Swensson RG, Szulc M. Lesion detection and signal-to-noise ratio in CT images. *Medical physics* 1981;8:13-23.

26. Bauhs JA, Vrieze TJ, Primak AN, et al. CT dosimetry: comparison of measurement techniques and devices. *Radiographics* 2008;28:245-253.

27. Computed tomography radiation dose education slides. *American association of physicists in medicine (AAPM)*; 2012. (See: http://www.aapm.org/pubs/CTProtocols/documents/EducationSlides.pptx).

28. McCollough CH, Leng S, Yu L, et al. CT dose index and patient dose: they are not the same thing. *Radiology* 2011;259:311-316.

29. Coursey CAF, D. P. CT and radiation: What radiologists should know. *Applied radiology* 2008;37.

30. Fleischmann D, Boas FE. Computed tomography--old ideas and new technology. *European radiology* 2011;21:510-517.

31. Slovis TL. The ALARA concept in pediatric CT: myth or reality? *Radiology* 2002;223:5-6.

32. Shrimpton PC, Edyvean S. CT scanner dosimetry. *The British journal of radiology* 1998;71:1-3.

33. Report 54: Medical imaging – Assessment of image quality. *International commission on radiation units and measurements (ICRU)*, Bethesda, MD; 1996.

34. Barrett JF, Keat N. Artifacts in CT: recognition and avoidance. *Radiographics* 2004;24:1679-1691.

35. White paper: Iterative Reconstruction in Image Space (IRIS). *Siemens healthcare*, Forchheim, Germany; 2009.

36. White paper: CT dose reduction using SharpView CT. *SharpView AB*, Linköping, Sweden; 2007.

37. White paper: iDOSE^4 iterative reconstruction technique. *Philips healthcare*, Koninklijke, The Netherlands; 2011.

38. Hsieh J. Removing the Barriers (ASIR). *CT clarity - The magazine of CT (GE healthcare)*, Buckinghamshire, UK; 2009.

39. White paper: Veo - CT model-based iterative reconstruction. *GE healthcare*, Buckinghamshire, UK; 2010.

40. White paper: SAFIRE - Sinogram Affirmed Iterative Reconstruction. *Siemens healthcare*, Forchheim, Germany: Siemens AG; 2010.

41. White paper: AIDR 3D - Reduces dose and simultaneously improves image quality. *Toshiba Medical Systems*, Tokyo, Japan; 2011.

42. Beister M, Kolditz D, Kalender WA. Iterative reconstruction methods in X-ray CT. *Physica medica* 2012;28:94-108.

43. Willemink MJ, de Jong PA, Leiner T, et al. Iterative reconstruction techniques for computed tomography Part 1: Technical principles. *European radiology* 2013;23:1623-31.

44. Ledenius K, Stalhammar F, Wiklund LM, et al. Evaluation of image-enhanced paediatric computed tomography brain examinations. *Radiation protection dosimetry* 2010;139:287-292.

45. Renker M, Nance JW, Jr., Schoepf UJ, et al. Evaluation of heavily calcified vessels with coronary CT angiography: comparison of iterative and filtered back projection image reconstruction. *Radiology* 2011;260:390-399.

46. 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. *American journal of neuroradiology* 2012;33:218-224.

47. Pontana F, Duhamel A, Pagniez J, et al. Chest computed tomography using iterative reconstruction vs filtered back projection (Part 2): image quality of low-dose CT examinations in 80 patients. *European radiology* 2011;21:636-643.

48. Kalra MK, Woisetschlager M, Dahlstrom N, et al. Radiation dose reduction with Sinogram Affirmed Iterative Reconstruction technique for abdominal computed tomography. *Journal of computer assisted tomography* 2012;36:339-346.

49. Yamada Y, Jinzaki M, Hosokawa T, et al. Dose reduction in chest CT: comparison of the adaptive iterative dose reduction 3D, adaptive iterative dose reduction, and filtered back projection reconstruction techniques. *European journal of radiology* 2012;81:4185-4195.

50. Katsura M, Sato J, Akahane M, et al. Comparison of pure and hybrid iterative reconstruction techniques with conventional filtered back projection: Image quality assessment in the cervicothoracic region. *European journal of radiology* 2013;82:356-360.

51. Kulkarni NM, Uppot RN, Eisner BH, et al. Radiation dose reduction at multidetector CT with adaptive statistical iterative reconstruction for evaluation of urolithiasis: how low can we go? *Radiology* 2012;265:158-166.

52. Moscariello A, Takx RA, Schoepf UJ, et al. Coronary CT angiography: image quality, diagnostic accuracy, and potential for radiation dose reduction using a novel iterative image reconstruction technique-comparison with traditional filtered back projection. *European radiology* 2011;21:2130-2138.

53. Mansson LG. Methods for the evaluation of image quality: A Review. *Radiation protection dosimetry* 2000;90:89-99.

54. Young WF, Jr. Clinical practice. The incidentally discovered adrenal mass. *The New England journal of medicine* 2007;356:601-610.

55. Hsieh J. Computed Tomography: Principles, design, artifacts and recent advances. *SPIE press*; 2009:167-172.

56. Catphan ® 500 and 600 manual. *The phantom laboratory*, Greenwich, NY; 2009.

57. Bath M. Evaluating imaging systems: practical applications. *Radiation protection dosimetry* 2010;139:26-36.

58. Knapp TR. Treating ordinal scales as interval scales: an attempt to resolve the controversy. *Nursing research* 1990;39:121-123.

59. Bath M, Mansson LG. Visual grading characteristics (VGC) analysis: a non-parametric rank-invariant statistical method for image quality evaluation. *The British journal of radiology* 2007;80:169-176.

60. Brealey S, Scally AJ. Bias in plain film reading performance studies. *The British journal of radiology* 2001;74:307-316.

61. Obuchowski NA. Special Topics III: bias. *Radiology* 2003;229:617-621.

62. Becker HC, Augart D, Karpitschka M, et al. Radiation exposure and image quality of normal computed tomography brain images acquired with automated and organ-based tube current modulation multiband filtering and iterative reconstruction. *Investigative radiology* 2012;47:202-207.

63. Kilic K, Erbas G, Guryildirim M, et al. Lowering the dose in head CT using adaptive statistical iterative reconstruction. *American journal of neuroradiology* 2011;32:1578-1582.

64. Rapalino O, Kamalian S, Payabvash S, et al. Cranial CT with adaptive statistical iterative reconstruction: improved image quality with concomitant radiation dose reduction. *American journal of neuroradiology* 2012;33:609-615.

65. Obuchowski NA, Zepp RC. Simple steps for improving multiple-reader studies in radiology. *American journal of roentgenology* 1996;166:517-521.

66. Obuchowski NA. How many observers are needed in clinical studies of medical imaging? *American journal of roentgenology* 2004;182:867-869.

67. Swinscow TDV. Statistics at square one. *BMJ books*, London, UK; 1996.

68. Adult Routine Head CT Protocols (Version 1.1). *The American association of physicists in medicine (AAPM)*; 2012. (See: http://aapm.org/pubs/CTProtocols/documents /AdultRoutineHeadCT.pdf).

69. Tricarico F, Hlavacek AM, Schoepf UJ, et al. Cardiovascular CT angiography in neonates and children: Image quality and potential for radiation dose reduction with iterative image reconstruction techniques. *European radiology* 2013;23:1306-1315.

70. Martinsen AC, Saether HK, Olsen DR, et al. Reduction in dose from CT examinations of liver lesions with a new postprocessing filter: a ROC phantom study. *Acta radiologica* 2008;49:303-309.

71. Kropil P, Lanzman RS, Walther C, et al. [Dose reduction and image quality in MDCT of the upper abdomen: potential of an adaptive post-processing filter]. *RoFo : Fortschritte auf dem Gebiete der Rontgenstrahlen und der Nuklearmedizin* 2010;182:248-253.

72. Rajiah P, Schoenhagen P, Mehta D, et al. Low-dose, wide-detector array thoracic aortic CT angiography using an iterative reconstruction technique results in improved image quality with lower noise and fewer artifacts. *Journal of cardiovascular computed tomography* 2012;6:205-213.

73. Kligerman S, Mehta D, Farnadesh M, et al. Use of a hybrid iterative reconstruction technique to reduce image noise and improve image quality in obese patients undergoing computed tomographic pulmonary angiography. *Journal of thoracic imaging* 2013;28:49-59.

74. Utsunomiya D, Weigold WG, Weissman G, et al. Effect of hybrid iterative reconstruction technique on quantitative and qualitative image analysis at 256-slice prospective gating cardiac CT. *European radiology* 2012;22:1287-1294.

75. Hou Y, Xu S, Guo W, et al. The optimal dose reduction level using iterative reconstruction with prospective ECG-triggered coronary CTA using 256-slice MDCT. *European journal of radiology* 2012;81:3905-3911.

76. Niu YT, Mehta D, Zhang ZR, et al. Radiation dose reduction in temporal bone CT with iterative reconstruction technique. *American journal of neuroradiology* 2012;33:1020-1026.

77. Dobeli KL, Lewis SJ, Meikle SR, et al. Noise-reducing algorithms do not necessarily provide superior dose optimisation for hepatic lesion detection with multidetector CT. *The British journal of radiology* 2013;86:20120500.

78. Vardhanabhuti V, Loader R, Roobottom CA. Assessment of image quality on effects of varying tube voltage and automatic tube current modulation with hybrid and pure iterative reconstruction techniques in abdominal/pelvic CT: a phantom study. *Investigative radiology* 2013;48:167-174.

79. Katsura M, Matsuda I, Akahane M, et al. Model-based iterative reconstruction technique for radiation dose reduction in chest CT: comparison with the adaptive statistical iterative reconstruction technique. *European radiology* 2012;22:1613-1623.

80. Shuman WP, Green DE, Busey JM, et al. Model-Based Iterative Reconstruction Versus Adaptive Statistical Iterative Reconstruction and Filtered Back Projection in Liver 64-MDCT: Focal Lesion Detection, Lesion Conspicuity, and Image Noise. *American journal of roentgenology* 2013;200:1071-1076.

81. Machida H, Tanaka I, Fukui R, et al. Improved delineation of the anterior spinal artery with model-based iterative reconstruction in CT angiography: a clinical pilot study. *American journal of roentgenology* 2013;200:442-446.

82. Draft guidance for industry, clinical investigators, and Food and drug administration staff - Design considerations for pivotal clinical investigations for medical devices. *U.S. Department of health and human services; Food and drug administration (FDA); Center for biologic evaluation and eesearch*; 2011. (See: http://www.fda.gov/MedicalDevices /DeviceRegulationandGuidance/GuidanceDocuments/ucm265553.htm)

Appendix