Duke Radiology
Duke University School of Medicine

DUKE UNIVERSITY
MEDICAL PHYSICS
GRADUATE PROGRAM

Update from TG 233: Performance Evaluation of Computed Tomography Systems

Justin Solomon

Duke University Medical Center
Department of Radiology

CARL E. RAVIN
ADVANCED
IMAGING
LABORATORIES

AAPM Spring Clinical Meeting 2019

TG 233 Membership

Donovan Bakalyar

Kirsten L Boedeker

Samuel Brad

Jiahua Fan

Shuai Leng

Michael McNitt-Gray

Kyle J. Myers

Lucretiu Popescu

Juan Carlos Ramirez

Frank Ranallo

Ehsan Samei

Justin Solomon

Jay Vaishnav

Jia Wang

What am I going to talk about today?

- Overall philosophy behind TG 233
  - What need is being filled by the report?
  - Why doesn't TG 233 tell me exactly how to do a CT annual?
  - What TG 233 does not address
- Summary of the report
  - Pre-test inspection
  - Basic system performance
  - Operational performance
    - Automatic exposure control
    - Task-based image quality
    - Iterative reconstruction
- Tools provided along with the report
  - imQuest
  - iQModelo

Why does TG 233 exist?

- The way physicists evaluate the performance of CT systems needs to be updated
- New technologies
  - Tube current modulation
  - Iterative reconstruction (IR)
- New metrics needed
  - Traditional IQ metrics such as CNR are inadequate in some scenarios
- Changing roll of a Dx physicist
  - Pass/Fail specification -> clinical performance/utilization/optimization

TG 233 is a toolbox, not a recipe

- Report does NOT give a step by step guide on how do an annual CT physics survey
  - A one size fits all approach does not work with the complexity and diversity of modern CT systems
- Report does introduce new testing methods and metrics that a physicist can draw on
  - Idea is to standardize individual testing methodologies, not to standardize the entire testing process
  - Up to the expertise/discretion of the physicist to determined when it makes sense to apply a given test on a given system

Philosophy of TG 233

MEDPHYS 3.0

From	To
Specifications	Performance
Equipment	Operation
Quality check	Process consistency
Presumption	Actual utility
Compliance	Excellence

www.aapm.org/MedPhys30/

Areas not addressed by TG 233

- New dosimetry methods
  - TG 200 - CT Dosimetry Phantoms and the implementation of AAPM Report Number 111. Donovan will discuss this later in this session
  - TG 293 - Task Group on Size Specific - Dose Estimate (SSDE) for Head CT
- Dual energy CT (spectral CT)
  - TG 299 - Quality Control in - Multi-Energy Computed Tomography (MECT)
  - TG 291 - Task Group on Educational Report on Multi-Energy CT
- Cone-beam CT
- Cardiac CT
- Perfusion imaging

Outline of the report

- Pre-test inspection
  - Essentially a quick checklist
- Basic system performance
  - A brief tabular summary of testing methods that are already well established
- Operational performance
  - This is the bulk of the document
  - Introduces new testing methods and IQ metrics
  - Majority of talk will focus on this

Pre-test inspection

- Among other things…
  - Checklist of basic safety items related to the room construction (signage, shielding, etc)
  - Regulatory checklist*
  - Check for XR-29 compliance
  - Review QC program

	Task	Detail
1	"Caution Radiation Area" warning sign at all entrances to the scan room and other required postings	Verify presence off warning signs if required by your state regulations or institutional policy. Also check for the presence of other postings required by your state, accreditation bodies, and institution. Note that the code of federal regulations (CFR) regarding radiation warning signs (10 CFR 20.1901-1905) from the Nuclear Regulatory Commission (NRC) are intended to apply only to radioactive materials licensed by the NRC and thus do not directly apply to radiation producing machines. However, individual states may have posting requirements for radiation producing machines such as CT scanners. Independent of regulatory requirements, such postings are recommended by this task group at all entrances to the CT scan room.
2	Functioning of X-ray indication light at room entry	If you have X-ray indication lights, confirm their proper operation. Consult your state regulations or institutional policy to determine if they are required and comply with regulations and policies.
3	Doors to room closed when making an exposure	If the room has doors, confirm their proper operation. Consult your state regulations or institutional policy to determine if they are required and comply with regulations and policies. Note: This can create difficulties when transporting patients through the doors requiring "workarounds" from the staff.
4	Labeling, visibility and access to emergency stops	Emergency stops must be present, operable and accessible. This last point is extremely important. For example, wall mounted emergency stops can inadvertently be made inaccessible by a moving cart or by rearranging the room. See CFR rule 21CFR1020.33.
5	X-ray warning label at the control panel	See CFR rule 21CFR1020.30 (j): "Warning label. The control panel containing the main power switch shall bear the warning statement, legible and accessible to view: "Warning: This x-ray unit may be dangerous to patient and operator unless safe exposure factors, operating instructions and maintenance schedules are observed."
6	Production of a clearly noticeable signal during x-ray generation	This signal can be visible, audible, or both. See CFR rule 21CFR1020.
7	Direct line of sight to patient during procedure	Operator should have a direct line of sight to patient during CT examination. See CFR rule 21CFR1020.
8	Visual determination of a reference tomographic plane or reference plane offset	See CFR rule 33, 21CFR1020.33 (g). For CT, the rule reads as follows: "For any multiple tomogram system, means shall be provided to permit visual determination of the location of a reference plane. The relationship of the reference plane to the planes of the tomograms shall be provided to the user in addition to other information provided according to 1020.30(h). This reference plane can be offset from the location of the tomographic planes." The physicist should confirm adherence to this rule. If lasers are used, they must be usable under ambient lighting conditions of 500 lux.
9	Operator initiation of scans	Initialization of exposure must require positive, deliberate action by the operator. See Canada's Safety Code 35: "Safety Procedures for the Installation, Use and Control of X-ray Equipment in Large Medical Radiological Facilities".
10	Oral communication between operator and patient	Operator should be able to orally communicate with the patient from within the control room. See CFR rule 21CFR1020.
11	Display of technique factors before scan	The anticipated technique factors (e.g., kV, mAs) must be clearly indicated to the operator prior to the actual scan. See CFR rule 21CFR1020.
12	Pre-scan display of prescribed CTDI$_{vol}$, DLP, and size of phantom	The pre-scan (i.e., anticipated) CTDI$_{vol}$ and DLP should be displayed to the operator. The size of the CTDI phantom (i.e., 32 cm diameter or 16 cm diameter) should be indicated.[23]
13	Post-scan CTDI$_{vol}$, DLP and phantom size recording	The post-scan (i.e., delivered) CTDI$_{vol}$ and DLP should be displayed in the patient's dose report.[23]
14	Operator and service manuals present	Verify presence of operator, service, and other necessary technical manuals provided by the manufacturer. See CFR rule 21CFR1020.
15	MITA Smart Dose	National Electrical Manufacturers Association (NEMA) XR-29 Standard[4], also known as MITA Smart Dose, requires CT scanners to incorporate automatic exposure control (NEMA XR-28-2013), DICOM-compliant radiation dose structured reporting (NEMA XR-29-2013), dose check features (NEMA XR-25-2010[5]), and reference pediatric and adult protocols (NEMA XR-28-2013). Each manufacturer has a vendor certification web portal on the MITA Smart Dose website (http://www.medicalimaging.org/policy-and-positions/mita-smart-dose/). Medical physicists should contact the manufacturer of each CT scanner to upgrade software, if needed, and to obtain a verification of compliance. Without compliance (effective Jan. 1 2016), CMS reimbursement is reduced by 5% for outpatient scans in 2016 and by 15% in 2017. Note that compliance with XR-29 is established based on CT scanner capabilities, not on usage. When used appropriately, these features improve the safety of CT scans for the patients and can improve the image quality.
16	Dose data connectivity	Following MITA Smart Dose provision above, the scanner should ideally not only be capable of producing a DICOM radiation dose structured report, but also send the report to a destination for interpretation (e.g., via a dose monitoring system). This should ideally be integrated with a system by which the institution's technologists and radiologists properly understand the use and interpretation of dose results, and dose notification and alert values in accordance with institutional policies.
17	Presence of adequate shielding	If a new room has been constructed for your machine, a medical physicist, radiation safety officer or other qualified personnel should visually confirm the presence of shielding during construction, whenever possible. The best direct assessment technique uses a radioactive source and detector (in accordance with appropriate safety measures). An alternative technique is to use a portable x-ray source and detector. A third technique to confirm the presence of adequate shielding is to use the scattered radiation emitted when scanning a phantom (e.g., the body CTDI phantom) and a high-sensitivity radiation measurement device. If a new machine is being installed, a medical physicist, radiation safety officer or other qualified personnel should check the integrity of the shielding, since this could have been compromised during the construction process. One of the techniques outlined above should be used. It is imperative to ensure compliance with state regulations, which may involve design and verification of shielding integrity by a physicist and communicating the results of that verification to a state agency.[6]
18	Presence of an appropriate QC program	A continuous quality control program (QC) is required as part of CT accreditation programs (e.g., ACR, IAC, TJC) and may also be required by other accreditation bodies, the state, or by the institution. The presence and appropriate implementation for such a program should be verified as part of this pre-test inspection and if a QC program does not exist or is insufficient, one should be properly instituted.

* We can't promise that going through this checklist will ensure regulatory compliance!!!

Basic system performance

- Tabulates standard testing methods
- These tests are all familiar to clinical physicist
- Tests come from a number of sources:
  - American Association of Physicists in Medicine (AAPM) report 748 and report 39.
  - American College of Radiology (ACR) CT quality control (QC) manual for the ACR CT Accreditation Program (CTAP).
  - European Commission (EC) report 162 on CT quality assurance.
  - Food and Drug Administration Code of Federal Regulations (CFR) part 120.
  - International Atomic Energy Agency (IAEA) Human Health Series No. 1912 and Human Health Report No. 5.
  - International Commission on Radiation Units & Measurements (ICRU) report 87.
  - International Electrotechnical Commission (IEC) report 61223-3-5.
  - Vendor quality control documents.

Basic Performance| Geometric

- Laser alignment accuracy
- Table indexing accuracy
- Image position accuracy
- Image thickness accuracy
- Gantry tilt accuracy

Basic Performance | Radiation Output

- HVL
- Exposure reproducibility
- Exposure time reproducibility
- Exposure time accuracy
- Exposure linearity
- Tube potential accuracy
- Beam profile
- Localizer radiograph dose
- CTDI accuracy

Basic Performance | Image quality

- CT number accuracy
- CT number uniformity
- Artifact assessment
- High contrast (line pair) resolution
- Noise magnitude
- Contrast-to-noise ratio
- Slice Sensitivity profile

Just to clarify

- **TG 233 is not saying that all these tests need to be done!**
- We just compiled the standard tests that are available
- Up to a physicist to decide which tests make sense for the systems they test
  - Of course, regulation and accreditation guidelines will dictate a lot of that

Operational performance

- This is the bulk of the document
- Focus is on clinical performance
- Towards protocol optimization, operational consistency
- Section includes
  - Tube current modulation
  - Spatial resolution
    - For potentially non-linear CT systems
    - In-plane and z-direction
  - Noise
    - Noise power spectrum analysis
    - Non uniformity in noise
  - Task-based image quality
    - Task-based IQ from Fourier domain calculations
    - Task-based IQ from spatial domain calculations
    - Task-based IQ from human reader experiments

Basic vs Operational

- Basic performance
  - Does the system meet the manufacturer's specs?
  - Is is working as designed?
  - Is it safe to use?
  - Pass/fail?
  - Acceptance testing
- Operational performance
  - Are we using the system optimally?
  - How does it compare to another system?
  - Are we getting consistent clinical performance?
  - Commissioning

Would you rather be scanned a system with excellent specs that is used poorly, or a system with mediocre specs that is used optimally?

Operational performance

- This is the bulk of the document
- Focus is on clinical performance
- Towards protocol optimization, operational consistency
- Section includes
  - Tube current modulation
  - Spatial resolution
    - For potentially non-linear CT systems
    - In-plane and z-direction
  - Noise
    - Noise power spectrum analysis
    - Non uniformity in noise
  - Task-based image quality
    - Task-based IQ from Fourier domain calculations
    - Task-based IQ from spatial domain calculations
    - Task-based IQ from human reader experiments

What is tube current modulation?

- Noise in CT images depends on how much signal get to the detector
- For a fixed tube current (mA)
  - Larger patient -> less signal at detector -> noisier images
  - Smaller patient -> more signal at detector -> less noisy images (overdose?)
- TCM adapts the tube current for different patient sizes
- Goal is more consistent image quality for different patient sizes
- Different manufacturer's have different implementations and philosophies

Why test TCM?

- TCM systems have a very strong impact on patient dose and image quality
- Traditional testing has ignored this important feature
- TCM is software driven, and software updates can change how TCM works on a system (often without obvious notification from manufacturers)
- When designing protocols, need to predict how the system will respond to patients of different sizes
- TG 233 describes 2 primary methods to test a TCM system:
  - Discrete sizes: Mimics the case of scanning two patients of different size with the same settings
  - Continuous sizes: Mimics the case of scanning a single patient whose size changes along the scan direction

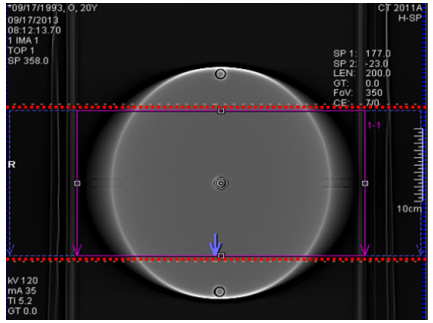
TCM| Discrete adaptation

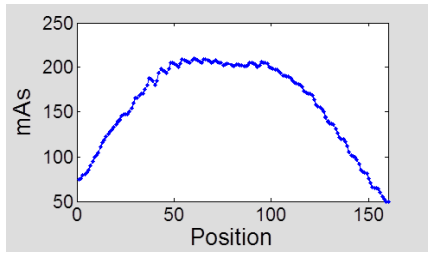
- Question: How does the system respond to patients of different sizes with the same scan settings?
- Use 2+ phantoms of different sizes
- For each phantom
  - Perform localizer (scout)
  - Scan with TCM engaged using the default settings for the protocol of interest
  - Record dose and image noise

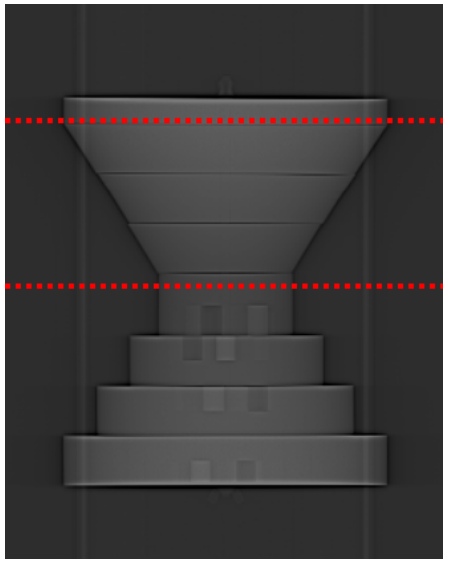
	Scan	kV	mAs / ref.	CTDIvol* mGy	DLP mGycm	TI s	cSL mm
Topogram	1	120	35 mA	0.13 L	10	7.8	0.6
CTDI32	2	120	215 / 210	14.48 L	124	1.0	0.6
Topogram	1	120	35 mA	0.13 L	4	3.2	0.6
CTDI16	2	120	43 / 210	2.90 L	25	1.0	0.6
Topogram	1	120	35 mA	0.13 L	5	3.9	0.6
CTDI10	2	120	23 / 210	1.55 L	13	1.0	0.6

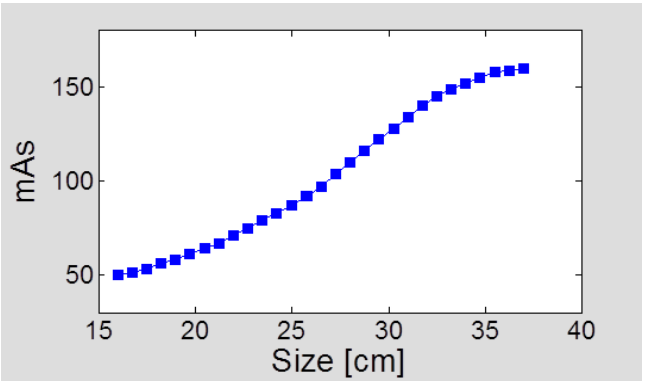
TCM| Continuous adaptation

- Question: How does the system respond along the longitudinal axis of a patient with varying size?
- Use a single variable sized phantom (or line up multiple phantoms of different sizes)
  - Perform localizer (scout)
  - Scan with TCM engaged using the default settings for the protocol of interest
  - Record dose and image noise as a function of phantom size



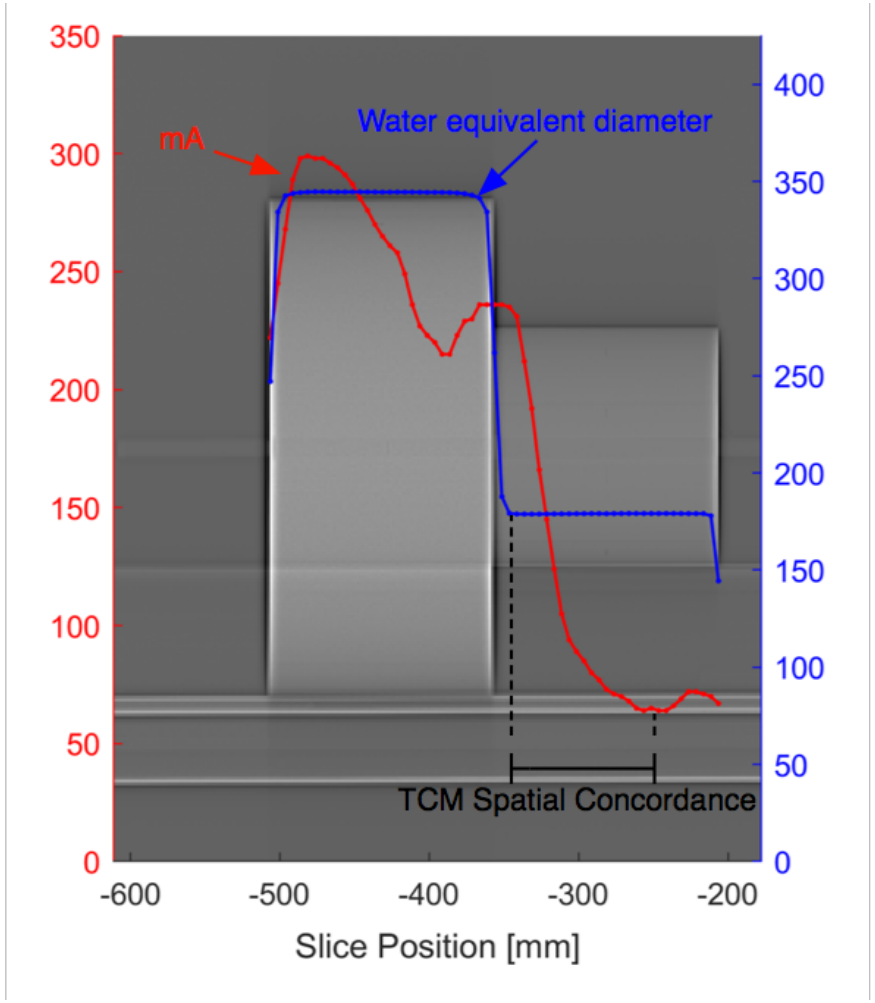






TCM| Analysis

- TG 233 describes several summary statistics to be calculated based on the collected data
- Will not discuss details for most of those statistics today, you'll need to wait for report ☺
- Spatial concordance:
  - How well does the system adapt to an abrupt change in patient size within a single scan (e.g., neck-to-shoulder transition)



Example of clinical utility| TCM changes

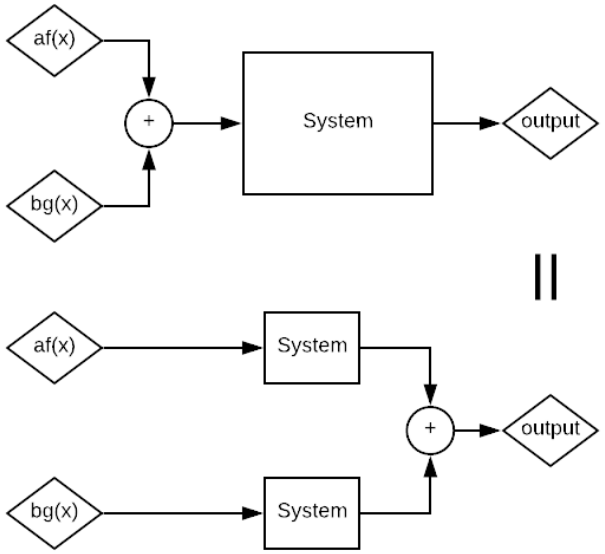
- During my annuals, I scan a multi-sized phantom under consistent conditions from year to year using TCM
- Check the tube current profiles to see if they have changed
- I also do this after software upgrades

Operational performance

- This is the bulk of the document
- Focus is on clinical performance
- Towards protocol optimization, operational consistency
- Section includes
  - Tube current modulation
  - Spatial resolution
    - For potentially non-linear CT systems
    - In-plane and z-direction
  - Noise
    - Noise power spectrum analysis
    - Non uniformity in noise
  - Task-based image quality
    - Task-based IQ from Fourier domain calculations
    - Task-based IQ from spatial domain calculations
    - Task-based IQ from human reader experiments

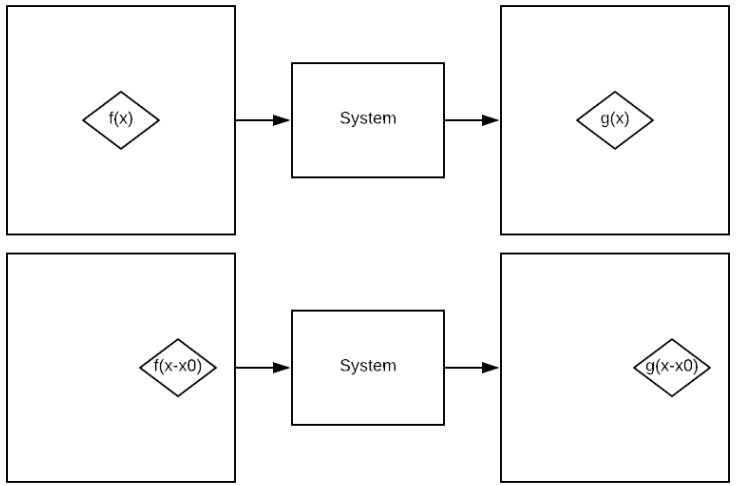
Definitions | Linearity

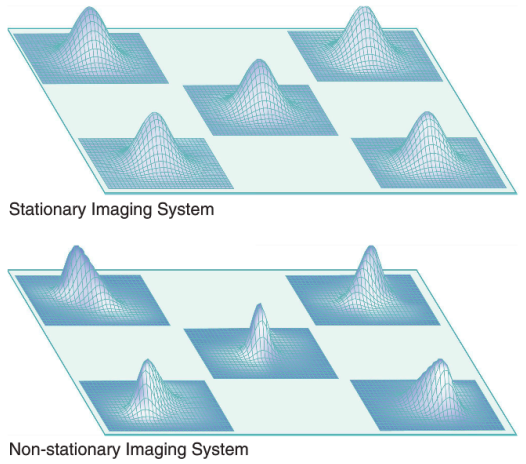
- What does it mean for a system, S, to be linear?
  - Mathematically: $S[a \cdot f(x) + b \cdot g(x)] = a \cdot S[f(x)] + b \cdot S[g(x)]$
    - (superposition principle)
  - Colloquially: If you add two (or more) signals together and then put them through the system, you'll get the same result as putting each signal through the system individually and then adding the result.



Definitions| Shift Invariance

- What does it mean for a system, S, to be shift-invariant?
  - Mathematically:
    - Let $g(x) = S[f(x)]$ (define output of system)
    - Let $f_{x_0} = f(x - x_0)$ (define shifted input)
    - -> $S[f_{x_o}] = g(x - x_0)$
  - Colloquially: an arbitrary translation to the input results in an identical translation to the output

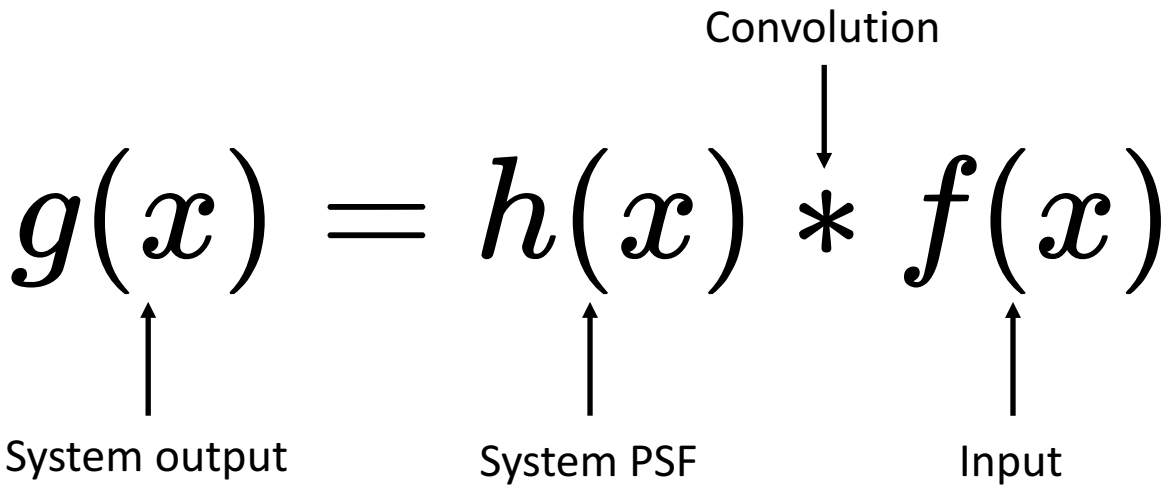




*Bushberg The Essential Physics of Medical Imaging

Definitions| LSI system

- If a system is linear and shift-invariant (LSI) then…
  - The output of the system to ANY input is given by a convolution with the system's point spread function (PSF)



Question: Are CT systems linear?

- Answer: Sometimes (ish)
  - For FBP reconstruction, a CT system usually behaves mostly linear (with some exceptions)
  - Iterative reconstructions typically make the system non-linear

Question: Are CT systems shift invariant?

- Answer: No
  - The output depends of location within the field of view of the input
  - Spatial resolution is slightly variable throughout the FOV

Is Linear Systems Theory of any use in CT?

- Yes (IMO)
- We can use concepts from linear systems theory to describe the resolution properties an a CT system
- But…
  - We need to understand its limitations and be careful about not generalizing our results too much. More on that in a few slides.

Factors that affect x-y resolution

- System optics
  - X-ray focal spot size and motion
  - Detector element size/spacing
  - Detector cross-talk + afterglow
  - Geometry (SID)
- Acquisition settings
  - Number of angular samples (i.e., “views”)
- Reconstruction
  - In-plane pixel size
    - This is determined by the reconstructed FOV and matrix size
  - **Kernel (if using FBP)**
  - **Regularization scheme (if using IR)**
    - Local noise and contrast conditions can affect local resolution in IR

x-y resolution: Iterative reconstruction

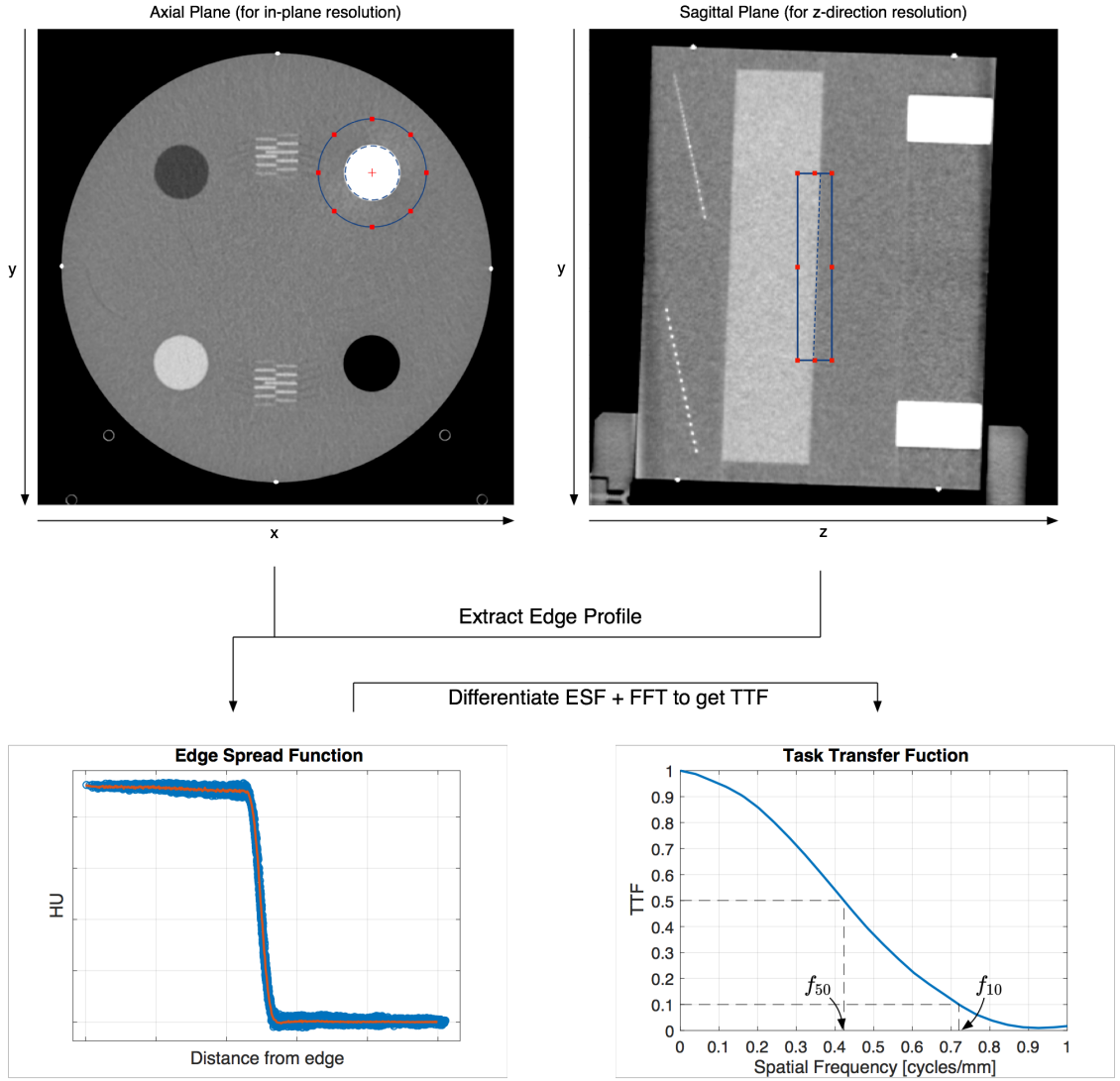
- Iterative reconstructions (IR) don't technically have a "kernel" like FBP
- IR is non-linear
  - "Regularization" schemes used that try to smartly penalize/smooth noisy data
  - From a statistical point of view, this is how a Bayesian "prior" is included into the reconstruction process. We basically assume the data should have some degree of smoothness
  - IR algorithms have parameters that let the user choose how much smoothness they want to impose (more smoothness = less noise but worse resolution)
- Colloquially speaking, IR algorithms are trying to minimize noise in what are perceived to be uniform image regions, and enhance resolution in image regions with a lot of "edges".
  - Net affect is that resolution in IR is highly complex and difficult to fully characterize.
  - High-contrast features have different resolution properties compared to low-contrast features
  - **Resolution becomes dependent on the local noise and contrast conditions of an image**

Factors that affect z-resolution

- System optics
  - X-ray focal spot size
  - Detector element size/spacing
  - Detector cross-talk + afterglow
  - Geometry (SID)
- Acquisition settings
  - Focal spot motion (i.e, double sampling)
  - Pitch
- Reconstruction
  - Slice Thickness
    - Similar affect as the in-plane pixel size
- We don’t backproject across z so we don’t really have blurring due to reconstruction processing in the z-direction
  - One exception is that IR algorithms sometimes use “cross-slice” information to help reduce in-plane noise, this has the effect of reducing z-direction resolution

Measuring spatial resolution

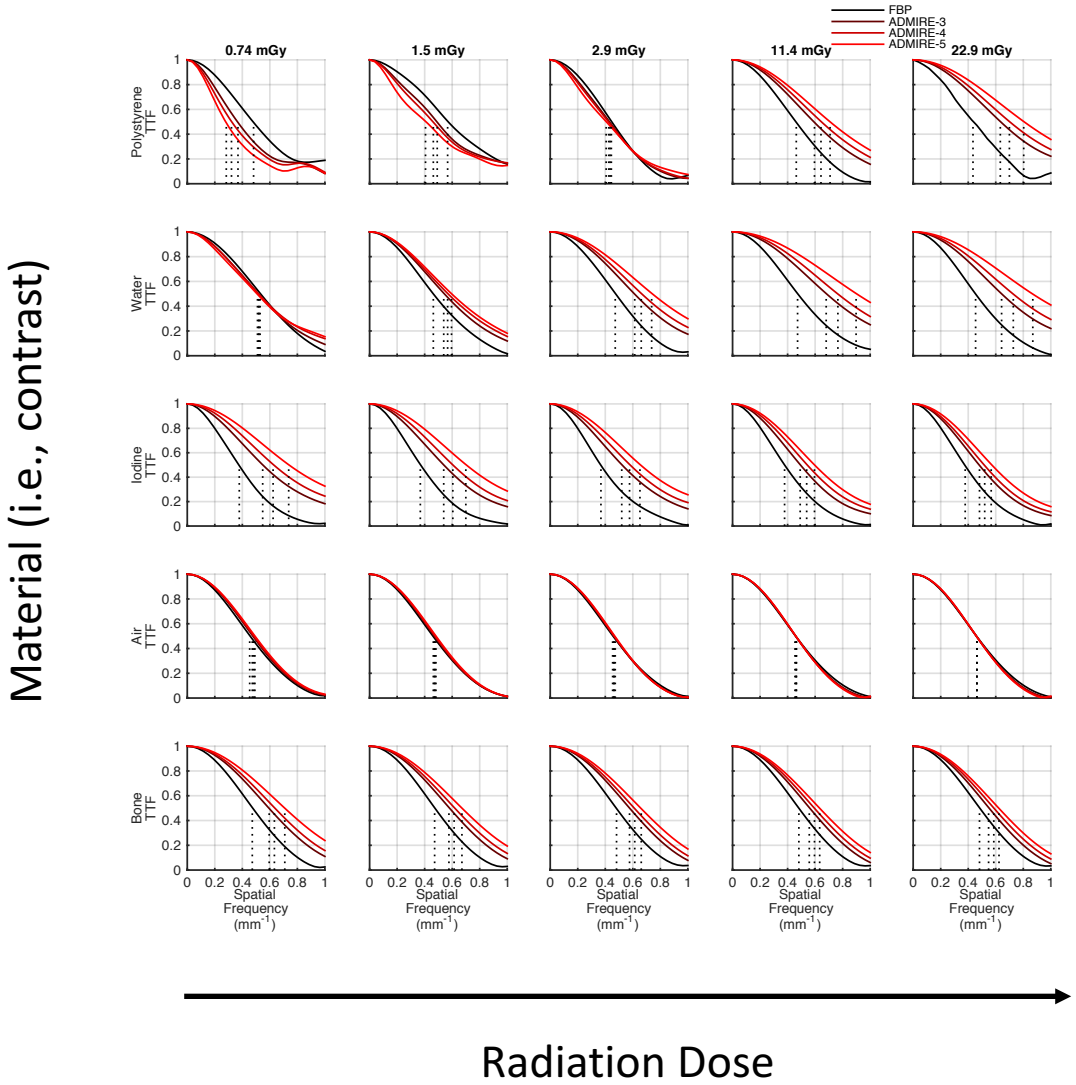
- Several potential methods
  - Line pair patterns
  - High contrast BB or wire -> PSF -> MTF
  - Edge spread function (ESF) -> PSF -> MTF



Why not just use high contrast LP patterns?

- **When IR is used, the resolution of high-contrast features is usually not the same as low-contrast features!!!**
- Also, the local noise conditions affect resolution
- LP patterns overestimate the actual resolution for low-contrast features.

*Solomon, Med Phys 2014

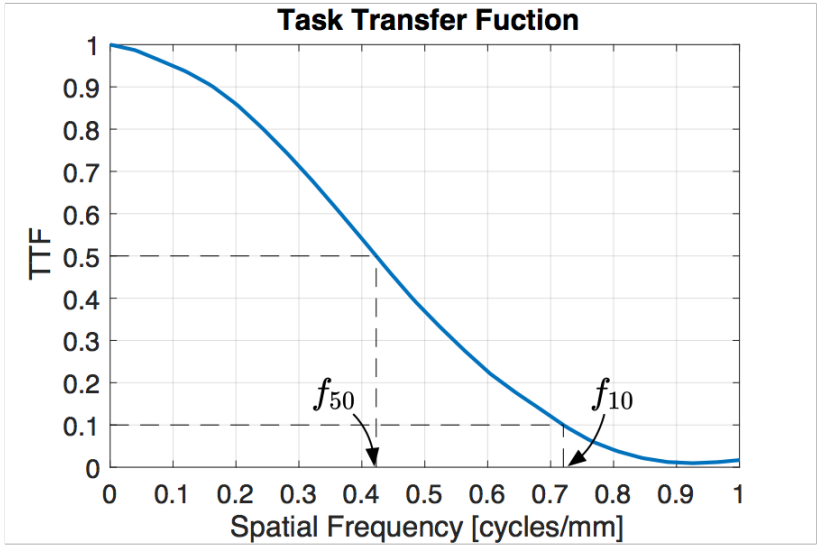


The ”task transfer function” (TTF)

- TG 233 adopts this terminology
- TTF is analogous to an MTF, but we call it a TTF to acknowledge that it really only predicts the resolution of objects of a given contrast, with a given level of image noise
- i.e., we’re acknowledging that the system is non-linear, but we’re going to use linear systems analysis anyways, but limit the generalizability of the results.

Resolution Analysis

- We recommend doing a TTF measurement for protocol(s) of interest
- Make the measurement using a cylinder with an appropriate contrast for the protocol being assessed
- Report discusses how much image data you need to get a reliable measurement
- TTF can be summarized by the 50% and/or 10% frequency
- Software is being provided along with the report to help make these measurements (more on that later).



Operational performance

- This is the bulk of the document
- Focus is on clinical performance
- Towards protocol optimization, operational consistency
- Section includes
  - Tube current modulation
  - Spatial resolution
    - For potentially non-linear CT systems
    - In-plane and z-direction
  - Noise
    - Noise power spectrum analysis
    - Non uniformity in noise
  - Task-based image quality
    - Task-based IQ from Fourier domain calculations
    - Task-based IQ from spatial domain calculations
    - Task-based IQ from human reader experiments

What do I mean by noise?

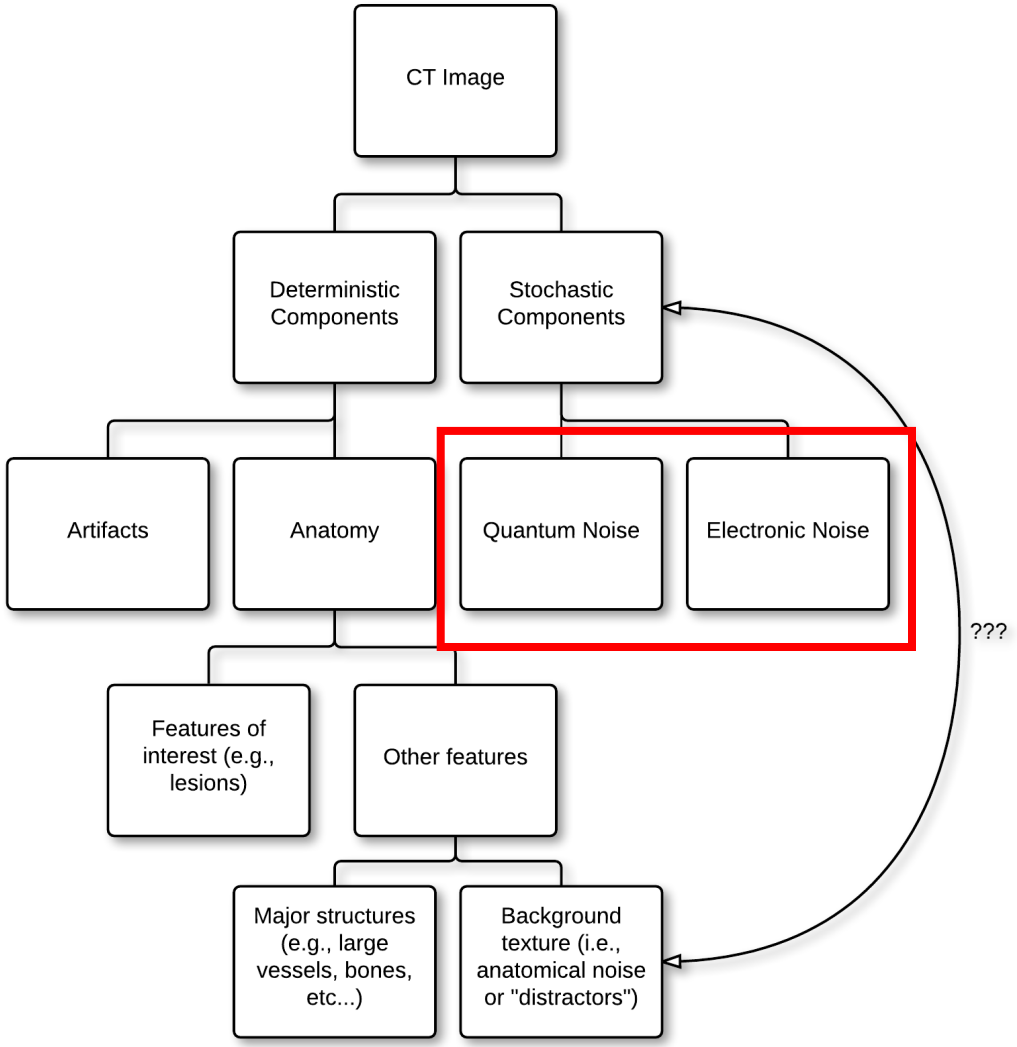


Image = Deterministic + Stochastic

$I(x,y,z) = S(x,y,z) + N(x,y,z)$

Noise is characterized by computing statistics of the random components of the image

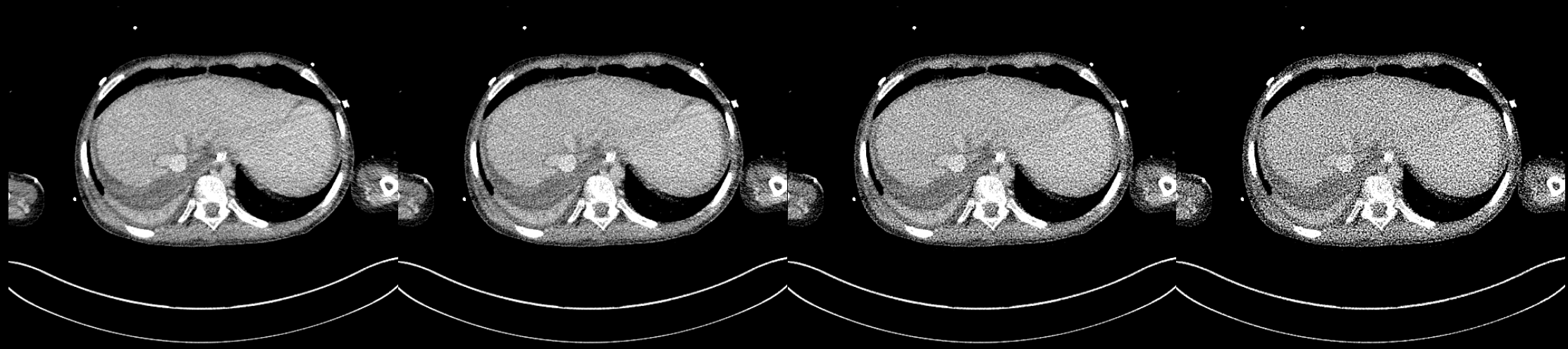
CT noise depends on

- Dose
  - More dose -> more signal at the detectors -> less noise
  - Any technical scan factors that affect dose will also affect noise
- Detector efficiency
  - More efficient detector -> less noise
- Reconstruction
  - Kernel
    - Sharper kernel -> more noise (but better resolution)
  - Slice thickness
    - Using a thicker slice is like averaging signals across detector rows. More signals averaged together is like having more photons.
    - Thicker slices -> more photons per image -> less noise (but worse z-direction resolution)
  - Iterative reconstruction
    - Typically IR reduces noise compared to FBP for the same dose (but can also cause a marginal losses in resolution, especially for low-contrast features)
- Patient size
  - For the same scan technique (mAs, kV, pitch), fewer photons reach the detector for larger patients which results in greater image noise
    - Note that in practice, AEC methods such as tube current modulation are designed to deliver higher dose for larger patients with the goal of consistent noise across patient sizes

Noise has ***magnitude*** and texture

- Think of noise as a spatial dependent random field (i.e., a collection of random numbers), N(x,y,z)
- Noise magnitude is quantified by the standard deviation of N, $\sigma(x, y, z)$.

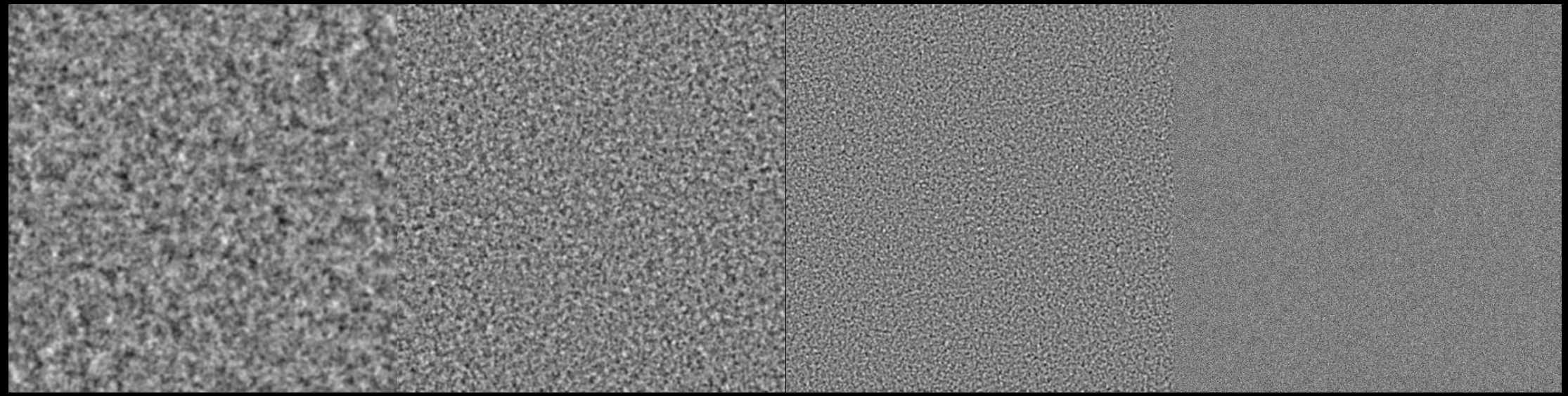
Increasing noise magnitude (i.e., decreasing dose)



Noise has magnitude and ***texture***

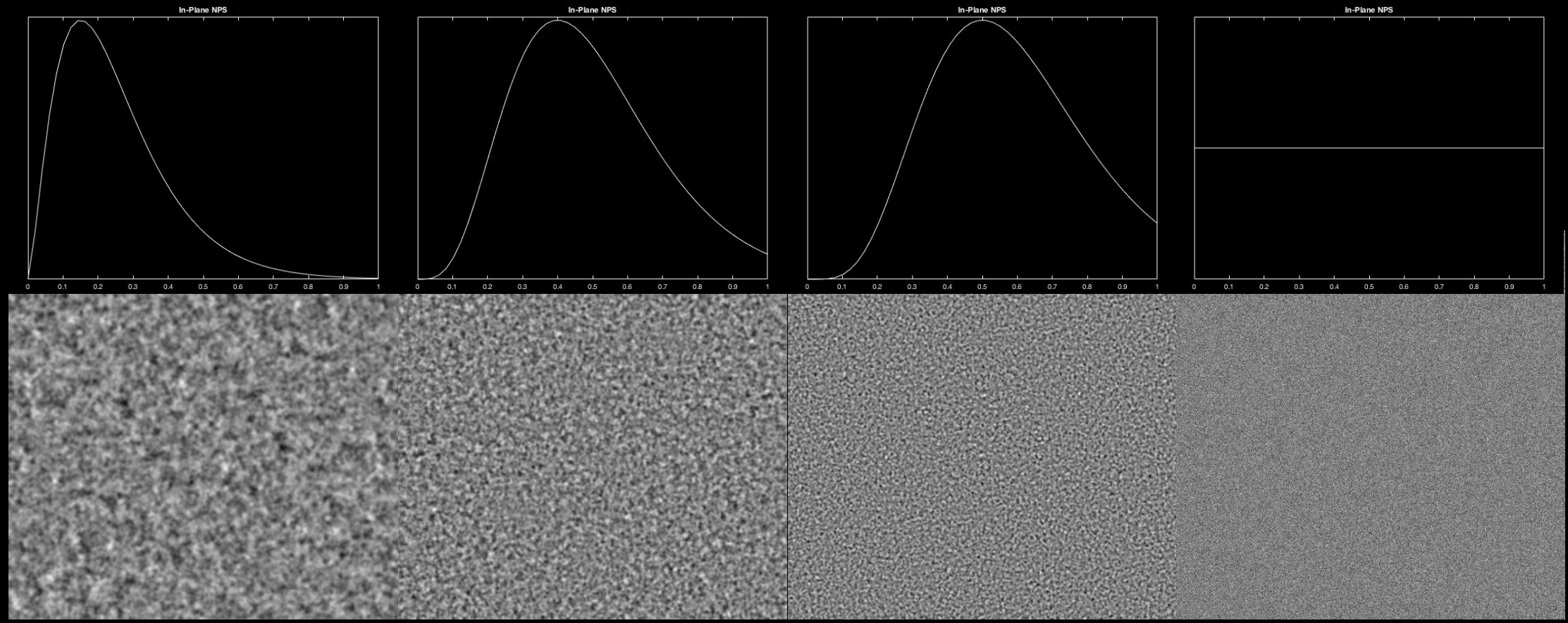
- Noise texture is desribed by the correlations between noise in different voxels. $E[N(x_1, y_1, z_1)N(x_2, y_2, z_2)]$
- Describes how neighboring pixels tend to fluctuate with each other

These all have the same noise magnitude!



Noise magnitude and texture quantified by NPS

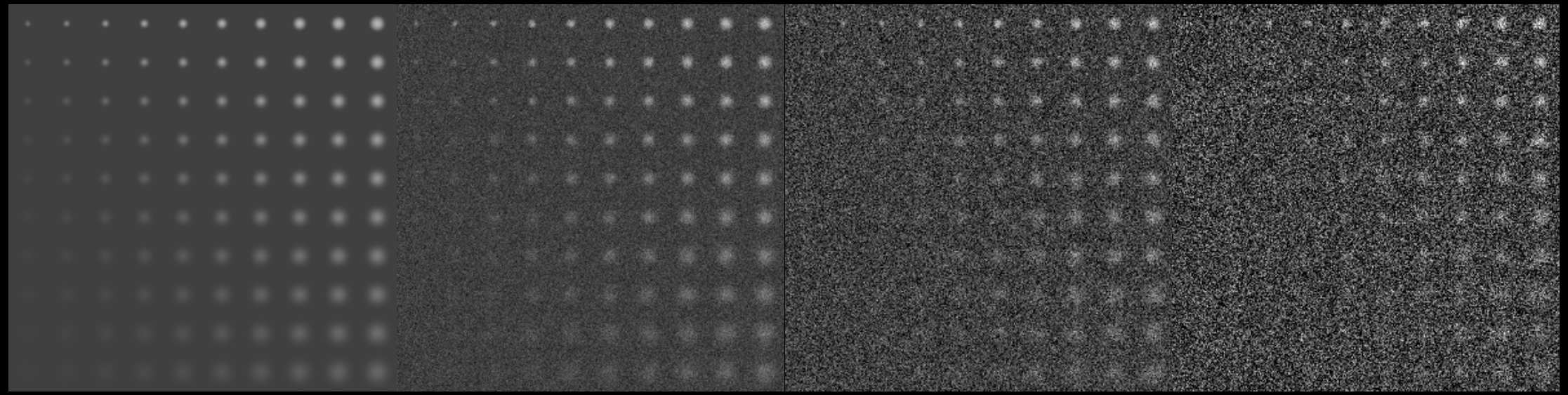
- NPS = Noise Power Spectrum
- Describes noise correlations in frequency domain



More noise is bad!

- Detection decreases as noise increases
- Confidence decreases as noise increases

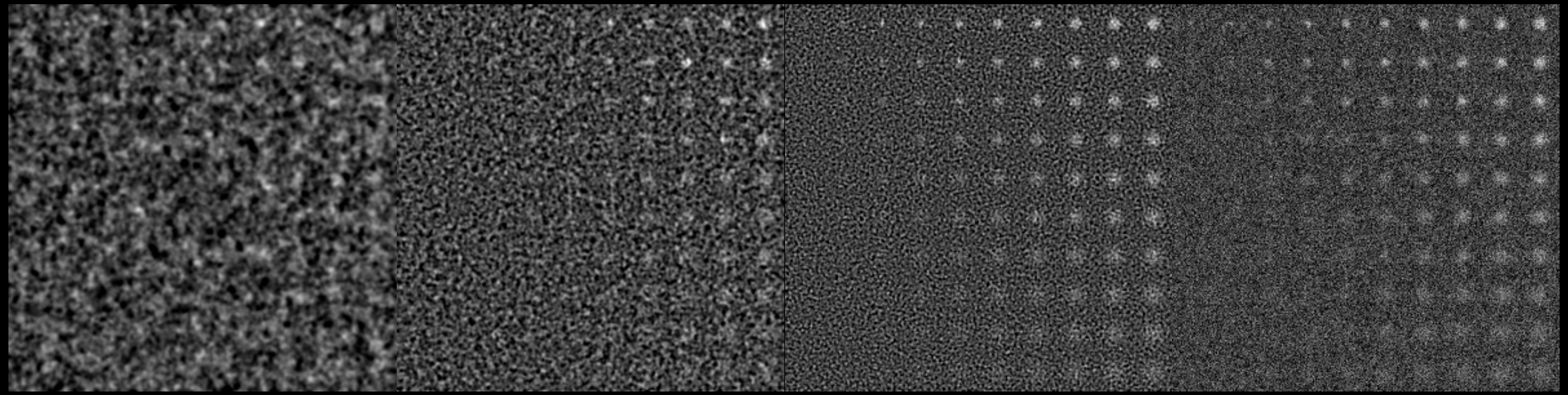
Increasing noise magnitude



Noise texture affects detectability!

- If frequency content of feature to be detected ”overlaps” a lot with the NPS, it will be difficult to detect.

Same noise magnitude, different texture

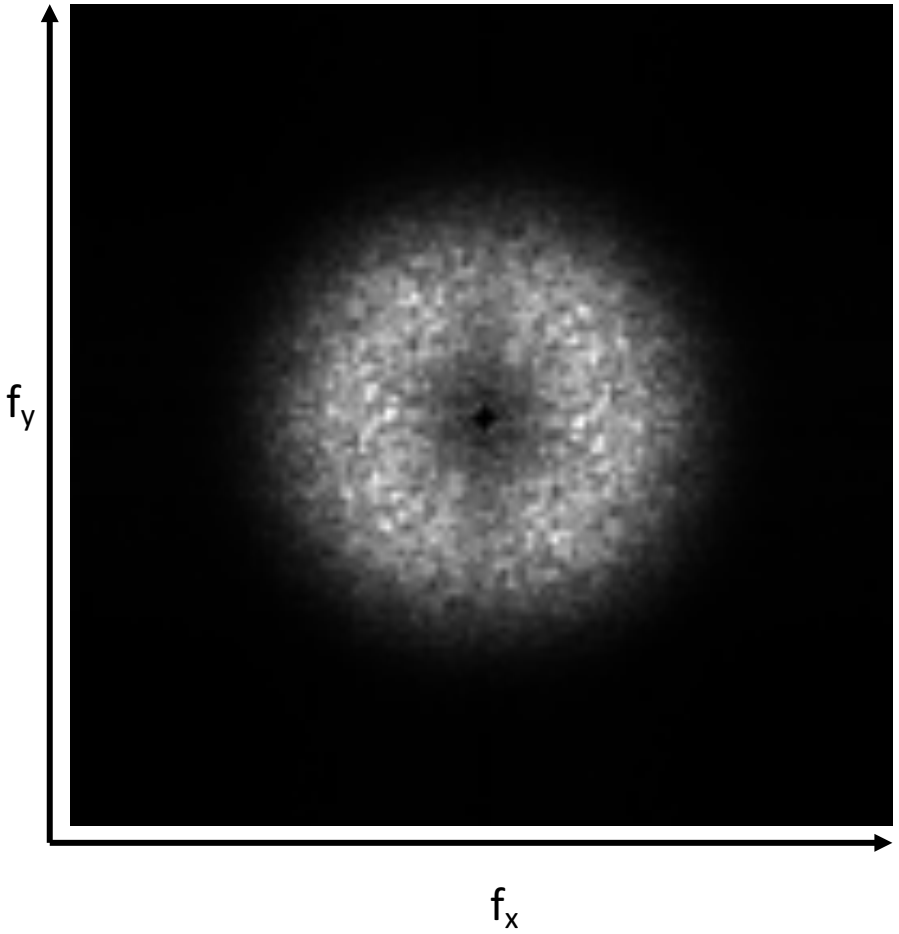


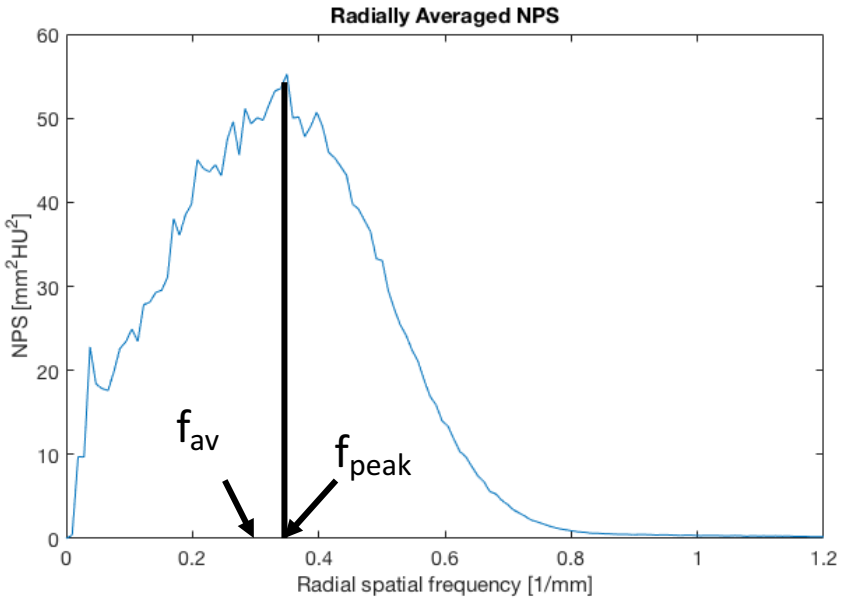
What is the noise power spectrum (NPS)?

- NPS is the Fourier transform of the noise auto-correlation
  - $NPS(f_x,f_y,f_z) = F[A(\Delta x,\Delta y,\Delta z)]$
- Describes noise autocorrelations in spatial frequency space
  - i.e., how much noise power is contained by each spatial frequency
- Integral of the NPS is equal to the pixel variance
  - Thus the NPS characterizes both the magnitude and texture of noise!

Anatomy of an NPS

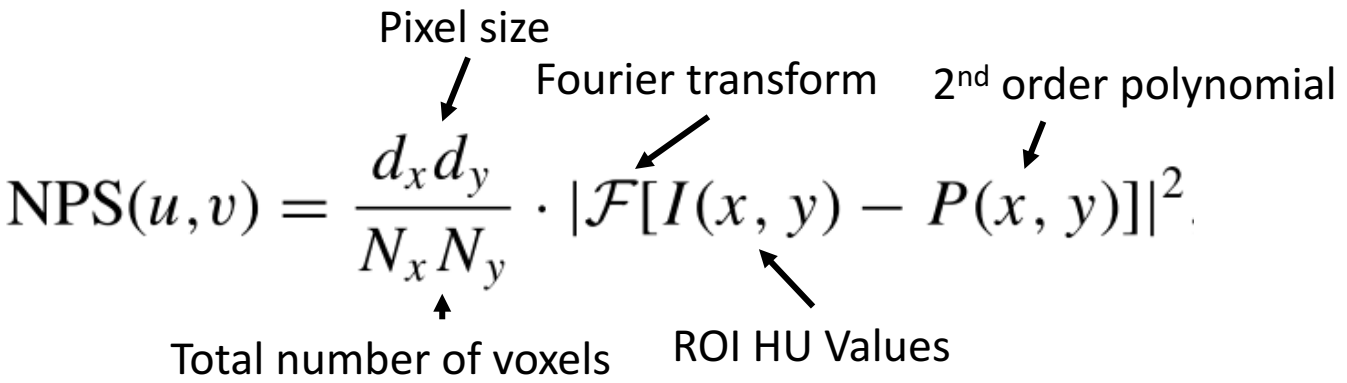
2D NPS





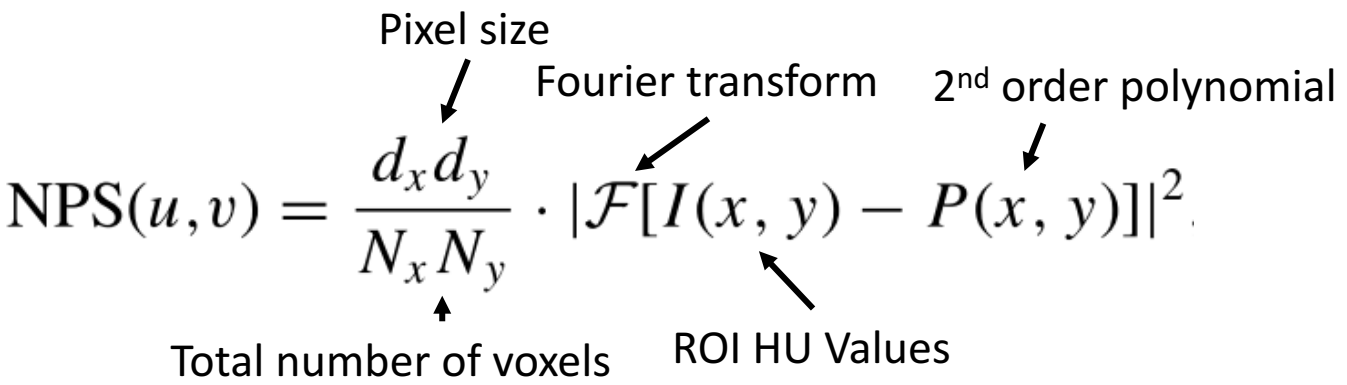
Measuring NPS| Phantom

- Place a number ROIs in uniform region of the phantom (can use multiple slices if the phantom size doesn't change across slices).
- For each ROI, subtract out a 2nd order polynomial, P(x,y) to remove DC component and low-frequency structured noise.
- Compute squared magnitude of Fourier Transform of subtracted ROI.
- Normalize by ratio of voxel size (area if 2D, volume if 3D) to total number of pixels



Measuring NPS| Phantom

- Do this procedure for each ROI and take the average NPS across all ROIs
- Can compute noise magnitude by either integrating the NPS or by computing voxel STD (average across all ROIs)
- Compute summary statistics of NPS
  - f_{av} and f_{peak}
- Tools will be provided by TG to help with these measurements

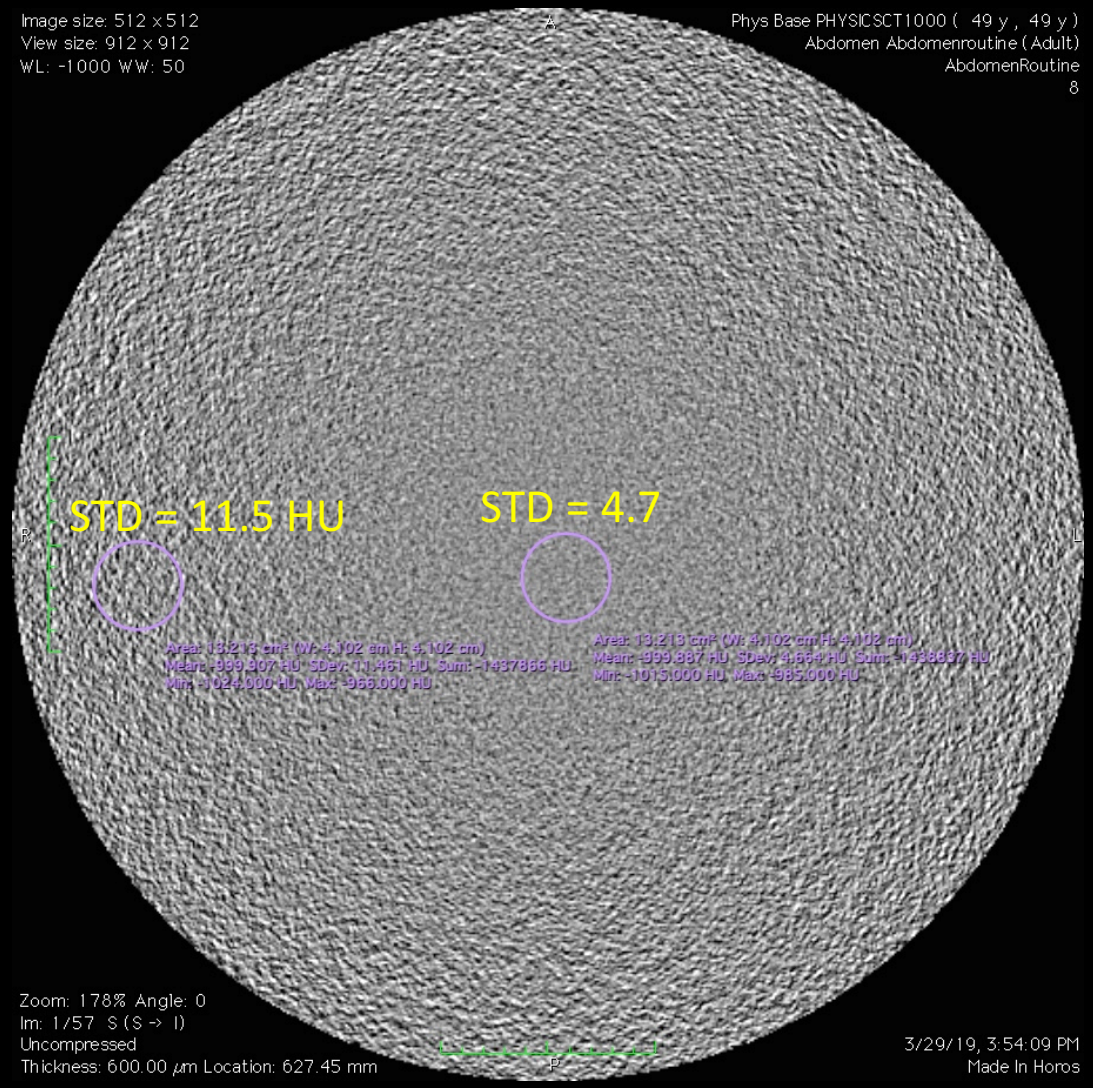


Definitions | Stationarity

- Wide-sense stationarity (WSS):
  - A statistical condition describing the properties of the noise
  - A random process (e.g N(x)) is considered to be wide sense stationary if
    - Its first order statistics (e.g. mean, STD) are constant (same for all x,y,z)
    - Its second order statistics (e.g,. Autocorrelation) depend only on the distance between points, not their absolute positions: $A_N(x_1,x_2) = A_N(\Delta x)$
- By definition, the NPS is the Fourier transform of the noise auto-correlation function, assuming wide sense stationarity
  - $NPS(f_x,f_y,f_z) = F[A(\Delta x,\Delta y,\Delta z)]$
- Thus NPS analysis only makes sense if the noise is WSS

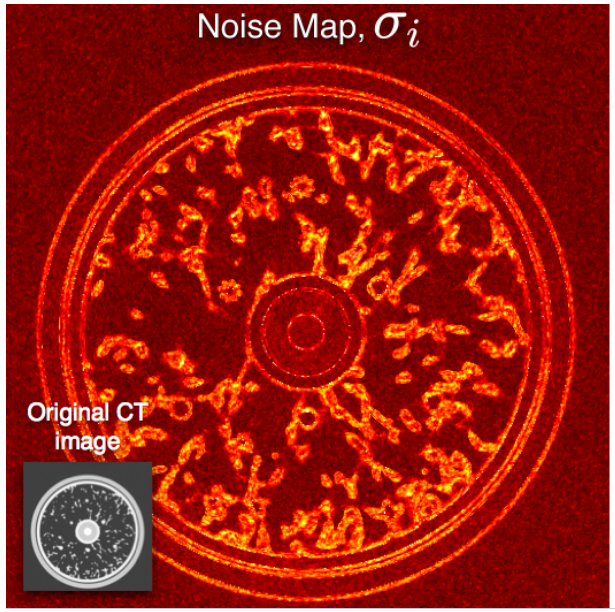
Question: Is CT noise wide-sense stationary?

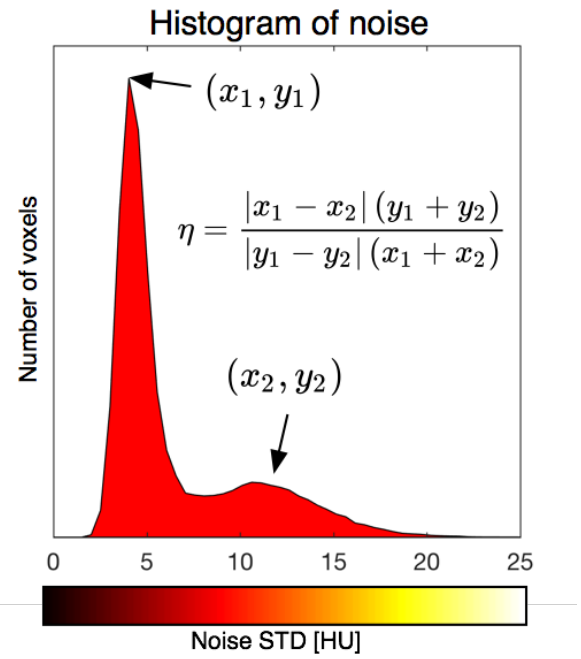
- Answer: Globally, no
  - Both noise magnitude and texture change throughout the FOV
- Locally, yes (usually)
  - If we limit measurements to small local ROIs, the analysis provides meaningful characterization of the noise



Noise non-uniformity

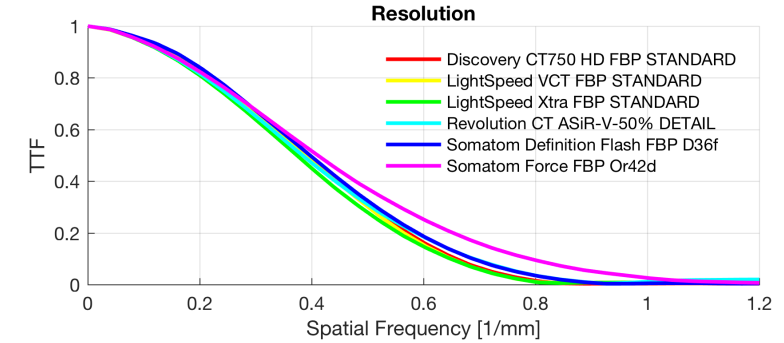
- Iterative reconstructions have unique and potentially highly non uniform noise properties
- Report describes some methods to characterize both the global non uniformity of noise and the potential local non-uniformity for IR reconstructions

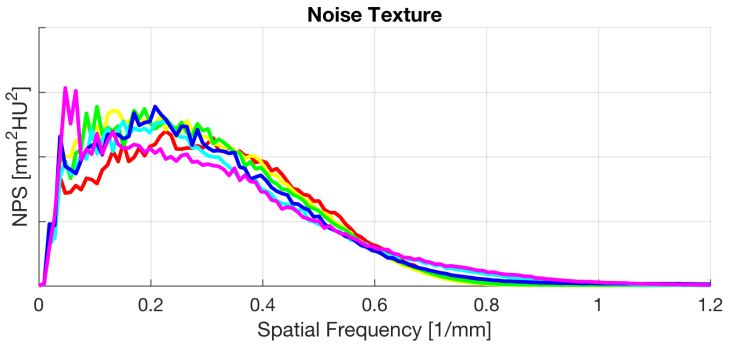


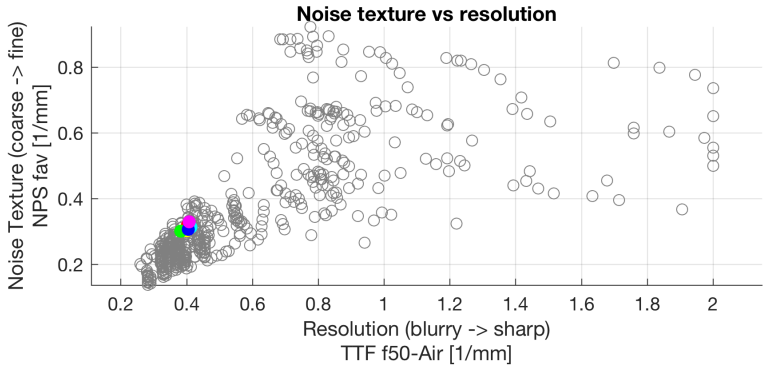


Example of clinical utility| IQ Matching

- We have many different makes/models in our health system
- We want to achieve images with consistent noise and resolution properties
- To figure out which recon settings to use, we measured TTFs and NPSs for a bunch of recon settings
- Found the settings that had similar noise texture and resolution across scanner models

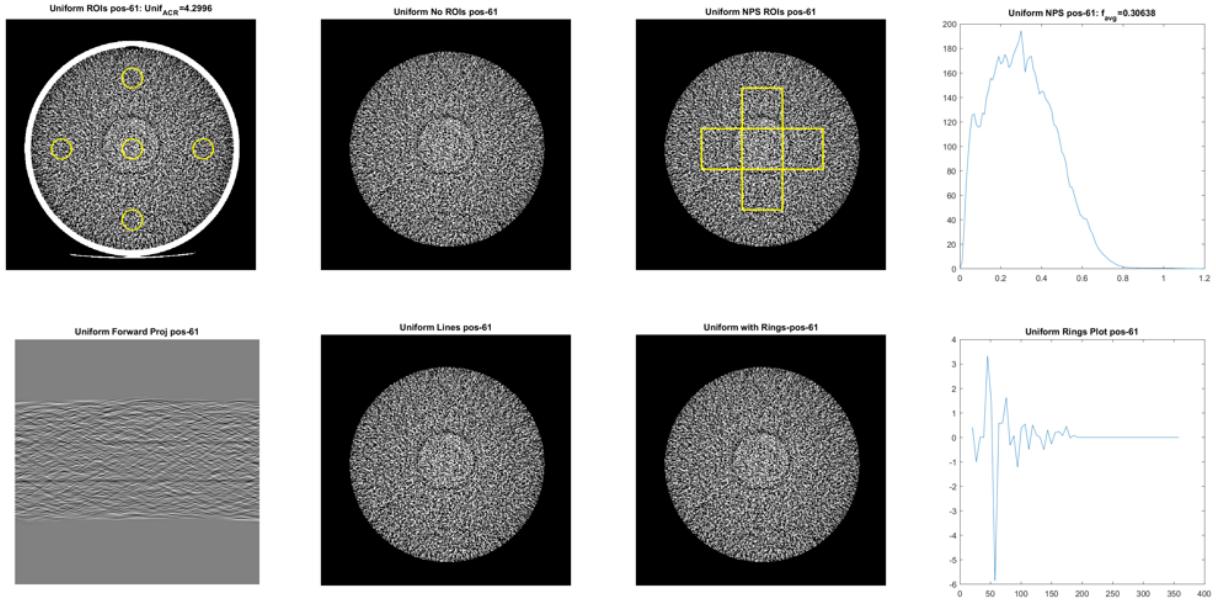






Example of clinical utility| Daily QC Analysis

- We have daily water phantom QC images sent to a dedicated server for automated analysis
- The system detects artifacts using NPS analysis
- Sends me an email if it thinks there's an artifact so I can investigate



Operational performance

- This is the bulk of the document
- Focus is on clinical performance
- Towards protocol optimization, operational consistency
- Section includes
  - Tube current modulation
  - Spatial resolution
    - For potentially non-linear CT systems
    - In-plane and z-direction
  - Noise
    - Noise power spectrum analysis
    - Non uniformity in noise
  - Task-based image quality
    - Task-based IQ from Fourier domain calculations
    - Task-based IQ from spatial domain calculations
    - Task-based IQ from human reader experiments

Definitions| Task-based image quality

"Any general definition of **image quality** must address the **effectiveness** with which the image can be used for its **intended task**."

-ICRU Report 54

Definitions| Task-based image quality

Therefore, a task-based IQ metric tries to quantify how well the images can be used for some medical purpose

Traditional vs task-based IQ

How does it taste when you put them all together?

Ingredients

Task-based IQ Metrics	Traditional IQ Metrics
• Detection accuracy	• Noise standard deviation
• Sensitivity	• Noise power spectrum
• Specificity	• Resolution
• Estimation accuracy	• Modulation transfer function
• Classification accuracy	• Contrast
• Detectability index	• Contrast-to-noise ratio

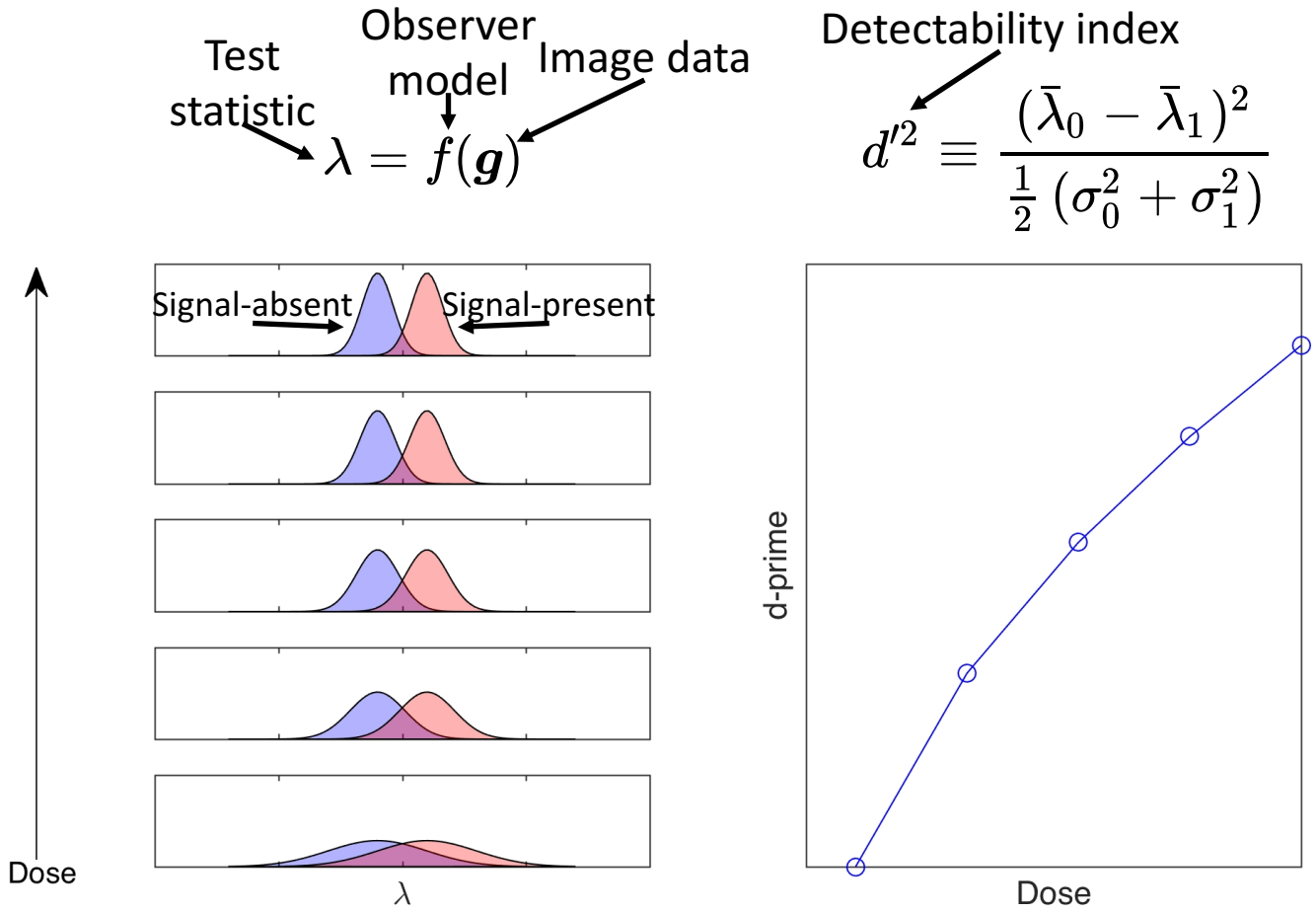
What is needed for task-based IQ?

- 3 things
  1. A well defined imaging task
     - Detection of a subtle lesion
     - Classification
     - Estimation
  2. A “reader”
     - Could be a real human reader
     - Or could be a mathematical model
  3. A way to estimate the ability of the reader to perform the task on the images in question
     - Could be done with a reader study (show radiologists images with known lesions and see how well they can detect them)
     - If using a mathematical observer, performance can be estimated in many ways!

Mathematical observer models

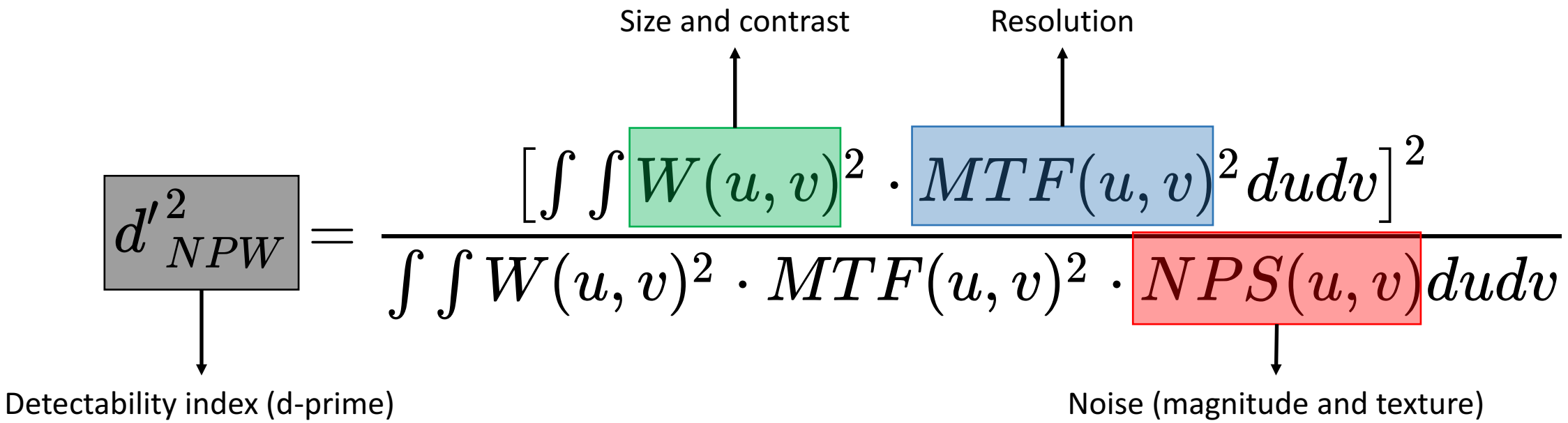
- Using radiologists for reader studies is cumbersome
- Mathematical observer models can act as a surrogate
- Based on mathematics of statistical signal detection theory
- An observer model is just a “processor” for which you input an image, and it computes a scalar statistics related to the imaging task at hand
  - E.g. how likely is a lesion present in the image
- If the images are of good quality, the observer model should be able to perform the imaging task well (and vice versa)
- We can choose observer models that try to mimic how humans perceive images and make decisions
- We use the performance of the observer model as a metric of image quality

A detectability index



A detectability index is a scalar statistic related to how likely one would expect a given observer (human or mathematical) to detect a subtle signal in a noisy background.

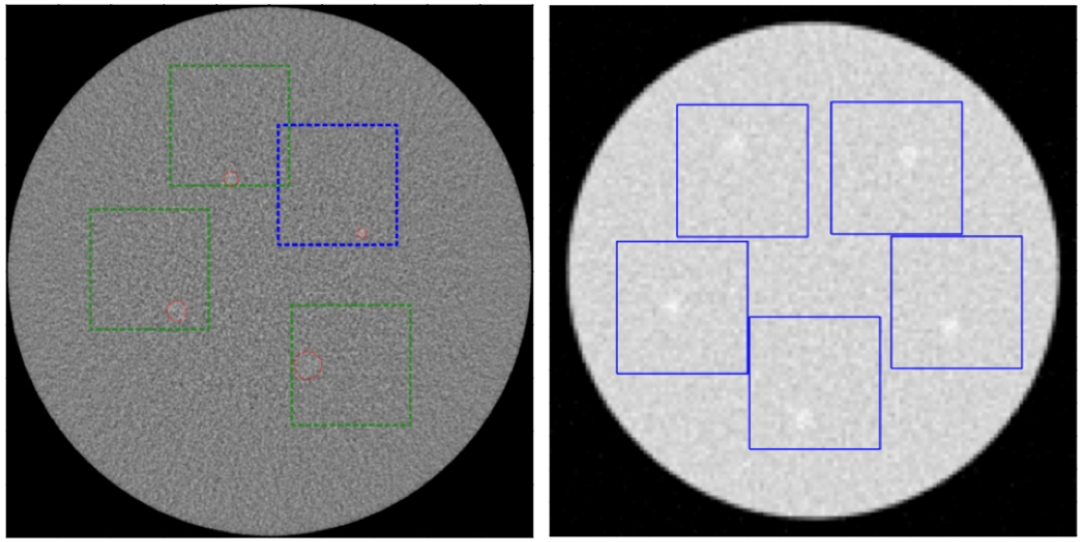
Relating traditional IQ and task-based IQ?



- Turns out for some observer models, the detectability index can be predicted directly if you know the noise and resolution properties of the images in question
- This example is based on a non-pre-whitening matched filter observer model
- This model has been shown to correlate will with human detection performance
- We can use the noise/resolution measurements from before and calculate a d'!
- Can think of it like a CNR measurement which accounts for resolution, noise texture, and the detection task of interest

Human reader studies

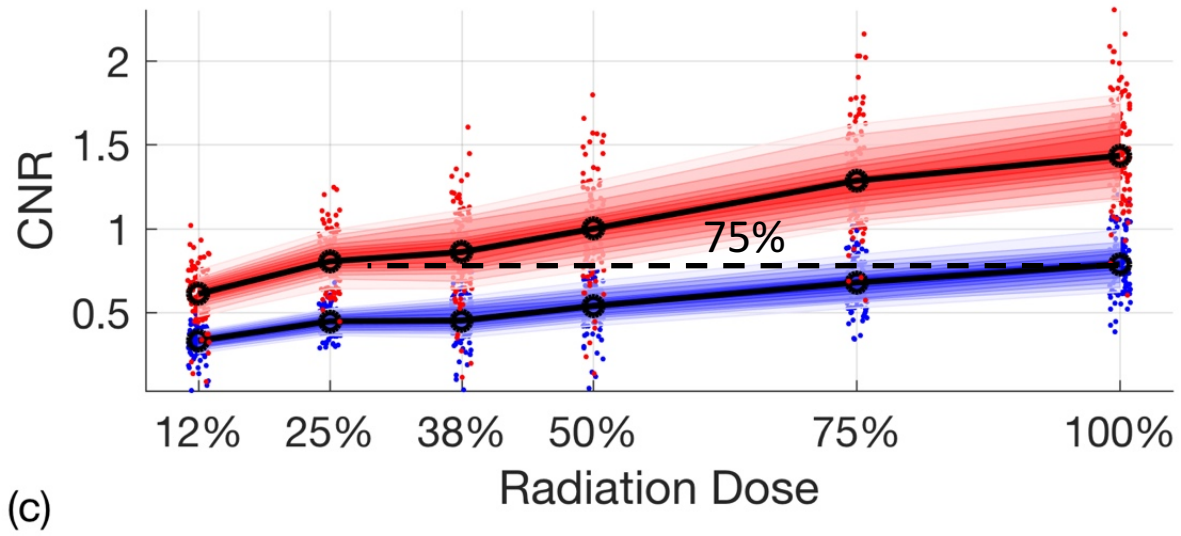
- Sometimes the models are too simplistic or require assumptions that are not true about the images
- In those cases it may be necessary to show the images to humans and have them perform a task
- Report discusses how to do such experiments and how to analyze the results
- Based on a large amount of academic work and on methods being used at the FDA



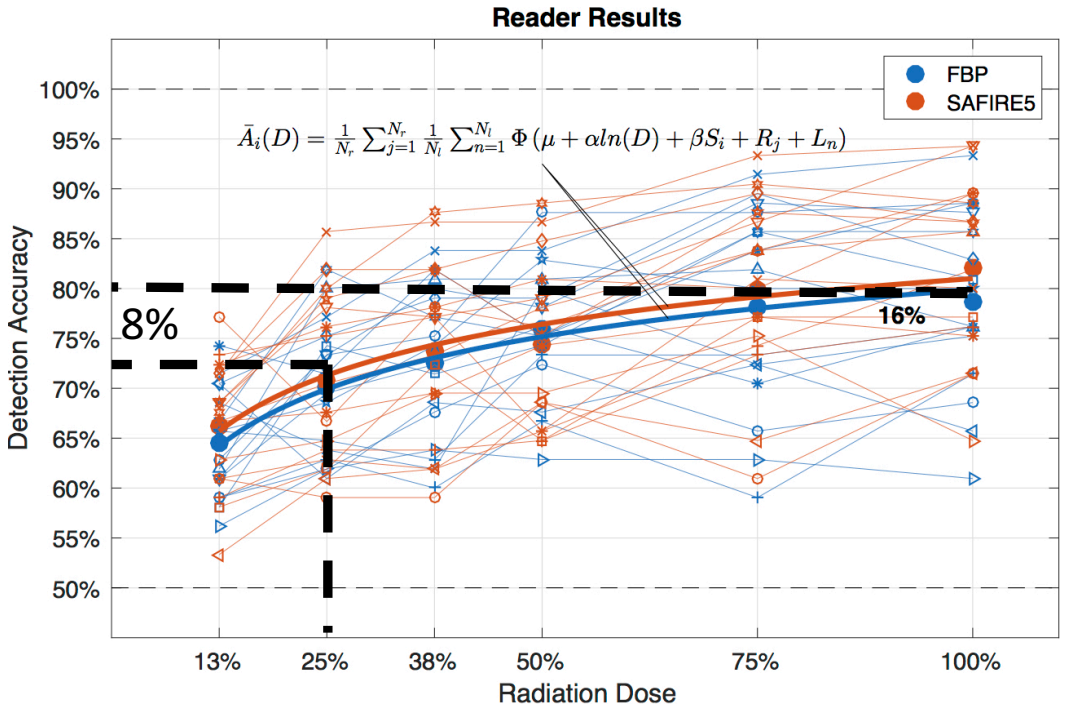
An example of why task-based IQ is helpful

If you trusted CNR and reduced dose by 75%, detectability would suffer by about 8%

Dose reduction based on CNR = 75%



Dose reduction based on human detectability = 16%

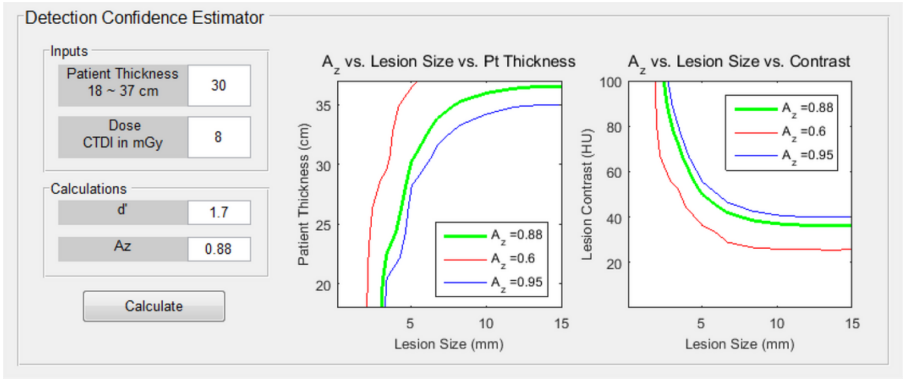


*Solomon, RSNA 2016

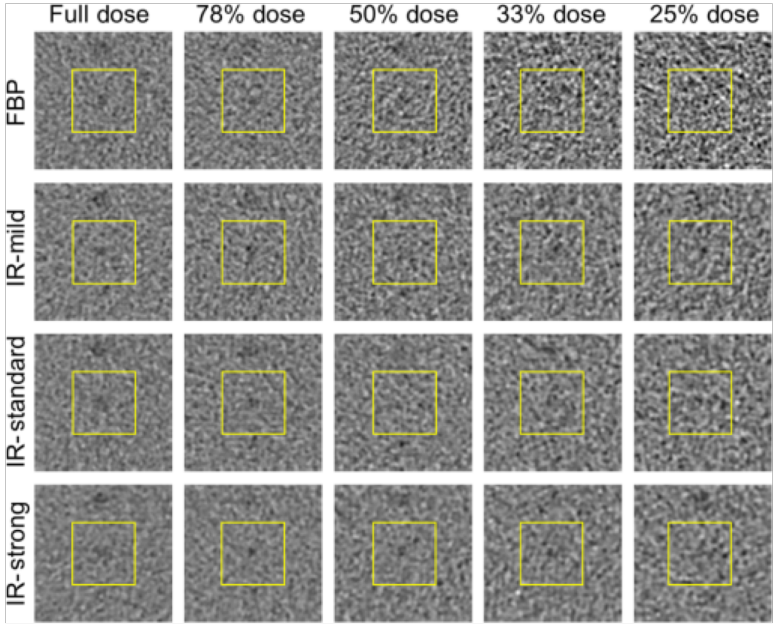
Example of clinical utility| Daily QC Analysis

- Zhang et al (Med Phys 2017) built an interface to help balance dose and image quality in terms of a detectability index based on measurements from a multi-sized phantom

- Favazza et al (JMI 2017) used task-based IQ metrics to determine dose reduction potential of a new IR algorithm



(a)



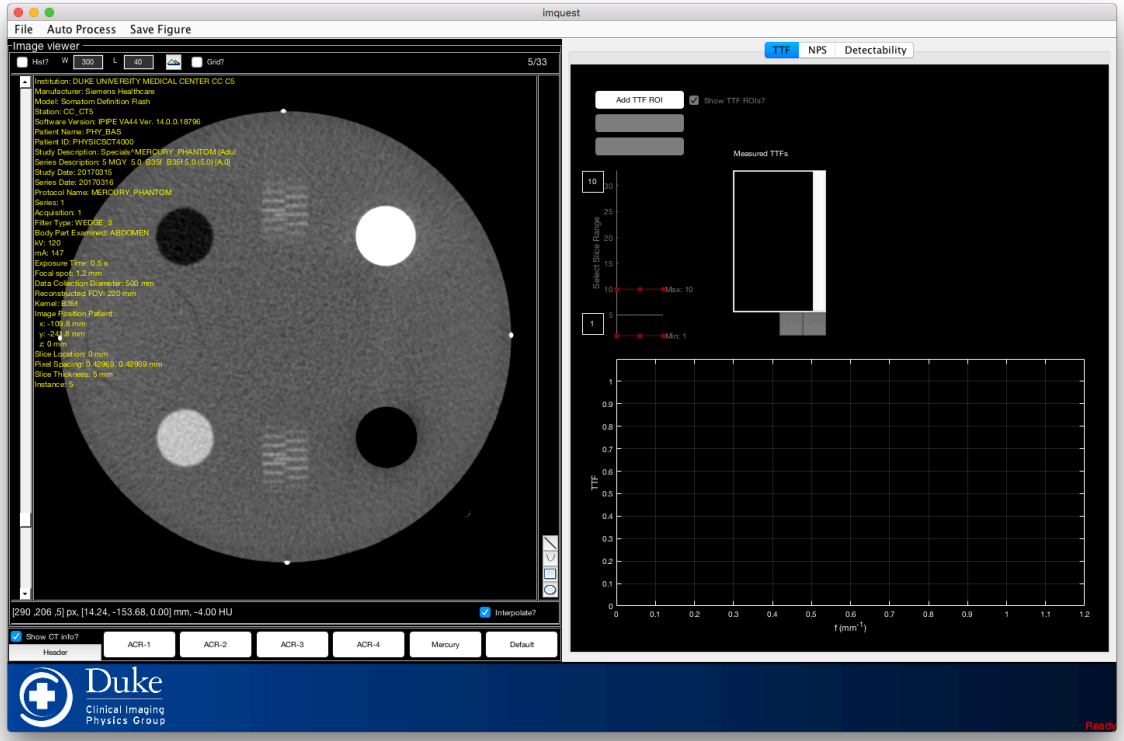
(b)

Software provided by TG

- imQuest
  - Matlab-based CT image analysis tool used to help make a number of the TG 233 measurements
  - Currently in beta form. Standalone version available
  - Source code will be available soon (waiting for Duke lawyers to clear some disclosure language)
- iMRMC
  - Statistical analysis Matlab code provided by the FDA for spatial domain task-based performance assessment. The tool helps size and analyze multi-reader multi-case (MRMC) reader studies.
- IQModelo
  - Matlab code provided by the FDA for parametric statistical methods for ROC performance analysis of linear model observers.

imQuest implementation

- GUI built in Matlab (distributed as stand-alone program)
- Utilizes a library of routines for task-based IQ analysis



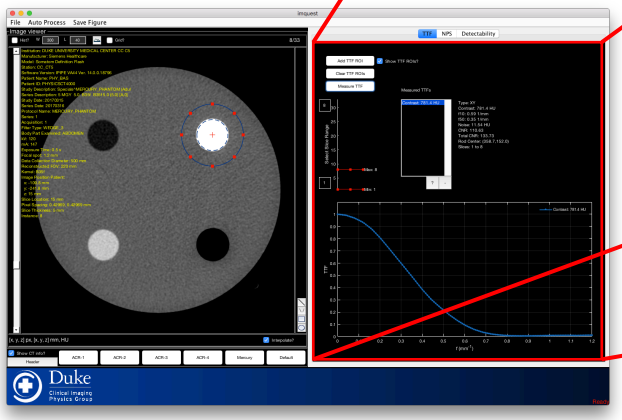
- Maidment et al, Med Phys, 2003
- Boedeker et al, Phys Med Biol, 2007
- Wilson et al, Med Phys, 2013
- Chen et al, Med Phys, 2014
- Solomon et al, Med Phys, 2015

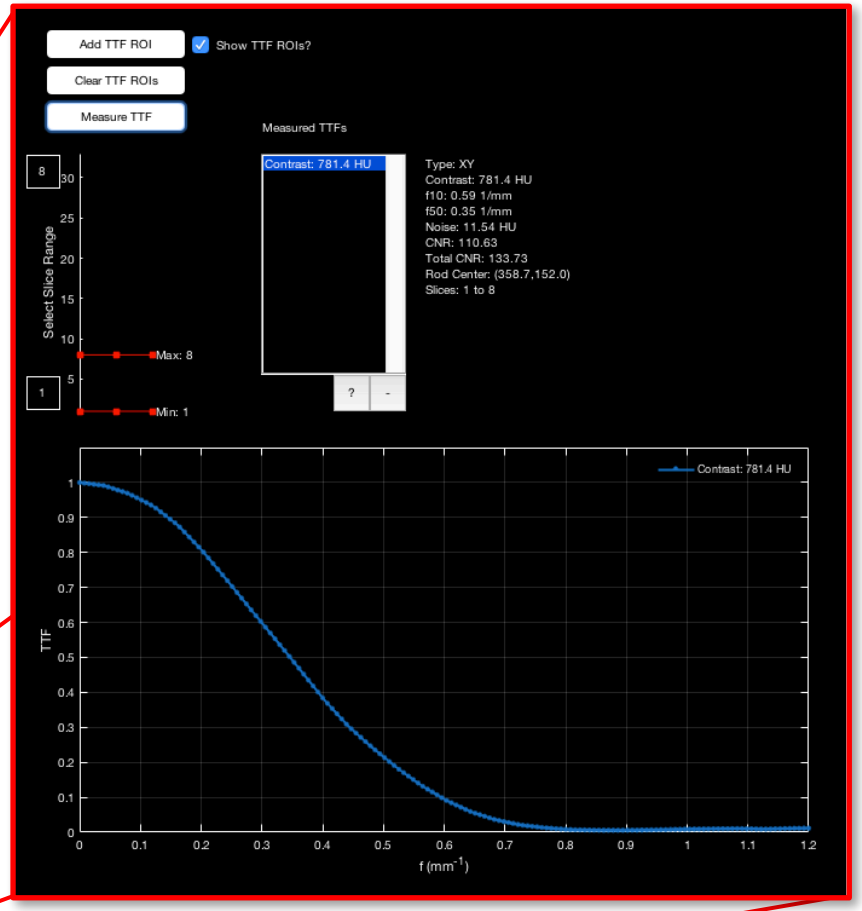
d’ is easy as 1, 2, 3

1. Measure resolution
2. Measure noise
3. Compute d’

d’ is easy as 1, 2, 3

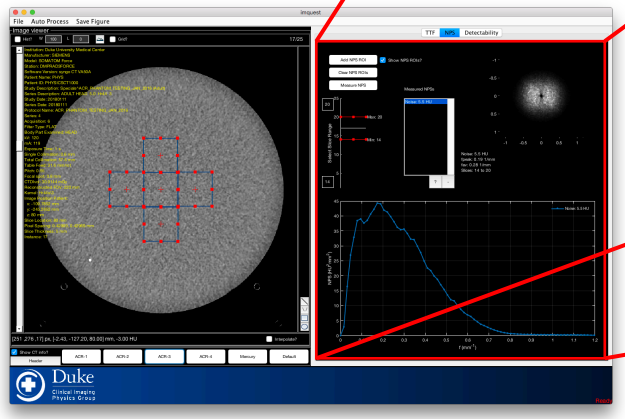
1. Measure resolution
2. Measure noise
3. Compute d’

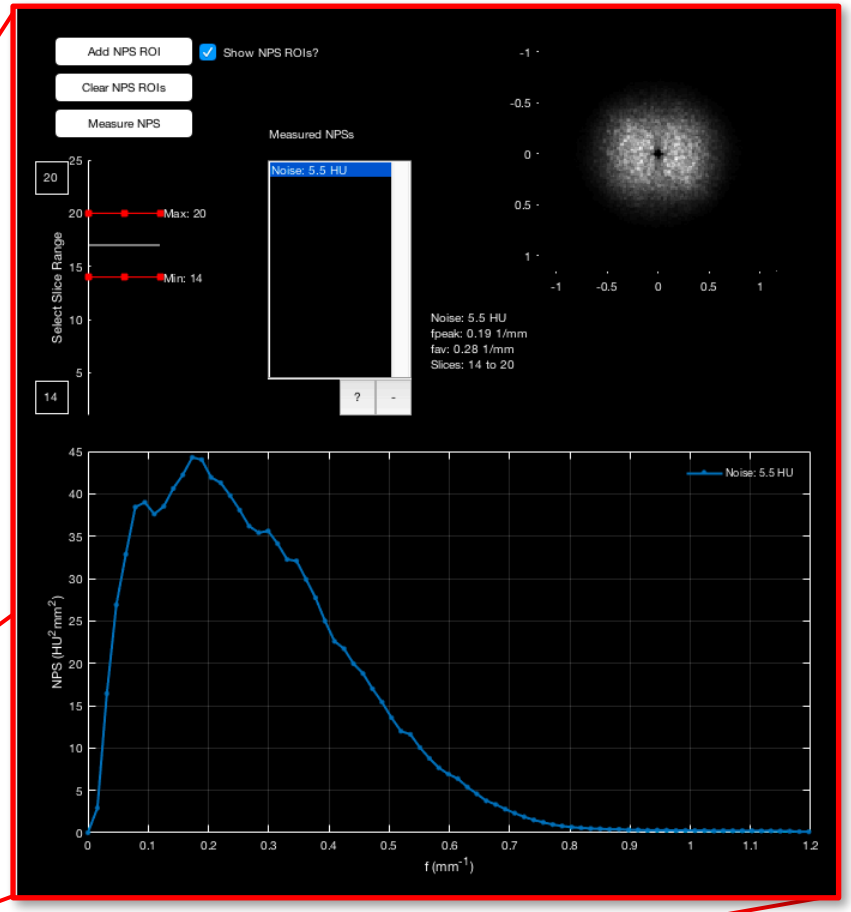




d’ is easy as 1, 2, 3

1. Measure resolution
2. Measure noise
3. Compute d’

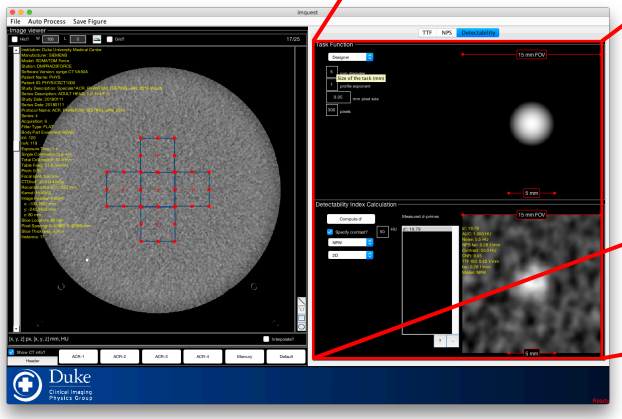


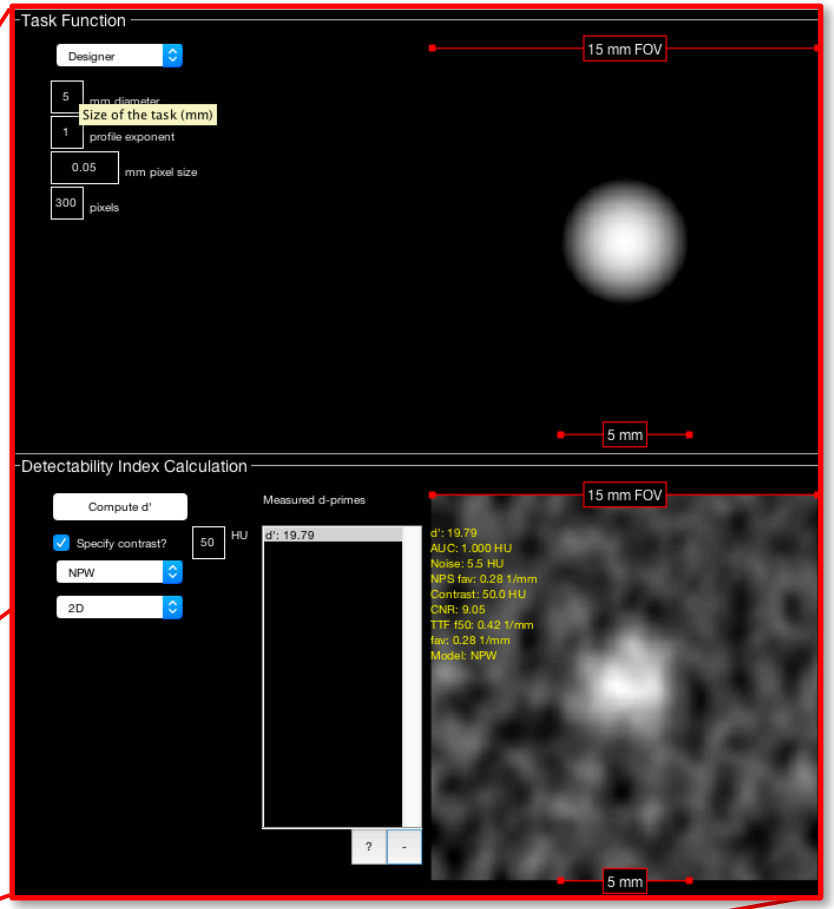


d’ is easy as 1, 2, 3

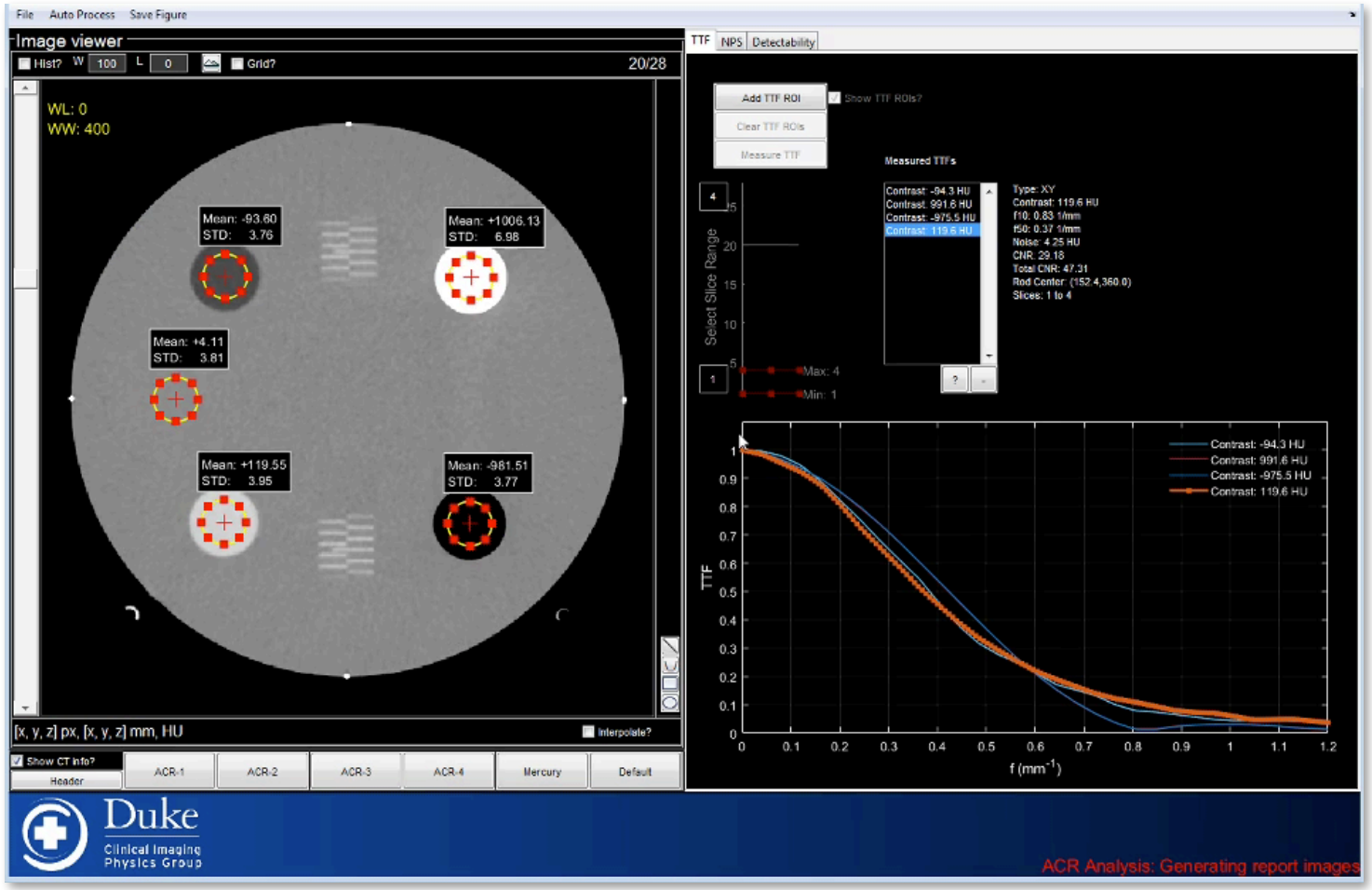
$$d'^2_{NPW} = \frac{\left[\int\int W(u,v)^2 \cdot TTF(u,v)^2 dudv\right]^2}{\int\int W(u,v)^2 \cdot TTF(u,v)^2 \cdot NPS(u,v) dudv}$$

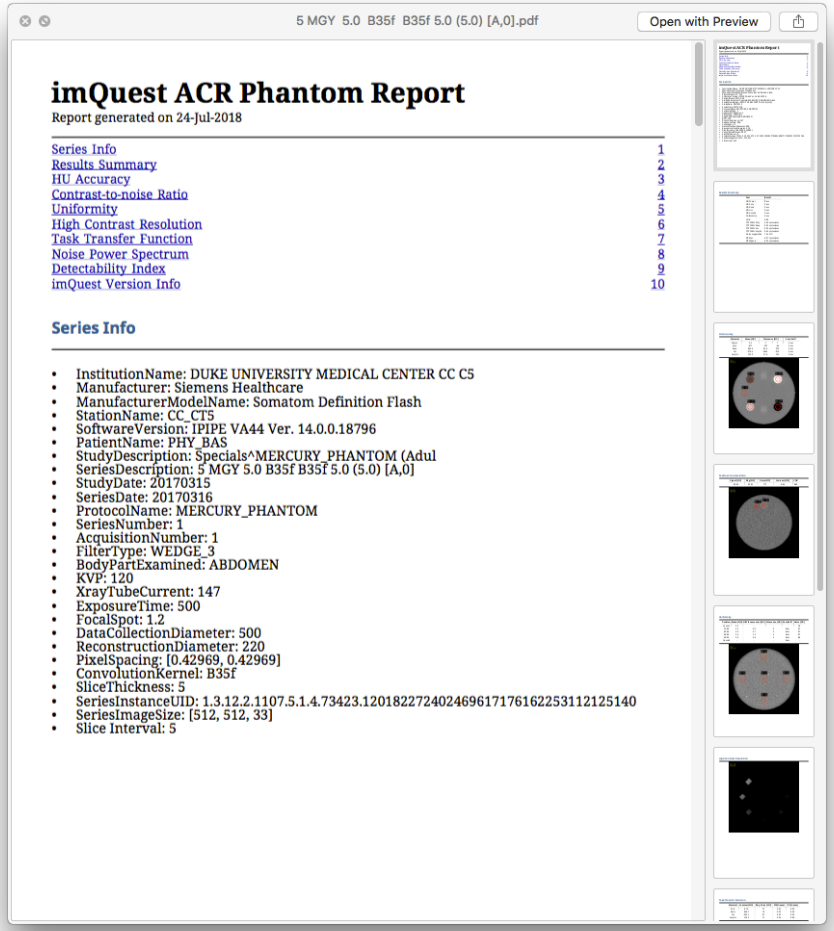
1. Measure resolution
2. Measure noise
3. Compute d’





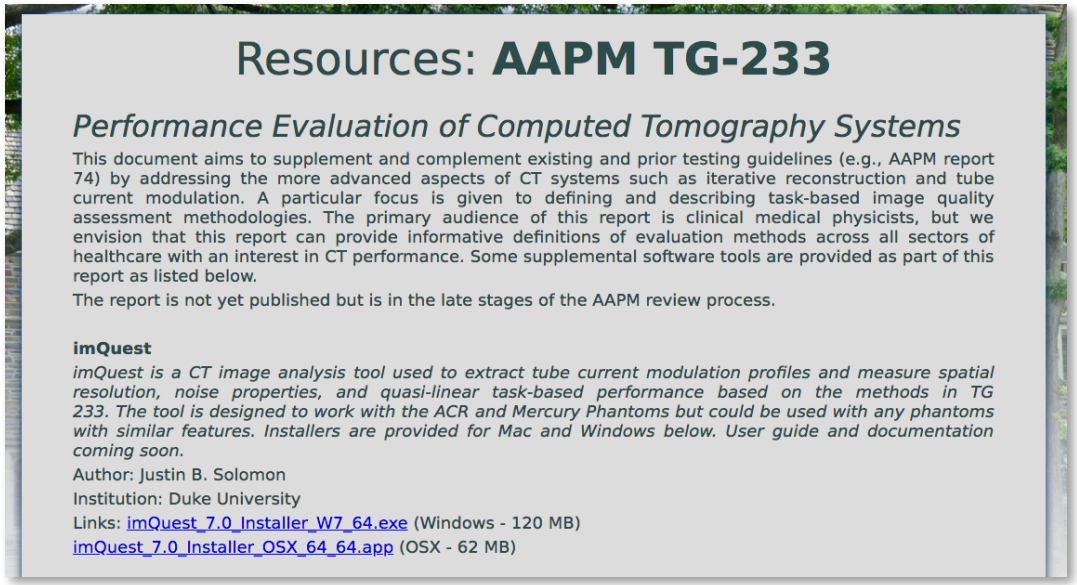
Automated analysis





Where to get imQuest

- Executable available for download here:
  - http://www.railabs.org/~samei/tg233.html
- Currently in beta form (there are some known and unknown bugs)
- Will be continuously updating over time.
- Source code will be posted very soon! (1-2 weeks)



References

1. Zhu, T.C., A. Ahnesjo, K.L. Lam, X.A. Li, C.M. Ma, J.R. Palta, M.B. Sharpe, B. Thomadsen, and R.C. Tailor, Report of AAPM Therapy Physics Committee Task Group 74: in-air output ratio, Sc, for megavoltage photon beams. Med Phys, 2009. 36(11): p. 5261-91.
2. Huda, W. and F.A. Mettler, Volume CT dose index and dose-length product displayed during CT: what good are they? Radiology, 2011. 258(1): p. 236-42.
3. McCollough, C.H., S. Leng, L. Yu, D.D. Cody, J.M. Boone, and M.F. McNitt-Gray, CT dose index and patient dose: they are not the same thing. Radiology, 2011. 259(2): p. 311-6. PMCID:PMC PMC3079120
4. The Association of Electrical Equipment and Medical Imaging Manufacturers, MITA Smart Dose CT XR-29: Standard Attributes on CT Equipment Related to Dose Optimization and Management, 2013.
5. The Association of Electrical Equipment and Medical Imaging Manufacturers, NEMA XR 25-2010: Computed Tomography Dose Check, 2010.
6. National Committee on Radiation Protection, NCRP Report No. 147: Structural Shielding Design for Medical X-Ray Imaging Facilities, 2004.
7. American College of Radiology. CT Accredidation Program. 2014 [cited 2014 July 16]; Available from: http://www.acr.org/Quality-Safety/accreditation/CT.
8. Zhu, T.C., A. Ahnesjo, K.L. Lam, X.A. Li, C.M. Ma, J.R. Palta, M.B. Sharpe, B. Thomadsen, R.C. Tailor, and A.T.P.C.T. Group, Report of AAPM Therapy Physics Committee Task Group 74: in-air output ratio, Sc, for megavoltage photon beams. Med Phys, 2009. 36(11): p. 5261-91.
9. American Association of Physicists in Medicine, AAPM Report No. 39: Specification and acceptance testing of computed tomography scanners, 1993.
10. European Commission, Criteria for Acceptability of Medical Radiological Equipment used in Diagnostic Radiology, Nuclear Medicine and Radiotherapy, 2012.
11. Food and Drug Administration, Code Of Federal Regulations Part 120: Performance Standards for Ionizing Radiation Emitting Products, 2013.
12. International Atomic Energy Agency, Quality Assurance Programme for Computed Tomography: Diagnostic and Therapy Applications, 2012.
13. International Atomic Energy Agency, Status of Computed Tomography Dosimetry for Wide Cone Beam Scanners, 2011.
14. International Commission on Radiation Units and Measurements, ICRU Report No. 87: Radiation dose and image-quality assessment in computed tomography, 2012.
15. International Electrotechnical Commission, Evaluation and routine testing in medical imaging departments - Part 3-5: Acceptance tests - Imaging performance of computed tomography X-ray equipment, 2004.
16. Gang, G.J., J. Lee, J.W. Stayman, D.J. Tward, W. Zbijewski, J.L. Prince, and J.H. Siewerdsen, Analysis of Fourier-domain task-based detectability index in tomosynthesis and cone-beam CT in relation to human observer performance. Med Phys, 2011. 38(4): p. 1754-68. PMCID:PMC 3069989
17. Gifford, H.C., M.A. King, P.H. Pretorius, and R.G. Wells, A comparison of human and model observers in multislice LROC studies. IEEE Trans Med Imaging, 2005. 24(2): p. 160-9.
18. Leng, S., L. Yu, Y. Zhang, R. Carter, A.Y. Toledano, and C.H. McCollough, Correlation between model observer and human observer performance in CT imaging when lesion location is uncertain. Med Phys, 2013. 40(8): p. 081908. PMCID:PMC 3724792
19. Christianson, O., J.J. Chen, Z. Yang, G. Saiprasad, A. Dima, J.J. Filliben, A. Peskin, C. Trimble, E.L. Siegel, and E. Samei, An Improved Index of Image Quality for Task-based Performance of CT Iterative Reconstruction across Three Commercial Implementations. Radiology, 2015. 275(3): p. 725-34.
20. McCollough, C., D.M. Bakalyar, M. Bostani, S. Brady, K. Boedeker, J.M. Boone, H.H. Chen-Mayer, O.I. Christianson, S. Leng, B. Li, M.F. McNitt-Gray, R.A. Nilsen, M.P. Supanich, and J. Wang, Use of Water Equivalent Diameter for Calculating Patient Size and Size-Specific Dose Estimates (SSDE) in CT, AAPM Repprt 220, 2014.
21. Medicines and Healthcare products Regulatory Agency, MHRA Report 05016: CT scanner automatic exposure control systems, 2005.
22. Solomon, J.B., J. Wilson, and E. Samei, Characteristic image quality of a third generation dual-source MDCT scanner: noise, resolution, and detectability. Med Phys, 2015. 42(8): p. 4941-53.
23. Wilson, J.M., O.I. Christianson, S. Richard, and E. Samei, A methodology for image quality evaluation of advanced CT systems. Medical Physics, 2013. 40(3): p. 031908. PMID: 23464323
24. Chen, B., O. Christianson, J.M. Wilson, and E. Samei, Assessment of volumetric noise and resolution performance for linear and nonlinear CT reconstruction methods. Med Phys, 2014. 41(7): p. 071909.
25. Richard, S., D.B. Husarik, G. Yadava, S.N. Murphy, and E. Samei, Towards task-based assessment of CT performance: system and object MTF across different reconstruction algorithms. Med Phys, 2012. 39(7): p. 4115-22. PMID: 22830744

References

26. Greene, T.C. and X.J. Rong, Evaluation of techniques for slice sensitivity profile measurement and analysis. J Appl Clin Med Phys, 2014. 15(2): p. 4042.

27. Maidment, A.D. and M. Albert, Conditioning data for calculation of the modulation transfer function. Med Phys, 2003. 30(2): p. 248-53.

28. Cruz-Bastida, J.P., D. Gomez-Cardona, K. Li, H. Sun, J. Hsieh, T.P. Szczykutowicz, and G.H. Chen, Hi-Res scan mode in clinical MDCT systems: Experimental assessment of spatial resolution performance. Med Phys, 2016. 43(5): p. 2399. PMCID:PMC PMC4841803

29. Yu, L., T.J. Vrieze, S. Leng, J.G. Fletcher, and C.H. McCollough, Technical Note: Measuring contrast- and noise-dependent spatial resolution of an iterative reconstruction method in CT using ensemble averaging. Med Phys, 2015. 42(5): p. 2261-7. PMCID:PMC PMC4401802

30. Li, K., J. Tang, and G.H. Chen, Statistical model based iterative reconstruction (MBIR) in clinical CT systems: experimental assessment of noise performance. Med Phys, 2014. 41(4): p. 041906. PMCID:PMC PMC3978426

31. Solomon, J. and E. Samei, Quantum noise properties of CT images with anatomical textured backgrounds across reconstruction algorithms: FBP and SAFIRE. Med Phys, 2014. 41(9): p. 091908.

32. Boedeker, K.L., V.N. Cooper, and M.F. McNitt-Gray, Application of the noise power spectrum in modern diagnostic MDCT: part I. Measurement of noise power spectra and noise equivalent quanta. Phys Med Biol, 2007. 52(14): p. 4027-46.

33. Solomon, J.B., O. Christianson, and E. Samei, Quantitative comparison of noise texture across CT scanners from different manufacturers. Med Phys, 2012. 39(10): p. 6048-55.

34. Solomon, J., A. Ba, F. Bochud, and E. Samei, Comparison of low-contrast detectability between two CT reconstruction algorithms using voxel-based 3D printed textured phantoms. Med Phys, 2016. 43(12): p. 6497.

35. Li, K., J. Garrett, Y. Ge, and G.H. Chen, Statistical model based iterative reconstruction (MBIR) in clinical CT systems. Part II. Experimental assessment of spatial resolution performance. Med Phys, 2014. 41(7): p. 071911. PMCID:PMC 4106476

36. Solomon, J., J. Wilson, and E. Samei, Characteristic image quality of a third generation dual-source MDCT scanner: Noise, resolution, and detectability. Med Phys, 2015. 42(8): p. 4941-53.

37. Siewerdsen, J.H., I.A. Cunningham, and D.A. Jaffray, A framework for noise-power spectrum analysis of multidimensional images. Med Phys, 2002. 29(11): p. 2655-71.

38. International Commission on Radiation Units and Measurements, ICRU Report 54: Medical Imaging- The Assessment of Image Quality, 1995.

39. Solomon, J., D. Marin, K. Roy Choudhury, B. Patel, and E. Samei, Effect of Radiation Dose Reduction and Reconstruction Algorithm on Image Noise, Contrast, Resolution, and Detectability of Subtle Hypoattenuating Liver Lesions at Multidetector CT: Filtered Back Projection versus a Commercial Model-based Iterative Reconstruction Algorithm. Radiology, 2017: p. 161736.

40. Solomon, J., A. Mileto, J.C. Ramirez-Giraldo, and E. Samei, Diagnostic Performance of an Advanced Modeled Iterative Reconstruction Algorithm for Low-Contrast Detectability with a Third-Generation Dual-Source Multidetector CT Scanner: Potential for Radiation Dose Reduction in a Multireader Study. Radiology, 2015. 275(3): p. 735-45.

41. Samei, E., M.J. Flynn, and W.R. Eyler, Simulation of subtle lung nodules in projection chest radiography. Radiology, 1997. 202(1): p. 117-24.

42. Chen, B. and E. Samei. Development of a phantom-based methodology for the assessment of quantification performance in CT. in SPIE Medical Imaging. 2013. Orlando, FL. 8668 86681E-86681E-7

43. Popescu, L.M. and K.J. Myers, CT image assessment by low contrast signal detectability evaluation with unknown signal location. Med Phys, 2013. 40(11): p. 111908.

44. Zhang, Y., S. Leng, L. Yu, R.E. Carter, and C.H. McCollough, Correlation between human and model observer performance for discrimination task in CT. Phys Med Biol, 2014. 59(13): p. 3389-404. PMCID:PMC 4057982

45. Solomon, J.B. and E. Samei, Correlation between human detection accuracy and observer model-based image quality metrics in computed tomography. Journal of Medical Imaging, 2016. 3: p. 12.

46. Abbey, C.K. and H.H. Barrett, Human- and model-observer performance in ramp-spectrum noise: effects of regularization and object variability. J Opt Soc Am A Opt Image Sci Vis, 2001. 18(3): p. 473-88. PMCID:PMC PMC2943344

47. Yu, L., S. Leng, L. Chen, J.M. Kofler, R.E. Carter, and C.H. McCollough, Prediction of human observer performance in a 2-alternative forced choice low-contrast detection task using channelized Hotelling observer: impact of radiation dose and reconstruction algorithms. Medical physics, 2013. 40: p. 041908.

48. Barrett, H.H., K.J. Myers, C. Hoeschen, M.A. Kupinski, and M.P. Little, Task-based measures of image quality and their relation to radiation dose and patient risk. Phys Med Biol, 2015. 60(2): p. R1-R75. PMCID:PMC 4318357

49. Favazza, C.P., A. Ferrero, L. Yu, S. Leng, K.L. McMillan, and C.H. McCollough, Use of a channelized Hotelling observer to assess CT image quality and optimize dose reduction for iteratively reconstructed images. J Med Imaging (Bellingham), 2017. 4(3): p. 031213. PMCID:PMC PMC5624940

50. Popescu, L.M., Nonparametric ROC and LROC analysis. Med Phys, 2007. 34(5): p. 1556-64.

References

50. Popescu, L.M., Nonparametric ROC and LROC analysis. Med Phys, 2007. 34(5): p. 1556-64.

51. Wunderlich, A. and F. Noo, A nonparametric procedure for comparing the areas under correlated LROC curves. IEEE Trans Med Imaging, 2012. 31(11): p. 2050-61. PMCID:PMC 3619029

52. Gallas, B.D., One-shot estimate of MRMC variance: AUC. Acad Radiol, 2006. 13(3): p. 353-62.

53. Popescu, L.M., Nonparametric signal detectability evaluation using an exponential transformation of the FROC curve. Med Phys, 2011. 38(10): p. 5690-702.

54. Vaishnav, J.Y., W.C. Jung, L.M. Popescu, R. Zeng, and K.J. Myers, Objective assessment of image quality and dose reduction in CT iterative reconstruction. Med Phys, 2014. 41(7): p. 071904.

55. Becchetti, M.F., J.B. Solomon, W.P. Segars, and E. Samei. Synthesized interstitial lung texture for use in anthropomorphic computational phantoms. 2016. 9783 97835Z-97835Z-5

56. Bolch, W., C. Lee, M. Wayson, and P. Johnson, Hybrid computational phantoms for medical dose reconstruction. Radiat Environ Biophys, 2010. 49(2): p. 155-68. PMCID:PMC PMC2855752

57. Lee, C., D. Lodwick, J. Hurtado, D. Pafundi, J.L. Williams, and W.E. Bolch, The UF family of reference hybrid phantoms for computational radiation dosimetry. Phys Med Biol, 2010. 55(2): p. 339-63. PMCID:PMC PMC2800036

58. Norris, H., Y. Zhang, J. Bond, G.M. Sturgeon, A. Minhas, D.J. Tward, J.T. Ratnanather, M.I. Miller, D. Frush, E. Samei, and W.P. Segars, A set of 4D pediatric XCAT reference phantoms for multimodality research. Med Phys, 2014. 41(3): p. 033701. PMCID:PMC PMC3987726

59. Segars, W.P., H. Norris, G.M. Sturgeon, Y. Zhang, J. Bond, A. Minhas, D.J. Tward, J.T. Ratnanather, M.I. Miller, D. Frush, and E. Samei, The development of a population of 4D pediatric XCAT phantoms for imaging research and optimization. Med Phys, 2015. 42(8): p. 4719-26. PMCID:PMC PMC4506297

60. Solomon, J., A. Ba, A. Diao, J. Lo, E. Bier, F. Bochud, M. Gehm, and E. Samei. Design, fabrication, and implementation of voxel-based 3D printed textured phantoms for task-based image quality assessment in CT. SPIE Medical Imaging 2016. 9783 978328-978328-11

61. Ma, C., L. Yu, B. Chen, C. Favazza, S. Leng, and C. McCollough, Impact of number of repeated scans on model observer performance for a low-contrast detection task in computed tomography. J Med Imaging (Bellingham), 2016. 3(2): p. 023504. PMCID:PMC PMC4886187

62. Tseng, H.W., J. Fan, M.A. Kupinski, P. Sainath, and J. Hsieh, Assessing image quality and dose reduction of a new x-ray computed tomography iterative reconstruction algorithm using model observers. Med Phys, 2014. 41(7): p. 071910.

63. Winslow, J., Y. Zhang, and E. Samei, A method for characterizing and matching CT image quality across CT scanners from different manufacturers. Med Phys, 2017. 44(11): p. 5705-5717.

64. Zhang, Y., C. Smitherman, and E. Samei, Size-specific optimization of CT protocols based on minimum detectability. Med Phys, 2017. 44(4): p. 1301-1311.

65. McCollough, C.H., M.R. Bruesewitz, M.F. McNitt-Gray, K. Bush, T. Ruckdeschel, J.T. Payne, J.A. Brink, and R.K. Zeman, The phantom portion of the American College of Radiology (ACR) computed tomography (CT) accreditation program: practical tips, artifact examples, and pitfalls to avoid. Med Phys, 2004. 31(9): p. 2423-42.

66. Kruger, R.L., C.H. McCollough, and F.E. Zink, Measurement of half-value layer in x-ray CT: a comparison of two noninvasive techniques. Med Phys, 2000. 27(8): p. 1915-9.

67. Geleijns, J., M. Salvado Artells, P.W. de Bruin, R. Matter, Y. Muramatsu, and M.F. McNitt-Gray, Computed tomography dose assessment for a 160 mm wide, 320 detector row, cone beam CT scanner. Phys Med Biol, 2009. 54(10): p. 3141-59. PMCID:PMC PMC2948862

68. O'Daniel, J.C., D.M. Stevens, and D.D. Cody, Reducing radiation exposure from survey CT scans. AJR Am J Roentgenol, 2005. 185(2): p. 509-15.

69. Schmidt, B., N. Saltybaeva, D. Kolditz, and W.A. Kalender, Assessment of patient dose from CT localizer radiographs. Med Phys, 2013. 40(8): p. 084301.

Thank you!| justin.solomon@duke.edu