Master Erasmus Mundus in
Color in Informatics and Media Technology (CIMET)

Evaluation of Image Quality of State-of-art CT Vendors in the Norwegian Market.

Master Thesis Report.

Presented by

Antonio Pelegrina Jiménez

and defended at

Gjøvik University College,

13 June 2014.

Academic Supervisor:

Dr. Marius Pedersen, Høgskolen i Gjøvik

Jury Committee:

- Dr. Alain Treméau
- Dr. Dimitry Semenov - врач Дмитрий Семенов

Acknowledgements:

Dag Waaler, Anne Catrine Martinsen, Hilde Kjernlie Andersen

Abstract

This study focus in the low-contrast detectability properties of two scanners examples in the state-of-the-art from Oslo Hospital Interventional center, as examples of different technologies, being one of them Dual-Energy source CT. We considered different noise filter possibilities and several iterative reconstruction options for the images from both vendors.

It is phantom-based (so does not use real patients) but it considers some of its related challenges, as it is the overweight trend in population. Beside the increased difficulty for and adequate image quality scan, this is an important risk factor in medicine. Several rings where added to phantom at each of the CT scans protocols used.

Image quality in diagnostics, refers to Diagnostic quality. First we considered some mathematical metrics and noise studies, and pursued measurements of Contrast to Noise Ratio.

Later, we wanted to check how this parameter relaters to observers perception, and we pursued several experiments. We took special attention onto the confidence intervals and the standard deviation value among observers.

The goal is to deeply study the acquisition methods of CT devices, then consider the workflow of data from CT to computers (at the selection of useful images), and finally, to have an insight onto the factors that can affect the transition from "*vision to decision"* at the medical sector.

Two different interpretations of the same image, could lead onto a different diagnose by a Doctor. This study explores the differences between the answers of Image Processing students, and Radiography students, and the distribution of data according to a normal (Gaussian) distribution, through some statistical methods as *Mann-Whitney* U-Test, *Anderson-Darling* Test, and *Student* t-Test.

Results presented here could contribute to some dose reduction decisions, that is the one of main dilemmas in Radiology nowadays: To obtain an adequate image quality for a sucessful diagnose, while keeping dose contribution ALARA (as low as reasonably achievable).

In terms of low contrast detectability, *Toshiba Acquilion* performance is good for the phantom without ring, while the iterative reconstruction algorithm *VEO* by GE Works better for the phantom with medium rings, with the drawback of a higher computational time. For the big rings, results depend on the particular dose level used. Exploration of new AiDR-3D image reconstruction by Toshiba is advisable for future work.

The author declares not to have conflict of interest.

The author is subject to a previous Confidentiality Agreement in Scintillation Physics research, as part of a stage held at a French corporation in Shanghai, China.

1 CONTENT TABLE

2 INTRODUCTION

2.1 PROBLEM DESCRIPTION

CT is medical imaging procedure that uses x-rays and digital computer technology to create cross-section images of the body. There has been an increased usage of this procedure to obtain diagnose information.

When we consider the cumulative effects of radiation onto the equation of patient health, it is not always responsible to proceed with CT, unless justified enough. We need to discern between *useful* and *justified*.

CT will always provide an information output, but to authorize it, we need to obtain a net benefit from it with views to the diagnostic, that could trade-off the negative side of the use of radiation. As well stated by ICRP (International Commission of Radiation Protection) every scan needs to be justified, the method optimized, and the dose range limited.

The increasing trend of cancer cases worldwide, and the necessity to study deeper the long term effects of radiation, makes this study especially sensitive. The individual risk of a person is not so high when it does not belong to a particular risk group (as to mention an example, smokers) and overall the trade-off between CT scan and the benefits from the obtained information uses to be positive.

But that does not mean we can use CT as a harmless method. The trend some years ago was to obtain the best image quality, by increasing the dose, but nowadays all manufacturers are researching onto reduce the dose to patient, while maintaining enough quality of the image for a successful diagnose. [1] [2] [3]

At nowadays scanners, X-ray Tube Current can be modulated [4] so we can adjust CT scanning dose depending on the clinical task, patient size, anatomical location and clinical practice. In order to reduce the dose, iterative reconstruction algorithms can be applied, to improve the quality of the original image. But where is the lower limit for an accurate diagnose?

To determine the optimal dose value under an ALARA strategy (as low as reasonably achievable) we need to study some image parameters as contrast and noise, those that influence the image perception. Several metrics had been intended to measure the quality, (in example, noise and contrast related). They provide a mathematical approximation, but they don´t account the wide range of visual illusions humans experience, more far than simultaneus contrast illusion between foreground and background. Image quality in medicine is related to diagnose quality, so observers experiments became the perfect complement to usual metrics.

The hypothesis is that noise magnitude and texture differs between the vendors of state-of-the-art CT scanners on the market, and that this affects the low contrast detectability.

2.2 MOTIVATION

According to Samei [5] one billion imaging examinations are performed worldwide annually. In particular, the use of CT scan has increased dramatically over the last two decades in many countries, and it constitutes in average 67% of the radiation contribution in the medical imaging sector [6].

In 2010 Norway carried out more CT scans per habitant than any other Nordic Country, overcoming Iceland at the top of ranking in 2009. [7] It could be risky to stay on top of this rank - unless we have certainty that each of them was correctly justified. Some recent

publication [8] mentions the generalized use of CT technology two times, in the top ten of 2014 hazards regarding the health sector.

In this study, we focus in a target group of special relevance, that is overweight people. Dose strategies can be less effective, and in addition, they are statistically prone to suffer from other diseases.

The differences at each country regulations or guidelines, and the presence of diverse agency recommendations, leaves the low dose awareness to be a choice of the CT manufacturer, according to each brand social corporate responsibility. The operator has an extent degree of freedom at the scanning parameters choice, nowadays just limited for very high dose levels by a password. But seems the market is becoming more conscious of dose risks, and the low dose challenge is each day more present at the R&D strategies. To select the best manufacturer for each possibility, does not avoid the dilemma for the doctor, but it simplifies it.

2.3 RESEARCH QUESTIONS

The initial hypothesis (and research question) is that noise magnitude and texture differs between the vendors of state-of-the-art CT scanners on the market, and that this affects the low contrast detectability for the scanners.

The first part of the hypothesis could be sometimes seen as an straight-forward question [9] from image processing perspective. We can agree that the filters, detectors technology and photoconverter devices, are a particular technology of each brand [10] and that texture difference is the most probable output, and a feasible assumption.

Here we will focus especially on the reconstruction step and the dose measurements, for different patient size simulation, in order to have an insight onto which factors cause differences in low contrast detectability and noise properties between scanners from different vendors. It will also studied under which strict conditions, we could cluster the data from two diverse observers group, interpreting the same medical images, for the statistical analysis of their answers.

2.4 RESEARCH METHODS

This study aims to establish a method for quantifying and comparing low contrast detectability between CT scanners brand and each of the possible reconstruction method, in order to determine its image quality through two complementary schemes: numerical measurements (contrast to noise ratio) and observers experiment. Answers here are of quantitative nature.

The work has also included practical parts, such as the acquisition of images using different CT scanner. Results of the analysis have been done using statistical methods.

Scans were done in collaboration with the Interventional Centre of Oslo University. There we found two different manufacturers available, at different sections of the Health City. We decided to compare two different brands on the Norwegian market, as example of different technology approaches for the scan (in example, one of them includes Dual-Energy Technology). Processing of the images took place in Oslo Hospital and Gjøvik University, while the perceptual experiments were designed and implemented at the Norwegian Colorlab.

2.5 LIMITATIONS OF THE STUDY

We didn´t worked with real patients, but with a non-anthropomorphic model named phantom Capthan 500, used both for calibration and research purposes. Their different modules can allow us to study in detail characteristics as noise and contrast detectability, without exposing a patient, according to OIEA guidelines. Due to Oslo Hospital limitations, we were able only to use two different scanners brands, by General Electric™ (Milwaukee, US) and Toshiba ™ (Tokyo, JP)

The different post-processing options of images by the integrated hardware and software are until a high extent confidential. The goal here is not to improve those iterative methods (in example the named VEO ™ by GE) but to explore in depth its Image Quality, in particular noise and contrast properties.

3 BACKGROUND

3.1 CT HISTORY HIGHLIGHTS

First clinical CT system was built by Godfrey Hounsfield at EMI Research (London) – the first human patient was scanned in 1971 [11]. Image reconstruction from projections was studied earlier and independently by physicist A. Cormack, based in a mathematical demonstration by J. Radon in Austria, 1917.

As a tomography system, it allows to obtain depth information, providing an improvement of accuracy over traditional plate radiography. Attenuation of X-rays is expressed in Hounsfield Units [12], and displayed in shades of grey. [*more info about HU can be found in Materials & Methods chapter]. We are in the named fourth generation since the first commercial scanner. Significant improvement has been realized in axial resolution, and more recently, in low contrast detectability. [13]

In this study we will compare MULTISLICE CT devices, where multiple detector rows are installed at gantry, for a faster scanning procedure. [14] DUAL CT technique, the one used by GE Discovery device, started in 1977, originally in the scope to obtain high temporal resolution in cardiac imaging. It uses two energies of x-ray, in order to resolve a better Contrast to Noise ratio, by switching low and high energies at tube.

3.2 DUAL ENERGY CT AND DOSE IMPLICATIONS

Dual Energy CT devices, is the equivalent to a two-wavelength spectral camera in the range of energies used at CT. It requires an x-ray tube source, able to provide two different x-ray energies.

No source of x-ray energy is strictly monochromatic, among those available for medical practice. At its earliest stage of development, two consecutive scans were required (with the risk of patient movement) but since 2006 has become a much more efficient process, being some commercial devices able to perform both scans by rapid switching of frequencies at tungsten tube, and thanks to detectors development.

At previous technology generations tube technology was not able to provide an adequate output of quanta at low dose configuration. Nowadays this problem has been solved, but careful monitoring of tube parameters is recommended. At FBP stage, a Dual CT scanner could imply a higher dose for an equivalent image quality, but reconstruction step can overcome this problem.

As well stated by Jiang [15]: for most materials "x-ray attenuation coefficients are a function of energy", and hence low contrast differentiation is influenced mainly by low-energy x-ray photons. Some studies can consider the physics of interaction events behind CT, but to describe in example scattered radiation interaction is not a straight forward issue, and probabilistic methods as *Montecarlo* simulation are used for this purpose.

3.3 COMPUTER TOMOGRAPHY FUNDAMENTALS

An x-ray circular scan is performed. Projections performed at different angles, are back-projected and joined onto a single image. To illustrate this, imagine you do a sampling at each degree (360) of gantry. Actually CT does more, in the order of one thousand [1] [12]. Back-projection is defined by Jiang [15] as "a process in which the intensity of a sample is distributed uniformly along the path that formed the sample". At Filtered-back-projection (FBP) profiles are filtered and then back-projected onto a single image.

Radon transform [16] sometimes receive the name of *sinogram*. The reason is, that Radon transformation of a Dirac delta function (that represents an ideal pulse of infinite height, and negligible width) resembles a sinusoid form.

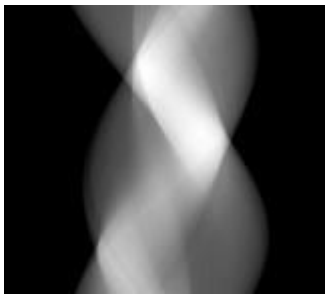


Figure 1. Image of Sinogram [17]

Two-dimensional Radon transform (for a discrete space), is closely related to Hough Transform (for the continuous space). The last was patented in 1962, as an efficient way to detect features in images (originally straight lines, later generalized to circles and ellipses)

The start point of both methods is to relate a Cartesian representation *(x, y)* with a polar one *(r, θ)*. Radon transform could be also related with Fourier transform (that converts signals from time space onto the frequency one). According to [15] Chapter 2, page 49: "any object can be considered a collection of very small points" referring here to a superposition theorem.

Spatial resolution: At the original CT prototype, developed by EMI ™ Corporation, images were of 80x80 pixels. Nowadays is usual to work with 512x512 matrix [1] and some systems can produce 1024x1024 [18]

Noise can hide pathology and valuable diagnostic information. As projected data is not free of quantum noise and electronic noise, CT device proceeds with high-pass filtering [19], enhancing higher frequency elements (edges). The process is named Filtered back-projection. Mathematically, it could be seen as a convolution process.

CT embedded filtering performs two main tasks as is edge enhancement and noise reduction, and can be varied - depending on the anatomy we want to study - among a predefined set. In example, Discovery™ CT750 can use up to 5 reconstruction kernels for some functions of cardiac imaging. But as well stated by Hsiang [15] *"the design details of these filters are generally considered proprietary*".

3.3.1 Reconstruction algorithms: Iterative and non-iterative

Each brand has different reconstruction algorithms, of confidential nature: some act as mathematical filters, others want to reduce noise and artifacts through iterative reconstruction.

The latter, assumes that smooth changes between consecutive images, are most probable than big ones [19]. FBP propagates and sometimes amplifies noise into patient images, hiding pathology and valuable diagnostic information. With the use of Iterative methods, image perceptual quality is improved, through a tradeoff between noise cancelling and edge enhancement.

Adaptive Iterative Dose Reduction AiDR™ by Toshiba, compares each pixel value with those surrounding it. Large differences are removed iteratively. Hence, we can deduce, noise will be more likely reduced more at higher frequencies (fine details) while the texture of low frequency noise will remain as background. It uses to be used in a weighted scheme to be linked to the original image.

The more advanced methods, as VEO reconstruction by General Electrics, are based on the physics of data acquisition, including some statistical approaches onto reconstruction. The drawback is to be more computationally expensive, and hence time-requiring, in the context of the medical sector.

According to [20] each individual model (noise statistics, system optics, or detector characteristics) is developed onto a mathematical function that describes with accuracy the way data is back-projected. For a better output reconstruction, the VEO function was optimized by comparing computer-simulated outputs with those real from CT scan. Solution here is not analytical-based, but a numerical one (iterative reconstruction). It also consider some metrics from the original reconstructed image.

3.3.2 Multi-slice CT Evolution.

From 16 row detectors, technology evolved onto, 32, 64, up to 320, thanks to the evolution of scintillators technology (from xenon gas chambers, onto more compact and efficient high performance ceramics. This improved the performance of CT scanners, particularly for helical scans [21] . However, beside the intrinsic advantages of having a wider detector gantry in order to have quicker scan procedure - and hence reduce radiation dose to patient - there is a limit on the benefits or having more detector rows than 64.

The wider is a detector area, we will get more over-scanning at the top & bottom of patient, hence providing extra dose. In addition, the more detectors area it implies more mathematical reconstruction, and then we could get more artifacts at image.

This is known as the cone-beam problem. Some authors as Hsieh, Jiang [15] has proposed to label scans with a number of rows equal or higher to 64 as “cone-beam CT” to highlight the fact they could not operate correctly without advance calibration and reconstruction technologies. In fact some practitioners [1] [21] assert that images can be slightly better at *64-slice* devices, than in newer ones with a higher number of rows.

In the following picture, we can see cone-beam effect exaggerated, for better understanding.

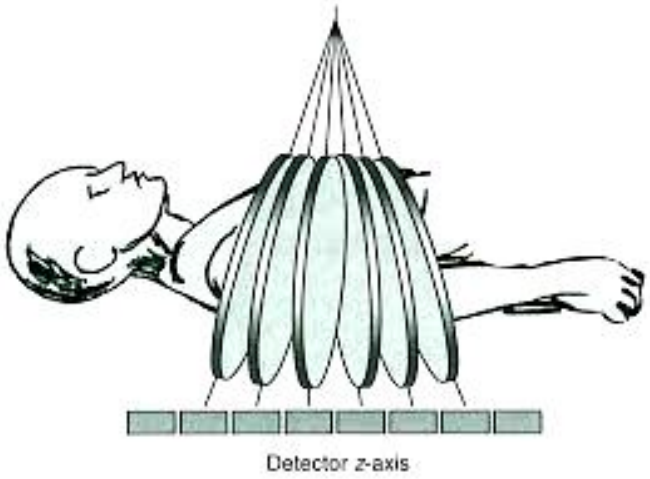


Figure 2. Image cone-beam [21]

3.3.3 Beam hardening

The beam hardening is due to the fact materials absorbs easily low-energy radiation. It can be explained by Beer-Lambert law.

$I = I_0 e^{-\int \mu(E) dS}$ Equation 1. Beer-Lambert Law.

μ denotes the linear attenuation coefficient, dependent of Energy.

The mean energy of x-rays becomes higher, and according to BOAS [19] it produces "dark streaks between two high attenuation objects, with surrounding white streaks". Another of its usual effects, is a black shadow when scanning the brain [1]. For radiation-scatter basic understanding, Photoelectric, and Compton Effect can be referred.

Some authors, as E. Boas [19] – having this one a financial disclaimer regarding GE - refers that “Dual Energy can reduce beam hardening, but not scatter effects”

Actually, Beam Hardening could be minimized in spite of more noise at picture – using more energetic x-rays. We can expect beam hardening to be less present at low dose images, but those will have more noise, inconvenient to detect low contrast objects. A trade-off between beam hardening and noise is needed. For the interested reader, from a detectors perspective: [22]

Another strategy is to use *Bowtie Filters*, to provide a higher dose towards the center of Field of View [19] (in order to compensate the fact, X-rays passing through the middle portion of a uniform cylindrical phantom are hardened more than those traversing the edges). Using “hard” (high energetic) X-rays in medical imaging, leads to reduced dose received by the patient.

However, at higher energies the conventional scintillators quickly become the limiting factor (they must be thin in order to provide reasonable spatial resolution) [23]

In this study, we will make use of iterative algorithms that take into account beam hardening effect, to be compared to others filters were is not considered (sometimes available at the same device). Toshiba has a wide set of filters [1] and we can chose among "beam hardening correction" or not.

3.3.4 Noise and dose

The lack of photons when dealing with low dose (photo starving), increases substantially the noise at image. The noise increases by the square root of x, as the dose decreases 1/x. Once said noise is proportional to the inverse of the square root of the number of photons, the operation settings can affect it are Tube current, Scan time and Slice thickness. In example, if we reduce onto half the tube current - or equivalently the acquisition time – this will increase the noise by square root of 2.

Special reconstruction algorithms are needed, in order to recognize low contrast objects.

3.4 NORWEGIAN MARKET AND DOSE REGULATIONS GUIDELINES

The extensive use of radiology scans, has overcame the amount of radiation received by natural background - until then the largest source -at several developed countries [24]. There is a collection of good practice recommendations by several institutions on the field, being the most significant those by ICRP (International Commission on Radiation Protection) [25]

In Europe, the European Commission established the High Level expert Group MELODI "*Multidisciplinary European Low Dose Initiative*" that involves 5 member states: Finland, France, Germany, Italy and United Kingdom. Europe has more strict guidelines than those used in the United States [15] [8, Chapter 12, page 470] [26]. A new EU legislation is expected for February 2018 [3]

In Norway, the Radiation Protection Authority (NRPA), dependent on the Ministry of Health and Care Services, initiated in 2000 to develop the program KVIST for Quality Assurance in Radiotherapy.

According to NRPA website [27] the ongoing Project "aims to stimulate collaboration by focusing on clinical, technical and administrative problems that can be addressed and solved on a National Level". In conjunction with other Nordic Countries, NRPA has stated its will to establish the IAEA Concept of "Triple A" [27]: Awareness, Appropriate-ness (or proper-ness) and Audit. But in difference with other medical procedures, there is no binding regulation in Norway about CT dose management [27].

3.5 IMAGE QUALITY

Image quality in radiology is diagnostic quality: to detect those details the radiologist perceive as important. As stated by Dr. Kundel [28] [29] *"the images of highest IQ are those that enable the observer to most accurately report diagnostically relevant structures and features"*. Hence, we set complementary initiatives to measure image quality, starting through metric measurements of contrast to noise ratio, to be linked later with the results of some observers experiment.

3.5.1 Sharpness

If we were working with anatomical images, we could define sharp visualization - where anatomical details are clearly defined - versus *standard* visualization, were characteristic features are detectable, but details are not fully reproduced [30] In this study no anatomic structures are involved, as we are working with a phantom.

3.5.2 Artifacts: how dose affects image quality.

Some image quality artifacts, as the patient motion ones (in example by breathing) do not relates with dose parameters. Others as *cupping* and *capping* are Dose Related.

For a collimation up to 40 mm (as it will be our case) geometrical artifacts as *stair-step* and *cone-beam* effect are not relevant. Collimation here refers to integrate an area of detectors onto a single output, to be linked onto photo-converter input. Helical CT can reduce motion artifacts (when comparing it with axial scanning) and is the start point of some advanced reconstruction methods, as VEO by GE.

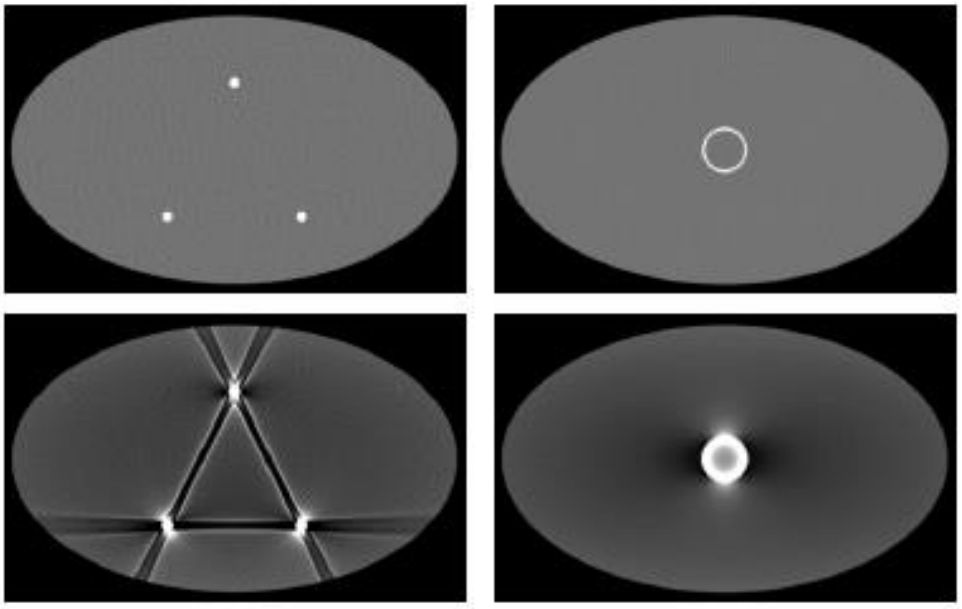


Figure 3 [19] Simulated beam hardening and Cupping artifact

3.6 MEASURING IMAGE QUALITY

While at a normal radiography plate, the image can be easily seen as hyper-exposed, due to detector saturation, the more complex technology in CT includes a step of normalization of data [4], so we will not see a picture more dark (as in photography) [31], but we will see a different level of noise in the picture.

According to [32] photon noise obeys a Poisson distribution, while electronic one is "almost always" Gaussian.

In order to measure spatial resolution, Modulated Transfer Function is used. To measure contrast resolution, the most basic parameter is Signal to Noise Ratio (SNR). Rose Criterion [33] , [34] establish a value of 5:1 as a threshold of detectability for human eye.

Sharpness can be measured with the Line Spread Function (LSF) that according to Dr. Cunningham [35] "describes the spatial-resolution characteristics in terms of how a sharp line appears at the image.

3.6.1 Quantum Detector Efficiency

A parameter named QDE (Quantum detector efficiency) with a value between zero and one, is sometimes used, as related to the number of photons that are attenuated by the detector material. However, QDE not says that much of at a CT device overall performance. It is not a reliable parameter for image quality, as some geometrical factors, and filtering plates (that scatter radiation) can affect its value. But we can keep in mind the fact higher the QDE value is, having constants the other factors, higher is the SNR.

3.6.2 Noise Power Spectrum

As Sagara [36] well defines, NPS is *"a graphical depiction of image noise variance, expressed as a function of its frequency"*. According to Jiang [15], Noise Power Spectrum is obtained by squaring the magnitude of Fourier transform of the noise image. In practice, it can be obtained by scanning a water phantom.

There is a recent Thesis that state a NPS shift toward higher frequencies at GE Discovery Scan, in comparison with Toshiba Acquilion [37]. This fact was also stated by the head of Oslo University Interventional Center, Dr. Martinsen [1]. We can see a detailed figure at ANNEX 2.

3.6.3 Contrast to Noise Ratio

CNR is defined as the mean value difference between CT numbers (a measure of attenuation) in background and surround, divided by the standard deviation, among a region of interest. As all values in the equation are expressed in Hounsfield Unit, the ratio has no units.

$$CNR = \frac{CT\ number\ of\ object\ (HU) - CT\ number\ of\ background\ (HU)}{\frac{(SDo\ (HU) + SDb(HU))}{2}}$$

The use of this metric, allows our work to be compared to a broad catalogue of previous studies within the Medical Sector.

3.6.4 Color representation and contrast

Dr. Samei [5] states that "in radiology at least, increasing dynamic range by going to color may actually be confounded by color-luminance dependencies". At medical practice, doctors use to move between consecutive CT slices in order to discern better medical features by looking onto 3D information. This movement can create new perception illusions, usually subtle to appreciate. In a hypothetical color representation, we can refer some examples by E. Dixon and Dr. Shapiro [38]. Even if Just Noticeable Differences are higher in color than in grayscale, "the visual system has lower spatial resolution in the color channels than in the luminance channels". [39]

Pseudo-color is occasionally used for CT outside the medical sector (gamma-ray based devices) - See ANNEX 5 – and in other medical imaging areas, but within radiology, it is the exception and not the rule.

3.6.5 Observer experiments

D. Manning explain [40] "medical images are two dimensional, but generally incorporate cues that lead to the perception of depth. This third dimension is inferred by density differences and structural overlap".

3.7 PREVIOUS STUDIES IN LOW CONTRAST DETECTABILITY

According to [41] there are few studies about dose performance of Dual Energy devices. To use Dual CT does not require necessarily higher dose, but here I see words are used in analogy with mathematical language: to be necessary, is different as to be sufficient. Some authors [42] had stated an average increase of 20% in dose performance with respect non-dual CT devices.

Regarding detectors and contrast imaging research, we can find some interesting hints from last SPIE Conference of Medical Imaging [43]. For Image Processing solutions, an example by Chen, Y. can be illustrative [44]

According to Jiang [15] Impact of Noise Power Spectra on the detectability of low contrast-objects "is complex, and many studies had been devoted to this subject". Here I would recommend [37] as introduction, as it is based on the same real-devices of this study.

4 MATERIALS & METHODS

4.1 EXPERIMENTAL ENVIRONMENT: CT SCANNERS

The two devices we work with, are examples of different CT technologies.

First one we used was a Toshiba *Acquilion ™ (Toshiba Corporation, Tokyo, Japan)*:

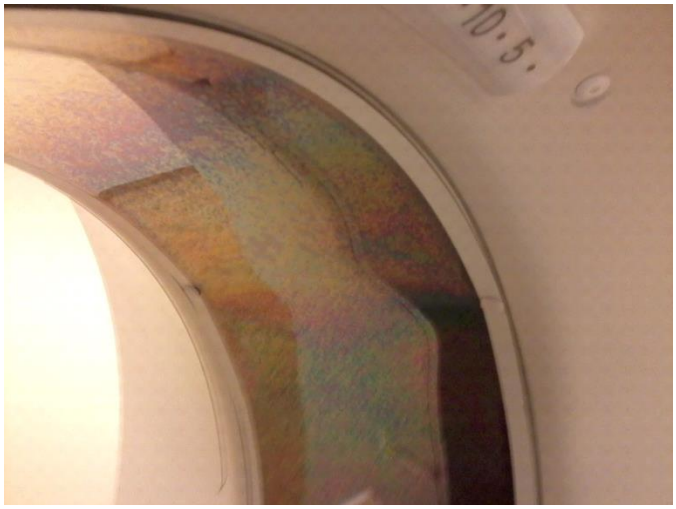


Image 1. Detector gantry detector detail. Toshiba Acquilion Scanner [45]

While GE Discovery ™ CT750HD (GE Healthcare, Milwaukee, US) is a Dual-Energy CT System.

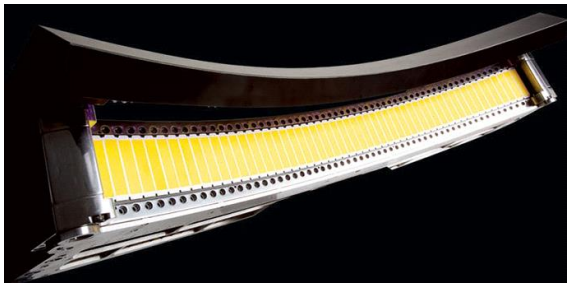


Image 2. Gemstone detector [46] of GE Discovery.

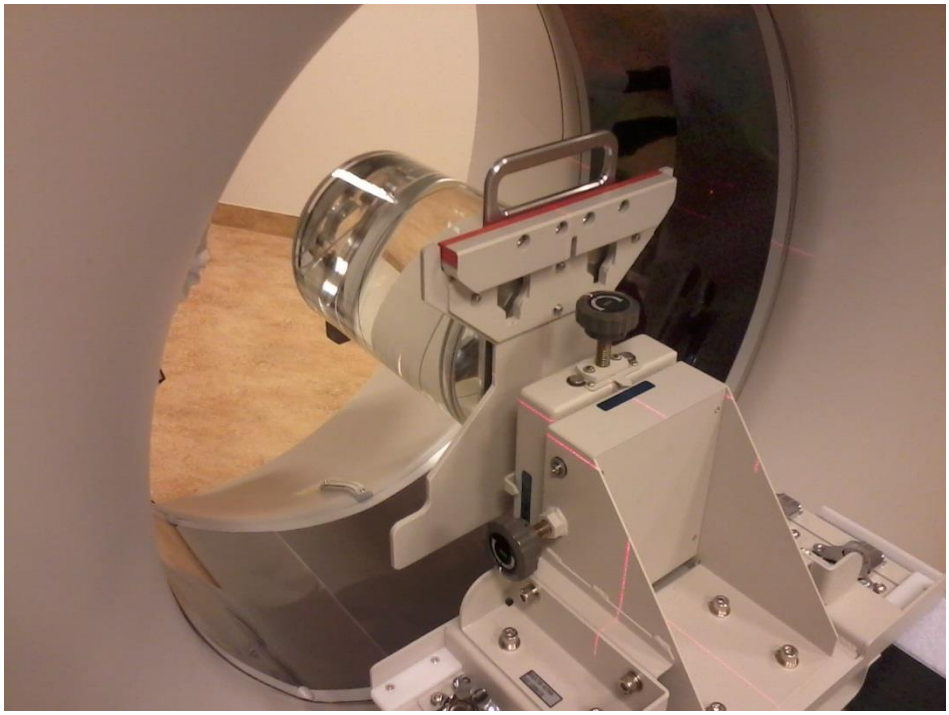
Both installed in 2011, and not used more than 3 times a week by radiographers (as average) [1].

4.2 PERIODICAL QUALITY TESTS

The scanners tested were calibrated each of the mornings we proceed with testing. Both devices, the one by GE and the other by Toshiba, were comparable in the lifetime of operation (installed in 2011) and had been upgraded - in software - as much as possible.

At the actual CT generation (4th) all the detectors are arranged in a unique circular gantry: as x-ray source is closer to the patient than detectors, the CT device needs to deal with high geometrical magnification [21]. Here we can see the importance of a calibration process protocol.

As stated in section 3.6.2, Noise power spectra can be obtained by scanning a water phantom.



Images 3.. Water Phantom Test. [45]

Quantum noise has been referred [34] to be the most important parameter affecting image quality. As well stated by ZARB, it is the "statistical fluctuation or standard deviation of CT numbers".

With reference to water attenuation value, we can define CT-numbers, as a linear transformation of the original linear attenuation coefficient μ [Eq.1]. A CT-number is expressed in Hounsfield Units.

$$CTnumber\ (HU) = 1000\frac{\mu_x - \mu_{water}}{\mu_{water}}$$

A substance denser than water, leads to a positive CT-number. The operational range of H.U is usually (-1000, +1000) where -1000 is for the air.

In example, bones have usually a CT number of 900 H.U. while Lung will have a negative value. If we have an image where CT number equals zero, it will have no grays, just black and white.

Blood can range between +30 to +45, and water has the value of zero by definition [else we will have a mathematical indetermination in the equation, [0/0]. In practice, due to stochastic nature of interaction radiation – matter, the noise distribution of the water phantom scan, will have an average value around zero [47], but usually ranging within the values [-4, +4].

CT numbers, do not vary very much with beam energy, when we deal with soft tissues [48]. But they vary noticeably for high Z (atomic number) materials. This is the principle used by Dual Energy CT in order to discern materials.

4.3 CAPTHAN PHANTOM

Capthan 500 is a non-anthropomorphic phantom. It has a cylindrical shape and it is composed of several modules, to study characteristics as noise power spectra, or low contrast detectability.

It allow us to proceed with in-deep investigations about Image Quality, without the risk of exposing a patient onto radiation. As a drawback, it could be stated [49] that has not much to do with the anatomy of a patient.

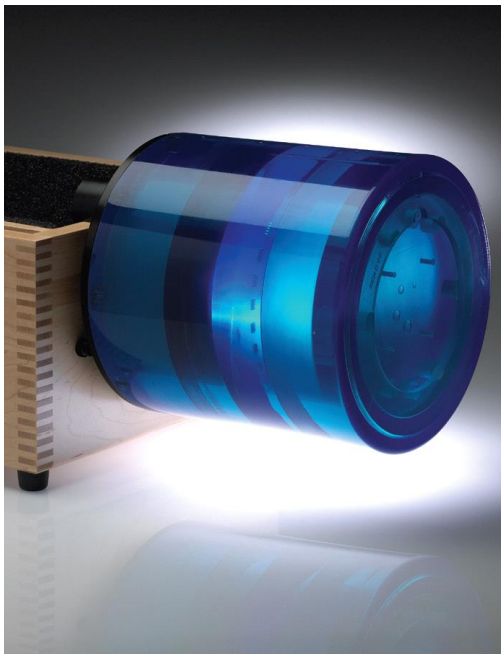


Image 4. Phantom device - photo by manufacturer [50]

Either if the following statement could sound obvious, better to underline. If the Hospital were you pursue the experiments has several phantom items available, be sure to use always the same one. Even if they belong to same model, differences could arose from manufacturing process.

Image 6. [45] This image shows a transversal view of several modules (containing diverse density materials and shapes) assembled inside the phantom cylinder. Given the fact the color of the cylinder is blue, here an underwater-suited camera (Olympus TG-1) was used to obtain this artistic view, where we can clearly see the details inside. In the clinical practice, however, those modules are scanned by x-rays, and displayed as grey levels.

In this study, we will focus in the low-contrast module of phantom.

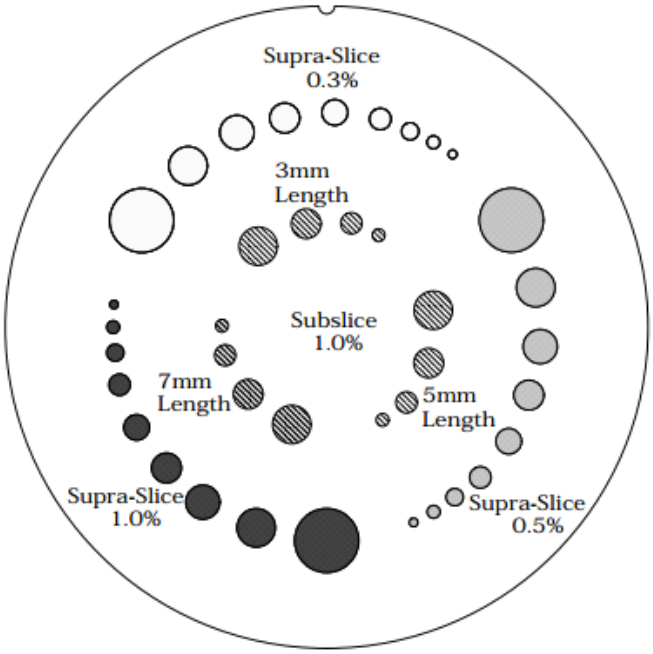


Illustration 1. Low contrast Module of Capthan ™ 500 phantom [50] Page 28

Nominal target contrast levels (0.3%, 0.5%, 1.0%) will be referred in further discussions as contrast serial 3, 2 and 1, respectively.

Subslice module are made by spheres, and *Supraslice* module are made cylinder-shaped. The intersection area at the center of the contrast module - marked by an upper dot - produces an equal circle area for both options. Subslice module is useful for the comparison of contrast properties between thinner and wider CT slices.

Supra-slice target diameters
15.0 mm
9.0 mm
8.0 mm
7.0 mm
6.0 mm
5.0 mm
4.0 mm
3.0 mm
2.0 mm

Table 1. Supra-slice target diameters

4.4 CONFIGURATION OF PROTOCOLS AT WORKSTATION

Manufacturer software, allows to save protocols, for an easier replay at our phantom study. This however, can represent a benefit in time of operation when dealing with high-risk patient, but on the same time, a radiation-dose risk if parameters are not carefully revised before each scan.

Image 6. Working station Computer Display. [45]

4.4.1 Collimation of detectors

The key point to compare both scanners, is to match the collimation. Each device has different detector size, but we can play with easy mathematics, integrating their outputs in order to obtain a similar detector coverage. For every collimation, several energies are selected, and for each Energy, several filters. There might be also some differences in SPECTRA between the two scanners (as mentioned in *Section 3.6.2*)

GE has a collimation of 40 mm (as 64 detectors of 0.625mm) while Toshiba has a collimation of 32 mm (as 64 detectors of 0.5mm). GE Detector coverage of 40mm, allows the device to cover all the area of our phantom module in just one rotation.

Some author could state that the use of a narrow collimation mean decreases geometric efficiency, but this drawback is no more present from 16-slice MDCT [1].

The lack of prevision regarding this point, made us to repeat some scans at Toshiba - in order to obtain new data of 5mm slices with a different collimation scheme, feasible for comparison with GE data.

4.4.2 Dose parameters

We will trust the values provided by console data. Other studies, can incorporate incorporated thermo-luminescent detectors at phantom, but its operation is quite risky for practitioners. High dose levels are protected by an administrator password at software.

Patient dose can be modulated at every rotation (along z axis) or at each sequence of rotations, depending on the scanning area (in example head, hips). In any case, we suffer from a 3D effect from scanner geometry: the phantom (or a patient) is over-exposed in the outer area and in the middle.

Carefully monitor of tube current setting is advisable, as the photon flux will improve with increased x-ray exposure. Tube current diverse range of values could lead to confusion, when comparing axial and helical scanning methods.

In the successive, when showing graphics comparing diverse scanning methods, we will use the *CTDI-vol* [24] [52] dose index, instead to refer onto tube current values. The reason behind, is that for helical and axial configurations, we have tube current settings quite different, that however correspond to the same *CTDI-vol*. This index (a mean value) represents the radiation dose of a single CT slice [53] and is calculated by "dividing the integrated absorbed dose by the total beam width" [54]. In practice, our selected opearting *CTDI-vol* values of (10, 15 and 20 *mGy*) could also be referred by a nominal scale: *low, medium and high dose level.*

4.4.3 Scanning mode: Axial or Helical

"Adaptive collimation" was first done by Toshiba [1]. After that, other brands adopted the method, but under a different name.

Spiral (also named helical) mode needs additional rotations at the beginning and the end of the scan - in order to obtain data, to reconstruct images over the prescribed volume – BUT has some possibilities of dose saving when the patient does not cooperates, because scan time is much shorter. [30]. Toshiba uses an "active collimation" method in order to reduce the over-scanning at end, claiming to reduce patient dose up to 20% [55].

For a reduced scan length, the helical mode will be inefficient regarding dose reduction. For every rotation, there is an excess radiation (so is not advisable to do so many rotations). Helical CT can also reduce motion artifacts, and is needed for some reconstruction methods (as VEO by GE, that will be detailed later).

4.4.4 Filtering setup

At Toshiba, FC18 vs. FC08 (the first corrects the beam hardening, while the later not). As stated in Section 3.3.3, when X-rays passes through the phantom, the photons with lower energy are

absorbed more rapidly. Hence, X-rays passing through the middle portion of a uniform cylindrical phantom are hardened more than those traversing the edges. At GE, we can reconstruct with two different filters: lung or standard. In this study, we will use only Standard data.

4.4.5 Slice thickness

Comparison between thinner and wider slices can be studied using the *subslice* module of the Capthan phantom (and we could see that those wider are less noisy)**.** If we look onto Illustration 1, we can see that *Subslice* module are spheres, and *Supraslice* module are cylinders.

According to Boas [19], for a conventional FBP image, the the standard deviation due to Poisson noise is proportional to: sqrt (1/(slice thickness * tube current)). Noise hence decreases with slide thickness. In this study, CNR data is obtained for 2.5 and 5mm width slices.

4.5 ENTERING THE CT ROOM

4.5.1 Safety measures

Exposure limitation: The limit of radiation exposure for medical workers is limited to 20 times the one for external users (being this last 1 *miliSievert* per year) and monitoring is compulsory. We used a personal Film-based Dosimeter [56] provided by NRPA. Even if the CT scanner is not working while we are in the room, safety measures requires to wear it.

4.5.2 Alignment procedure of Phantom

With the help of alignment Lasers, and some indent marks, painted in color white at Phantom

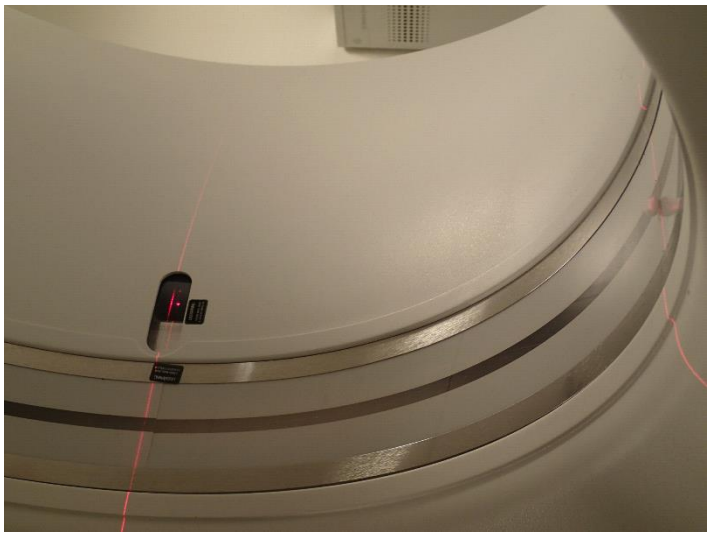


Image 7. Lasers for phantom alignment [57]

It is very important to have phantom placed at gantry isocenter. Once the phantom is placed, we need to go out of the room, be sure to close the door, and activate the first protocol at computer.

4.6 ACTIVATION OF PROTOCOL SETTINGS

When working with Toshiba *Acquilion*, we should have the following setup, to be set in the software console (AutoQALite v.3.00)

First we choose "phantom" instead of "body"
Axial scan: 120 kV. 1s/rot. 3mm thick. 2cm distance to each side of white dot.
Start location S18.750, end I18.750
As Scan FOV we select "Large body". Most scanners only have this option: by choosing this, we can compare data between brands.

When operating with GE Discovery, The selection of an slice width, corresponding with the button Interval (mm) at console, affects reconstruction step, but not the scanner itself [In example, we settled 0.625mm as original slice width in GE scan]

4.7 REPETITION FOR THE PHANTOM WITH RINGS

We use some annulus ovals (in the successive, referred as rings) simulating fat properties. Manufacturer Codes are: CTP579, CTP651, and CTP599. [50]. In order to assembly them at phantom, a water-based gel is used – so it does not affects scanning.

4.7.1 Increasing Field of View

The scanner uses a different FOV (increasing) for a different scan size (increasing). For each of the discs, we use DFOV 200, 360, 400 and 500, respectively. Exact DFOV numbers may slightly differ among our scanner brands: in example 210, 365.

As well stated by BOAS [19] when a Field of View is set smaller than the object size, "most of modern scanners set the values of the band of interest onto a constant boundary value" in order to avoid artifacts once the image is reconstructed from *sinogram* (else the final image would be surrounded by a white circumference). Another solution, could be to scan a slightly higher FOV than the one of interest, but this one implies more dose to the patient. In our case, as working with a phantom, both options could be feasible.

Figure 5. Sinogram images of a scan with higher FOV than the object size. Left one is the original processing, while at right one an easy trick has been applied for future convenience. [19]

At Image 8, we can see the phantom with the smallest of rings () that will be scanned with a DFOV200.

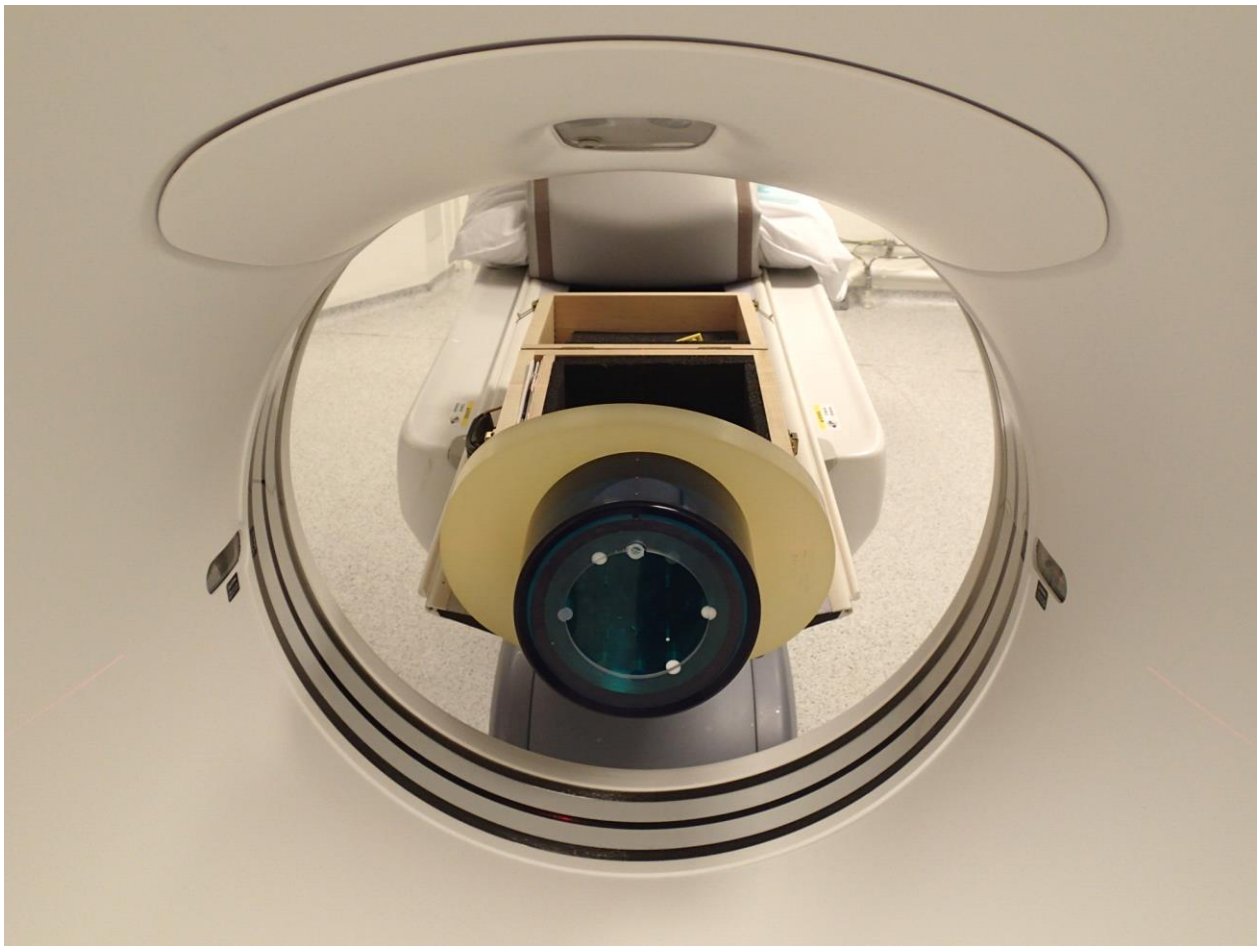


Image 8. Phantom with the smallest Ring [57].

For the acquisition of data, we leave each time the operations room, being sure to close well the door.

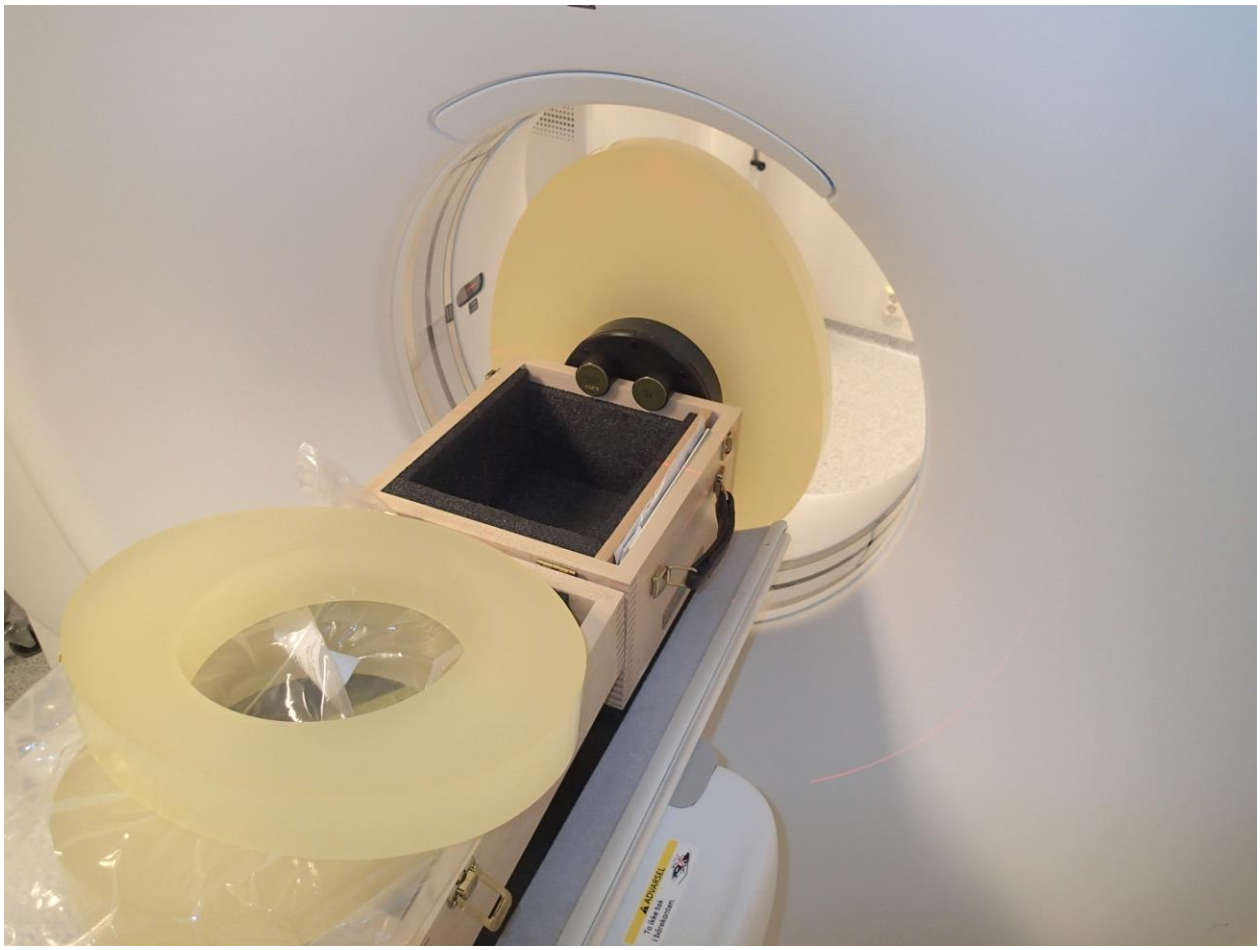


Image 9. Operation with the egg-shaped ring (DFOV400) [57]

When working with bigger rings, is important to do balance of weight – with the help of phantom box - when operating with the big ring module: nobody wants the phantom to accidentally fall onto the detector gantry, the most expensive component at a CT device.

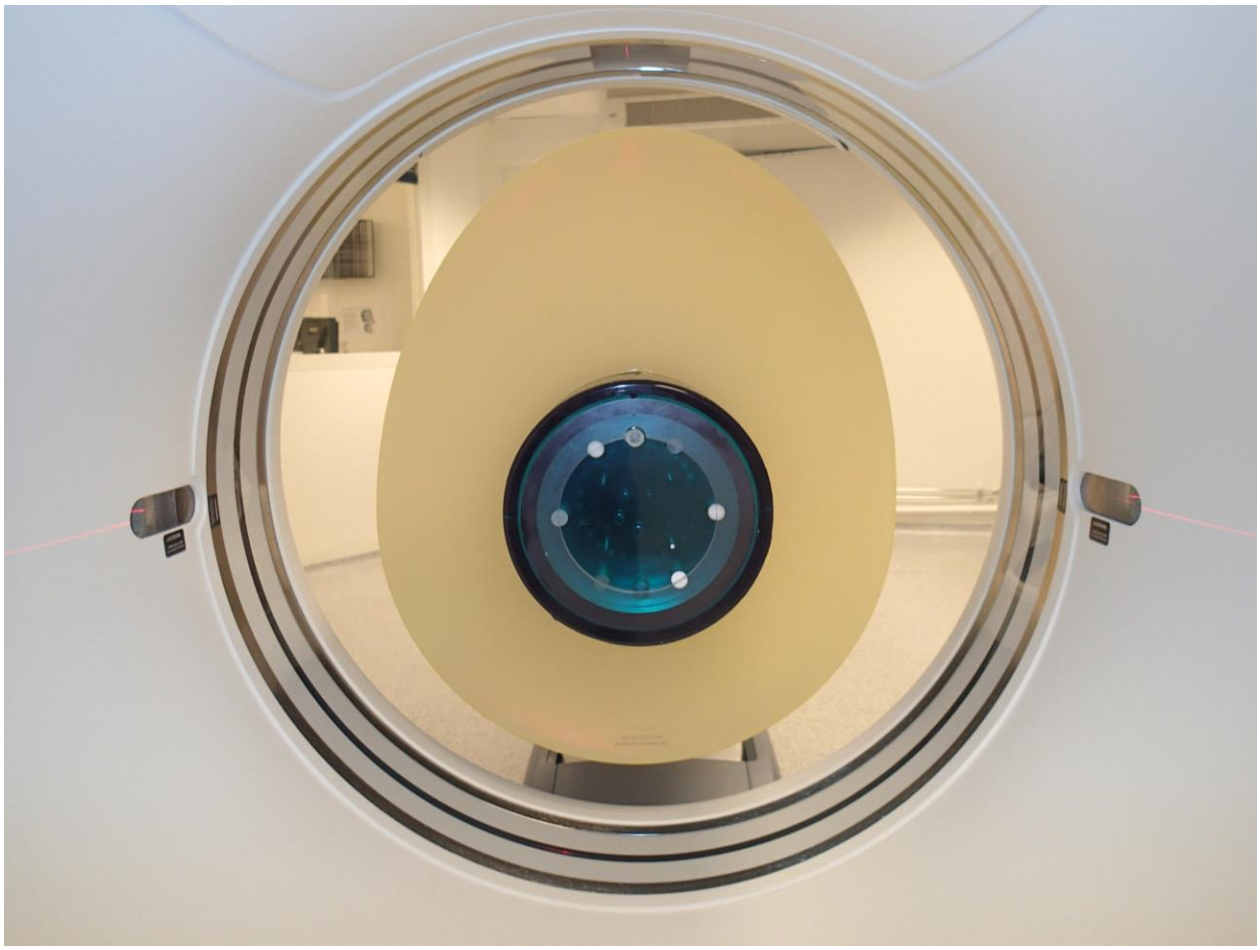


Image 10. Front view. [57]

4.8 ITERATIVE METHODS

In addition to the usual FBP reconstruction scan, we performed the following methods:

AIDR (Adaptive Iterative Dose Reduction) 3D at Toshiba, that works in the processed image. [55] In this project, we combine 30% iterative / 70% FBP, usual configuration for abdominal scan. [15] (Hsiang book, 3.6)

VEO at GE. This is the first Model-Based iterative reconstruction product, and it works with RAW DATA instead of the final image. It applies a statistical model of the physics behind the scanner (as can be, the particular detector technology, or optoelectronic converters), as a noise reduction strategy, before doing the back-projection of data.

It is a computationally expensive algorithm, as it includes a statistical model of the physics behind the GE scanner. In order to work with VEO, we need helical scan. It is not possible to proceed with those images obtained with axial. – During the development of this project, we obtained axial images with GE scanner that were later discarded. It is important to pursue an adequate Design of Experiment (DOE) from the start.

As advantage, VEO may enable improvement in low contrast detectability and spatial resolution [58] but as a drawback, it takes much more time to get the image ready (around 30 minutes, for one phantom scan).

Image 21. Protocol picture, when working with VEO. [45]

4.9 POST PROCESSING

Once images are ready, from the available track of images, we will only use those containing the low-contrast slice center, marked by the white dot. This means a discard of around (8 images of each 9 scanned). Several DVD-discs were reduced onto a few MB of valid data.

4.9.1 Contrast to Noise Ratio

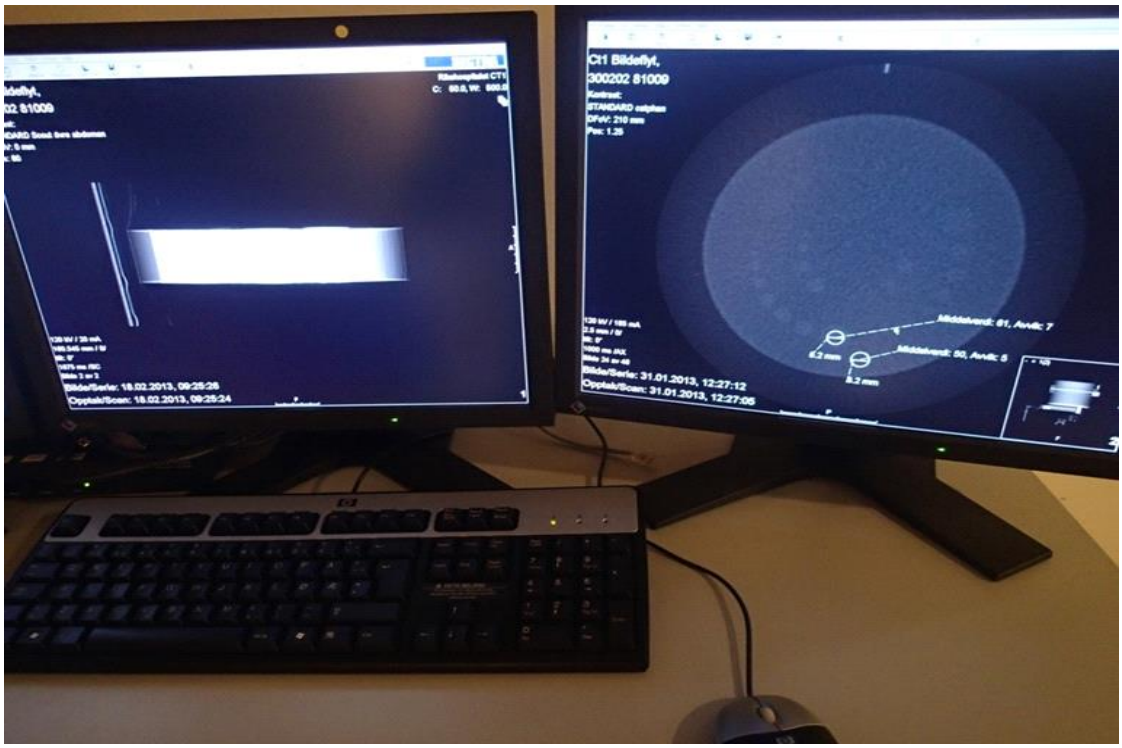


Image 13. Caption of CNR calculation at Oslo Hospital computers. [57]

PACS-SECTRA ™ software was used [59] [60] at Oslo Hospital. At this stage, we discarded to use those programs working with DICOM files but not approved by law for medical use.

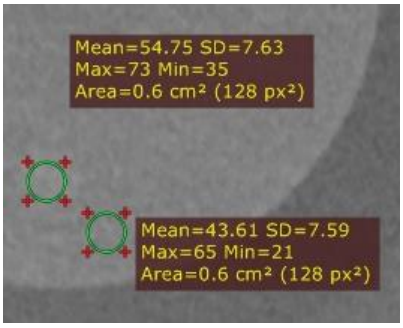


Image 14. More didactic Illustration, from another commercial software.

CNR Formula is detailed at Section 3.6.3. It is calculated for the biggest circle (either from a cylinder or a sphere projection, of each *supraslice* and *subslice* serials, respectively). The reader can here refer back to illustration 1. Having three different density material serials, this make a total of 6 CNR measurements per valid slice. *Capthan* manual states that the difference

between foreground and background should be 10 Hounsfield units, however inter-phantom and inter-scanner variations, can make this difference to be up to 12 HU [61] [62].

Regions of Interest (ROI) were set according to practitioners guidelines, always avoiding the edges area of circle. By my own initiative, and for higher coherence, I established a constant ratio between the areas, for the measurement of sub-slice areas (of a lower diameter). Some authors as [36] had reported the convenience to use an area of 200 squared mm for *supra-slice* serials, while sub-slice it could performed with a 60 squared mm ROI [63]. Nevertheless, the most important thing is to maintain the ROI diameter the same, for the background and foreground selection, and among the full dataset.

Some authors [64] normalize the image noise with regard to the effective dose, to obtain a figure of merit (FOM). Effective dose, is the one measured inside the scanning room, and makes use of a high-risk procedure based on fluoroscopy detectors. Oslo Hospital management, limited this possibility to some stages of calibration. For the repeated scanning of the phantom, we trust the values adjusted at the protocol screen, to be reproduced accurately enough by the scanning device.

4.9.2 Digital Imaging Communications in Medicine

DICOM standard [65] [66] (acronym of Digital Imaging Communications in Medicine) was studied, in order to obtain or review scanning parameters from its metadata. Most of manufacturers had accepted this standard since its creation in 1980.

DICOM developers are currently working in order to include a more detailed registry of dose parameters [67] at future actualizations, as a response to the increasing relevance of ALARA strategies in the medical imaging sector. [68]

4.10 SUBJECTIVE EXPERIMENT

Even at a simple display configuration, we cannot avoid to consider facts from human perception, not measured directly by CNR metrics. In our case, the observation of the images is prone - even when pursued in very controllable conditions - to the simultaneous contrast illusion, or to light adaptability process of human vision.

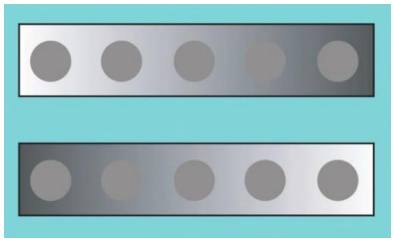


Image 15. Illustration of Simultaneous contrast effect [69]

As stated by Troscianko [69] "All grey circles have the same luminance, but those against a dark background look lighter and vice versa"

We decided to design an observer experiment, in order to complement CNR data, and joining Colorlab Facilities in Gjøvik, Norway. The initial scope, was to find the finest diameter observers are able to see, but this concept extended onto a deep study of statistics by different groups data, and a more complete brands comparison.

In our study, Toshiba and GE scanners internal configuration, allowed us to reconstruct slices of 5 mm in both, but for thinner slices, we would need to compare 2 mm slices from Toshiba width 2.5 mm from GE scan. Sagara [36] states that "difference less than 1 mm had been reported to be not relevant regarding image quality scores". In any case, the 0.5mm difference does not become automatically negligible, and for the comparison of images by observers we only displayed slices of 5 mm width. In order to verify Sagara statement, we should be able to obtain diverse width from the same manufacturer (and not from two diverse ones).

4.10.1 Display screen

A 12-bit depth can have 2 to the power to 12 (4096) pixel values (grey levels) [32]. CT x-ray images can range from 12 to 16 bits, so we need to consider special quality displays. Toshiba Acquilion One, in example, can provide up to 14-bit images output.

The selected display was of Liquid Crystal Display (LCD) technology. Model *ColorEdge™ CG246* by Eizo™J apan. Either if this screen incorporates a self-calibrating detector, an external and more precise module (EyeOne™ by X-Rite™) was used.

An advanced color display as the detailed is adequate, even if some monochrome versions had been designed for Hospital environment. Given the actual state of display technology, a color display can reach a number of intensity levels, equivalent to the best grey-level display just a few years ago. [70]

4.10.2 Illumination conditions

Some studies [70] for comparison of color and monochrome displays, used an illumination value of 23 lux. It could be considered high, in relation to the usual standard of darkness (2-3 lux) but is more close to the real operation environment of doctors. From a strict perspective, white coats Doctors use to wear could be also discouraged in the sense they reflect significantly in the digital screen [71].

To be in a total dark room (the ideal case) is far from reality of medical environment, were doctors usually analyze images at different illumination places among his journey, due to time limitations.

In this experiment, we are closer to the real conditions of medical daily practice, than other studies in radiography. Despite this differences, experiment was done at a closed room, with very low level of illumination - that could be said to be "almost dark".

4.10.3 Images configuration

We want observers not to about dose level, but neither the brands. When working with PACS SECTRA ™ Software, blinding of images can be done by accessing in the menu "Visualize: Image Information: Anonymize". 45 blinded images were used (15 for each Field of View) in a pseudo-random order, and displayed without zoom. When playing with low contrast, windowing is determinant.

According to [60], window width adjust the range of pixels sent to the display system, and window level (named C in reference to center) sets the center of window width. Changing the display contrast parameters (W and C) alters the appearance of the CT image, without changing the reconstructed image data (Hounsfied Units at Picture) [12]. In this experiment the window settings were chosen for abdominal viewing. (Window level C= 40, window width W=300) [72]. Distance to screen was approximately 50 cm.

4.10.4 Procedure

Conditions of observers experiment were hence set according to the following:

- Color display EIZO (CG246)
- Display luminance set to the maximum value.
- 45 blinded images used (15 for each Field of View) in a pseudo-random order.
- Displayed without Zoom.
- Room luminance of 30 lux.
- 17 observers: 7 from 3rd course of Radiology, and 10 image processing students.
- FC08 and DR08 not included at observers experiment.

Sofware ViewDEX [73] was initially considered, as it includes time spent by the observer as a parameter, however was later discarded. Dr. Martinsen [1] suggested would be better used in a PhD study. MicroDICOM ™ for Windows [74] can be used as a simple but complete viewer interface. Macintosh OS options were also explored at Colorlab, but finally not implemented.

Observer´s visual acuity was checked with a "Dynamic Illegible" table. A *Landolt Ring* test is also valid for this purpose. They were given instructions in English about the experiment, including a drawing of the phantom module without further annotations. The justification to show the module image, was to reduce the possibility of observers could start to believe to see circles outside the real locations.

It happened in the history of Medicine, at the time it was believed to be known a number of lobes to be in the liver - that resulted to be wrong. Counting and perception was broadly influenced by this assumption until the later founding. For the reader interested in this field, an historical review about Medicine Illustration evolution is referenced [75].

The time spent by experiment participants in the lecture and fill in of form (about age, ocular pathologies if present, and some general questions as studies background) was also used to allow some dark adaptation of visual system. The total time employed to evaluate the set of 45

images took around 20 minutes for each observers. This time was taken as a threshold value in the design of experiment, to avoid eye fatigue and blurriness (apart from observer associated fatigue). Some recommendations refer to a maximum time of 30 minutes [76]

Carl Zylak mentions [77] [39] the following rule known as 20-20-20: "Look away from the monitor at least every 20 minutes and focus on something over 20 feet away for 20 seconds" In our experiment, three short 10-second pauses where established, where the observer does not look into the screen. Another possibility– that was occasionally used - is to show the same image twice - in two different moments of the experiment. However this option, is limited by the 20-20-20 mentioned rule.

4.10.5 Statistical Analysis of data

It includes size adjustment, according to the actual overweight trend in population in some countries. I applied different configurations of rings to the phantom, in order to simulate different patient weight. It also considers different noise filter possibilities (as beam hardening reduction), and several reconstruction options for the images, as can be iterative reconstruction from the constituted image, or directly from raw data.

Image quality in diagnostics, refers to Diagnostic quality. We take special attention onto the confidence intervals and the standard deviation value among observers. Even if the last is a simple statistical parameter, here is very indicative as two different interpretations of the same image, could lead onto a different diagnose by Doctor. We explored the differences between the answers of Image Processing students, and Radiography students, and the distribution of data according to a normal (Gaussian) distribution, through some statistical methods as *Mann-Whitney* U-Test, *Anderson-Darling* Test, and *Student* t-Test.

5 RESULTS & ANALYSIS

5.1 CONTRAST TO NOISE RATIO

5.1.1 Overview

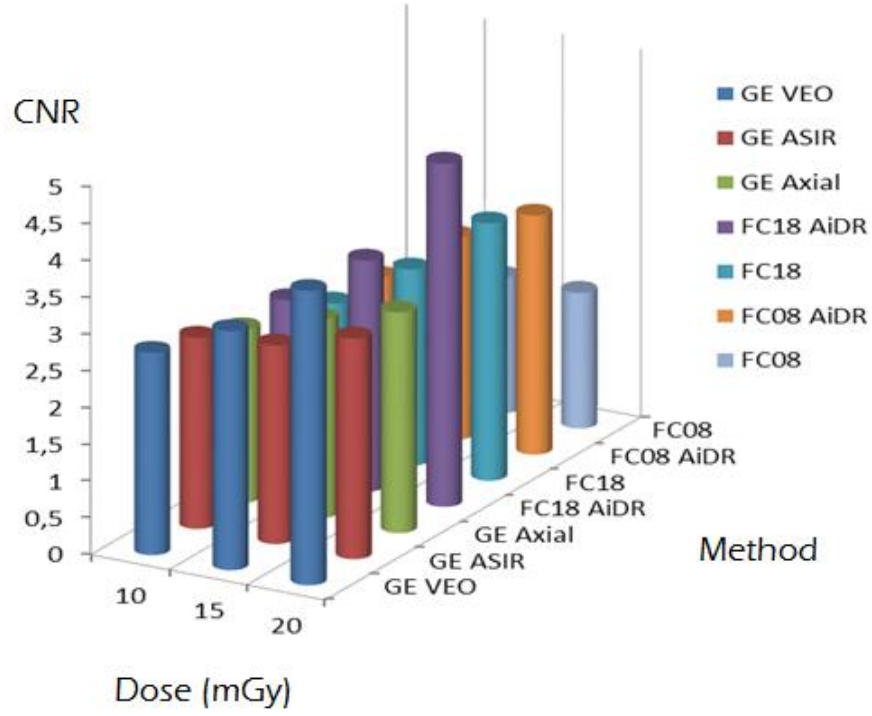


Figure 6. DFOV200 Contrast Media 1. 5mm slice width.

The biggest CNR is obtained with FC18 AIDR. But this information is not the goal of our research, as it denotes the better method for the higher dose. When our final scope is to reduce the dose, we need to look onto methods performance at a medium dose level and the lowest one. Either we could wonder, do we need so high CNR in order to discern the relevant features?

For contrast media 2:

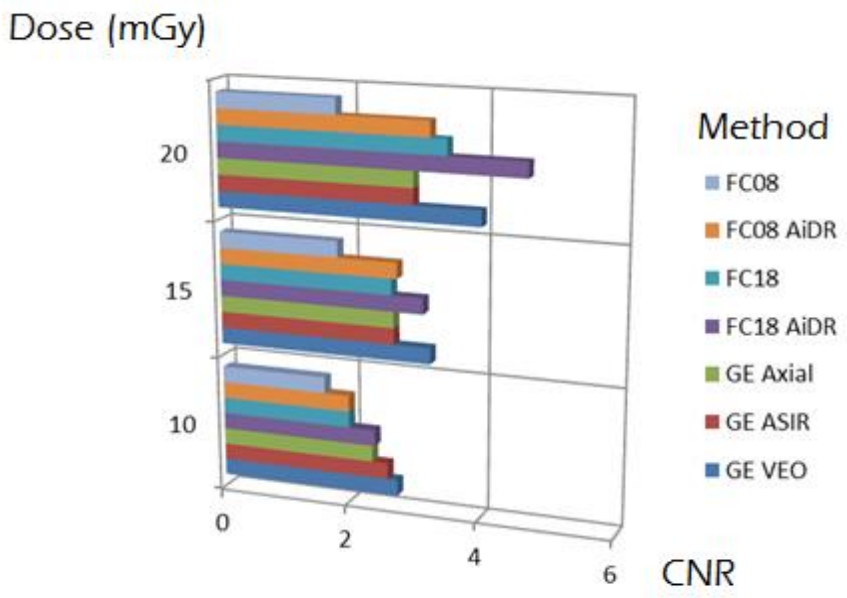


Figure 7. CNR values. DFOV200 Contrast media 2. 5mm slice width.

In an alternative display:

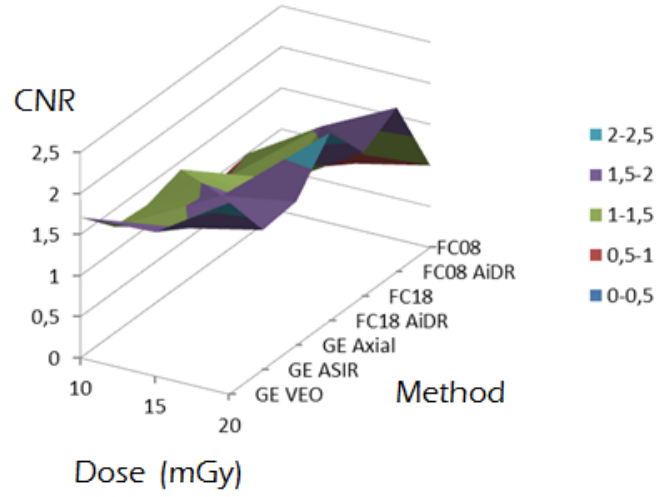


Figure 8. CNR values. DFOV200 Contrast media 2. 5mm slice width. Alternative display with CNR displayed in different colors according to interval range.

In the case of NO RING, Toshiba provides the higher CNR with 20*mGy* (with the beam hardening filter FC18 and Iterative reconstruction AIDR) while GE does it locally at 10*mGy* dose index. For the medium dose case, methods performance is similar.

If we display CNR values of contrast media 1 and contrast media 2 as topographic map, we can see similar peaks distribution, but a 50% increase at CNR scale for contrast media 1.

5.1.2 Comparison between iterative and non-iterative methods

If we cluster iterative and non-iterative methods:

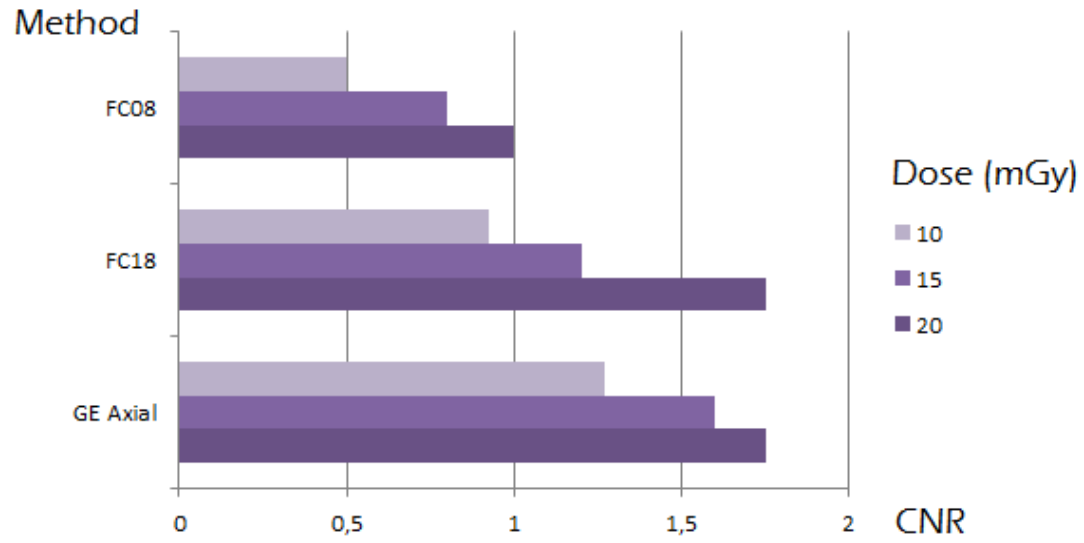


Figure 9. 5mm slice width. DFOV200. Contrast media 2. Non-iterative methods, clustered by method

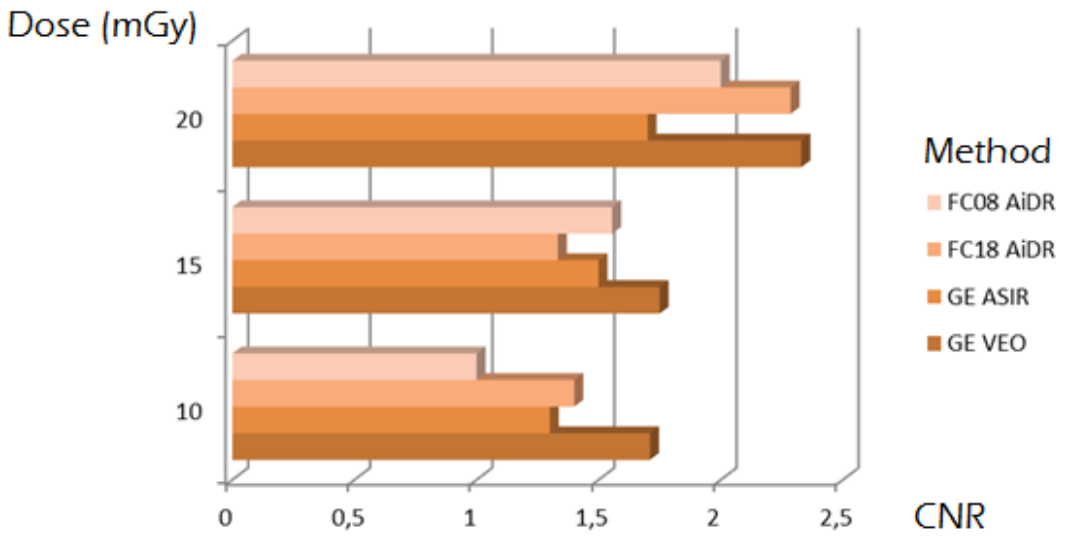


Figure 10. 5mm slice width. DFOV200 Contrast media 2. Iterative methods. Data clustered by dose level.

Toshiba with Beam Hardening, overcomes GE ASIR. This two methods are comparable in time efficiency. Highest CNR is for VEO, well intended, with the drawback to be a method computationally expensive. In the medical sector, where time is precious, having just that little difference in CNR, we could consider to use FC18 AIDR, being this most efficient. Regarding 10*mGy* data, we can see we don´t have the nice downstairs shape as happened for non-iterative methods.

Just as an example, if a CNR of 1.5 would be enough for our detectability purposes, FC08 AiDR by Toshiba and GE ASIR, both at medium dose, would be a reasonable choice. But if we have the time for processing images, GE VEO at low dose will be more considerate with patient health.

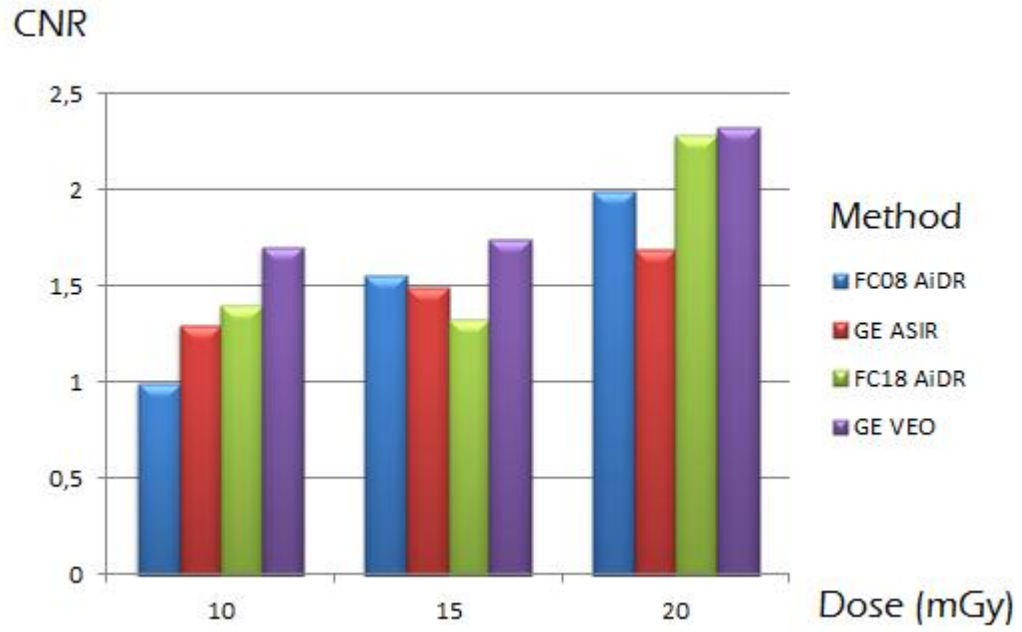


Figure 11. 5mm slice width. DFOV200 5mm. Contrast Media 2. Methods ordered according to its performance at low dose case.

If we order the methods at increasing CNR for the lowest dose case, can be seen the relation is not maintained for medium and high dose. One reason is that some of this methods had been specifically designed to improve low-dose configuration, another is stochastic nature of radiation interaction with matter.

Phantom with the little ring (DFOV360)

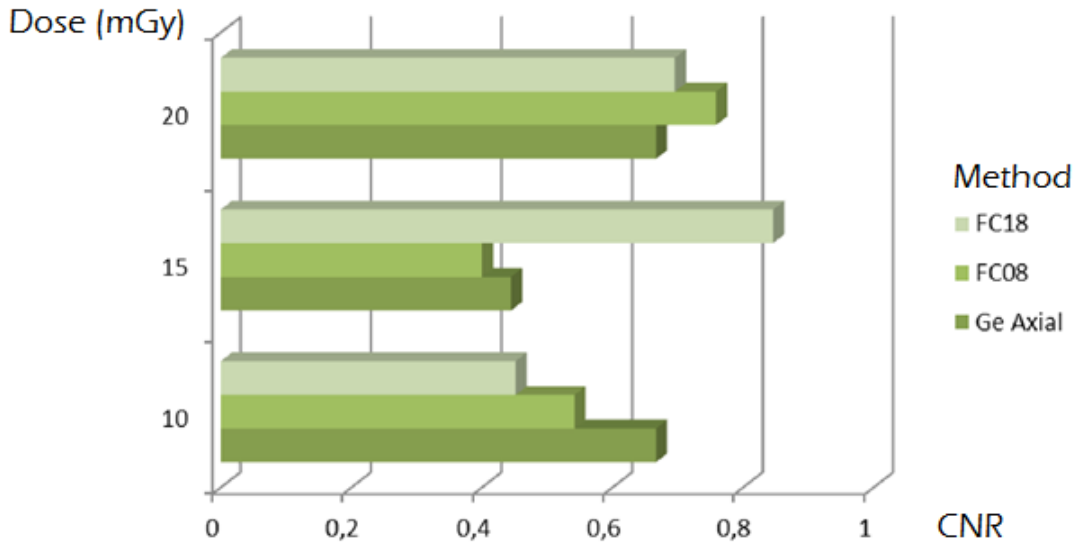


Figure 12. DFOV360. Contrast media 2. 5mm slice width. Non-iterative methods.

According to *Figure 12*, Toshiba with beam hardening (FC18) provides the highest CNR value at medium dose, among non-iterative methods.

The step difference of CNR values from ASIR to VEO is much higher at medium and high dose, than it is for low dose. In other words, the change from ASIR into VEO at low dose, gives us an incremental in CNR value of less than 0.5. Knowing the time required for VEO processing is considerable, is that worthy? In order to contribute to some of this question´s solution, an observers experiment was designed.

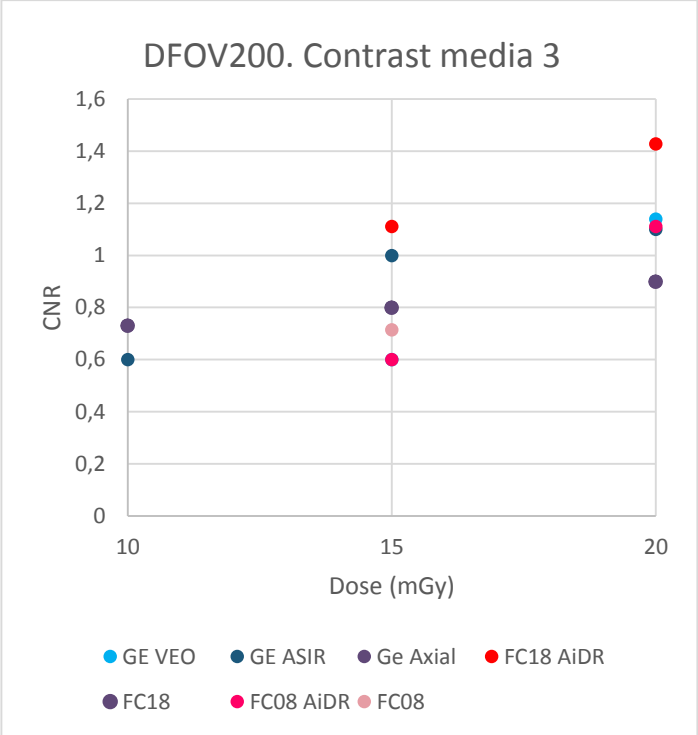


Figure 13. CNR Values for contrast media 3. Phantom without ring.

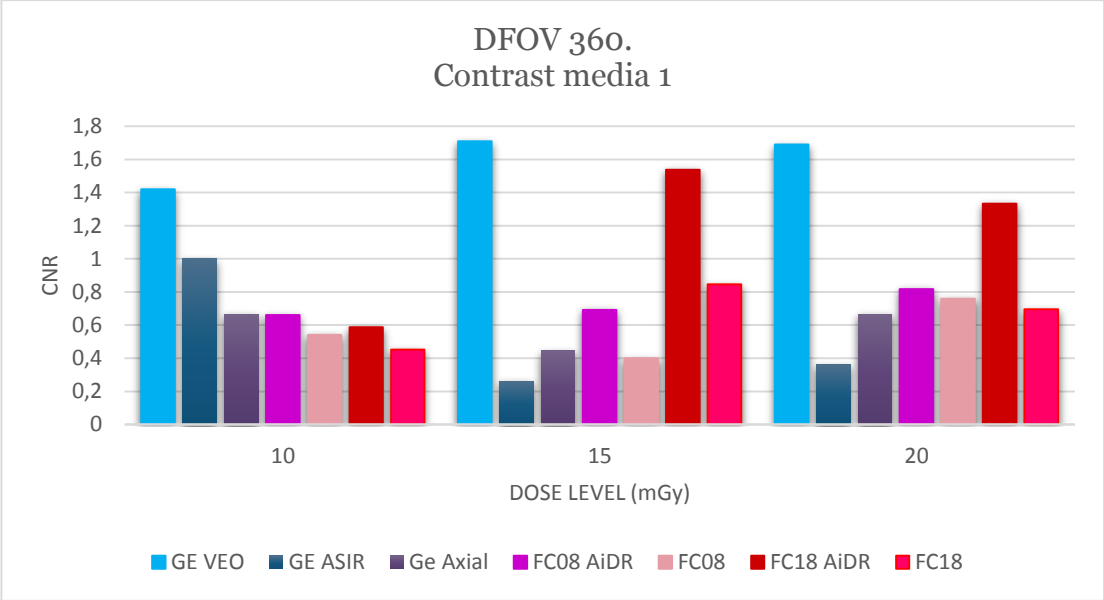


Figure 14. CNR values for contrast media 1. Phantom with littlest ring.

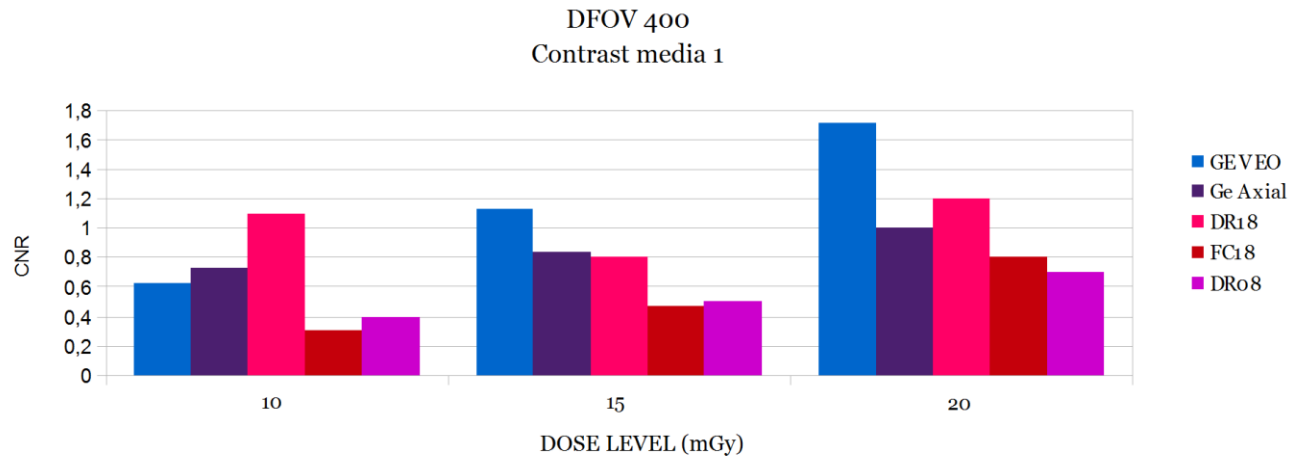


Figure 15. CNR values for Contrast media 1. Phantom with medium ring (DFOV400)

As shown in *Figure 15*: For low dose, DR18 (contraction used for FC18 AiDR) provides a higher CNR than VEO of GE. CNR values by VEO improve for medium and high dose level images.

When exploring data of DFOV400, we found interestingly that higher dose case does not perform better than medium dose one, for detecting contrast 2 and contrast 3 materials.

5.1.3 DFOV 500

We discarded those slides for processing after consulting with Oslo University Hospital practitioners. We performed every scan configuration, but at the evaluation step we founded even with the best algorithm, photo starving was too high. See Annex 4 for details and example.

5.1.4 Slice width and CNR

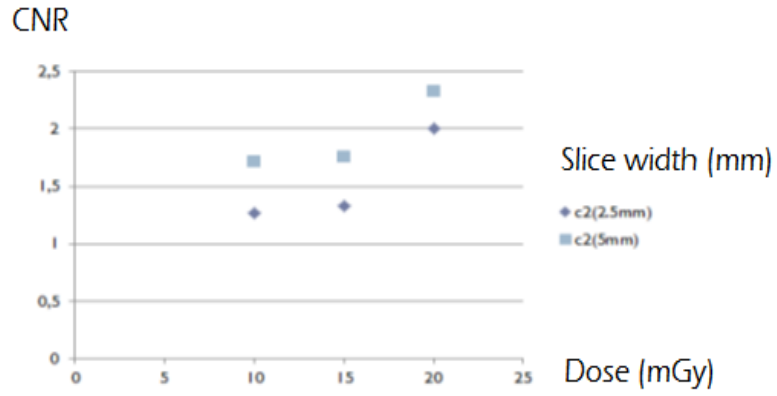


Figure 16. DFOV200. Contrast material 2. VEO method

At *Figure 16* it can be seen the straight-forward relation: "more thickness, less noise" (detailed at section 4.4.5) for VEO model-based reconstruction, and a width difference of 2.5mm.

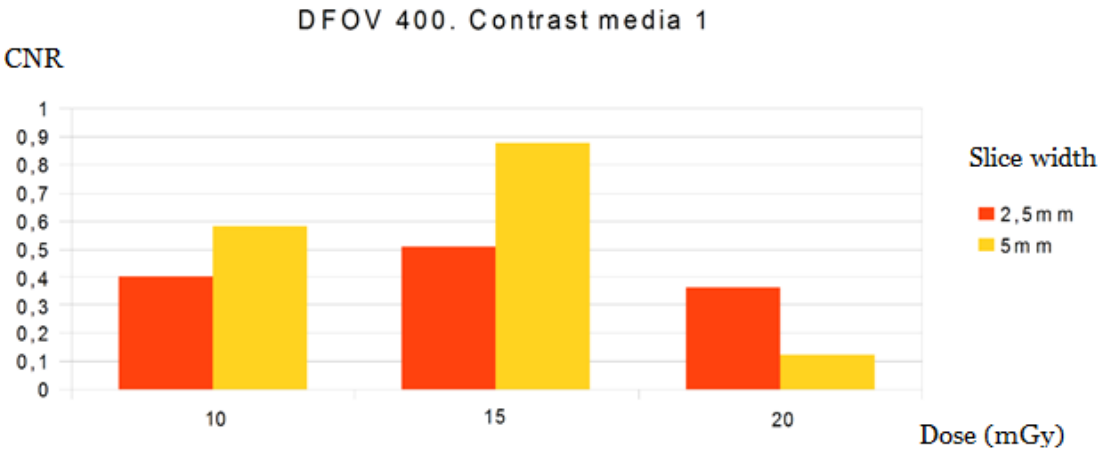


Figure 17. Standard Axial Data. DFOV400. CNR values for different dose levels and slice width.

5.1.5 Axial and Helical options. Subslice and Supraslice data.

For a phantom without ring, CNR values at helical case are lower than CNR values at axial. Here is shown the data for 2.5mm slices. A denotes "Subslice Helical" B: Subslice Axial. C: Supraslice Helical. D: Supraslice Axial.

Contrast media 1

	A	B	C	D
10 *mGy*	1,18	1,71	1,33	2,6
15 *mGy*	1,33	1,86	1,47	2,16
20 *mGy*	1,43	2,18	1,84	2,66

Contrast media 2

	A	B	C	D
10 *mGy*	1,37	1,57	0,75	1,4
15 *mGy*	1,33	2,16	1,07	1,45
20 *mGy*	1,71	2	0,86	1,4

Contrast media 3

	A	B	C	D
10 *mGy*	0,36	1,65	0,47	1
15 *mGy*	1,29	2,4	0,29	0,72
20 *mGy*	1,57	1,33	0,67	1

It can be seen how CNR(A) < CNR (B) and CNR (C) < CNR (D) at every case.

For the supraslice case, the lowering factor is a 0.50 [CNR(C) is around the half of CNR(D)], while for subslice it is generally around 0.70. This relation is maintained for DFOV360 at low and medium dose. At DFOV360, contrast 1, high dose, the factor is 1 instead.

		DFOV360		**DFOV400**	
Dose level	Contrast serial	Helical	Axial	Helical	Axial
Low	1	0,45	0,875	0,4	0,17
	2	0,34	0,28		
	3	0,15	0,23		
Medium	1	0,43	0,77	0,51	0,44
	2	0,19	0,42		
	3	0,14	0,23		
High	1	0,84	0,8	0,37	0,6
	2	0,26	0,24		
	3	0,22	0,27		

Table 1: CNR Data processed from 2.5mm width slices

At DFOV400, the relation is inverse, having helical better CNR values than axial (where detectable, in contrast media 1), with the exception of high dose case.

5.2 OBSERVERS EXPERIMENT

5.2.1 Standard Deviation values

It can be interesting to look, for each of the methods, how behaves the Standard Deviation on the number of counted circles. Two different interpretations of the same image, could lead to a different diagnose.

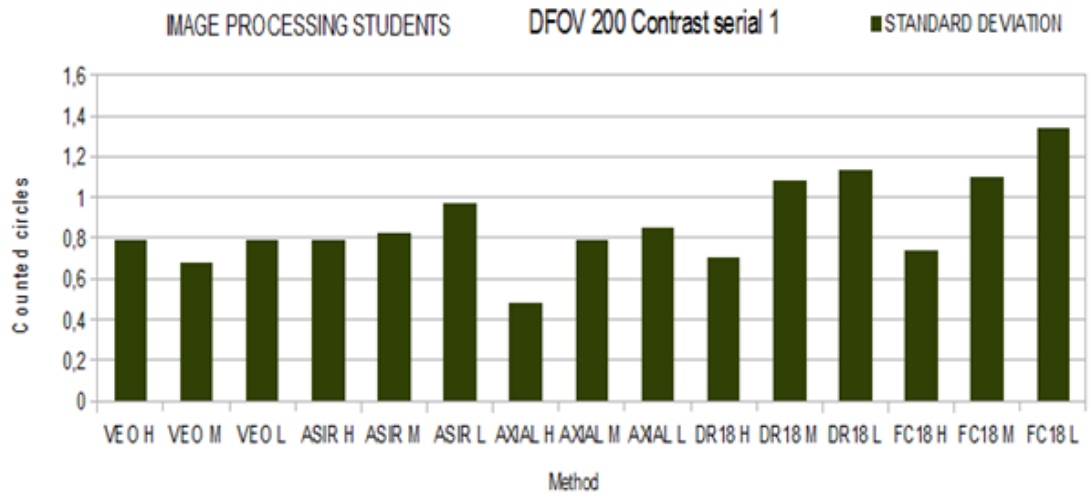


Figure 18. Standard Deviation of counted circles (DFOV200, Contrast serial 1, 10 observer dataset)

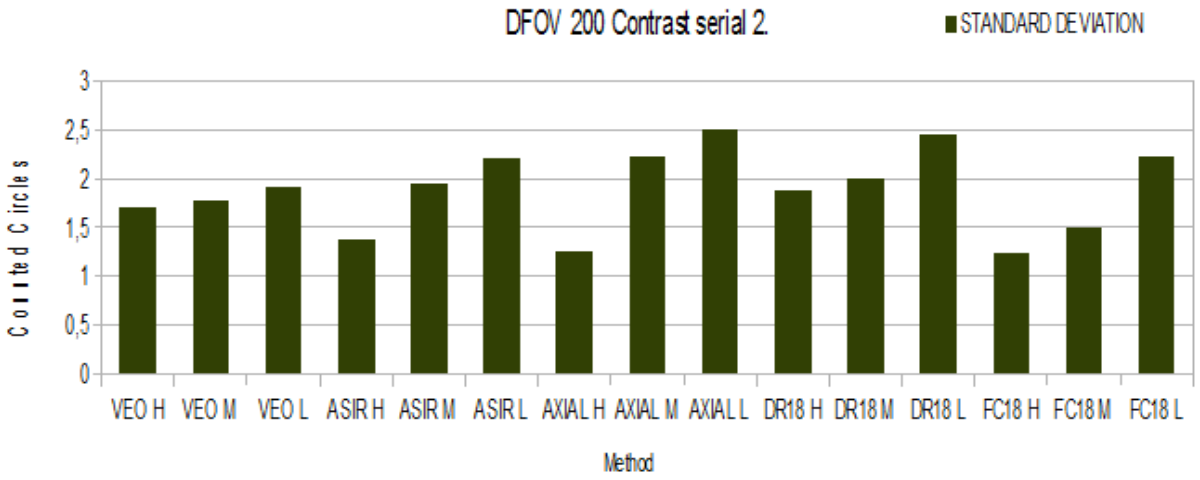


Figure 19. Standard deviation graphic (DFOV200, Contrast serial 2, 10 Observers: Image processing students)

At Figure 20, we can see a monotonous relation between dose level and standard deviation value - that in Figure 19 is also present, with just an exception for VEO method that could be considered inside the error margin.

Anderson test, calculated for the SUM of counted circles along the full experiment, for a particular contrast serial, revealed a normal distribution of the answers for both figures, as expected when we use a set of random observers.

In the case of contrast serial 3, the one of lowest density material, most of observers use to agree in seeing nothing or just one or two circles (those of bigger diameter).

Figure 16 shows how standard deviation is lower for the high dose case, at each of the methods. In other words, for high dose we have the highest agreement at the answers. This behavior, where expert observers agree much more in high contrast images, had also been referenced for color images [78].

In absolute value, standard deviation values are higher at the contrast serial 2: as it is manufactured by a material of half the density with respect to contrast serial 1, it uses to be more difficult to see.

What happens when we include an oval (also referred as ring) to the phantom?

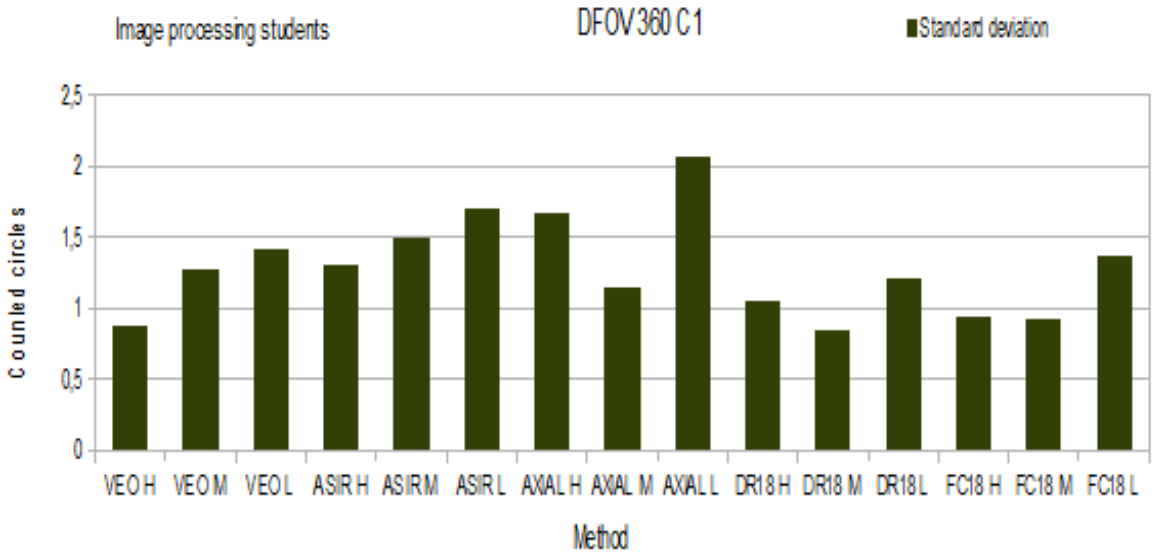


Figure 20. DFOV360, Contrast media 1. Standard Deviation values of counted circles. Image processing students.

Iterative methods by GE (VEO and ASIR) the highest agreement among observers happens for High dose case, while at basic method (ASIR) and those of Toshiba, the highest agreement among observers happens to be for the medium dose case.

For Toshiba methods, and for GE non-iterative, there is more agreement at medium dose than at high dose. Actually, some CNR measurements were better in the medium dose case than the high one case. Translated here, to have better detectability at medium low case than high dose case, implies more agreement between observers at medium low case, in other words, a lower standard deviation.

If we just look onto low-dose case:

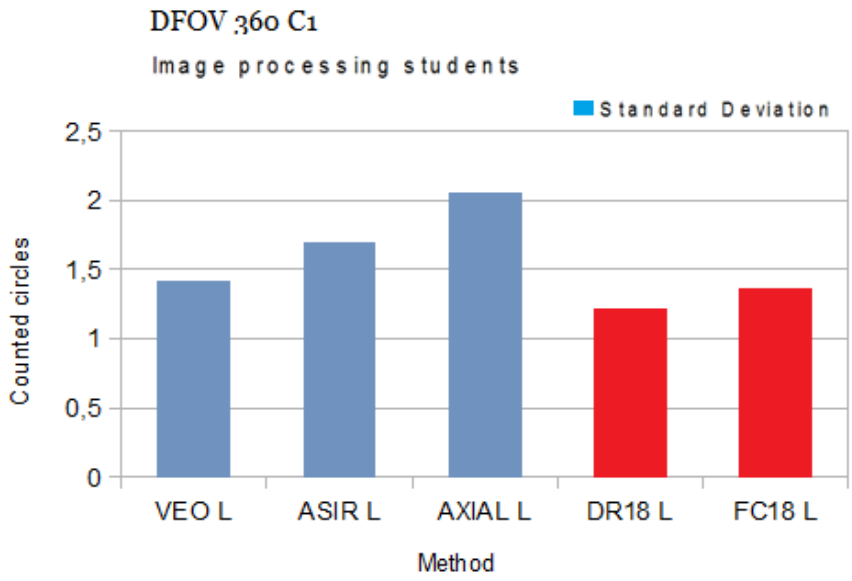


Figure 21. Standard deviation regarding low dose images. DFOV360.

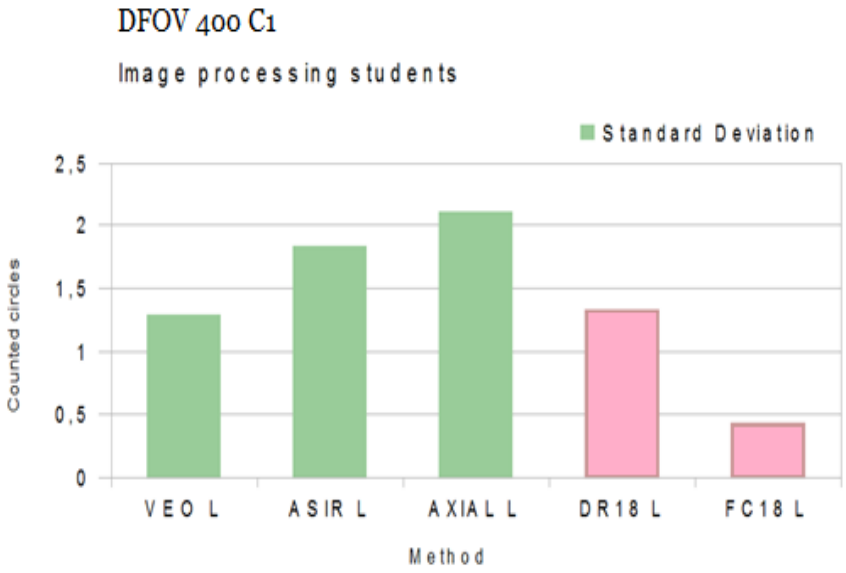


Figure 22. Standard deviation regarding low dose images. DFOV400.

DFOV400 data of standard deviation needs to be interpreted having in mind that for FC18, the majority of image processing students counted no circle at all or 1 circle, with a standard deviation value of 0.5 counted circles.

5.2.2 Cluster of datasets: Mann-Whitney Test

In a later stage of Thesis project, Radiography students joined the observers experiment. It arose the question: Until which point, there is statistical difference in their answers?

If we find no statistical difference (at a 95% of confidence) between the two datasets, we could join together all data, and to obtain statistical parameters of higher reliance. On the other hand, is also interesting to see, those cases where the answers could differ substantially.

In order to discern if radiographers provide different results - statistically speaking (that is, up to 95% confidence) – with respect to diverse background observers, I selected to use the Mann-Whitney U Test. It is often considered a non-parametric alternative to t-student test about similarity of samples, and as advantage, we don´t need to deal with the assumption of high number of data [51].

This method provides a Z-value (a measure of the distance to the normal distribution) and the U factor as a final output to compare with a threshold.

For the considered group of samples (10 observers from CIMET, 7 from radiography) we have from formula, that any U-value founded lower than 14, reveals not statistical difference between groups, at 95% confidence.

Very few cases had been founded (just one) where the U value goes below 14, but we can study the proximity or farness to the threshold, as a measure of the disagreement inside the agreement, between the 2 observer´s datasets.

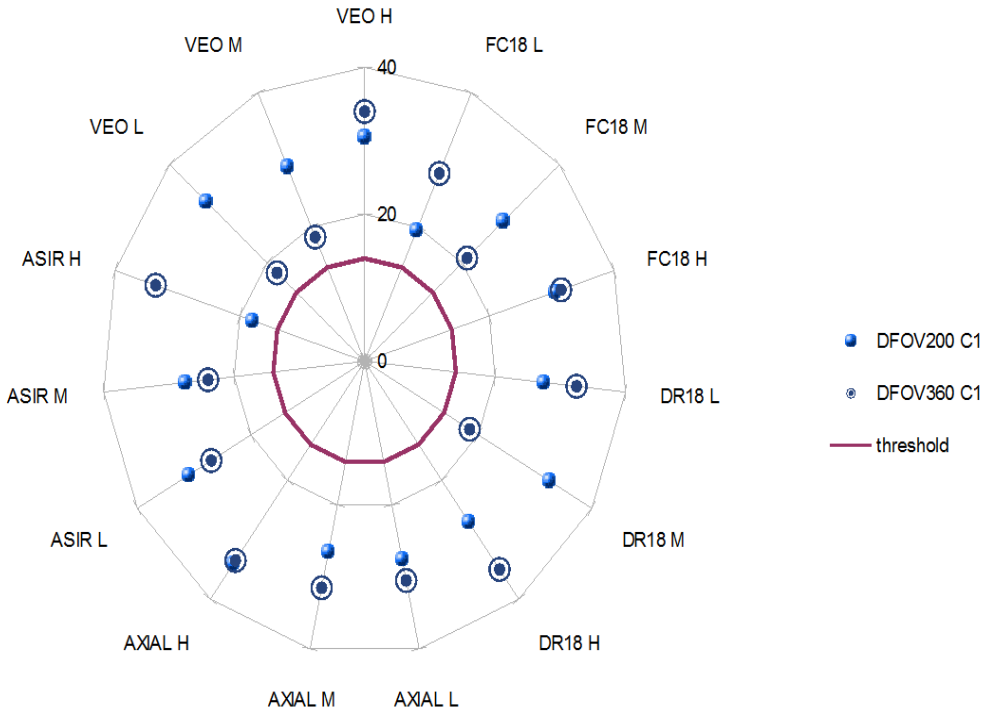


Figure 23. Here is displayed all data for contrast serial 1. It includes phantom with a ring, and without it. Symbol for a phantom with the ring, has been chosen to be a ring with an inner circle at graphic. We can see values near U=14 easily. There are no values over the purple line.

If we group the methods by manufacturer (GE or Toshiba) we can find interesting information.

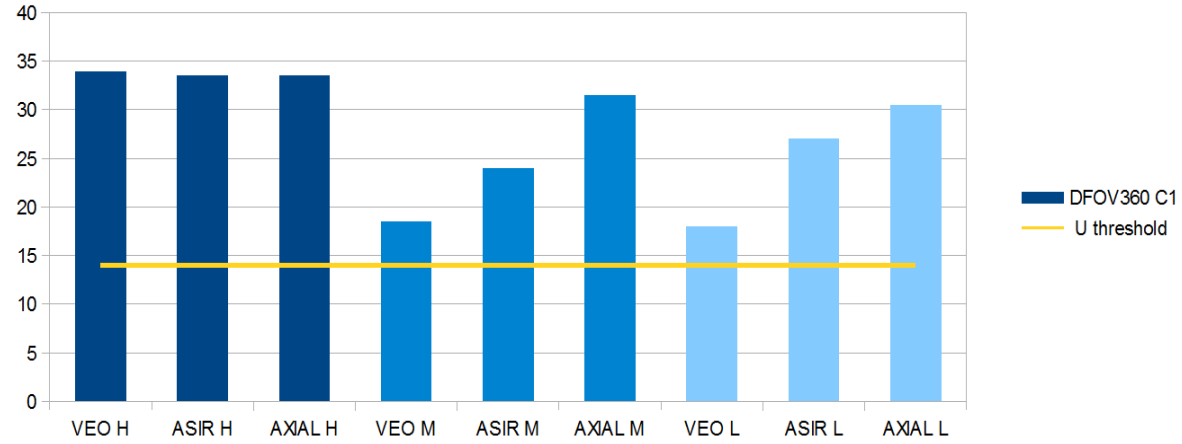


Figure 24. U-values for GE Methods. Phantom with One Ring. Contrast Serial 1.

Any U value over the Threshold line, reveals no statistical difference. We had used color code for dose (changing the saturation level of blue).

The fact to find no statistical difference, it can be interpreted as a good new. It will allows us to cluster the data, and that will provide more statistical reliance of results. But that this mean that CIMET students could be as good as the future radiographers, in the "attention to the detail"?

Maybe some of the observers could be. However, differences among the full dataset of observers are still remarkable: the standard deviation value provided by Image processing Students dataset is in general much higher than the one for radiography students.

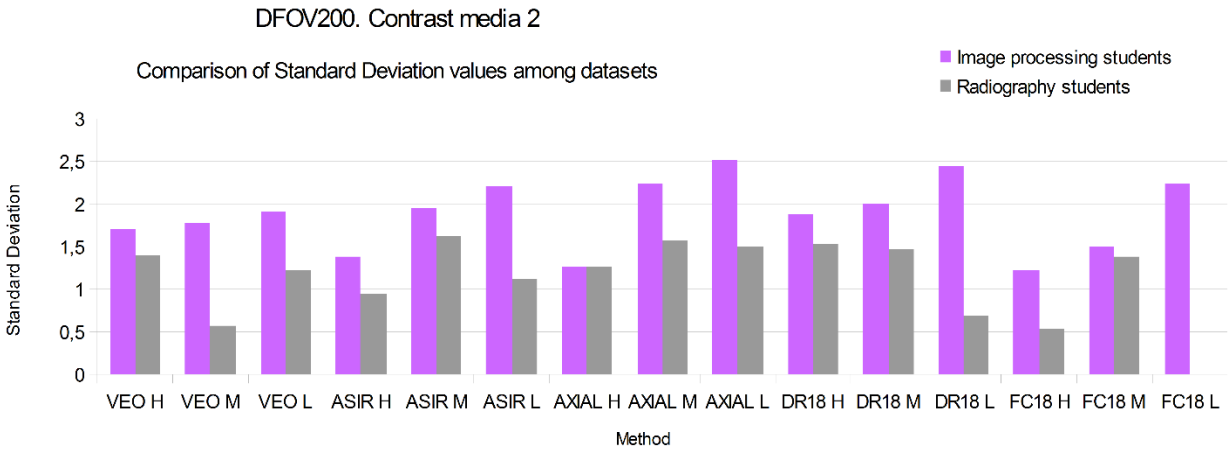


Figure 25. Comparison of standard deviation values among the two datasets. DFOV200 Contrast media 2.

For the most basic method included in observers experiment (FC18) and low dose case: Image processing Students have an standard deviation equal to zero, as all of them agreed to count 3 circles.

Hence, even if Mann-Whitney test can reveal no statistical difference between groups for most of cases, it could be interesting to go further, and study a measure of "the disagreement inside the agreement".

If we show the data by dose grouping, here both brands included, we can see how evolves the U value when increasing dose:

Figure 23. Evolution of U-values for each method, for low, medium and high dose images.

Just look how for ASIR Method, the U-value for low dose is outside the threshold, but progressively moves inside the threshold when the dose is high.

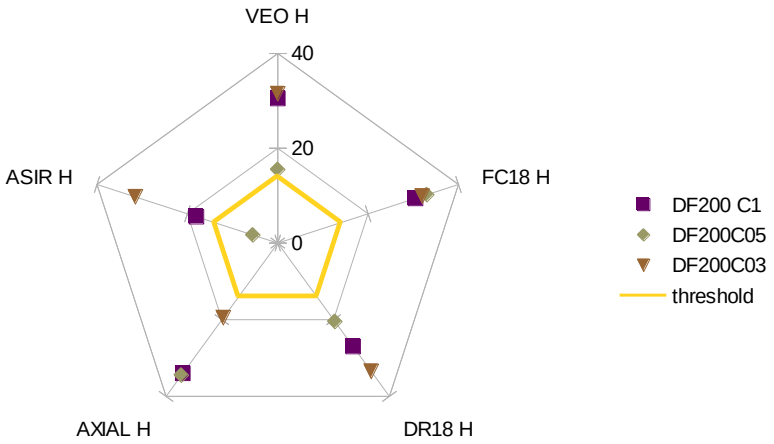


Figure 26. High dose images case. U values displayed without linking line.

At *Figure 25* lines serve to see the shape easier. Values are the points at axis. For contrast serial 2 (that is, density value 0.5) and high dose, we see there is a U-value below threshold.

Figure 27. Polar display of U values .We can also see how far is each U value to threshold, for each GE Method.

Figure 28. A line linking the points was selected, in order to see better we have some geometrical symmetry at polar display.

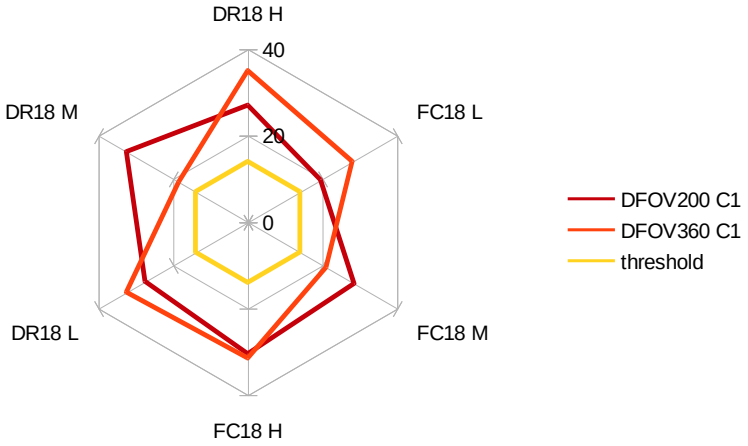


Figure 29. Example with Toshiba methods. Phantom with and without ring.

5.2.3 Comparison of brands: GE and Toshiba methods.

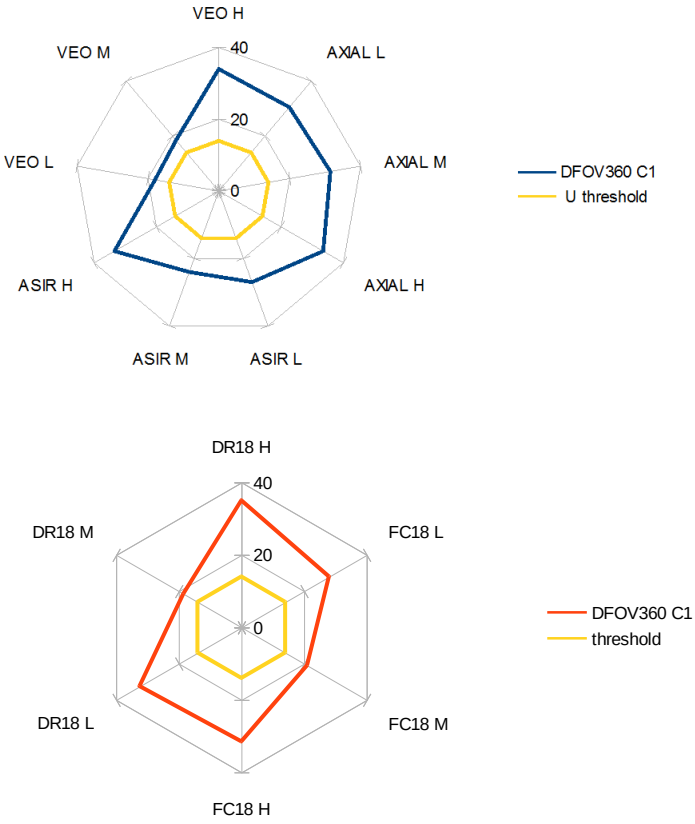


Figure 30. Comparison of GE methods and Toshiba ones.

We can see there is a similar shape between the two graphics. Is important to see that polar distribution is spaced differently (9 vortex in GE, 6 for Toshiba). We can see how observers from different backgrounds, behave similarly in their answers distribution for the more advanced method of Toshiba (DR18) and the more advanced method of GE (VEO) at high dose case, with a U value near 40.

Here the data of U-value, for phantom with littlest ring:

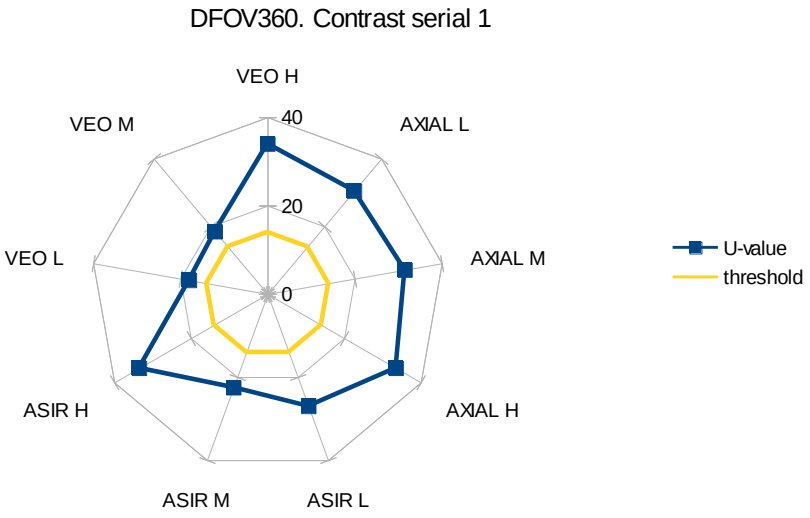


Figure 31. Methods by GE. Distance to U Threshold. DFOV 360 Contrast serial 1.

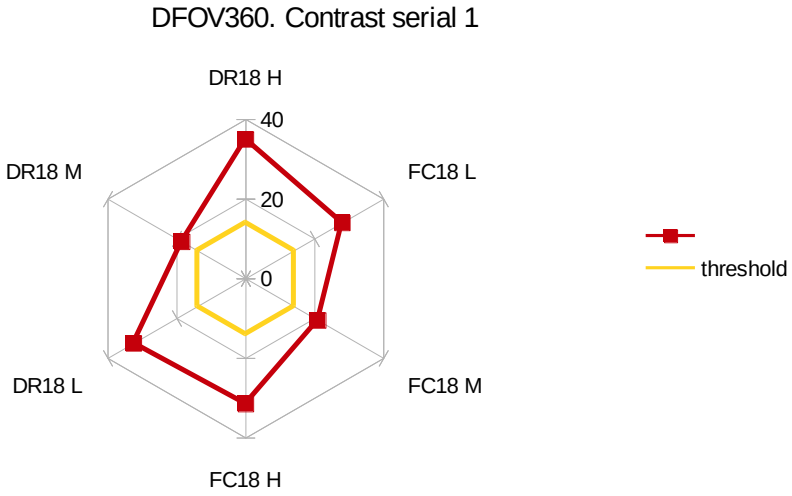


Figure 32. Methods by Toshiba. Distance to U Threshold. DFOV360 Contrast Serial 1

Figures *29* and *30* shows that no value goes below U threshold, so we can also group the datasets for DFOV360 as well.

If we go back onto Standard Deviation study, now we can see how data evolves when we consider all observers together. In this example of a phantom with the littlest ring, the same

comments apply as before, but now with more statistical relevance. Still the highest disagreement between observers, occurs at Axial for low dose. There is less disagreement for Toshiba standard method (FC18) than for GE standard Method (AXIAL).

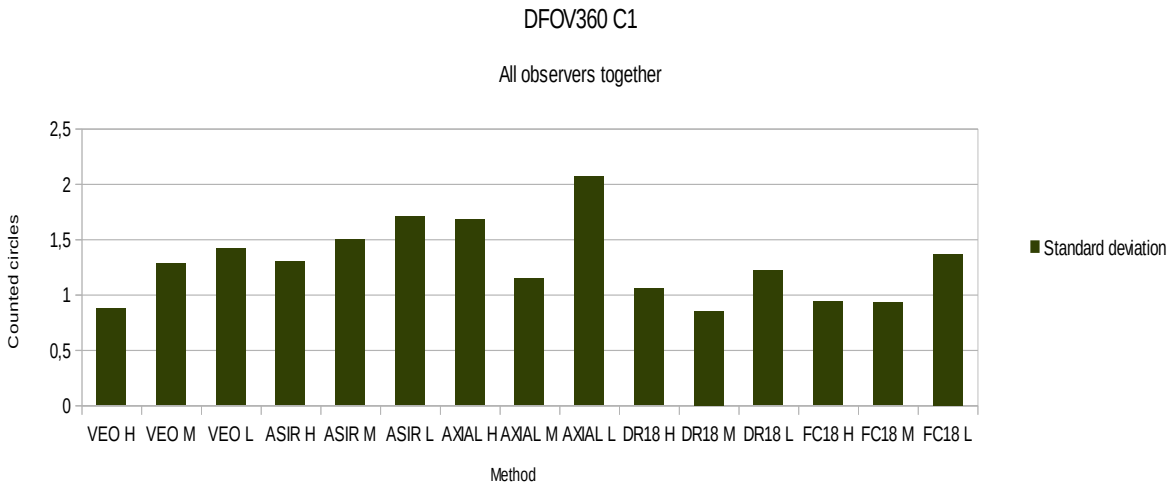


Figure 33. Standard deviation values. Observer´s data clustered (image processing and radiography students groups). Higher statistical reliance of data than Figure 21.

Data for the higher ring (DFOV400)

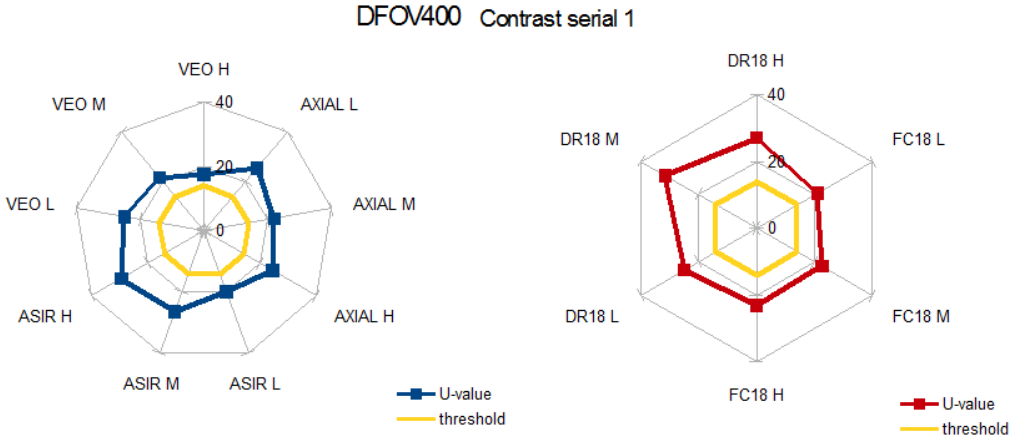


Figure 34. U values distribution for DFOV400 Contrast serial 1. GE methods (blue color) and Toshiba (red color)

For GE methods, we have no more the nice graphic shape that was displayed at DFOV200 and DFOV360 data, for same contrast serial one. Disparity among observers (even if no statistically relevant still) is here more independent of the particular method or dose. Now is remarkable more difficult to detect circles at the image, due to the big ring. Here to have disparity or agreement has less sense to be studied, as in some cases, all observers could agree in see no circle at all. But Mann-Whitney U test remains useful, in order to link the data from both observers group.

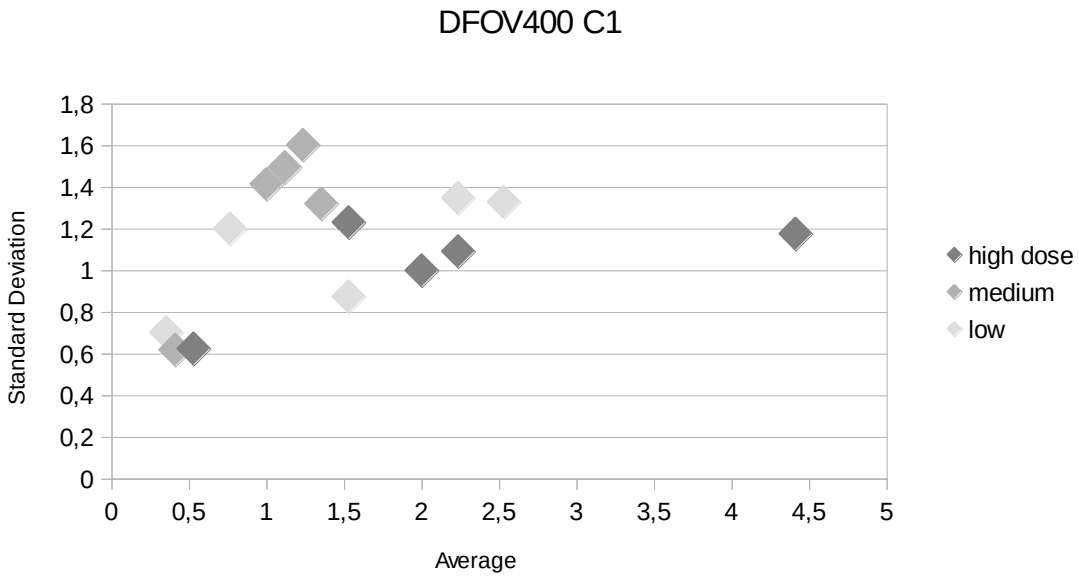


Figure 35. Standard deviation versus the average number of counted circles. DFOV400 Contrast serial 1.

At Figure 35 we can see the relation of standard deviation and the number of counted circles, for the phantom with DFOV400 Ring, all methods considered. If we just look onto high dose, Pearson correlation value is 0.636. For medium dose 0.882, and for low dose 0.739.

The right-side value (for an average of 4.5 counted circles) corresponds to the method VEO.

5.2.4 Calculation of confidence intervals

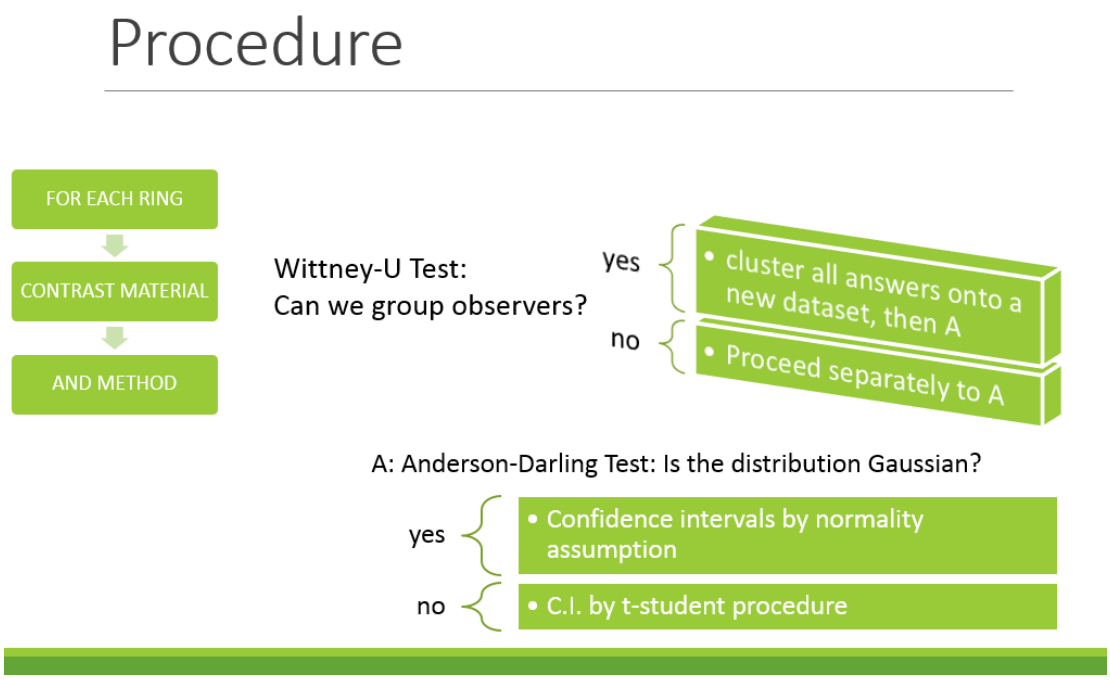
After studying if we can group observer data, comes the moment to calculate the confidence intervals for counted circles information. For this purpose, we will need to study normality of the distribution, by ANDERSON-DARLING Test [79]

In the cases data is not normal (or in other words, not Gaussian) we will employ the t-student distribution. As hypothesis, we could state that non-normal distribution is more likely to occur:

In contrast serial 1, advanced method, high dose, (circles are easy to see) and at contrast serial 3, basic method and low dose (circles too difficult to see, and most people counting zero circles)

When data is isolated by method and dose, sometimes (for an advanced iterative method i.e. VEO) we will have a high agreement in circles counted, that will be far to be a normal distribution of answers. In this case, we can use *t-student* Method to calculate Confidence Intervals.

Figure 36. Diagram illustrating the meta-algorithm employed for observers data statistical processing.



The named *t-student* distribution, was defined by Mr. W. Gosset, when he was working at *Guiness* ™ breweries as student. Is not a normal distribution, although is symmetrical around zero. It is said to have "n-1" degrees of freedom, where n is the size of the simple. Only when n tends to infinity the distribution will be equal to the normal [philosophical implications apart, about how to rest 1 to infinity].

We can state through an easier way: as degree of freedom is large, the t distribution approaches the normal one.

$$\lim_{n\to\infty}(t\ distribution) = normal\ distribution$$

The normal distribution is centered in zero and has standard deviation 1.

In practice, this infinity is quite close, and most of tables don´t show values up the degree of freedom of 50, where distribution is considered to be "enough close" to that normal.

The parameter $t_{\alpha/2}$ was calculated using the available tables at literature [80] for each of the observers group. For a given degree of freedom, the value $t_{\alpha/2}$ denotes the t-value "such that the area to its right, is equal to α/2, under the Student t-distribution" [].

The confidence interval is calculated through the formula:

$$\bar{X} \pm (t_{\alpha/2} \times \frac{s}{\sqrt{n}})$$

Where s is the standard error, and α depends on the level of confidence chosen. In our case, α = 0.05, for a 95% of confidence. Hence we will look onto the value $t_{0.025}$ at table, being care of using the value n-1, that is, the correspondent to degree of freedom, for the table search.

Some tables use quantile at the coefficient search. For a 95% confidence, 1- α = 0.95, and we look at table the quantile (1 – (α/2)) = 0.975

In example: For DFOV200, and for DFOV360 (check rest of DFOV) in the case of Contrast serial 1, there is no statistical difference (at 95% confidence) between observers, at every - and each - of the methods. So we can group observer data, onto a group of 17 observers, and check its normality adjustment for each method.

For the configuration displayed at *Figure 23*, there is no statistical between the observer´s answers to the images of a phantom without ring (at a 95% probability) despite of the particular observer educational background.

This applies for every method and every dose at contrast media 1. Hence, for the C1 data, a total of 17 observers can be considered to form a unique group, and we can proceed with confidence interval calculations (according to a normal or a t-student distribution, depending on each method and dose).

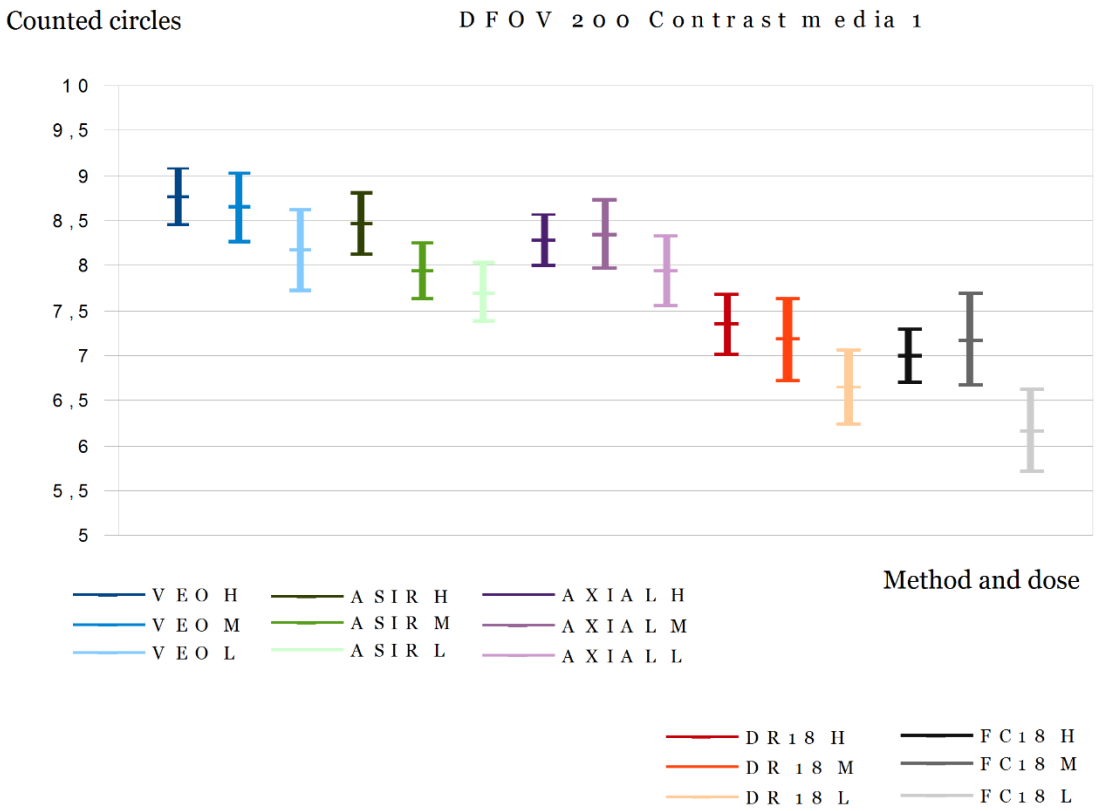


Figure 37. Counted circles with confidence interval. DFOV 200 Contrast media 1.

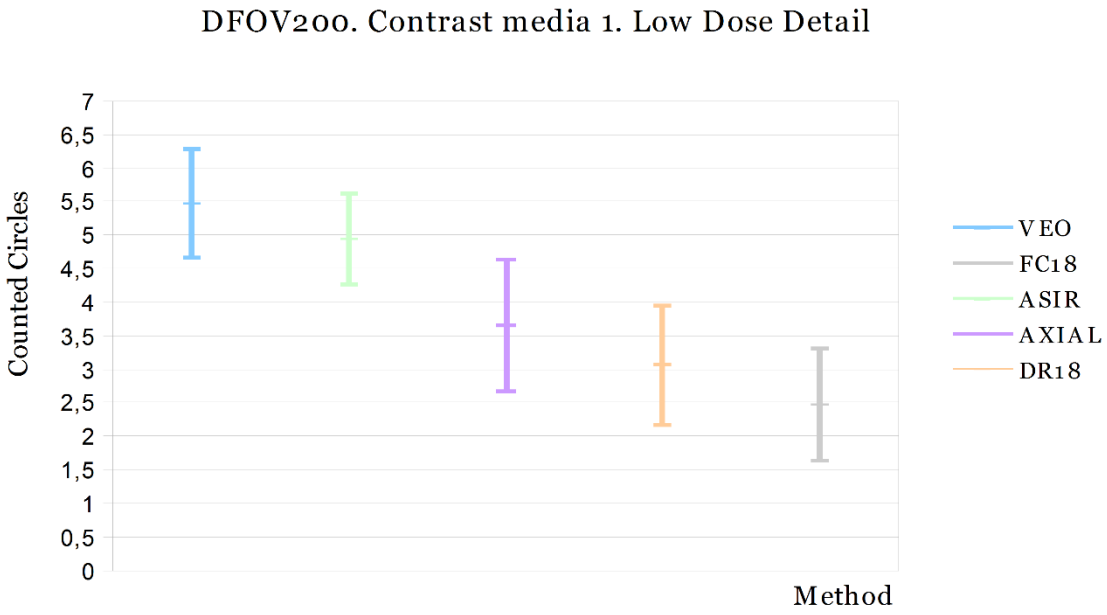


Figure 38. Counted circles with confidence interval. Low dose methods comparison. DFOV200 C1.

This figure shows that VEO has been designed for low-dose improvement.

5.2.5 Standard Deviation (revisited)

Now with the 2 datasets of observers linked, we can explore this parameter with more statistical reliance. Interesting results are found:

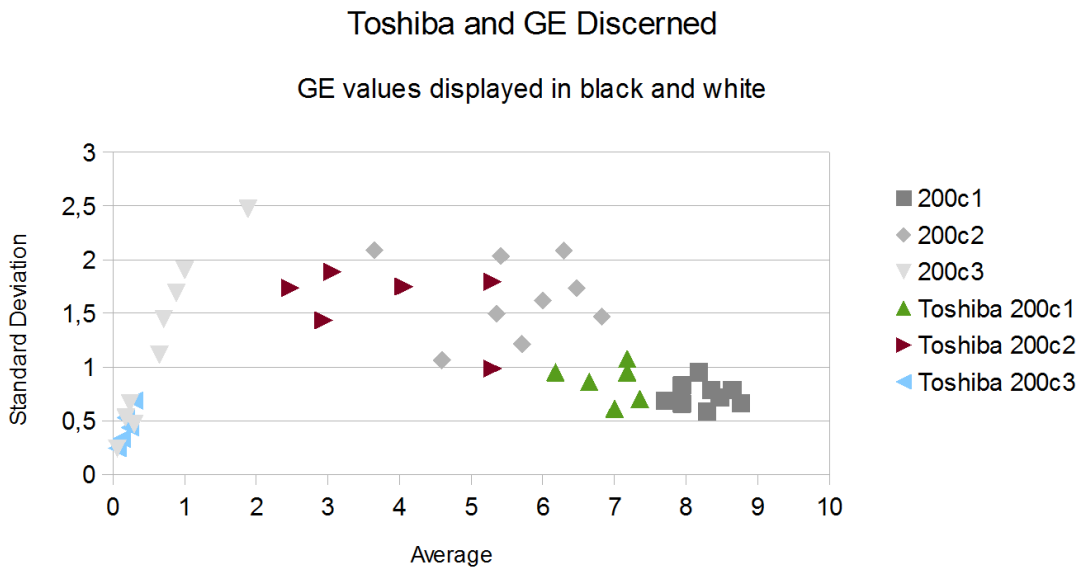


Figure 39. Toshiba and GE comparison of standard deviation and average of counted circles.

Can be easily seen from the data clusters, that when working with GE methods, in similar conditions, observers use to count more circles than in Toshiba case.

Interesting the C3 case, were relation between counted circles and standard deviation is linear (with a Pearson correlation of 0.97)

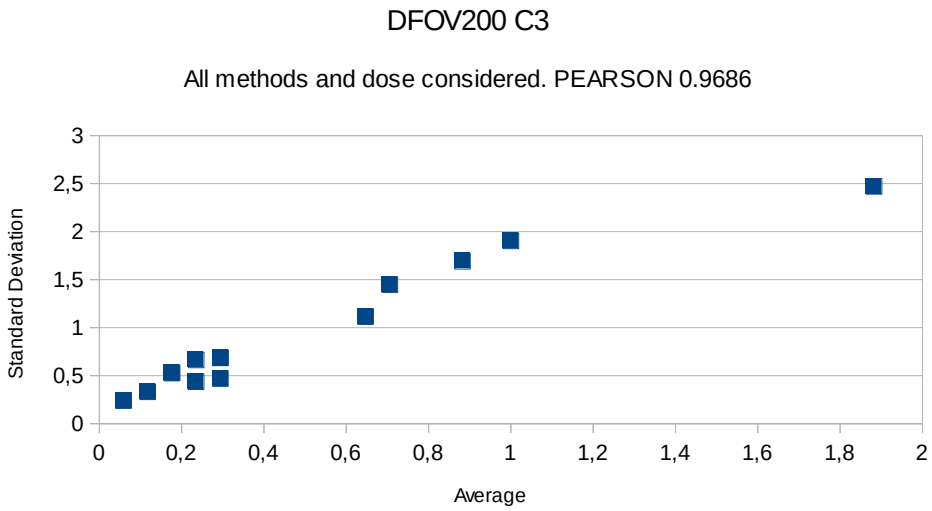


Figure 40. Standard deviation of counted circles versus the average counted. DFOV200 Contrast media 3

There is more agreement (less standard deviation) when people see almost ALL circles and when people see almost no-one.

5.3 CNR AND OBSERVERS DATA

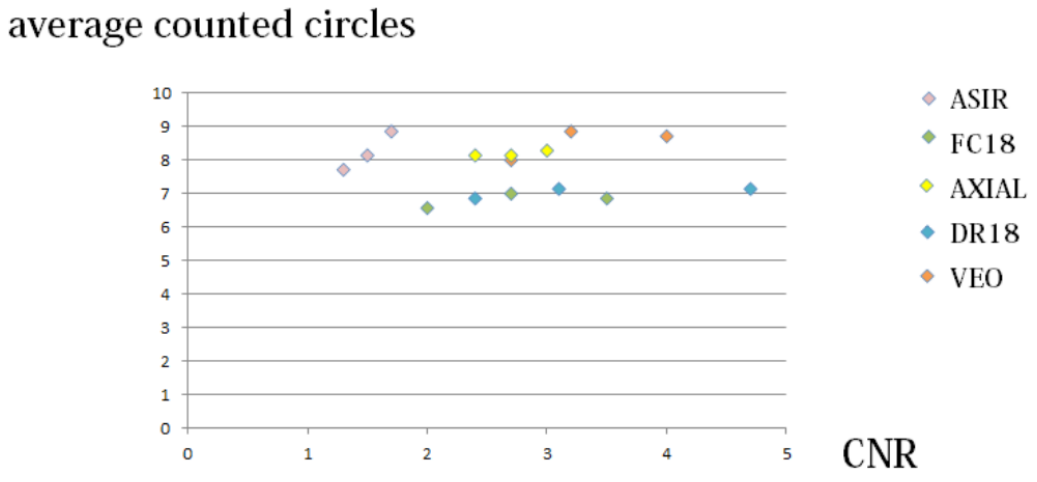


Figure 41. Average counted circles versus CNR. DFOV200 Contrast media 1. 5mm width. Dose level not detailed.

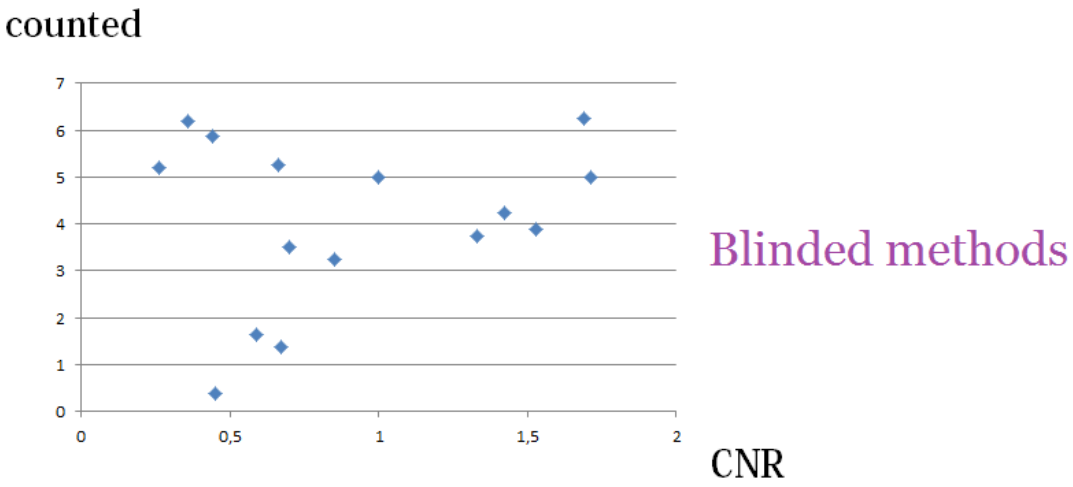


Figure 42. DFOV200. Contrast media 25mm width. Blinded methods (all dose considered)

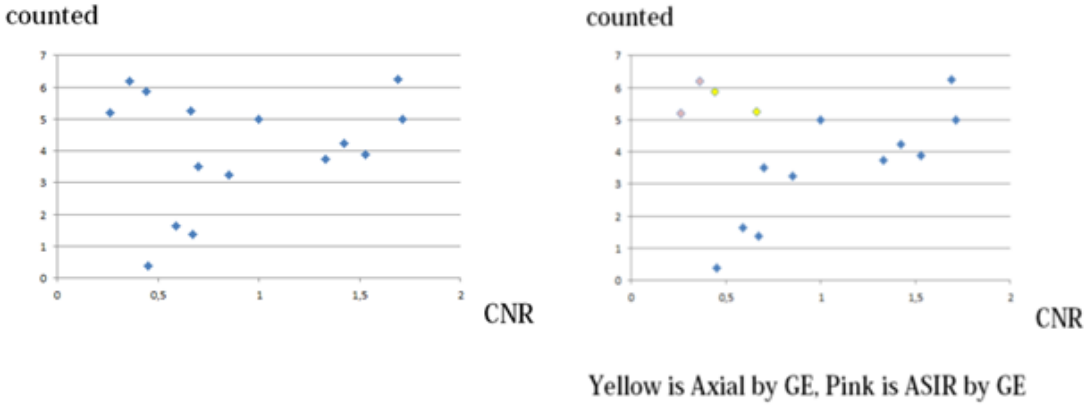


Figure 43. DFOV200 Contrast media 2. Some methods revealed

Points out the diagonal (x=y) line: it can be interpreted as a good new for GE. It means that for contrast media 2, even if some methods provide a CNR of approximately 0.5, observers counted as much circles as for other methods with CNR=2. In other words, they detected more little details (a better spatial resolution). However, if we look onto the relation between CNR and counted circles, this out-layers significantly affect the linear correlation factor.

Pearson correlation values:

For DFOV200 Contrast serial 1, if we restrict the calculations to ASIR and VEO methods, we have a correlation of 0.874 (a significant difference with the table value with all methods together). However, if we restrict for ASIR and VEO at DFOV360 Contrast serial 1, Pearson value is 0.793, almost no change compared the case we consider all data.

If we instead exclude ASIR

	Contrast serial 1	Contrast serial 2	Contrast serial 3
DFOV 200	0.336	0.744	0.549
DFOV 360	0.797	0.17	-
DFOV 400	0.856	-	-

In example, for DFOV360 contrast serial 2, people almost does not count circles, and we found even some problem to link observers data (due to some out-layer answer ASIR H value) If we calculate Pearson correlation there (with some data discarded, and hence less statistical value) value is very close to zero (0.17).

If we exclude ASIR and VEO from calculations, there is no significant change in the range of values with regard to the previous table. We see some improvement at Contrast serial 2, but a decrease in Pearson value at Contrast serial 3.

	Contrast serial 1	Contrast serial 2	Contrast serial 3
DFOV 200	0.20	0.78	0.22
DFOV 360	0,71	0.46	
DFOV 400	0.88		

We see from the previous tables that correlation values for DFOV200 contrast 1 and contrast 3 are quite low. However, if we consider <u>all contrast together</u> (ASIR method excluded) we find interesting results:

For the phantom without ring Pearson correlation value P= 0.812, and a *Spearman* Rank correlation S= 0.889. Spearman correlation coefficient is the Pearson correlation of the ranked variables, and it indicates until with extent the distribution is monotonous. For the phantom with little ring, P=0.77 and S=0.962.

5.3.1 Comparison between brands: using both measures jointly

As there is not a generalized linear relation between CNR and perceived details for every method, here arose the question: what we should prioritize in order to ranking the methods? As Image Quality is Diagnostic Quality: We should order methods first by counted circles, then by CNR. Counted circles will be translated to minimum diameter discernible.

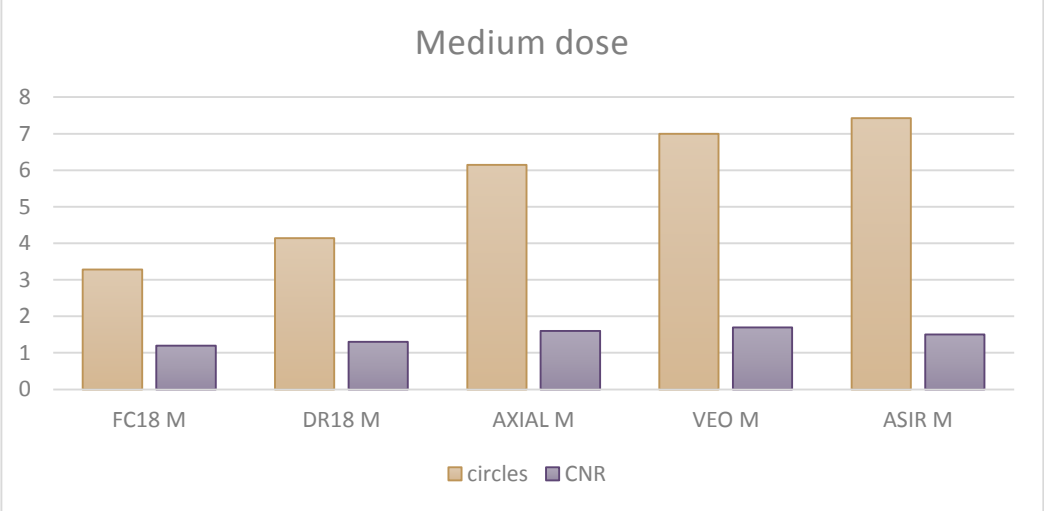


Figure 44. DFOV 200 Contrast media 2. Medium dose data.

At medium dose has no point to perform VEO (computationally expensive and then time required) as we can obtain similar results with ASIR. Here we see one valuable outcome of the observers experiment, as CNR values by itself would suggest us to use VEO or AXIAL.

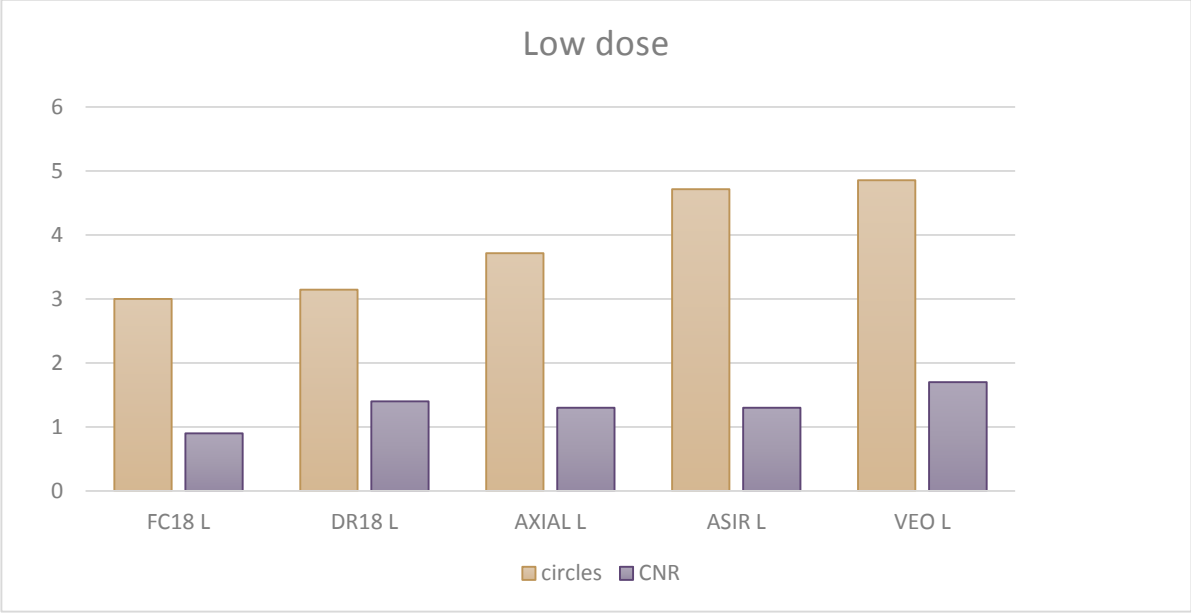


Figure 45. DFOV 2 Contrast media 2. Low dose data.

At low dose, DR18 shows improvement with respect to FC18 at CNR value, but this is not significant at the number of observer circles (here translated, on the smallest diameter discernible an observer could detect).

If we make no distinction between methods or dose, and we just consider the sum of counted circles for each configuration of phantom, we can see how detectability decreases for the phantom with rings, especially for the contrast serial 3 (as it corresponds to the lowest density material at module)

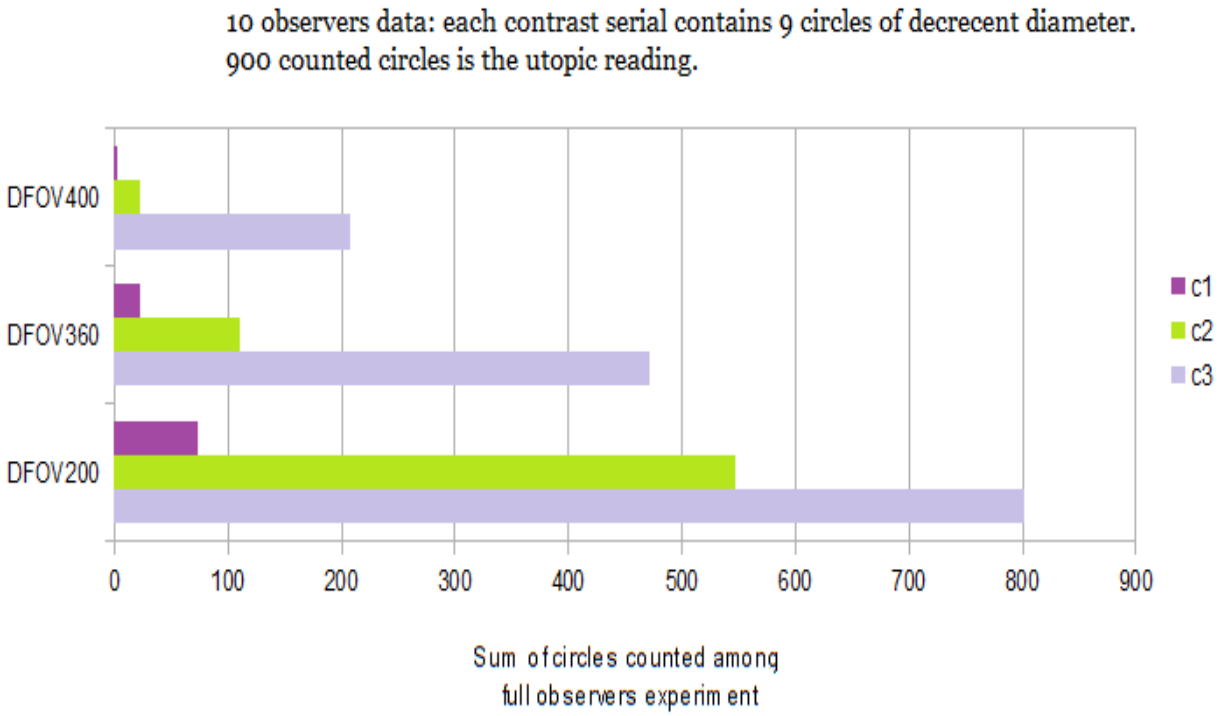


Figure 46. Total number of counted circles among experiment. It has been selected the image processing students dataset, for an easier visualization (dataset is composed of 10 observers, while radiography group were 7)

At *Figure 46* we see how decreases the capability to count circles at image, when adding rings to the phantom. Displaying with the dataset of 10 observers, is easy to see how the ideal count (if all circles were detected by observers) corresponds to a value of 900. There is a remarkable problem for detectability, when dealing with overweight patients. Just note that DFOV500 has not been included in the processing.

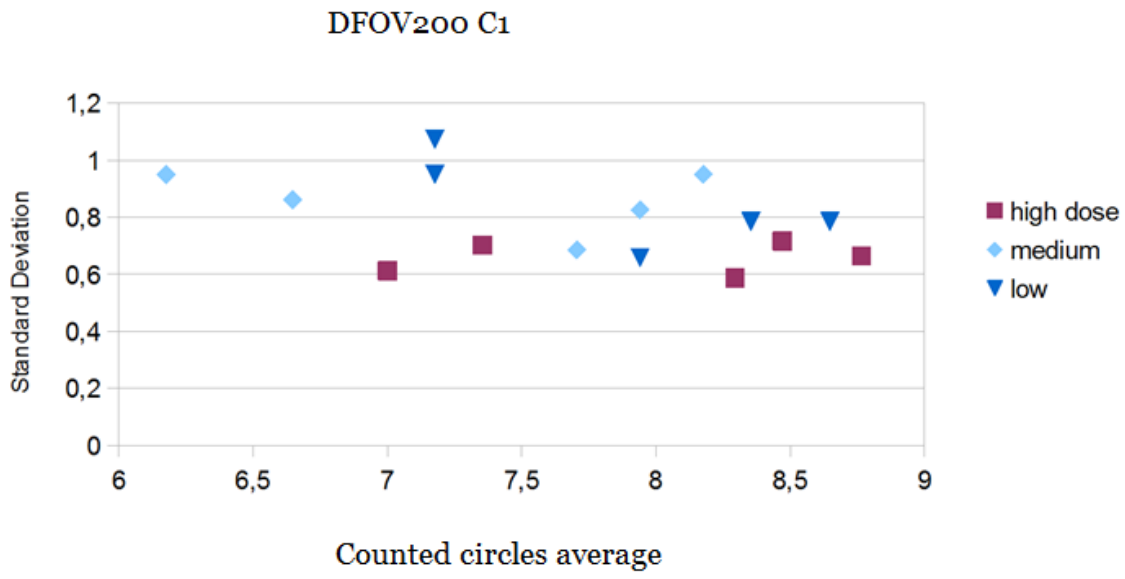


Figure 47. Threshold point at circles serial (decreasing in diameter from circle 1)

An average number of 7.5 counted circles, could be seen as a turning point in standard deviation values. Except for the case of high dose (as we are in this case in the best possible conditions for observation: a phantom without ring, material of density contrast 1%)

6 DISCUSSION

When dealing with Quality Test protocols, some authors as Kuttner [49] places the low contrast detectability as the last on a list of 23 steps, just after the spatial resolution one. That paper states that "If unavailable, the test of low contrast detectability may be rejected, as at this is related to noise, which can be measured in the noise test". However, other sources [50] recommends a weekly routine test specifically for this point. Either, Anders Tinberg stated [29] "*Contrast detail phantoms are good for Image Quality Analysis, but have little in common with anatomy of a patient*".

In agreement with some references about Dual layer-CT [81] (Siemens ™ related article) «Dual Energy CT is feasible without additional dose. There is no significant difference in image noise, while CNR can be doubled with optimized dual energy CT reconstructions". In GE (our Dual Energy example) we had found it happens for some cases, with DFOV 360 and bigger rings.

Filtering of tube makes Spectrum so different among manufacturers: GE has a lot of metal filtering, that removes a lot of "soft" (low energy) radiation. Toshiba instead, has less metal filtering, so it has more low energy radiation. CNR values for standard scans at GE, almost

doubles Toshiba values from DFOV 360 and higher rings. When the phantom without ring (that is, DFOV210) difference is not that high but still GE is better. Toshiba has better CNR values than GE for no ring and high dose.

As detailed in section 4.4.3, Spiral (also named helical) mode needs additional rotations at the beginning and the end of the scan - in order to obtain data, to reconstruct images over the prescribed volume – but it has some possibilities of dose saving when the patient does not cooperates, because scan time is much shorter. [30]

7 CONCLUSIONS

For a small patient (as can be a children) will be better to use Toshiba Acquilion ™ scanner, with AiDR18 Filter (Beam Hardening Correction Filter and Iterative reconstruction). For a patient suffering overweight, GE Discovery HD750 ™ generally performs better, in terms of low contrast detectability. However as drawback, GE most advanced method (VEO ™) is computationally expensive (needs 30 minutes of post-processing for a phantom-sized scan). According to *Figure 39*, when working with GE methods - in similar conditions - observers use to count more circles than with Toshiba methods.

Toshiba AiDR18 has the better value at low dose for DFOV400 (see *Figure 15*) However, particular CNR values are not always translated in a better detectability, it depends on the particular method. GE VEO is quite reliable for intermediate size rings. For DFOV360, ASIR can be enough in some dose configurations.

For a reduced scan length, the helical mode will be inefficient regarding dose reduction. For every rotation, there is an excess radiation (so is not advisable to do so many rotations). But helical CT can also reduce motion artifacts, and is needed for some reconstruction methods (as VEO by GE).

8 FUTURE WORK

The fact the desirable linear correlation between CNR and counted circles cannot be automatically extended onto some methods by GE, provides to main conclusions: it talks good for advanced GE methods in terms of observers detectability performance, but also calls for the development of new metrics that could be extensible to be applied onto this advanced algorithms output.

A newer version of AiDR algorithm by Toshiba, named *AiDR 3D*, is referred to work in both in the “raw data” and the reconstruction domains. [55] . The white paper by Toshiba, [82] claims AiDR 3D to be designed “to reduce patient dose without burdening clinical workflow”. This version however was not included at Toshiba device, at the time of scanning acquisitions.

Intra-observer variance among time could also be studied. The method of Forced Choice with Masked Images could be an alternative to the usual ROC, even if not used in Clinical Practice. As stated by D. Manning [40] "we are very familiar with the task of observing visual scenes [...] we are confident in this efficiency and we feel we have complete control over it [...] But the truth about vision is much more difficult".

From Physics perspective, there are opportunities to study the *Hystheresis* of detectors (by repeating each scan) or to pursue *Montecarlo* Simulations (to model the interaction of radiation with matter).

9 ANNEXES

ANNEX 1. ACRONYMS LIST

AIDR	Adaptive Iterative Dose Reduction
ALARA	As low as Reasonably Achievable
ASIR	Adaptive Statistical Image Reconstruction
CTDI	Volume Dose CT Index
CNR	Contrast to Noise Ratio
CT	Computer Tomography
DICOM	Digital Imaging Communications in Medicine
DFOV	Display Field of View
DVD	Digital Versatile Disk
FBP	Filtered Back Projection
GE	General Electric
HU	Hounsfield Units
MELODI	Multidisciplinary European Low Dose Initiative
NRPA	Norwegian Radiation Protection Agency
ROI	Region of Interest

ANNEX 2. NOISE POWER SPECTRA GRAPHICS [37].

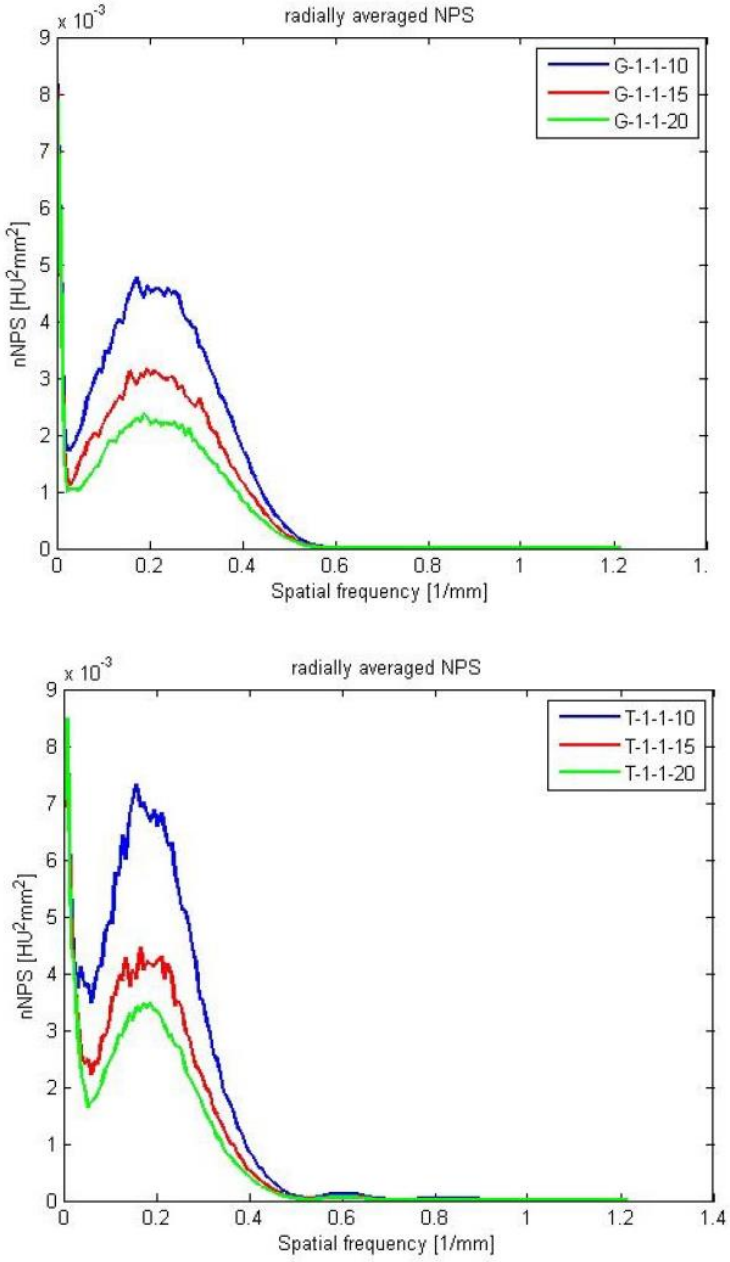


Figure. Upper graphic corresponds to GE. Lower one to Toshiba. [37]. The numbers 10, 15 and 20 refer to the CDTI-vol index (mGy). Graphic by Dawid Mozejko

ANNEX 3. SUBSLICE AND SUPRA-SLICE SECTIONS.

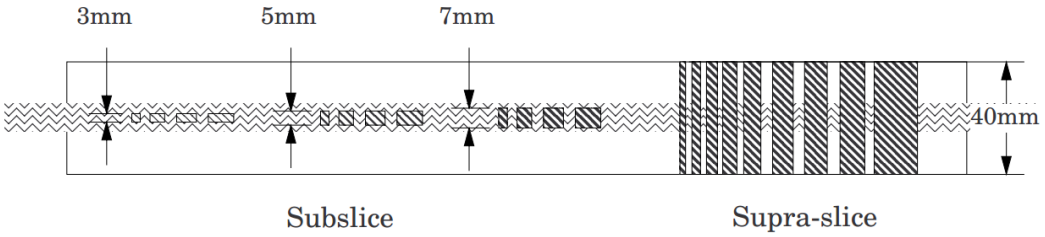
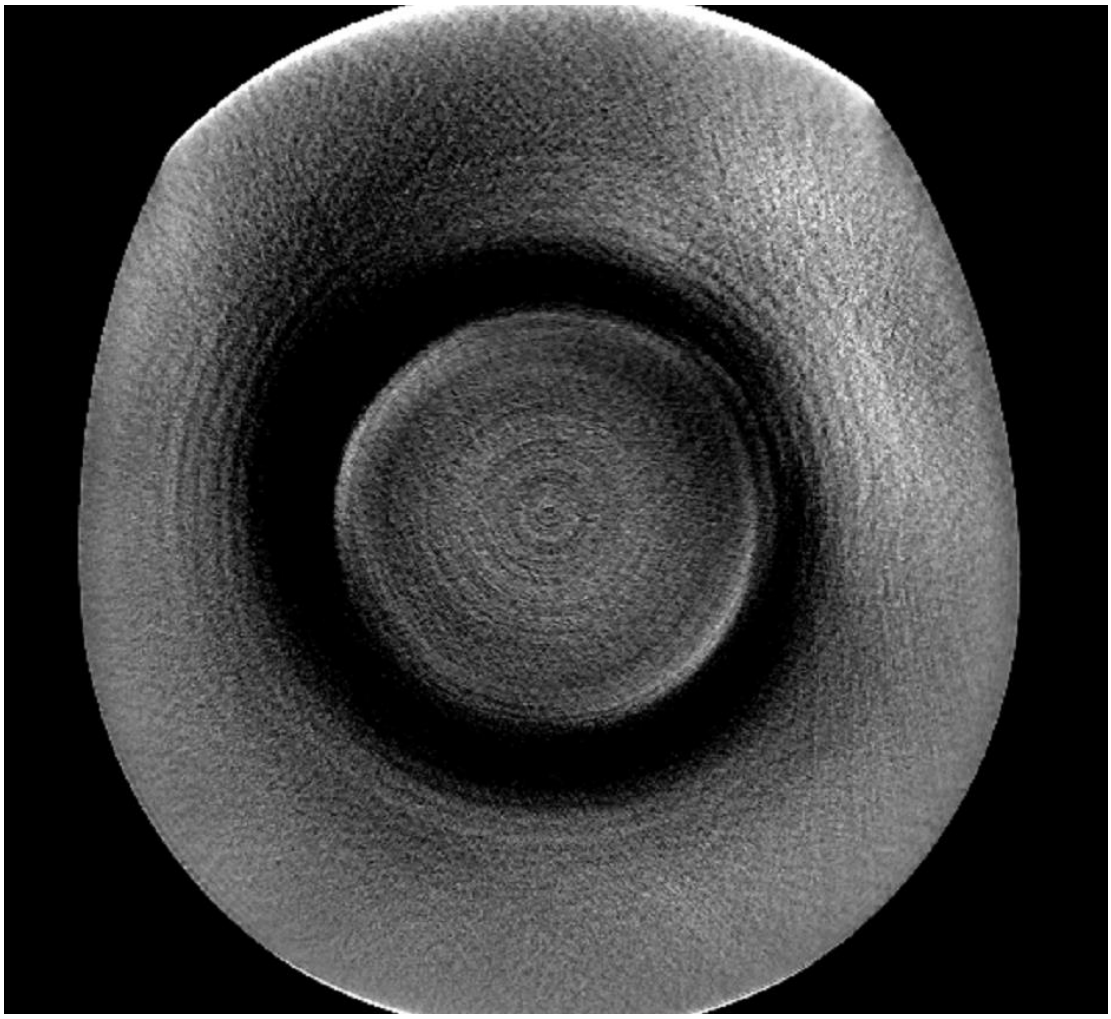


Figure. Subslice and supra-slice sections. [50]

Extract from the *Capthan* Manual [50] "The sub-slice targets are arranged in the inner circle of tests in the module. They are cast from the same material mix as the 1.0% supra-slice targets. The evaluation of its readability is helpful in understanding the scanner's different spiral imaging settings, and how those will affect the ability to visualize small objects with low contrasts from their background"

ANNEX 4. DFOV500



Picture: Image example of Phantom Scan with biggest ring. DFOV500. VEO method by GE. Disclaimer: the printed version will not be as reliable as the digital version displayed at a medical screen.

Images were also zoomed (either if this procedure includes interpolation). Even through this, was not possible to observe the super or supra slice modules of phantom, according to [62], [71].

ANNEX 5. New insights onto Norwegian Market.

In this Thesis, we focused on healthcare devices. However, CT systems had been recently presented (June 2014) at Norwegian Market for innovative applications. A company from UK has presented a CT scan in order to scan subsea pipelines, and have an estimation about the probability of hydrates formation inside. The device was tested in Bergen, Norway. It uses gamma-rays instead of x-rays as source (in order to deal with thicker material prospection).

10 REFERENCES

[1] Dr. A. C. Trædge Martinsen. Interviewee, *Verbal communication. The Intervention Centre, Oslo University Hospital.* 19 February 2013. Interviewee, *Verbal communication. The Interventional Centre, Oslo University Hospital.* 19 February 2013.

[2] Toshiba Medical Systems, [Available online] http://www.toshibamedicalsystems.com/tmd/english/products/dose/index.html. [Last access: June 2014]

[3] General Electric, «The Pulse Newsroom: Hospitals are Paying More Attention Than Ever to Radiation Dose from Medical Imaging in the Wake of new EU Legislation» March 2014. [Available online] http://newsroom.gehealthcare.com/hospitals-paying-attention-radiation-dose-medical-imaging-wake-eu-legislation/. [Last access: March 2014].

[4] A. Alenei, «Presentation "Trends in Radiation Protection of PET/CT Imaging" Slide 17. 2nd International Symposium on the System Radiological Protection. October 22-24, 2013. Abu Dhabi, UAE,» 4 March 2014. [Online].

[5] E. Samei, The Handbook of Medical Image Perception and Techniques, Cambridge, 2010.

[6] Healthcare Human Factors Group, «Technical Report "Computed Tomography Radiation Safety Issues in Ontario"» 3 March 2014. [Available online] http://www.health.gov.on.ca/en/common/ministry/publications/reports/disc_ct_mri/ct_report.pdf.

[7] The Nordic Radiation Protection co-operation, «Statement Concerning The Increased Use of Computer Tomography in Nordic Countries». [Available online]: http://www.nrpa.no/dav/db58f19fef.pdf. [Last access: 5 March 2014]

[8] ECRI Institute´s Health Devices Group, «Top 10 Health Technology Hazards for 2014. Adapted from Volume 42, Issue 11, November 2013». [Available online] www.ecri.org/2014. [Last access: 25 March 2014]

[9] J. Y. Hardeberg, Interviewee, *Radiology Workshop Høgskolen i Gjøvik*. June 2014.

[10] A. Pelegrina, Interviewee, *Thesis Public Defense*. [Interview]. 13 June 2014.

[11] Nobel Media AB 2014, «Perspectives "The Nobel Prize in Physiology or Medicine 1979"» Nobel Media AB 2014, Available online:

http://www.nobelprize.org/nobel_prizes/medicine/laureates/1979/perspectives.html . [Last access: 11 July 2014].

[12] W. Dr. Huda, «Physics of Medical Imaging Course» Iowa University, [Available online] http://www.uiowa.edu/hri/courses/physicsOfMedicalImagingReview/reviewSelfQuiz008_30q.html. [Last access: 10 May 2014].

[13] J. Hsieh, de *Computer Tomography: principles, design, artifacts, and recent advances. 2nd Edition. Chapter 1*, Wiley-Interscience 2009., 2009, p. 12.

[14] L. W. Goldman, «Principles of CT: Multislice CT doi: 10.2967/jnmt.107.044826» *J. Nucl. Med. Technol,* vol. 36, nº 2, 2008.

[15] J. Hsieh, «Computer Tomography: principles, design, artifacts, and recent advances. 2nd Edition» Wiley-Interscience, 2009.

[16] J. Radon, «On the determination of functions from their integral values along certain manifolds» *Medical Imaging IEEE, vol.5, no.4,* vol. doi:10.1109/TMI.1986.4307775., pp. 170 - 176, 1986.

[17] MATLAB Software Simulation, Mathworks, 2013.

[18] Universiy of Texas Health Science Center San Antonio , «Physics of Medical x-ray Imaging. Chapter 3. Research Imaging Institute» [Available online] http://ric.uthscsa.edu/personalpages/lancaster/DI-II_Chapters/DI_chap3.pdf. [Last access: May 2014].

[19] E. Boas, «CT artifacts: Causes and reduction techniques,» [Available online]: www.futuremedicine.com › Journal home › TOC. [Last access: May 2014].

[20] J. B. Thibault, «The white paper of VEO - The Model Based Paradigm» 2010. [Available online] http://www.gehealthcare.com/dose/pdfs/Veo_white_paper.pdf.

[21] P. D. P. Cattin, «Principles of Medical Imaging. University of Basel» 18 November 2013. [Available Online] http://miac.unibas.ch/PMI/03-ComputedTomography.html. [Last access: 18 April 2014].

[22] M. Reza Ay, «Experimental assessment of the influence of beam hardening filters on image quality and patient dose in volumetric 64-slice X-ray CT scanners» *Elsevier Physica Medica,* vol. 29, pp. 249-260, 2013.

[23] Z. Marton, «Benefits of Lu2O3 scintillator» SPIE Medical Imaging 2013.

[24] P. Khong, «ICRP Report 121. Radiological Protection in Paediatric Diagnostic and Interventional Radiology» *SAGE,* 2013.

[25] M. Rehani, «Managing Patient Dose in Computer Tomography» *ICRP,* nº 87. Chapter 1, 2000.

[26] European Commission, «Criteria for Acceptability of Medical Radiological Equipment used in Diagnostic Radiology, Nuclear Medicine and Radiotherapy» *European Commission Radiation Protection,* nº 162, 2012.

[27] Norwegian Radiological Protection Authority, «NRPA,» [Available online] http://www.nrpa.no/eway/default.aspx?pid=240. [Last access: 18 February 2014].

[28] H. Kundel, «Image quality and observer performance: new horizons for Radiology lecture» *Radiology,* vol. 132, nº (2), pp. 265-71, 1979.

[29] A. Tingberg, «Observer experiments – Theory and practice. Skäne University Hospital,» 13 11 2012. [Available online] http://www.radiofysik.org/userfiles/default/Anders_Tingberg_121113.pdf. [Last access: February 2014].

[30] H. Menzel, «European Guidelines on Quality Criteria for Computer Tomography. EUR16262» 2014. [Available online] www.drs.dk/guidelines/ct/quality/htmlindex.htm. [Last access: 22 April 2014].

[31] L. Lança, Digital Imaging Systems for Plain Radiography, New York: Springer Science Business Media, 2013.

[32] G. Dougherty, «Digital Imaging and communications in Medicine» Cambridge.

[33] European Radiology Conference, «Live Sessions Broadcast. Image Quality Discussion» from *https://www.myesr.org/cms/website.php?id=/en/ESR_ECR_news.htm*, Wien, 2014. 7 & 8 March.

[34] F. Zarb, «Image quality assessment tools for optimization of CT images» *Radiology by Elsevier,* vol. 16, pp. 147-153, 2010.

[35] A. P. Katherine, «Image Quality and Dose» de *Practical Digital Imaging and PACS.*, American Association of Physicists in Medicine. Medical Physics Monograph No.25.

[36] Y. Sagara, «Abdominal CT: comparison of low-dose CT with adaptive statistical iterative reconstruction and routine-dose CT with filtered back projection in 53 patients» *AJR Am J Roentgenol. doi: 10.2214/AJR.09.2989.,* vol. 195, nº (3), pp. 713-9, 2010 Sep.

[37] D. Mozejko, «Master Thesis "Image Texture, Uniformity, Homogenity and Radiation Dose in CT"» Gjøvik, 2013.

[38] E. Dixon, Washington University, 3 June 2011. [Available online] http://www.american.edu/media/news/20110603-Best-Illusion-2011-Second-Place.cfm. [Last access: 1 March 2014].

[39] E. Samei, Handbook of Medical Image Perception and Techniques, Cambridge, 2010.

[40] D. Manning, «Cognitive factors in reading medical Imaging. A survey of cognitive factors and models of medical image interpretation» from *The Handbook of Medical Imaging Perception and Techniques*, Cambridge, 2010.

[41] T. Johnson, «Physical Background Chapter» de *Dual Energy CT in Clinical Practice, Medical Radiology 2011.*, Springer-Verlag, 2011, p. 8.

[42] L. Yifeng, «Radiation dose reduction in computed tomography: techniques and future perspective» *Imaging Med,* vol. 1, nº 1, p. 65–84, 2009.

[43] T. e. a. Weidinger, «Threshold Optimization for efficient contrast imaging with quantum CT detectors» *Physics of Medical Imaging,* vol. Proc. SPIE, nº 86680Q, 2013.

[44] Y. Chen, «Improving low-dose abdominal CT images by Weighted Intensity Averaging over Large-scale neighbourhoods» *European Journal of Radiology,* nº 80, 2010.

[45] A. Pelegrina, *Author Caption,* Oslo: Samsung Mobile Camera.

[46] General Electric, «High Definition. The leading edge of CT clarity» 2013. [Available online] http://www3.gehealthcare.com/en/Products/Categories/Computed_Tomography/Discovery_CT750_HD/High_Definition.

[47] Hilde Kjernlie Andersen, *Verbal Communication,* 2013.

[48] T. Johnson, «Material Differentiation in dual energy CT: initial experience» *Eur Radiol 17. Springer-Verlag,* 2007.

[49] S. Kuttner, «A proposed protocol for acceptance and constancy control of computed tomography systems: a Nordic Association for Clinical Physics (NACP) work group report» *Acta Radiol. doi: 10.1258/ar.2012.120254.,* vol. 54, nº (2), pp. 188-98, 2012.

[50] D. J. Goodenough, The phantom laboratory. Capthan Instructional Manual, Greenwich, NY: The Phantom Laboratory, Incorporated, 2013.

[51] Gibbons, J.D, «Comparisons of the Mann-Whitney, Student's t, and Alternate t Tests for Means of Normal Distributions». *The Journal of Experimental Education* Vol. 59, N º 3, pp. 258-267, 1991.

[52] K. Jessen, «Dosimetry for optimisation of patient protection in Computer Tomography» *Pergamon. Applied Radiation and Isotopes,* vol. 50, pp. 165-172, 1999.

[53] L. KC, «Managing the radiation dose from pedriatic CT» *Appl. Radiol.,* vol. 35, pp. 13-20, 2006.

[54] AAPM, «The Measurement, Reporting and Management of Radiation Dose in CT» *Report of AAPM Task Group,* nº 96, 2008.

[55] R. Irwan, «AIDR 3D - Reduces Dose and Simultaneously Improves Image Quality» Toshiba Medical Systems Europe BV, 2011.

[56] IAEA-PRSM-1 Organismo Internacional de la Energía Atómica, «Manual Práctico de Seguridad Radiológica» Madrid, y Consejo de Seguridad Nuclear, 1996, p. 59.

[57] A. Pelegrina, *Author Caption,* Oslo: Camera: Olympus Tough TG-1.

[58] General Electric Healthcare, «Discovery CT750 HD» 2001. [Available online] http://www3.gehealthcare.co.uk/en-GB/Products/Categories/~/media/Downloads/uk/Product/Computed-Tomography/Veo/Discovery%20CT750%20HD%20Book.pdf.

[59] SECTRA Medical Systems AB, «RIS/PACS Diagnostic» 2013. [Available online] http://www.sectra.com/medical/diagnostic_imaging/solutions/ris-pacs/. [Last access: 2014].

[60] H. K. Huang, «PACS and Imaging Informatics. Basic principles and applications» Wiley, 2010, p. Table 20.2.

[61] A. C. Trægde Martinsen, «The possibilities of reducing radiation dose and improve image quality in CT diagnostics using advanced image processing» June 2011. AIT Oslo AS, 2011 [Available online] https://www.duo.uio.no/bitstream/handle/10852/28011/dravhandling-martinsen.pdf?sequence=3. [Last access: 29 May 2013].

[62] A. C. Trægde Martinsen, *Verbal Communication,* Oslo, 2013.

[63] A. C. Trægde Martinsen, *Reduction in dose from CT examinations of liver lesions with a new post-processing filter: a ROC phantom study. Acta Radiology, 2008. 49(3) 303-9*

[64] M. Söderberg, «Image Quality Optimisation and Dose Management in CT, SPECT/CT, and PET/CT» from *Thesis for the Degree of Doctor of Philosophy in Medical Science. Lund University*, Lund, 2012.

[65] The Association of Electrical Equipment and Medical Imaging Manufacturers, «The DICOM Standard 2013» [Available online] http://medical.nema.org/standard.html. [Last access: 2013].

[66] O. Pianykh, Digital Imaging and Communications in Medicine. A practical introduction and survival guide, Springer, 2012.

[67] E. R. Conference, «Live Broadcast. DICOM session» Wien, 2014.

[68] B. Aubert, «The work in IEC, DICOM and IHE to improve exchange of dose information between imaging modalities» 13th European ALARA Network Workshop, Oslo, 2011.

[69] T. Troscianko, «Basic Vision: An Introduction to Visual Perception,» Oxford, 2012, p. 62.

[70] H. Roehring, «Image Quality Evaluation of Medical Color and Monochrome Displays using an image colorimeter» *Medical Applications of Radiations Detectors II,* vol. Proc. SPIE 8508, 2012.

[71] *Dr*. W. Daaler, Interviewee, June 2013.

[72] M. W. Kusk, «Multislice CT. Billedkvalitet, Dosis & Teknik» 2011, p. Table 5.1.

[73] «ViewDEX software» Södra Älvsborgs Sjukhus, [Available online] http://sas.vgregion.se/sas/viewdex.

[74] MicroDicom, «MicroDicom Software,» [Available online] http://www.microdicom.com/.

[75] N. Sjostrand, «The medical illustration as the expression of illusion and imagination. The liver as an example from history» *Sven Med Tidskr.,* vol. 11, nº 1, pp. 17-51, 2007.

[76] International Telecommunication Union. Géneve, «Methodology for the subjective assessment of the quality of television pictures.,» Vols. 1 - 2. Recommendation ITU-R , nº BT.500-13, p. 13, 2012.

[77] C. Zylak, «Ergonomic Factors for the Radiologist's. Workspace. MD. RC425».

[78] M. Pedersen, *Doctoral Dissertation "Image quality metrics for the evaluation of printing workflows",* AIT Oslo AS, 2011.

[79] T. Anderson, «Stanford University. Anderson-Darling Tests of Goodness-of-Fit,» 18 February 2010. [Available online] statweb.stanford.edu/~ckirby/ted/papers/2010_Anderson-Darling.pdf.

[80] Universitat de Lleida., «Curso de Bioestadística. Presentación: ¿Cómo se utiliza la tabla de la t de student?» [Available online] web.udl.es/Biomath/Bioestadística/Dossiers/…. [Last access: 1 February 2014].

[81] J. Schenzle, «Dual energy CT of the chest: how about the dose?» *Invest Radiol. ,* vol. 45, nº 6, pp. 347-53, 2010.

[82] E. Angel, «AIDR 3D iterative reconstruction: Integrated, Automated and Adaptive dose reduction» Toshiba Medical Systems.