Radiation Doses in Mammography Exams: effects of oncological treatments

Ana Isabel Leite Ribeiro
Mestrado em Física Médica
Departamento de Física e Astronomia
2018

Orientador
Sandra Sarmento, Técnico Superior de Saúde, ramo Física Hospitalar,
Instituto Português de Oncologia do Porto, EPE

Coorientador
Bruno Mendes, Técnico Superior de Saúde, ramo Física Hospitalar,
Instituto Português de Oncologia do Porto, EPE

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, _____/_____/_____

Aknowledgments

First of all, I would like to thank my advisors, Sandra Sarmento and Bruno Mendes, for accepting me to do this dissertation and their crucial guidance. Their constant availability and their constructive criticism allowed me to learn and grow over the last year. I would also like to thank the opportunity to elaborate this dissertation at IPO-Porto.

I would like to acknowledge the researchers of the Epidemiology Service at IPO-Porto for some guidance in the sampling of the collected data and for useful discussions and insights into the statistical significance of results.

Lastly, I say thanks:

To my colleagues, for always helping each other out.

To my brother Bruno, and Ana, for always helping me when I needed, even though I avoided asking.

To my parents for giving me the whole world and believing in me more than anyone else.

Abstract

Automatic exposure control (AEC) allows digital mammography units to determine the optimal acquisition conditions, based on the detected breast composition, and the equivalent breast thickness. With automatic radiation dose index monitoring (RDIM) systems it is today possible to compare exam doses and establish diagnostic reference levels (DRLs) in much larger scales.

In oncology institutes, mammography exams will include non-standard breasts, such as patients who were previously submitted to oncological treatments. Surgery and radiotherapy can cause changes that are visible in later exams. The goal of this study was to determine if those changes are detected by the AEC system, and to compare dosimetric data for treated and untreated breasts, in order to determine if the two sets need to be considered separately when comparing exam doses.

The GE digital mammography unit at IPO-Porto was studied using phantoms, with particular emphasis on the applied image processing, and it was decided to use only processed clinical images in this study. This simplified the data collection process, since processed images are archived in Picture Archiving and Communications System (PACS) and can be obtained retrospectively.

Data from 1872 mammography images were retrieved from PACS, of exams performed in two mammography units at IPO-Porto (GE Senographe DS and GE Senographe Essential). Dosimetric data was compared to estimated dose values, through Dance's method, as suggested by EUREF [1]. The differences found were within around ±10% and ±30%, for Entrance Surface Air Kerma (ESAK) and Mean Glandular Dose (MGD) respectively. The larger differences in MGD values are probably related to the glandularities measured by the mammography unit, which are very different from those considered in Dance's method. The results also suggest that the two mammography units measure glandularity in a different way, despite being from the same manufacturer.

A preliminary study of ten patients with repeated mammography exams between 2009 and 2017, was used to establish a baseline of normal variability of parameters. The variability of radiation doses was small (<10%), and good positioning reproducibility was found. The most variable parameter was found to be the compression force.

Exams of 141 patients who were being monitored for benign conditions and had no treatment until the date of the exam were used to assess the normal difference between the left and right breasts. This variability was found to be very low, within ±3% in most cases.

Exams of 258 patients who had surgery and radiotherapy were studied. Comparing the treated and untreated breast of the same patient, mean radiation doses were found to be larger for treated breasts. (CC: 1.49mGy, 1.41mGy and 1.70mGy; MLO: 1.72mGy, 1.58mGy and 1.98mGy; for two subsamples of the DS system and a sample of Essential's unit, respectively). This was mostly related to the increased compressed breast thicknesses of treated breasts. The treatments probably cause the breast tissue to become more rigid and sensitive, which in turn influences the compression force applied.

Indicated values of glandularity are similar for treated and untreated breasts of the same patient, in both mammography units (mean differences were around 3% for the DS unit and 2% for Essential). Untreated breasts in the treated set of patients had lower mean glandularity values (~58% for the DS unit and ~26% for Essential) than the no treatment set of patients, for both mammography units (~73% for the DS unit and ~46% for Essential). This should be investigated further, using an automated method for independent estimation of glandularity for comparison with other studies.

The main conclusion of this study is that the automatic exposure control systems of the two direct digital mammography units studied do not appear to detect changes in breast composition when imaging treated breasts. The dependence of exam doses on compressed breast thickness appears to be very similar for treated and untreated breasts. Since DRLs are set for "standard" patients (or as a function of compressed breast thickness), this study indicates there is no need to separate or exclude treated breasts from the collected data when setting DRLs.

Key-words: *digital mammography, mean glandular dose, diagnostic reference levels, breast cancer, automatic exposure control, oncological treatments*

Resumo

O controlo de exposição automático (AEC) permite aos mamógrafos digitais determinar as condições de aquisição ótimas, baseando-se na composição da mama detetada e da espessura equivalente da mama. Com sistemas de monitorização automática de índices de radiação de dose (RDIM) é atualmente possível comparar doses de exames e estabelecer níveis de referência de diagnóstico (DRL) em grande escala.

Em institutos oncológicos, as mamografias vão incluir mamas que não são "padrão", como de pacientes que foram previamente submetidas a tratamentos oncológicos. A cirurgia e radioterapia podem causar alterações que são visíveis em exames posteriores. O objetivo deste estudo foi determinar se essas alterações são detetadas pelo sistema AEC, e comparar dados dosimétricos de mamas tratadas e não tratadas, de modo a determinar se ambas têm de ser consideradas separadamente, quando comparamos doses dos exames.

Os equipamentos de mamografia digital da GE, do IPO-Porto, foram estudados com recurso a fantomas, com particular enfase no processamento de imagem aplicado, e decidiu-se usar apenas as imagens clínicas processadas neste estudo. Isto simplificou o processo de recolha de exames, já que as imagens processadas são armazenadas no Picture Archiving and Communications System (PACS) e podem ser obtidas retrospetivamente.

Dados de 1872 mamografias foram recolhidas do PACS de exames realizados em dois mamógrafos do IPO-Porto (GE Senographe DS e GE Senographe Essential). Dados dosimétricos foram comparados a valores de dose estimados pelométodo de Dance, como sugerido pela EUREF [1]. As diferenças encontradas foram à volta de ±10% e ±30%, para o air-kerma de entrada à superfície (ESAK) e para a dose glandular média (MGD), respetivamente. A maior diferença nos valores de MGD estão provavelmente relacionadas com as glandularidades medidas pelo equipamento que diferem bastante das glandularidades típicas consideradas no método de Dance. Os resultados também sugerem que os dois mamógrafos medem a glandularidade de maneiras diferentes, apesar de pertencerem ao mesmo fabricante.

Um estudo preliminar de 10 pacientes com vários exames realizados entre 2009 e 2017 foi feito para estabelecer uma referência de variabilidade normal dos parâmetros. A variabilidade das doses de radiação foi pequena (<10%), e verificou-se boa reprodutibilidade de posicionamento. O parâmetro mais variável foi a força de compressão.

Exames de 141 pacientes que estavam a ser seguidas por condições benignas e não tinham qualquer tratamento até à data do exame, foram usados para avaliar a diferença normal entre mama esquerda e direita. Esta variabilidade mostrou ser muito pequena, dentro de ±3% na maioria dos casos.

Exames de 258 pacientes que tiveram cirurgia e radioterapia foram estudados. Ao comparar a mama tratada com a não tratada da mesma paciente, verificamos doses de radiação médias maiores para as mamas tratadas (CC: 1.49mGy, 1.41mGy e 1.70mGy; MLO: 1.72mGy, 1.58mGy e 1.98mGy; para duas subamostras do sistema DS e uma amostra do sistema Essential, respetivamente). Isto esteve principalmente relacionado com a maior espessura comprimida das mamas tratadas. Provavelmente, os tratamentos tornam o tecido da mama mais rígido e sensível, o que em vez influencia a força de compressão aplicada.

Os valores indicados da glandularidade são semelhantes para mamas tratadas e não tratadas da mesma paciente, em ambos mamógrafos (diferença média de cerca de 3% para o sistema DS e 2% para o sistema Essential). As mamas não tratadas do conjunto de pacientes tratadas apresentaram uma glandularidade média menor (~58% para o sistema DS e ~26% para o Essential) do que no conjunto de pacientes não tratadas, para ambos os mamógrafos (~73% para o sistema DS e ~46% para o Essential). É preciso investigar isto melhor, usando um método automático para uma estimativa independente da glandularidade para comparar com outros estudos.

A principal conclusão deste trabalho é que os sistemas de controlo de exposição automático dos dois mamógrafos estudados não parecem detetar alterações na composição da mama quando examinam mamas tratadas. A dependência das doses da espessura comprimida da mama parece ser muito semelhante para mamas tratadas e não tratadas. Como os DRLs são definidos para pacientes “padrão” (ou como função da espessura comprimida), este estudo indica que não há necessidade de separar ou excluir as mamas tratadas dos dados recolhidos, quando se estabelece os DRLs.

Palavras – chave: *mamografia digital, dose glandular média, níveis de referência de diagnóstico, cancro da mama, controlo de exposição automático, tratamentos oncológicos*

Contents

List of Tables

List of Figures

Acronyms

AEC - *Automatic Exposure Control*
RDI - *Radiation Dose Index*
RDIM - *Radiation Dose Index Monitoring system*
PACS - *Picture Archiving and Communication System*
DRL - *Diagnostic Reference Level*
DICOM - *Digital Imaging and Communications in Medicine*
CC - *Cranio Caudal (mammographic view)*
MLO - *Medio Lateral Oblique (mammographic view)*
AOP - *Automatic Optimization of Parameters*
PV - *Premium View*
FV - *Fine View*
SNR - *Signal to Noise Ratio*
CR - *Computed Radiography*
DDR - *Digital Direct Radiography*
FFDM - *Full Field Digital Mammography*
RIS - *Radiology Information management System*
HIS - *Hospital Information System*
MGD - *Mean Glandular Dose*
ESAK - *Entrance Surface Air-Kerma*
HVL - *Half Value Layer*
STD - *Standard (AOP mode)*
L - *Left*
R - *Right*
RT - *Radiotherapy*
PMMA - *Polymethyl methacrylate*
MPV - *Mean Pixel Value*
SD - *Standard Deviation*
ROI - *Region of Interest*
Raw - *"for processing" images*
Proc - *"for presentation" images*
LDA - *Local Dense Area*

1 Introduction

1.1 Goal

To determine if breast cancer treatments cause changes that are detected as breast composition changes by the automatic exposure control systems of direct digital mammography units, and if these changes in “detected composition” are sufficiently high to justify separation of treated and untreated breasts when comparing mammography exam doses and setting diagnostic reference levels.

1.2 Motivation

Direct digital mammography units are equipped with automatic exposure control (AEC) systems that determine the optimal exposure conditions for each image acquisition, based on the equivalent breast thickness and the breast composition, estimated through the transmission of the radiation with a preliminary low dose exposure. Monitoring radiation dose indexes (RDI) and image quality are crucial to ensure that the system is being adequately used and working properly.

Manual collection of data for dose audits usually focuses on reference situations (untreated breasts). With radiation dose index monitoring (RDIM) systems that collect dosimetric data automatically from PACS (Picture Archiving and Communication System), it is now possible to compare exam doses and establish diagnostic reference levels (DRLs) based on very large datasets.

In an oncology institute, dosimetric data pertaining to mammography will include breasts previously submitted to surgery and radiotherapy, in addition to data from untreated breasts. Surgery and radiotherapy as breast cancer treatments can cause changes that are visible in mammography exams. Therefore, it seems timely to compare dosimetric data for treated and untreated breasts, to determine if the two need to be considered separately in a large-scale analysis.

1.3 Summary

This work began with a bibliographic revision of some basic concepts such as, the importance and purpose of mammography exams, how a direct digital mammography unit works, the standards for storing and sharing of digital images, the methods to calculate the dose of a mammography exam, and the importance of diagnostic reference levels (DRLs). A quick review of the breast density concept and types of breast cancer treatment was also conducted.

Secondly, simple phantom images were studied to better understand how a direct digital mammography system works. These images were available in the raw “for processing” images and the processed “for presentation” ones. Some of the images had further processing, named “Premium View”. Comparing these versions allowed for an understanding of the processing that would be applied to the final clinical images which were the object of this study, what changes might result from such processing and whether these changes might need to be taken into account during further computational processing.

Throughout this work, computational processing of the images, such as contour delineation of the imaged breast area and detection of surgical clips and a treated side, as well as a preliminary comparison of the mammography units, were conducted. Mean Glandular Dose (MGD) and Entrance Surface Air Kerma (ESAK) were estimated with the method suggested by EUREF [1] .

Several different sets of clinical images were studied, to evaluate variations of exposure, dose and image parameters. Comparisons were made:

- To establish the normal variations of parameters along the years, in the same facility and to the same patients, with no treatment;
- To see the typical variability between the left and right breast of individual patients with no treatment;
- To determine normal differences between the two mammography units
- To study the variations of the parameters between treated and untreated breasts.

1.3.1 Structure

The structure of this work initially introduces the basic concepts of mammography, the functioning of a digital direct mammography unit, and other basic concepts like DICOM, PACS, DRLs, breast density and dose estimations, in section 2.

Section 3 contains the characterization of the collected data samples, the mammography systems considered in this work and the software used. The methods used throughout this study are also described in this section. These are separated for the study of clinical images, and the study of phantom images. The methods of computational processing and the independent estimation of radiation doses are included as well.

The results obtained are presented and discussed in section 4). This section is divided into subsections corresponding to the subcategories from the previous chapter.

Finally, conclusions and future works are presented in section 5.

2 Basic Concepts

2.1 Why is mammography necessary?

Breast cancer is the second most common cancer overall and is the most common in women. As of 2012, breast cancer accounted for about 25% of all diagnosed cases [2] and it was estimated to be the fifth most common cause of death among all cancer deaths. In Portugal, about 17 new cases are diagnosed per day and over 1600 women die every year [3]. Most developed countries have a higher incidence rate, and undeveloped countries have higher mortality rates. The causes of breast cancer cannot always be precisely determined, even though there are established risk factors. The mortality rate, on the other hand, can be reduced with access to developed healthcare systems. These systems usually involve prevention and screening policies, which could explain, in part, the lower mortality rates in developed countries [4].

To reduce the mortality rate, it is important to detect the disease as early as possible. This way, there are more treatment options and better chances of survival. Screening for breast cancer with mammography can be more effective in detecting the early stages of a pathology and smaller tumours than with basic breast exams [5].

A mammography exam exposes women to ionizing radiation and could, in the long term, induce breast cancer. It has been shown that the mortality reduction due to screening outweighs the risk [6]. This risk decreases with the age of exposure [7]. Many studies have recommended biennual mammography screening for women between the ages of 50 and 69 years [8]. Mammography is also crucial in diagnosis, follow up of patients after treatments and assessing the evolution of benign situations, which are the type of mammography exams performed at the IPO-Porto.

2.2 What is mammography?

Mammography is a radiographic examination that produces medical images, specified to detect pathologies in the breast, predominantly breast cancer. There are two types of mammography examinations performed: screening mammography and diagnostic mammography. Screening mammography aims to detect early stages of breast cancer in asymptomatic women. Diagnostic mammography exposes symptomatic women to evaluate suspicious structures previously identified. Mammography can also play an important role in planning the adequate treatment, when necessary [9] and in follow up and assessment of treatment outcomes.

Mammography is a bidimensional image and lacks the three-dimensional localization of structures. Therefore, the procedure usually consists of four acquisitions: for each breast, two different view modes are performed:

- Craniocaudal (CC), which is a top-to-bottom view and should show as most of the breast as possible;
- Mediolateral Oblique (MLO), which is a side view taken at a certain angle and should show the whole breast including the pectoral muscle and the inframammary angle.

Both views should have the nipple in profile, as presented in Figure 2.1, as examples.

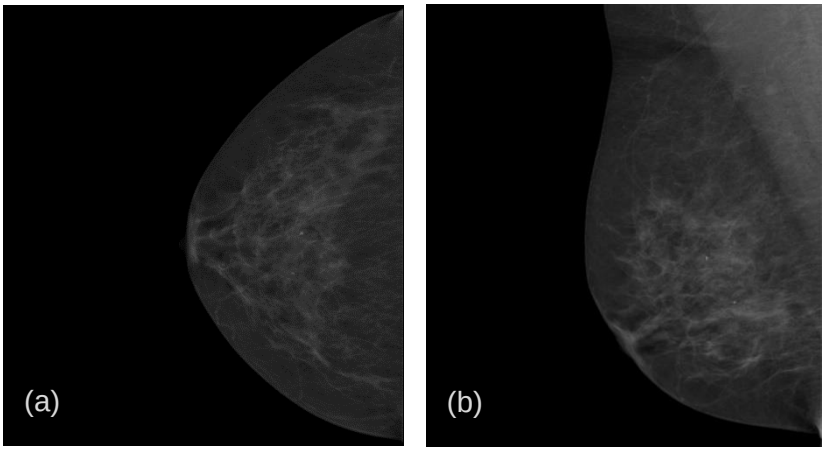


Figure 2.1 - Example of mammography images of a right breast: (a) the craniocaudal view; (b) mediolateral oblique view. (selected from the IPO – PORTO database).

An important step in the mammographic examination procedure is the positioning. The breast should be centred and its positioning reproducible as much as possible. Normally, the MLO view implicates higher compressed breast thicknesses and higher doses [10], [11], possibly due to the presence of the pectoral muscle [12]. Sometimes other specialized views are used, such as magnification, to better assess smaller details.

The detection of lesions relies mostly on the quality of the image. One imTportant aspect to improve image quality, which differentiates from a conventional radiography, is the compression of the breast.

Compression is used to spread out and separate the different tissues and structures, reducing overlapping of anatomy. Radiation dose can be lowered, and the exposure dynamic range can be lessened because the spreading of the tissues allows a reduced and more uniform attenuation [13]. It also results in fewer scattered X-rays, enhancing image contrast. The distance to the image receptor is decreased, resulting in less geometrical artefacts, and the pressure that is applied prevents blurring due to patient

movement. Even with all these benefits, compression can be uncomfortable and sometimes painful for the women being examined, and it should only take a few seconds for each view.

The maximum compression force applied should not exceed 200 newtons [1]. This parameter was shown to be one that varied significantly between practitioners, between devices and between institutions, and could cause some inconsistency of compressed breast thickness values [14]. According to a preliminary force variability study by Mercer et al. [15], the general trend was to apply higher compression to larger breast volumes.

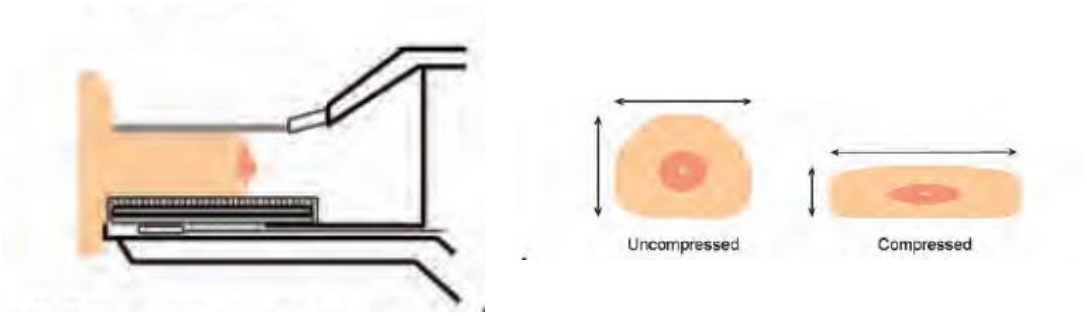


Figure 2.2 - Illustration of breast compression [17].

The appropriate compression is key to a more precise diagnosis. A compression paddle exerts pressure on the contact area between the breast and the paddle. The force measured is independent of the individual breast, as in, the same force applied to a small or large breast leads to different pressures. Too much pressure can reduce sensitivity *(the number of screen-detected cancers divided by the sum of screen-detected cancers and interval cancers diagnosed before the next screening round*) and would be more painful. Too little pressure would decrease specificity *(the number of true negative findings divided by the number of exams without cancer diagnosis)* [16].

2.3 Mammography unit

The mammography unit consists in an X-ray tube and an image receptor on opposite sides, and this apparatus should be able to rotate about a horizontal axis. The whole unit can move vertically to adjust height for patients of different statures.

A characteristic feature of the mammography system is the adjustable

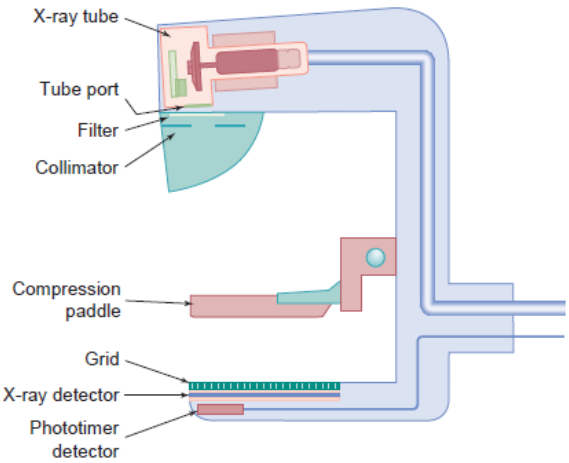


Figure 2.3 - Part of the mammography unit. [17]

compression paddle, operated through a foot-pedal. The compression of the breast is crucial, as already referenced. A compressed breast thickness is estimated through the position of the compression paddle during the exposure.

Energetic electrons, generated in a heated filament, are accelerated in the X-ray tube and led to hit a target (anode) generating bremsstrahlung and characteristic radiation. A beam of X rays leaves the tube port, goes through filters and collimator and is then transmitted to the breast. Radiation can be absorbed, scattered and transmitted through the different tissues. The radiation that leaves the breast can be first incident on an anti-scatter grid, and then it reaches the detectors, that measure the attenuated intensity and eventually form an image.

A collimator is positioned right after the X-ray output window and a face protection shield prevents patient's other body parts, like the head, from being exposed and appearing on the image, while trying to maximize the amount of breast tissue being imaged. Consequently, there is a "half-field" X-ray beam geometry, as seen in Figure 2.4. This way, the heel effect could be used as an "advantage". The X-ray intensity is higher on the cathode side and it decreases towards the anode side. Since the compressed breast thickness is higher on the chest wall and it decreases towards the anterior part of the breast (the nipple), the X-ray tube is positioned with the cathode over the chest wall, to obtain better uniformity of the transmitted X-rays through the breast.

The mammography equipment should be built ergonomically, meaning it should be of easy usage and the machine should not scare the patient, and it is also important that is easy to clean [1]

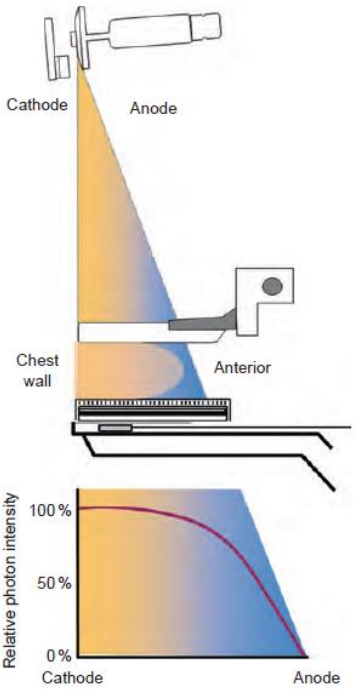


Figure 2.4 - Orientation of the cathode-anode direction of the x-ray tube, and the heel effect. [17]

2.3.1 X-ray spectrum

Breast tissues, either normal or cancerous structures, have similar attenuation coefficients. Moreover, both the attenuation coefficients and the difference between them are lower at higher energies (Figure 2.5 (a)). At higher energies, the necessary dose to produce an image is lower but so is the contrast (Figure 2.5 b)).

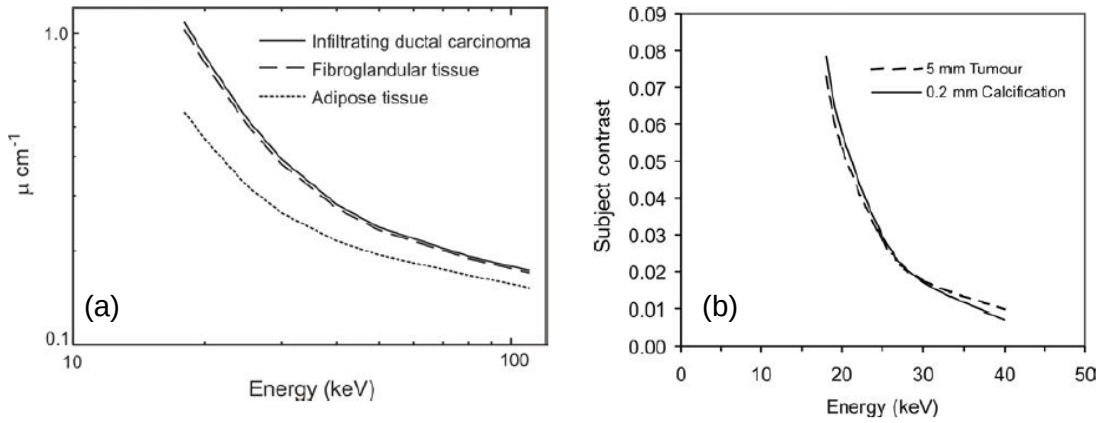


Figure 2.5 (a) - Dependence of linear attenuation coefficient with X-ray energy; (b) - Dependence of image contrast on X-ray energy. [9]

The X-Ray tube of a mammography unit produces a low energy spectrum. This spectrum will be adjusted for the specific thickness and composition of individual breasts.

The X-ray spectrum is a combination of bremsstrahlung and characteristic radiation, which represent a continuous spectrum and discrete energies, respectively. The optimum energy for film imaging is around 18-23keV [9]. Materials with characteristic X-ray production on that range, like Molybdenum (Mo) and Rhodium (Rh), are used as targets. With digital detectors, it is possible to adjust contrast during image display, and materials with higher atomic number and higher melting point, such as Tungsten (W) can be used.

To optimize the beam shape, filter materials are added. Generally, the tube port itself is made of beryllium, with a low atomic number and small thickness, allowing the transmission of all but the lowest energy X-rays. Adding extra filtration can improve the energy distribution of the X-ray spectrum, by selectively removing unwanted low or high energies. The added filtration is more attenuating at the lowest energies, which is crucial because at lower energies the radiation dose can be given to the patient without reaching the detector and consequently, not forming an image. This would result in unnecessary dose to the patient. To achieve the optimal and adequate spectrum, the materials used as filters can be Mo, Rh and silver (Ag), that have K-absorption of higher energies, between 20 and 27 keV [17].

To obtain the optimal effective X-ray energy and contrast on the image, one must select the adequate anode and added filtration materials, plus the tube voltage, kV, depending on the type of breast being imaged. The typical target/filter material combinations used by the General Electrics mammography units considered in this work are Mo/Mo, Mo/Rh and Rh/Rh (Figure 2.6). In digital mammography, combinations with other materials such

as Aluminium, Tungsten and Silver can be used as well (Rh/Al, W/Rh. W/Al) [1]. The selection of a Mo target and a Mo filter is related to thinner breasts. For denser breasts, it can be selected a target of Mo combined with a Rh filter. To obtain a more penetrating X-ray beam, a target of Rh can be used, but never with a filter of Mo, since it would strongly attenuate the characteristic energies, and a filter of Rh is normally used.

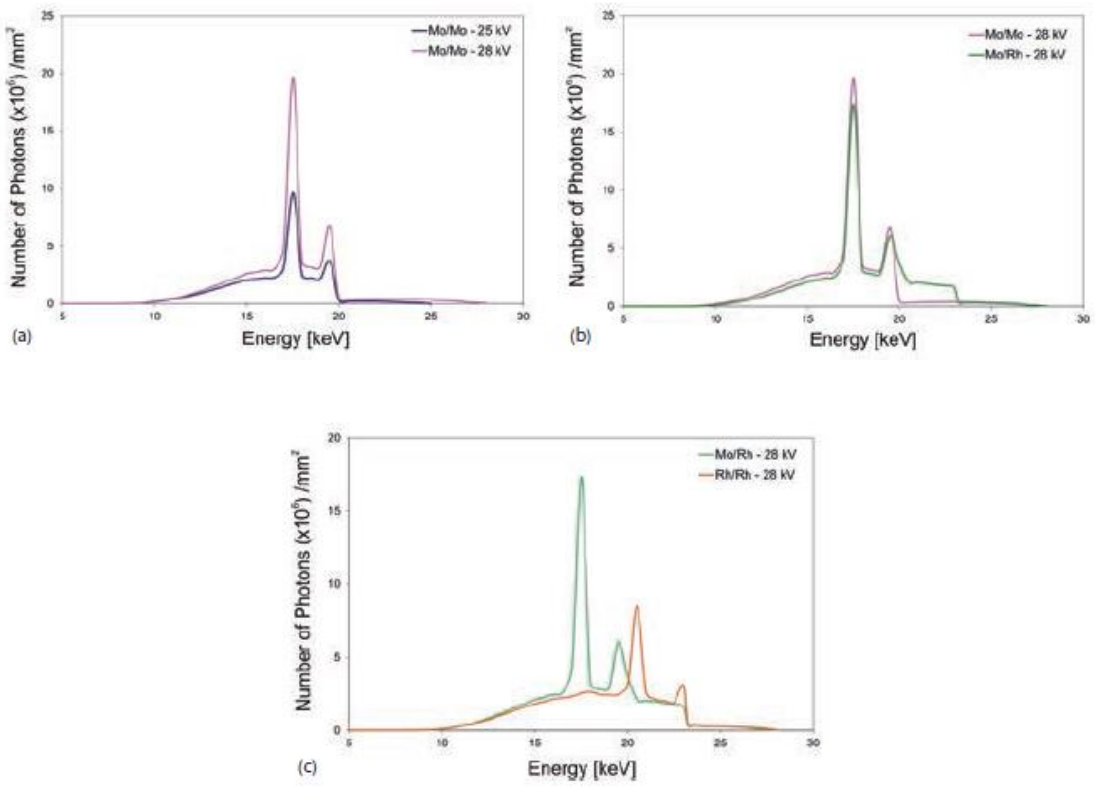


Figure 2.6 - X-ray spectrum for different target/filter combinations. [9]

The typical voltage supplied by the X-ray generator is below 40 kV, which differentiates from a conventional X-ray tube. Higher tube voltages are used for thicker and denser breasts, to obtain a more penetrating beam. Normally the combinations of the materials and the tube voltage are lower tube voltage for the Mo/Mo materials and increasing tube voltage for Mo/Rh and Rh/Rh. To not overheat the target there is a limit to the tube current, depending on focal spot size and material of the target. The smaller focal spot size is used for magnification. The smaller the focal spot size, the less maximum tube current and accordingly, the longer the exposure times. [17]

The radiation transmitted through the breast contains scattered X-rays. With increasing breast thickness, the number of scattered X-rays is higher. The scattered radiation contributes to adding random noise to the image and degrading its contrast. Even though contrast can be adjusted in digital images, the added noise degrades the signal to noise

ratio. Typically, an anti-scatter grid is present in a mammography unit. This is used to reject scattered radiation before it reaches the image receptor. The grid is located between the breast and the detector. Nevertheless, its use may impose an increase of dose, because less radiation is received in the detector which might not be enough to form the image. The use of the grid is essential for denser breasts, but with smaller breasts, reducing the scattered radiation might not compensate for the higher dose necessary [9].

2.3.2 Automatic Exposure Control

An essential component of modern mammography systems is the Automatic Exposure Control (AEC), usually incorporated with the image receptor. Originally, this function employed a set of sensors that measured the X-ray fluence and if it reached a pre-set threshold it would send a signal to stop the exposure, in order to provide the adequate optical density for films. This has now evolved into a more sophisticated, fully automatic feature, implementing microprocessors, which makes it possible to make adjustments of the technique factors during the exposure. This concept is also referred to as Automatic Optimization of Parameters (AOP) [18]. Depending on the thickness and exposure parameters, the penetration of radiation in the breast varies, and thicker breasts being irradiated with lower tube voltages require longer exposure times. Longer exposure times result in patient discomfort and possible movement, thus, the AEC system should consider the acquisition time when optimizing the exposure parameters.

The AEC system does a low dose pre-exposure. This trial exposure is very fast, typically <100 ms, and in some digital mammography units it can form a complete low dose image, so it can determine the transmission through the breast. Some AEC systems even find an area of highest attenuation within a defined area of the detector during the pre-exposure. The system then applies an algorithm, that based on the compressed breast thickness and the X-ray transmission, infers the breast composition and automatically selects the tube loading (tube current x exposure time in units of mAs) and the combination of tube voltage (kV), target and filter materials to apply on the actual exposure. All these parameters are selected to achieve a predetermined dose to the detector, contrast, SNR or any image quality parameter set by the manufacturer [19]. So it is possible to optimize the image according with the needs: lower doses, better image quality or a combination of both. The system should have algorithms to weight the parameters, taking into consideration the dose reference levels as well as the technical restrictions of the machine.

2.4 Direct digital imaging

Breast cancer can be detected through masses, with or without spiculations, microcalcifications and distortions of breast structures. Particularly, the microcalcifications, require detectors with a high spatial resolution. Therefore, mammography images must have high spatial resolution, which requires large numbers of pixel per image [17].

For many years, screen-film was the technology used as the detector and archiving of mammography images. Its small dynamic range was one of the major problems, preventing the detection of some lesions. Digital mammography can overcome that issue, while allowing to acquire, visualize and store images, separately.

In digital mammography, it is important to achieve a certain signal-to-noise ratio (SNR) at the detector. With these digital systems, it is possible to adjust the contrast while displaying the image. This results in the reduction of the dose of digital systems in comparison to screen-film.

Digital images can be accomplished through an indirect conversion method, like computed radiography (CR). It consists of reading out a photostimulable storage phosphor (PSP) plate that is kept in a cassette after it is exposed to radiation. The direct conversion method of direct digital radiography (DDR) replaced phosphor with a photoconductor material that absorbs X-rays (usually amorphous selenium). Studies have found that the performance of CR systems is lower than of DDR systems, having higher doses [10], [12], poorer image quality [20] and lower cancer detection rates [21].

Direct digital imaging in mammography can be accomplished with Full-Field Digital Mammography (FFDM). The detector can contain many detector elements (dels) arranged in an array of around 2000-3000 columns and 3000-4000 rows [9]. These dels are photoconductors that collect the signal as electrical charge. The charge is then converted to digital values which are transferred to a matrix that produces the image.

Advantages of digital systems go from improving the resolution, contrast and SNR, to the lower radiation doses and the ability to acquire instantaneous images that can be stored and transmitted electronically. Images can also be submitted to processing techniques and manipulation of settings to better visualize specific regions.

The potential loss of subject contrast is compensated by the improvement of display contrast during the visualization and some image processing. While aiming to obtain images of better quality there was a tendency of increasing doses, a phenomenon referred to as “dose creep”, because the image itself no longer presented signs of

overexposure (as it did for screen-film images) [19]. This can be prevented with the indication of dose indicators attached to the image information.

2.5 DICOM

With digital images, it is useful to assign to each image information related to the patient, the acquisition technique and dosimetry. This kind of information is important for comparison of data from different manufacturers and institutions.

To determine a set of standards to help the transfer of radiological images and related information, the American College of Radiology (ACR) and the National Electrical Manufacturers Association (NEMA) joined and defined a mechanism to encode pixel data together with information about the images, in a list of data elements, as well as the means to exchange data with point-to-point communication (ACR-NEMA 300). In 1993 an updated version was created, taking advantage of the developments of local area networks, and the name was changed to DICOM (Digital Imaging and Communications in Medicine) [22]. The standard was later applied to other medical images, and even other areas, and is, to this day, always being improved.

The standards may differ for each specific modality such as DICOM CT (computed tomography), DX (digital x-ray), MG (digital mammography), US (ultrasound), MR (magnetic resonance) and NM (nuclear medicine).

The standard describes IODs (information object definitions), that are specific to each type of modality of the image. There is a different IOD for mammography and ultrasound, for example. The different modalities still share common information about the study and the patient, but different information associated with the acquisition technique and encoding of the pixel data. These types of information are organized into Data Sets. The patient Data Set includes the identification of the patient, that sometimes must be anonymised, and the study Data Set includes information about the date and time of the procedure, for example.

A Data Set consists of a Data Element tag (identifies the Data Element), the Value Representation (specifies the type and format of the data), Value Length (contains the length of the Data Element) and the Value [23]. In a display of the DICOM header like the one seen in ImageJ2 (Figure 2.7), it only shows the Data Element tag, its name and value.

The Data Element tags are represented as (gggg,eeee), where gggg corresponds to the group number and eeee to the element number. For example, the patient's name is attached to the tag (0010, 0010) and the study date is assigned to the tag (0008,0020),

which belong to two different groups of Data Sets. The list of registered DICOM elements can be consulted on the DICOM Data Dictionary [24].

Raw pixel data is encoded in Data Element (7FE0,0010), and the rest of the information is usually in the “header”, that can be of different lengths.

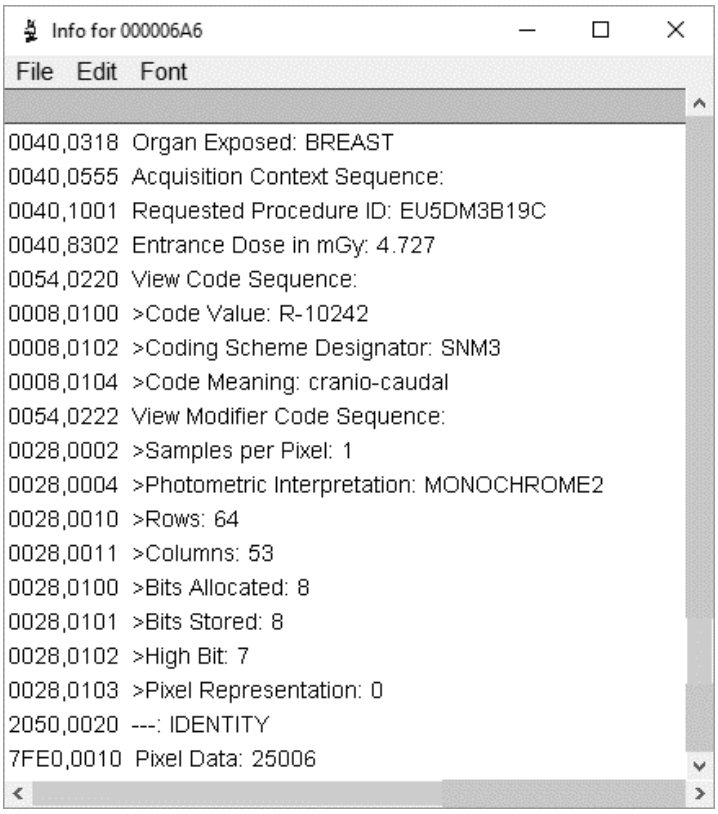


Figure 2.7 - Example of DICOM header, presented through ImageJ, and including the Pixel Data tag.

In order for the pixel data to be correctly displayed, the DICOM standard also offers parameters such as the Presentation State, the Grayscale Standard Display Function (GSDF) and the Pixel Intensity Relationship.

The mammography modality (MG) usually provides two types of DICOM image, the ‘for processing’ (Raw) image and the ‘for presentation’ (Processed) image.

The raw image should be available to allow evaluation of the image receptor. These images usually present a linear relationship between the dose of the data element and the corresponding pixel value [1]. Generally, some preliminary operations are still applied, like a flat field or gain correction, to correct some non-uniformities of the detector or the X-ray field itself (Heel Effect).

Other processing operations are applied to form the processed image, and these vary among manufacturers and devices. The goal is to adjust the image for a more suitable display and interpretation, improving the perceptibility of clinically relevant information.

Some of those operations include thickness equalization at the edge of the breast, inversion of the greyscale, a non-linear transformation (e.g. logarithmic) of the pixel data, noise reduction and contrast optimization [19].

2.6 PACS and RIS

Digital mammography images can take up a lot of space, considering their size, the number of images per exam (at least four) and whether the “for processing” images are also stored.

Some important components of a department with digital radiological imaging, such as mammography, are the Radiology Information management System (RIS, that can be a subdivision of a Hospital Information System (HIS)) and the Picture Archiving and Communication System (PACS) [19].

PACS is the system that stores, transfers and displays digital radiological images. This way, it is more efficient for a radiologist or a physician to access a certain image, in different viewing stations, within the facility, assuming the image is correctly identified.

RIS is an information system, that among other functionalities, maintains a patient database. This database includes personal information about the patient as well as records of current and previous clinical conditions and other useful information.

There should be a synchronization between the PACS and the RIS/HIS, for the images to correspond to the right patient record. More relevant information is given to the one interpreting the image, and the new information can be sent back to RIS and/or HIS and they’re updated.

With these systems, images cannot be lost (there still should be backup systems), and they can be replicated and used for studies and teaching if they’re anonymised [9].

Retrieving images and dosimetric data automatically from PACS is now possible, which allows to perform quality assurance and dose audits in large datasets.

2.7 Dose

The radiation of a mammographic exam can impose a risk of cancer induction, so there is a concern, especially in the case of screening many asymptomatic women, to not give excessive and unnecessary dose. It is of the most importance to monitor and optimize doses in mammography exams, as is required by many regulations.

The main parameter that is used as a radiation dose index (RDI), in mammography exams, is the average dose absorbed by the glandular tissue of a breast, uniformly compressed, referred to as mean glandular dose (MGD) or average glandular dose

(AGD). The glandular tissues, during an exposure, will receive different quantities of dose depending on the depth. Estimating it helps determine the associated risk of induced cancer due to the mammography's radiation. This quantity depends on the beam quality, the breast thickness and its composition.

There are different methods to estimate MGD, and mammography units of different manufacturers display organ dose in many forms and are not always clear about its estimation. GE systems, such as the ones used in IPO-Porto and considered in this study, follow the Wu et al. approach [25]. If we want to compare dose levels of a series of mammography examinations between institutions, or even countries, there should be a reference method. European guidelines (EUREF [1]) use the Dance et al. method which shall be used here to estimate MGD.

The practical quantity useful for dosimetry in mammography is the entrance surface air kerma (ESAK), which is the air kerma at the upper surface of the breast, in the absence of scattering. It is used to estimate MGD and can be determined as:

$$ESAK = Output \times mAs \left(\frac{FDD}{FBSD - CBT}\right)^2 \ (mGy) \tag{2.1}$$

Where the output corresponds to ESAK per unit tube loading (mAs), which is determined for the mammography unit used. FDD is the focus-detector distance, FSBD corresponds to the focus-breast support distance and CBT is the compressed breast thickness.

The MGD to a typical breast of thickness and composition can then be estimated as:

$$\mathrm{MGD} = \mathrm{ESAK}\, \mathrm{g}\, \mathrm{c}\, \mathrm{s} \ \ (\mathrm{mGy}) \tag{2.2}$$

The factor g corresponds to 50% of glandularity, this is, a model breast with a central region of an equal mixture by weight of adipose and glandular tissues. The factor c is a correction factor to account for the composition of the breast, in comparison to the 50% glandularity. Dance et al. presented the typical values of breast composition in the age ranges of 40-49 and 50-64 [26]. The s-factor is a conversion factor related to the X-ray spectrum. These coefficients were estimated through Monte Carlo radiation transport simulations and a simple model of the breast being imaged. Their values are tabulated as a function of the half value layer (HVL), exposure parameters, breast thickness and typical glandularity (which varies with the age of the woman) and are published in EUREF [1].

Dose increases with compressed breast thickness and breast density, where compressed breast thickness has a greater effect on MGD [13]. Higher HVL (higher kV) increases the penetrability of the X-ray beam, reducing the MGD but also decreasing

contrast. The attenuation of the breast tissues is lower at higher energies, and to obtain images with good quality there's a minimum required energy that needs to be transmitted to and absorbed by the receptor. The necessary kerma will decrease, with higher energies, while the conversion coefficients will increase, but more slowly, so ultimately MGD will diminish as energy increases [9].

2.8 Diagnostic Reference Levels

A radiological medical exam, such as a mammography, does not actually have a dose limit. Instead, individual justification and optimisation are considered. As a diagnostic procedure, a mammography exposure should provide sufficient image quality to obtain the necessary diagnostic information, at the lowest reasonably possible dose.

In this context, the Medical Exposure Directive (97/43/Euratom) [27] defined the Diagnostic Reference Levels (DRL) as "dose levels in medical radiodiagnostic practices (…) for typical examinations for groups of standard-sized patients or standard phantoms for broadly defined types of equipment".

In standard procedures, these levels should not be exceeded. When they are systematically exceeded, the procedures and/or the equipment must be reviewed, and corrections should be made.

Dose values below the DRL do not necessarily correspond to good practices. These reference levels simply serve as guidance and comparing DRLs between different populations could allow to understanding what dose levels can be achieved. Ultimately, this could lead to decrease the number of exams with excessive dose.

The DRLs are not to be applied to an individual patient, but to a group of standard sized patients or a standard phantom and can be established as the rounded third quartile (75% percentile) of typical dose values of a population. Doses of standard sized patients or phantoms are measured and then compared with the DRLs. When the number of standard sized patients is insufficient, the average dose of all patients available is taken.

DRLs are an important tool for clinical audits, which can provide a basis for a retrospective evaluation and for recommendations to improve procedures.

With these values established, doses can be compared between facilities, equipments and countries, and doses could be reduced while changing working procedures, seeing that other facilities can achieve lower doses.

Portugal has yet to define diagnostic reference levels for mammography [10]. The European DRL for both CC and MLO views is set at 10mGy of Entrance Surface Air Kerma [27].

One can look also into the (updated) achievable and acceptable maximum MGD that are suggested in European guidelines [28] for different PMMA thicknesses:

Table 2.1 - Calculations of detector dose and comparison to the DICOM indicated dose parameter.

Thickness of PMMA	**Equivalent breast thickness**	**Maximum average glandular dose to equivalent breasts**	
		acceptable level	achievable level
[cm]	[cm]	[mGy]	[mGy]
2.0	2.1	< 1.2	< 0.8
3.0	3.2	< 1.5	< 1.0
4.0	4.5	< 2.0	< 1.6
4.5	5.3	< 2.5	< 2.0
5.0	6.0	< 3.0	< 2.4
6.0	7.5	< 4.5	< 3.6
7.0	9.0	< 6.5	< 5.1

- c

At IPO-Porto, mammography examinations are not usually of the screening nature and are not necessarily of the standard sized patients. For the most part, these exams are performed on women that need to be carefully followed and even women who are undergoing treatment and its follow ups. So, should the measured doses in these cases have a different set of DRLs? If we want to implement the same principle and lower the dose values on, already, unhealthy patients, maybe this should be considered.
Normally, to conduct a retrospective study, images would have to be selected manually, so that only images of “standard” procedures were considered.
Nowadays there are systems that automatically collect all dosimetric data from the PACS. These are called Radiation dose index monitoring (RDIM) systems. With these automatic systems, dosimetric data is collected, in large scale, whether it pertains to standard procedures or not. It is important to know what effect treatments may have on exam doses, and if this effect is sufficiently high to justify separate DRLs for treated breasts.

2.9 Breast Density

Breast density usually describes the composition of the breast as having a certain percentage of glandular tissue and a complementary fraction of adipose tissue (fat) [29]. The glandular fraction of the breast is the nonfatty, dense tissue that appears as the white, radio-opaque, regions of the imaged breast. The breast density is sometimes

described as glandularity as it refers to the percentage of the breast imaged area occupied by the glandular tissue [30].

Denser breasts have a higher percentage of glandular tissue and are usually related with higher breast cancer risks. Its measurement could be considered for breast cancer risk prediction and prevention strategies [31]. Breast density decreases, on average, with increasing age [32].

Breast density can be assessed with mammography imaging and there are many suggested methods to estimate and classify this quantity. Originally it was visually estimated by radiologists, accounting for the proportion of "white" glandular tissue area on the imaged breast area. It is considered to range from 0% to 100% (from mostly fat tissue to mostly glandular tissue). Mammographic density is now sometimes classified with a qualitative scale known as BI-RADS® (Breast Imaging Reporting and Data System) [33], that categorizes the breast composition in four categories, described in the next Table.

Table 2.2 – Breast Composition Categories as defined by the American College of Radiology [33].

Breast Composition Categories
a. The breasts are almost entirely fatty
b. There are scattered areas of fibroglandular density
c. The breasts are heterogeneously dense, which may obscure small masses
d. The breasts are extremely dense, which lowers the sensitivity of mammography

These categories no longer represent intervals of percentage glandularity to "emphasize the text descriptions of breast density, which reflect the masking effect of dense fibroglandular tissue on mammographic depiction of noncalcified lesions, because the Committee on BI-RADS® concludes that the association of subjectively estimated breast density with changes in the sensitivity of mammography is clinically more important than the relatively smaller effect of percentage breast density as an indicator for breast cancer risk".

But again. the visual assessment of breast density can be subjective. With digital mammography systems there were different developments on quantitative automated methods to estimate breast density [34]–[38]. Digital mammography units can estimate breast density while performing the exposure. Namely, the GE Senographe models calculate glandularity percentage values based on a pre-exposure image and the signal levels on a detected denser area [39].

The fact that in mammography, the breast density is being estimated from projected two-dimensional images could be limiting. Even with estimating a volumetric density, these

estimations are based on assumptions and models that may not apply for individual patients [31].

2.10 Breast cancer treatment

IPO-Porto is the institution that treats the highest number of breast cancer patients in Portugal [40].

Depending on the breast cancer stage and the tumour's type and size, the treatment options and its planning will vary. The patient's age and clinical history also plays an important role on deciding the best suitable treatment. All cancer treatments aim to eliminate cancer cells, but sometimes remove healthy surrounding cells as well.

Breast conservation treatment is a localized treatment and is usually a combination of surgery and radiotherapy, which are the oncological treatments that some of the patients considered in this study were submitted to. Scars and surgical clips can be identified in postoperative mammography images. There are reports of histological changes in the breast's tissues, after the radiotherapy [41]. These changes may affect later diagnostic and follow up imaging performed months or years later [42].

3 Materials and Methods

3.1 Mammography units

There are two mammography units at IPO-Porto: both are digital mammography units from General Electrics (GE) - one Senographe DS model acquired in 2005 and one Essential acquired in 2016.

These digital mammography systems have a flat-panel detector (CsI scintillator and amorphous silicon matrix), a X ray tube with two anode tracks (Mo and Rh), as well as two spectral filters (Mo and Rh).

These units provide an AOP (Automatic Optimization of Parameters) operating mode. The operator can choose manual exposure mode or one of the three automatic exposure modes available that prioritize different parameters:

Figure 3.1 – Senographe Essential of General Electrics. [39].

- CNT (contrast) – higher contrast to noise ratio (CNR) but higher dose;
- DOSE (dose) – lower dose but lower CNR;
- STD (standard) – compromise between CNR and dose.

The STD mode is most commonly used as it should satisfy most needs. In AOP mode, the system makes a pre-exposure, producing a low dose image. The equivalent breast thickness[1] and breast composition are determined from this image's signal levels, on the densest part of the breast. These values correspond to a set of the exposure parameters (kVp, target and filter combination), that are automatically set to acquire the final image [18]. The tube loading (mAs) is selected according to the AOP mode, the breast density and equivalent breast thickness. MGD is calculated from a computed ESAK value and using interpolations of the tables published by Wu et al. [39].

The Senographe DS is a normal Field of View (FOV) system while Senographe Essential is a large FOV system. This is accomplished with a larger detector that allows the examination of large breasts with a single exposure per view. The sizes of the images for each of these mammography units and their pixel size are indicated in Table 3.1.

[1] The equivalent thickness of an object is defined as the thickness of a reference material such as PMMA that, under certain radiological conditions and exposure parameters, would provide the same signal on the image receptor as the object [61]

When large breasts are imaged with a smaller detector, there is a need to make various exposures for the same view, which increases the dose to some regions. In addition, the interpretation process becomes more complex, with more images to consider.

Table 3.1– Image and pixel size of the two GE mammography units.

Model	Image size (in pixels)	Pixel size	Image size (cm)
DS	Width = 1914 Height = 2294	0.0941 x 0.0941 mm^2	18.01 cm x 21.58 cm
Essential	Width = 2394 Height = 3062	0.1 x 0.1 mm^2	23.94 cm x 30.62 cm

These units apply several image processing methods to the initial raw images. Namely, a black mask is applied to cover areas outside the useful image area that would appear as white and a “pseudo-log transformation” helps the manipulation of brightness and contrast levels. Other processing that these units offer are auto-contrast (optimizes the window level and width), thickness equalization (decreases the image dynamic range), fine view (FV, increases the sharpness of the image) and premium view (PV, increases the visibility of breast structures) [39].

3.2 Software

Throughout this work, there was need to use some software for visualization and analysis/manipulation of images. To rapidly visualise and analyse mammography images, the ImageJ2, an updated version of the software developed by Wayne Rasband at the National Institutes of Health [43], was used. To further process the images and retrieve information from them, Python 3 [44] scripts were developed.
All retrieved data was further analysed with recourse to MS Excel.

3.3 Preliminary characterization: Study of phantom images

Mammography exams in PACS consist of images that are only available in the “for presentation” version. A set of phantom images obtained during quality control tests, conducted on June 29th, 2017, in a GE Senographe DS system, were used to study and understand the functioning of a direct digital mammography unit. The first two sets of acquisitions considered both raw and processed images and were for the study of the response function and noise evaluation of the image receptor and the study of breast thickness and composition compensation. The third set of images was used to study the behaviour of the system when it images a local dense area and included two different acquisitions: with and without Premium View.

Both raw and processed phantom images were used, to investigate the effects of the image processing.

3.3.1 Response Function and Noise Evaluation

This test requires imaging a standard test block (uniform PMMA block of 45mm) with different tube loading (mAs) values. The setup is presented in Figure 3.2.

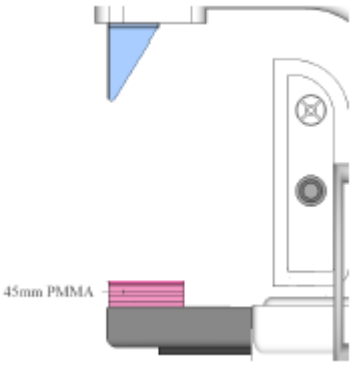


Figure 3.2 – Side view of the response function measurements. (Adapted from EUREF supplement [45])

Mean pixel values (MPV) and standard deviation (SD) for 12 raw images were measured in a 5mm x 5mm centred region of interest (ROI) using the *ImageJ2* software. The mAs and a dose parameter stored in the DICOM header were also retrieved. The signal to noise ratio was determined as follows:

$$SNR = \frac{MPV}{SD} \tag{3.1}$$

To understand the functioning of a digital mammography unit, the relation between the MPV and the mAs (and hence the dose) of the raw images was assessed. A linear relationship is to be expected [1].

The noise of those images was also studied, by looking into the relation of the standard deviation and the mean pixel value. The noise can be split into quantum, structure and electronic noise, by adjusting a 2nd degree polynomial trend line. This corresponds to:

$$SD^2 = k_e^2 + k_q^2 \times p + k_s^2 \times p^2 \tag{3.2}$$

- SD = standard deviation in the reference ROI
- k_e = electronic noise coefficient
- k_q = quantum noise coefficient
- k_s = structure noise coefficient
- p = mean pixel value in the reference ROI

The dominant component should be the quantum noise [45].

3.3.2 Detector Dose

After confirmation of the linearity between MPV and mAs, the dose to the detector was estimated with the phantom images of the response function study. The information needed to its calculations was provided and is presented in the Table 3.2.

The exposure factors were 29kV and the RhRh targer/filter combination, at 20mAs and with grid.

Table 3.2– Information necessary to determine the detector dose.

29kV RhRh 10mAs		**Value**
Yield (µGy/mAs) at 1m		31
MPV with grid		3896
MPV with grid + 2Al		505
MPV without grid + 2Al		707
Aluminium	Attenuates: (7.7x)	7,72
Grid	Attenuates	1,4

Considering a point at 1m from the source, if we have 9 mAs and a yield of 31µGy/mAs, the kerma at that point would be:

$$kerma = 9mAs \times \frac{31\mu Gy}{mAs} = 279\mu Gy = 0.28\, mGy \quad (3.3)$$

The source-detector distance is set at 66cm, and this includes the support of 1cm. Outside the source, at 20cm, there is attenuation from 2mm of aluminium, as the scheme in Figure 3.3 shows.

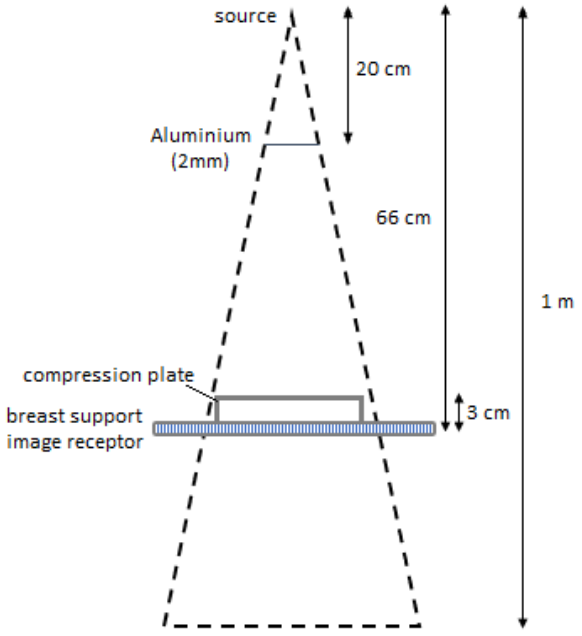


Figure 3.3 – Detector Dose measurement.

The dose at 66cm, neglecting the aluminium and the grid, should be calculated according to the inverse square law. Considering the attenuation factors presented at Table 3.2, these are multiplied to the dose at 66cm, to obtain the detector dose.
These values were then compared to the detector exposure indicated by the manufacturer.

3.3.3 Comparison of "for processing" and "for presentation" images

The MPV and SD of the same ROI, mentioned above, were also measured on the processed images. In a straightforward approach, one can plot the MPV of the processed images and the raw images together and draw a trend line to get a sense of what type of processing is being applied to the raw images.
To obtain a more precise trend line equation, the average of the terms obtained from 24 different ROIs were considered to measure the mean pixel value: 12 smaller ROIs (5mm x 5mm) and 12 bigger ROIs (30mm x 30mm) centred as shown in Figure 3.4.

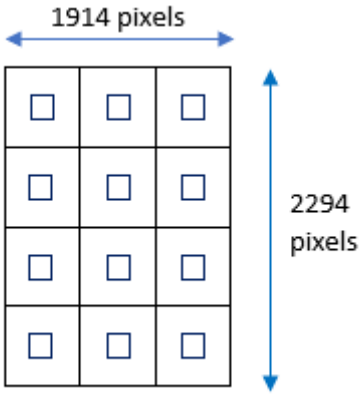


Figure 3.4 – Positions of regions of interest (ROI) considered to determine the relation between the PROC and RAW images.

To comprehend what further processing might be being done, a Python Script was written to apply the mathematical operation of the trend line to the raw images and then compare, pixel to pixel, the resultant images with the corresponding processed ones. For this, we subtract and divide these new images, to the processed images (PROC) and analyse what remained.

$$sub = processing(RAW) - PROC \quad (3.4)$$

$$div = processing(RAW)/PROC \quad (3.5)$$

If the mathematical operation is all that the mammography unit applies we should expect to obtain images of zeros and ones for the subtraction and division images, respectively.

These new images were saved as text files to keep the pixel values while visualizing them on *ImageJ2*. Parameters like MPV, SD, minimum, maximum, median and the histogram of these images were retrieved.

3.3.4 Image processing effects

Two different acquisitions were taken to study the behaviour of the mammography system when imaging denser areas and its features of image enhancing "Premium View" (PV) and "Fine View" (FV) [39]. The first acquisition was without PV[2], and the second one with PV.

The setup consists of three PMMA plates of 10mm thickness, with the compression paddle above 10mm spacers, as exemplified in Figure 3.5. On the compression paddle, different sets of stacks of smaller PMMA plates (2mm thick) are positioned in a central area within the AEC sensor area.

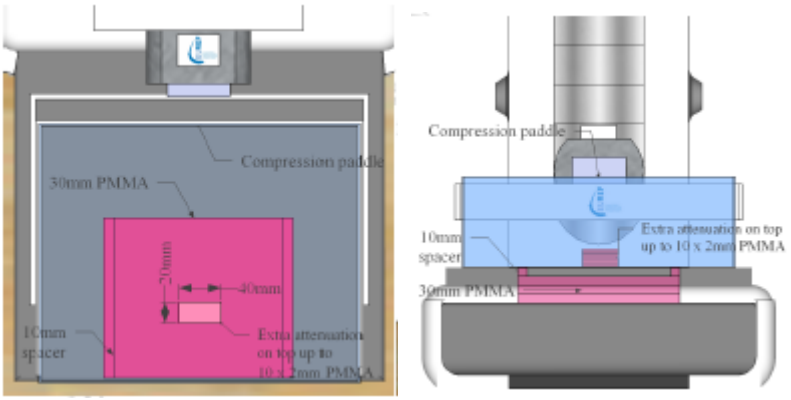


Figure 3.5 – Top view and front view of the setup for the local denser area measurement. [45]

In total, 11 images were retrieved, along with their exposure factors, of sets of small PMMA plates from zero plates to 10 plates (20mm), for each mode of acquisition (with and without PV). In fact, both raw and processed images were saved, and 22 images were then being considered for each mode.

3.3.4.1 Local dense area (LDA)

This test consisted in measuring the pixel values and standard deviation in a small ROI in the area of extra attenuation and then calculating the SNR of each image and the average of them all. The extra attenuation should be detected, so the exposure of the images with increasing thickness is expected to increase. EUREF set, provisionally, that the SNR values for each image should be within 20% of the average SNR [1].

[2] When referring to images with PV, FV is also present on those images.

3.3.4.2 Premium View

GE Medical Systems offer a post processing software, integrated in the units used in this study, that aims to improve the diagnostic performance. This algorithm is proprietary and it is supposed to help radiologists make a diagnosis, yielding higher cancer detection rates [46].

The mean pixel values and standard deviations were measured in the background and in the local dense area regions for both sets of images with and without Premium View.

A profile line was drawn to pass through the LDA, and the mean pixel value was measured.

To compare pixel to pixel the images with and without PV, the same method described in equations 3.4 and 3.5 was done (page 23). A direct comparison of the raw images with and without PV, and the processed images with and without PV, was not conducted pixel to pixel because the images are from different acquisitions.

3.4 Study of clinical images

3.4.1 Sample Characterization

The clinical images considered in this study are processed images with FV and PV, as raw images are not usually sent to PACS.

At IPO-Porto the mammography exams, as referred before, are not regularly performed for screening. The images retrieved for this study refer to patients who perform regular exams. Some patients who have been closely monitored to assess benign situations through the years and some patients who were being followed to assess treatment outcomes.

All exams considered consisted of the standard four views, in standard automatic exposure mode.

3.4.1.1 Normal variation between exams to the same patient.

To establish the baseline of normal variation of positioning and technique, dosimetric quantities in mammography to women over the years werecompared.

Ten patients were randomly selected among those who had multiple mammographic exams archived in PACS and no previous history of surgery or radiotherapy. These patients were submitted to between five and nine exams, performed in a GE Senographe DS mammographic unit and some with the GE Senographe Essential. The exams were taken between 2009 and 2017.

Table 3.3 – Number of exams per patient for the normal variations study.

Patient	Number of Exams	Number of Exams on Senographe Essential
A	7	1
B	5	0
C	7	1
D	9	1
E	9	0
F	6	1
G	7	2
H	6	0
I	8	1
J	5	1

In total 276 images corresponding to 69 exams were considered, for this study.

3.4.1.2 Normal variation between the left and right breasts and variation between treated and untreated breasts

3.4.1.2.1 First sample (DS 2014-15)

A series of over 400 mammography exams, performed in the same mammographic system (GE Senographe DS) at IPO-Porto were collected between February 2014 and January 2015 for an internal dose audit. These exams referred to women who were simply being monitored after diagnosis of benign conditions, as well as patients who had undergone treatments at the institution.

From that sample, exams from women who had undergone unilateral surgery alone and unilateral surgery plus radiotherapy were manually retrieved from the Picture Archiving and Communication System (PACS). Women that had correction of breast asymmetry and performed biopsies were excluded. For reference, exams of women who were not submitted to any treatment were also retrieved from PACS.

Exams with compressed breast thickness below 20mm were excluded (one patient). Cases with more than 4 images or with repeated views were also excluded. The final first sample consisted of 620 images from 155 anonymized patients, with ages between 33 and 81 years.

The mammography exams were classified according to the treatment of the patients being examined:

- Patients with no treatment, who were simply being monitored for benign conditions, or being diagnosed after detection of suspicious symptoms;
- surgery patients of left (L) and right (R) breasts;
- surgery and radiotherapy patients, of left and right breasts,

as presented in the next table. The patients that are labelled as surgery only are mostly ones that have had the mammography exam performed sometime after surgery and before radiotherapy.

Table 3.4 - Number of patients for each case of study of the first sample, DS 2014-15.

Type	No treatment	Surgery		Surgery and Radiotherapy		Total
		Left Breast	Right Breast	Left Breast	Right Breast	
Number of patients	62	20	10	29	34	155

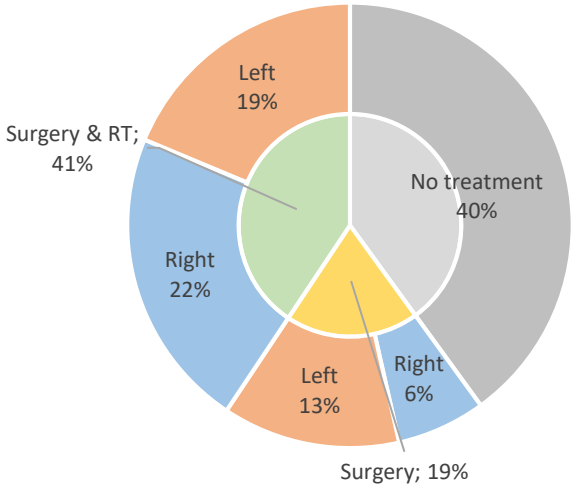


Figure 3.6 - Visual demonstration of Table 3.4.

Patients who had radiotherapy, usually had surgery performed on the same year or the year before. The time span between the radiotherapy treatment and the mammography exam in this sample is mostly 3 years or less. Some of the exams (about 10% of the sample, and 25% of the radiotherapy cases) were performed 4 or more years after the treatment, but in these cases the time of the treatment could not be precisely determined. This sample is referred to as “DS 2014-15” throughout this document.

3.4.1.2.2 *Second sample (DS 2017-18 and Essential)*

After an initial analysis and optimization of its methods, a second, larger sample was retrieved to complement the first.

Mammography exams performed between January 2017 and September 2018, in the newer mammography unit, the GE Senographe Essential, and some others performed in the GE Senographe DS, of IPO-Porto were considered. Exams to patients who had

radiotherapy treatment in 2013, following surgery, or no treatment at all for comparison, were retrospectively selected from PACS. This way we could study if the results from the first sample were reproducible and if the effects of the treatment were present still after 4 years.

Exams with extra or different views (neither CC nor MLO), had those views excluded, to have 4 views per exam. The criterium to reject the repeated views had to do with better positioning of the breast. Some patients with large breasts with exams performed in the Senographe DS unit had more than 4 views performed of incomplete breasts and were excluded. This reduced the number of exams with high compressed breast thicknesses for the DS unit. One exam performed in the Senographe Essential had a too high compressed breast thickness, of over 110mm, and was excluded.

Exams with initial misidentification were detected and corrected after applying a surgical clips detection method, explained later in this document.

The retrospection method of retrieving this data sample yielded exams in 2017 and 2018 of the same persons, sometimes on the same unit, sometimes on different ones. When a repeated patient existed the 2018 exam was used, except for the no treatment case of the Senographe DS which had less exams.

In the end, the second sample is resumed in the next table and graph, with a total of 244 exams and 976 images considered, of anonymised patients with ages between 33 and 86 years.

Table 3.5 - Number of exams for each case of study of the second sample, DS 2017-18 and Essential.

Mammography unit	No treatment	Surgery and Radiotherapy		
		Left Breast	Right Breast	**Total**
Essential	50	56	51	157
DS	29	30	28	87
Total	79	86	79	244

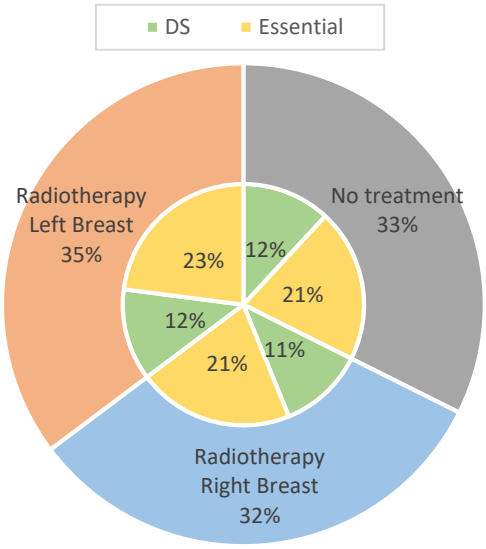


Figure 3.7 - Visual demonstration of Table 3.5.

The data from the Senographe Essential is referred to simply as “Essential” and the information corresponding to the data of this second sample and from the Senographe DS is referred to as “DS 2017-18”.

The 40 repeated patients who had exams on both mammography units[3] were set aside for consideration to compare more precisely the performance of the two mammography systems.

3.4.2 Studied Parameters

The *pydicom* library allows us to access the DICOM header. A Python script was written to retrieve information from all the images. The parameters considered for our study are presented in the Table 3.6, as well as the method of their retrieval.

All the relevant information was stored in a MS Excel file. Information from CC and MLO views were separated for appropriate analysis. Patients were also categorized with their clinical history regarding surgery and radiotherapy treatments to which breast. The detection of surgical clips helped confirm the treatment and treated side, since, even though rarely, it could be misidentified.

The relationships between some of these parameters were plotted, such as:

- MGD and Compressed Breast Thickness
- ESAK and Compressed Breast Thickness x Imaged Breast Area
- Glandularity and Compressed Breast Thickness

[3] Corresponding to 80 mammographic exams and 320 images.

Table 3.6 – Studied parameters and their extraction methods.

Parameters	**Method**
• Mammography unit • Pixel size • Image size • Patient's age • Acquisition date • Acquisition time • Compressed breast thickness • Compression force • View Position • Image Laterality • Exposure control mode • mAs • kV • Exposure Control Mode Description • Glandularity	Extracted from the DICOM header
• MGD (organ dose) • ESAK	Extracted from the DICOM header and estimated independently using the method recommended by EUREF [1]
• Contour • Area • Maximum width • Mean pixel value • Standard Deviation (of pixel values) • Minimum pixel value • Maximum pixel value • Detected surgical clips	Computational processing (explained further ahead)

3.4.3 Computational processing

The pixel data of mammography images can be extracted as an array, through the *pydicom* library [47] of the Python language, for example. This way manipulating it, detecting and measuring features can be done with computational programming.

3.4.3.1 Contour finding and calculation of imaged breast area

A method to delineate the breast contour was developed. The aim was to create a mask that applied to each image would give a new image with the background pixels as zero. After this, parameters like area and distances would be calculated, and other analysis could be done.

Using the Python library *skimage* [48], a first approach included applying some edge operators, such as the *sobel* and *canny* filters, to a smaller sample of three random patients. The results of applying these filters to obtain a mask weren't satisfactory and some examples are presented in Figure 3.8.

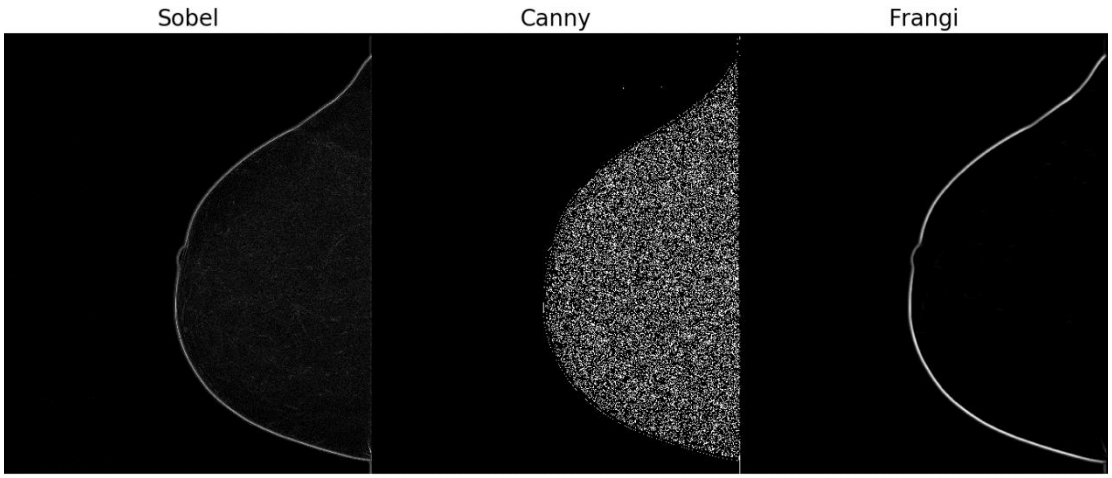


Figure 3.8 – Examples of application of the Sobel, Canny and Frangi filters.

Sobel presented a bright contour but also showed edges and other structures inside the breast. This edge operator performs discrete differentiation that computes "an approximation of the gradient of the image intensity function" [49] and the results are usually noisy, so applying a denoising filter first could help, but it wasn't the case here.

Applying the *canny* filter did not yield better results. As seen in the figure above, it results in a lot of thin and separated lines. The *canny* filter is an edge detector that first applies a derivative of a Gaussian filter to compute the intensity of the gradients, reducing the effect of noise. Varying the sigma (standard deviation) value of this filter only reduced

the number of structures detected. After applying the Gaussian filter the potential edges are thinned by removing non maximum pixels of the computed gradient [50].

Other filters were tested giving similar results, except the *frangi* filter that gave a better one. *Frangi* is usually used to segment and detect vessels or tube-like structures [51]. Maybe in Figure 3.8 it is not very clear, but there are still some structures appearing inside the breast.

The problem with these filters is that the many structures inside the breast are detected as well, and that is not our goal.

To delineate separate contours, a simple “find contours” function can be used. When applied to the original image or a filtered image, it did not yield the desired result. This function uses the “marching squares” method to compute contours of a 2D array. It works best when the contour is between “light” and “dark” values [52].

It was necessary to apply a threshold and obtain a binary image. There are many methods to determine the threshold value. Some methods gave thresholds too low that resulted in full black images (all pixel values equal to zero), like the *minimum* and *triangle* threshold methods.

Other methods were tried and compared in terms or visual results of the new image and the subtraction of the original image with the new image.

Some thresholds were too high causing “cuts” and “holes” where there were skin folds and overlying tissues (Figure 3.9), as it was with the *isodata*, *otsu* and *mean* methods. The *Li* method presented better results, having a lower threshold then the last three mentioned, and a smaller difference between the new image and the original.

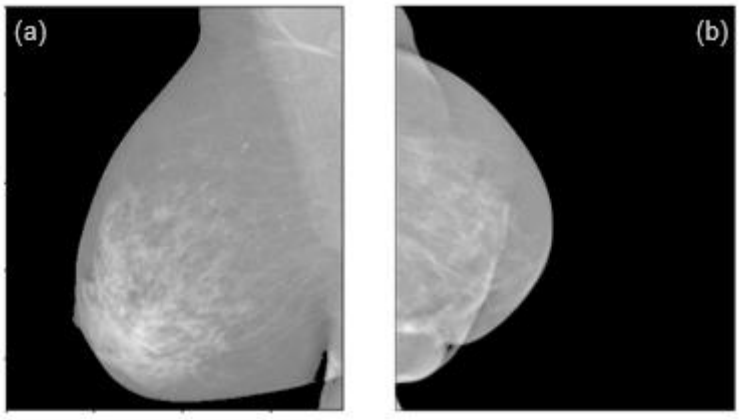


Figure 3.9 - A case of the mean threshold applied to an image and the cut it presents on a skin fold. (b) - A case, on another patient, where the isodata method shows a hole in an overlying tissues region.

The *skimage.filters.threshold_li* function returns a threshold value based on an adaptation of Li's Minimum Cross Entropy Thresholding [53] and it was applied to all the images.

The binary image has the background pixels equal to zero, and breast pixels equal to one. This way, it can serve as a mask: multiplying the binary image to the original image we get a new image with the true pixel values inside the breast and pixel values equal to zero on the background. The "find contours" function was then used to delineate the breast edge on these new images.

From the pixel values of the imaged breast, the minimum, maximum, mean pixel value and standard deviation were obtained, using Python's *numpy* library.

There's a function that gives labels to connected objects and it is possible to retrieve a lot of properties (*regionprops*) of each object including its area. Applying this to the original mammography image would present a lot of labels because the imaged breast presents a lot of structures. The binary image was used instead, and the imaged breast area was calculated.

Another way to calculate the area would simply be to count the non-zero pixels of the new image (original x binary). This would give the same result as with the *label* function, except when there are separated parts being imaged like it is shown in Figure 3.10. So, it was opted to calculate the area through the *regionprops* function.

The maximum perpendicular width of the area being imaged was considered as the arrow on the figure below. This consisted of the maximum count of consecutive non-zero pixels of all the rows of the image.

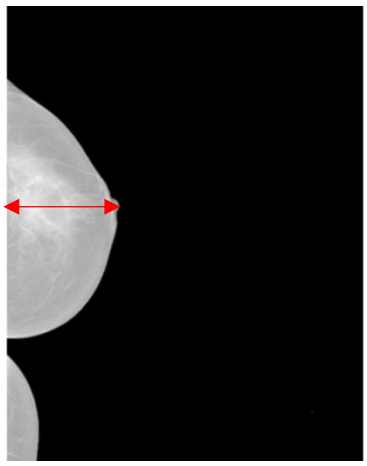


Figure 3.10 – Example of other tissue being imaged at the bottom. The arrow is representing the maximum width considered.

Both results of area and distance would be in number of pixels. To convert these parameters to the metric system one must have them multiplied by the pixel size, presented in Table 3.1 (page 20).

3.4.3.2 Surgical clips detection

The data used in this study required manually selecting mammography images from patients who had undergone treatments. This is time consuming, because it requires looking into patient's clinical history, individually.

To obtain more data for the statistical analysis (2nd sample), for this and future studies, there was a need to develop an automatic method to detect a treated side on a group of images.

One distinct aspect of most of the images of a mammography of a treated breast is the presence of surgical clips.

In a first approach, the characteristics of a surgical clip were studied by simply observing some examples. With this information and using the *skimage* library, a Python script was developed to try and detect the clips on the images of the first sample.

The clips are metallic and attenuate more radiation than the breast structures. The first attempt consisted in simply detecting connected objects in a binary image of the mammography with a higher threshold for the pixel values.

The properties of each detected object were assessed, namely: the area (number of pixels in the region), the eccentricity ("Eccentricity of the ellipse that has the same second-moments as the region." [54] Equals to 0 when it is a circle) and the solidity (ratio of pixels in the region to pixels of the smallest convex polygon that surrounds the region).

It is important to exclude other strange structures that sometimes appear on the images like electrodes, micro calcifications and others. To consider the object a surgical clip three conditions were imposed:

- Area between 50 and 800 pixels;
- Eccentricity > 0.95
- Solidity > 0.8

These values were chosen by experimenting running the script through the images of the DS 2014-15 sample.

If at least one image, of the 4 per exam, had a detected surgical clip, the patient pertaining to that exam would be classified as "treated", and the side of the treated breast assigned accordingly.

To evaluate the outcome, the result yielded by this program would be compared to the patient's category (no treatment, right breast treated, or left breast treated).
Furthermore, visual confirmation of the presence of the clips was needed. A script to save the images, grouping the CC and MLO views, in a format such as .tiff, was written. This allowed for a faster viewing of all the images and identification of the ones that had surgical clips.

3.4.4 MDG and ESAK estimation

The entrance surface air kerma and the mean glandular dose were estimated independently with the EUREF method described in subsection 2.7 (page 13), more precisely through equations 2.1 and 2.2, respectively.
The output of each mammography unit was available for the corresponding date from annual Quality Control measurements at IPO-Porto, for the different combinations of target/filter and kVp. When an output value was missing, interpolations were made from values of the previous and later years. The output values used are presented in Appendix D. For the exams taken in 2018, the output of 2017 was considered, as there was not a measurement for that year.
The tables published in the European guidelines are indicated for different thicknesses and HVLs (Appendix C). Linear interpolations were done to obtain c and g factors for the missing values of compressed thickness and HVL.
The values for the g factors are only established for two age groups, namely 40-49 and 50-64. In this case, ages below 40 were considered in the first group and ages above 64 were included with the latter one.
To evaluate the difference of the estimated values from the displayed values stored on the DICOM header, the percentage difference was determined as follows:

$$Dif\ \% = \frac{calculated_{value} - displayed_{value}}{calculated_{value}} \times 100\% \qquad (3.6)$$

3.4.5 Normal variability between exams to the same patient

For each patient, the mean deviation of each parameter was determined for different views, considering all the exams pertaining to that patient, in different years. The mean values of these deviations were also calculated for this patient group. Then overall mean values were also considered.

All these values were determined first excluding the images from the Senographe Essential unit, and later with all images considered.
The delineation of the contours of the same mammographic views through the years, to the same woman, were overlaid for visual comparison of the positioning technique.
This will serve as a reference for the comparisons between treated and untreated breasts.

3.4.6 Variability between left and right untreated breasts and between treated and untreated breasts

The cases of study were separated as right and left breasts (with no treatment), and treated and untreated breasts.
The absolute differences of each parameter studied between the two breasts of each patient were calculated, according to their treatment status. For patients with no treatment the difference “Right-Left” was considered. For treated patients, the difference “Treated-Untreated” was considered, regardless of the side.
The minimum, maximum, mean and median were computed and presented in “box plot” graphs, for the absolute values as well as for the differences of the different parameters. Having a small number of exams corresponding to patients who had undergone surgery alone, those patients are grouped together with those who also had radiotherapy.

3.4.7 Comparison of the mammography units

After all our analysis, we looked into patients who performed exams on both units at IPO-Porto in a span of a year, to understand the different results obtained for the two equipments.
The percentage differences for some of the image parameters, including the maximum and mean pixel value (PV) and the standard deviation (SD) of the imaged breast area were calculated as:

$$Dif\ \% = \frac{Essential - DS}{Mean(both)} \times 100\% \quad (3.8)$$

4 Results and Discussion

4.1 Preliminary characterization: Study of phantom images

4.1.1 Response Function and Noise Evaluation

The mean pixel values (MPV) of the raw images have a linear relationship with the tube loading (mAs), and subsequently, the dose. Figure 4.1 shows those plots and the corresponding linear trend lines with correlation coefficients $R^2 > 0.99$, as required by EUREF guidelines [1].

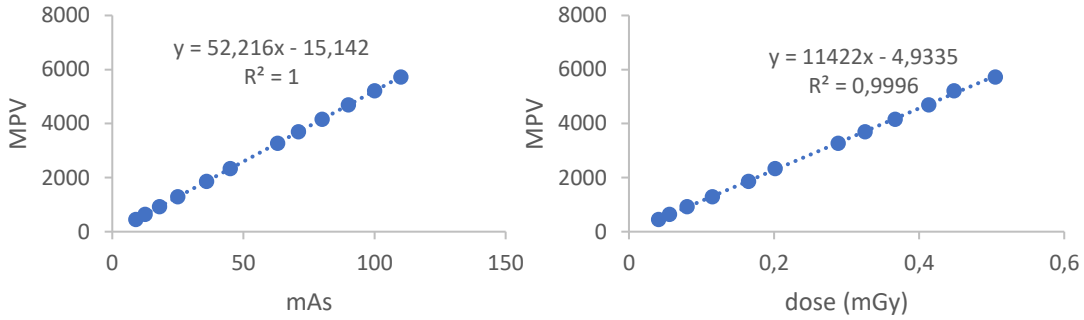


Figure 4.1 – MPV as a function of mAs and dose, of the response function images.

To evaluate the noise present in images produced in the digital mammography unit, the standard deviations squared and the mean pixel values of the reference ROI of each image are plotted in Figure 4.2.

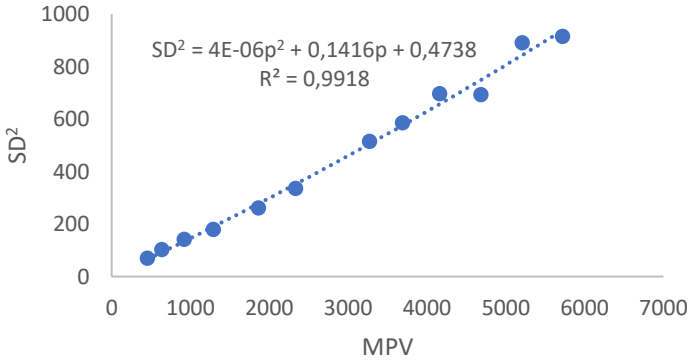


Figure 4.2 – Standard deviation (SD) squared as a function of mean pixel value (MPV), of the response function images.

Retrieving the coefficients and using equation 3.2 (page 21) we get the values presented in Figure 4.3.

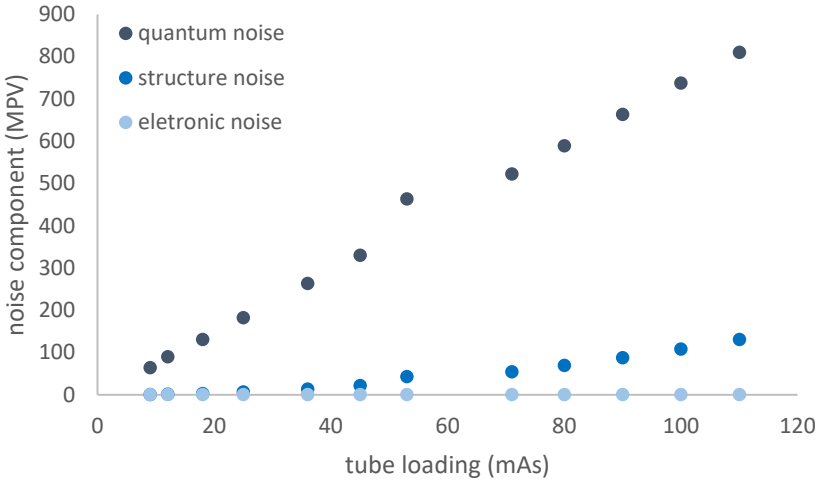


Figure 4.3 – Noise components corresponding to each tube loading value, of the response function images.

We can see that the dominant component is the quantum noise, as expected [1].

4.1.2 Detector Dose

The values of each step of the dose calculation method described in section 3.3.2, are presented in the next table. The final calculated values are plotted against the mAs values of each response function image, in Figure 4.4.

Table 4.1 - Calculations of detector dose and comparison to the DICOM indicated dose parameter.

DICOM			**Calculations**				**Relation with DICOM value**
Dose (nGy)	(mGy)	mAs	Dose at 1m (mGy)	Dose at 66cm (mGy)	+Al	+Grid	
40626	0,041	9	0,279	0,64	0,08	0,06	1,46
55948	0,056	12,5	0,388	0,89	0,12	0,08	1,47
80192	0,080	18	0,558	1,28	0,17	0,12	1,48
114941	0,115	25	0,775	1,78	0,23	0,16	1,43
164749	0,165	36	1,116	2,56	0,33	0,24	1,44
201224	0,201	45	1,395	3,20	0,42	0,30	1,47
288421	0,288	63	1,953	4,48	0,58	0,42	1,44
325193	0,325	71	2,201	5,05	0,65	0,47	1,44
366655	0,367	80	2,480	5,69	0,74	0,53	1,44
413402	0,413	90	2,790	6,40	0,83	0,59	1,43
447833	0,448	100	3,100	7,12	0,92	0,66	1,47
504930	0,505	110	3,410	7,83	1,01	0,72	1,44
							1,45

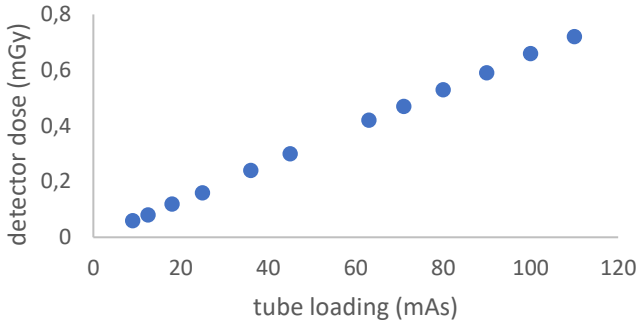


Figure 4.4 – Detector dose calculated as a function of tube loading (mAs)

It was found that the calculated value had a factor of around 1.45 in relation to the exposure dose indicated on the DICOM header (included in the proprietary GE DICOM tag of Exposure Control Mode Description (0018,7062)) [55], [56].

4.1.3 Comparison of “for processing” and “for presentation” images

Plotting the MPV of the raw images against the MPV of its corresponding processed images, suggests that the processing includes a logarithmic operation.

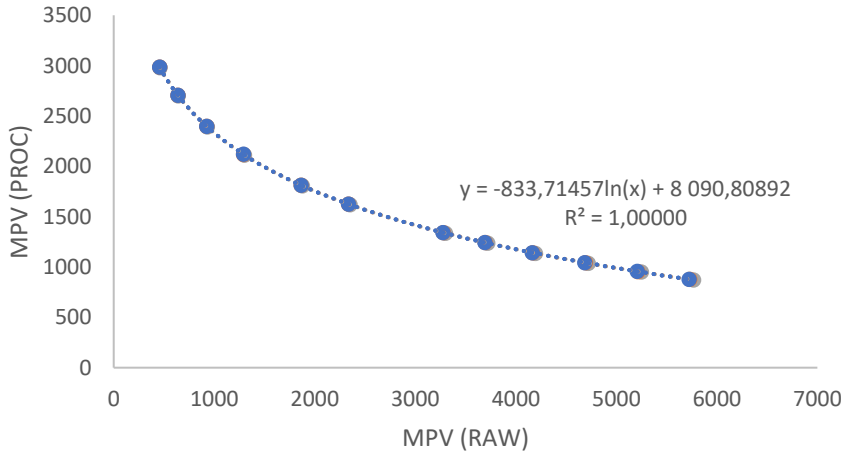


Figure 4.5 – Mean pixel values (MPV) of the “for presentation” response function images (PROC) as a function of the MPV of the “for processing” images (RAW).

A trend line was applied to all the 24 sets of values (of the ROIs referred Figure 3.4 in page 23), separately, and the average of the terms of the equations were taken.
The resultant equation was:

$$MPV_{PROC} = -833{,}67 \ln MPV_{RAW} + 8090{,}43 \quad (4.1)$$

With these terms applied on the raw images, the difference and division of that and the actual processed images was assessed, as described in equations 3.4 and 3.5 (page

23). Apart from an apparent *bad pixel* present in the lower mAs images, the results were practically as expected, where fluctuations attributed to the noise of the images are always present , and all the results are presented in the next two tables.

Table 4.2 – parameters referring to the subtraction images.

mAs	MPV	SD	Min	Max	Median
9	0,010	0,296	-0,494	125,862	0,011
12	-0,017	0,298	-0,498	117,438	0,001
18	-0,004	0,288	-0,490	8,941	0,001
25	0,006	0,288	-0,500	0,494	0,001
36	0,015	0,289	-0,496	0,495	0,035
45	-0,001	0,288	-0,492	0,498	-0,013
53	-0,049	0,286	-0,497	0,497	0,007
71	-0,016	0,288	-0,499	0,498	-0,018
80	0,028	0,287	-0,500	0,496	0,039
90	0,002	0,289	-0,501	0,499	0,006
100	0,009	0,288	-0,497	0,498	0,007
110	0,007	0,289	-0,496	0,499	0,016

Table 4.3 - parameters referring to the division images.

mAs	MPV	SD	Min	Max	Median
9	1	0,0001	0,9998	1,0346	1
12	1	0,0001	0,9998	1,0324	1
18	1	0,0001	0,9998	1,0026	1
25	1	0,0001	0,9998	1,0002	1
36	1	0,0002	0,9997	1,0003	1
45	1	0,0002	0,9997	1,0003	1
53	1	0,0002	0,9996	1,0004	1
71	1	0,0002	0,9996	1,0004	1
80	1	0,0003	0,9995	1,0004	1
90	1	0,0003	0,9995	1,0005	1
100	1	0,0003	0,9995	1,0005	1
110	1	0,0003	0,9994	1,0006	1

The division images are more precise, with pixel values closer to the unity, whereas in the subtraction images the standard deviation is higher.

It was confirmed that the system applies a log-transformation to the raw data.

4.1.4 Image processing effects

4.1.4.1 Local dense area (LDA)

Both raw and processed images without PV, had SNR values of the local dense area within the 20% of the average. The average SNR on the raw images was significantly lower than the SNR of the processed images. Considering the images with PV, the SNR values are lower in comparison to the images without PV, and the image with the maximum number of smaller PMMA plates (20mm) had a SNR value of the local dense area outside the 20% of the average. This correlates with a higher standard deviation found on the LDA of those images. All the average SNR values are presented in Table 4.4

Table 4.4 – Average signal to noise ratio (SNR) of the different modes of local dense area images.

	raw		**processed**	
	Without PV	With PV	Without PV	With PV
Average SNR	77	58	235	79

4.1.4.2 Premium View

The MPV of the background and of the local dense area of the images with and without PV are plotted in Figure 4.6.

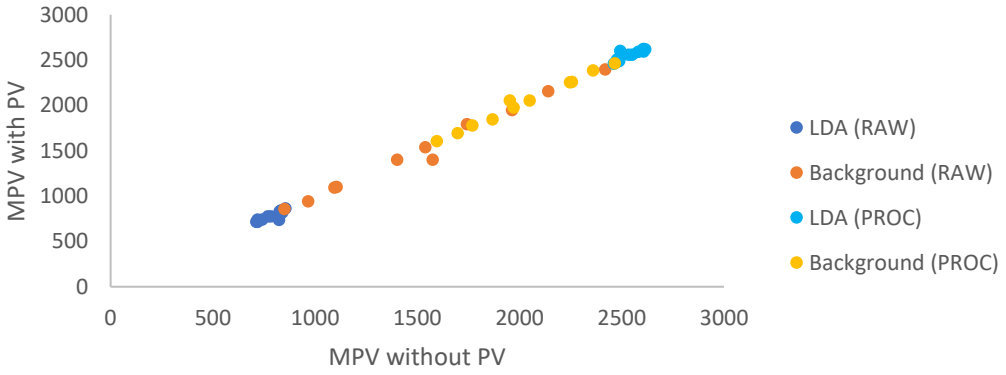


Figure 4.6 – Mean pixel values of the images with Premium View as a function of the MPV of the images without Premium View, for the background and the local dense area regions.

It can be seen that there is almost a linear correlation between the pixel values of the background of the images with and without PV with different local dense area thicknesses. A bigger difference is seen in the case of the images of LDA of 10mm thickness (5 smaller PMMA plates). The pixel values on the local dense area seem to be practically the same, regardless of the added PMMA plates.

The pixel values of the plot profiles passing through the LDA are presented in the next set of graphs (Figure 4.7).

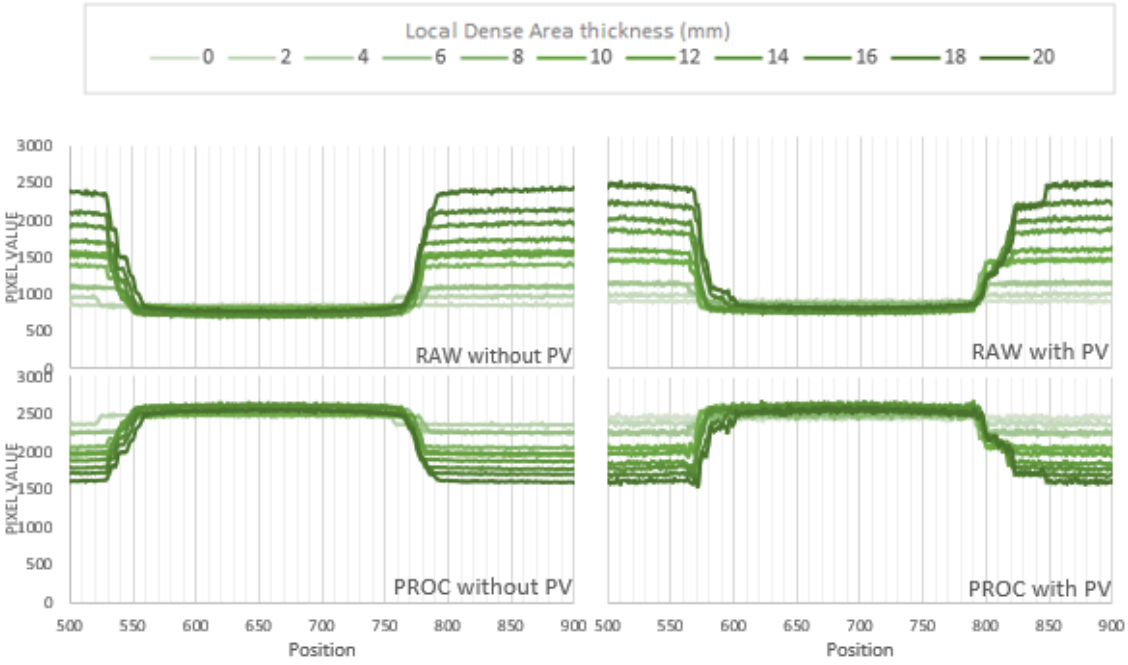


Figure 4.7 - Plot profiles (pixel values per position) of the different modes of local dense area images, around the local dense area.

Note that the acquisitions with and without PV are not the same, and the smaller PMMA plates, even as they were being stacked up, were not always precisely positioned on the same place, hence the slight displacement seen in the graphs.

It is easy to visualize the position of the smaller PMMA plates, with abrupter pixel value differences for images of higher thicknesses of LDA. In the raw images we have lower pixel values in the background and, with bigger LDA thicknesses there is a compensation that increases the pixel values in the background to maintain the pixels in the LDA at the “minimum”. In the processed images the opposite occurs, as expected from the comparison of the raw and processed images of the response function images. This correlates to what we see in the previous graph (Figure 4.6).

Comparing now the images with and without PV, there is a higher local variation of pixel values in the images with PV, i.e. these have more noise (higher SD). This is very clear in the processed images with PV. The noise also seems to be higher around the local dense area, i.e around denser tissues, that at the background (Figure 4.8).

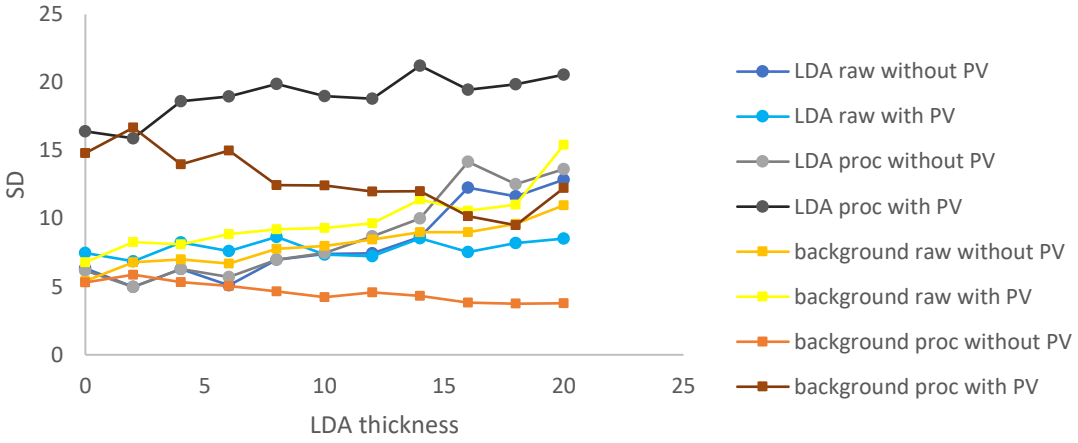


Figure 4.8 - Standard Deviation (SD) of the local dense area and background of the raw and processed images, with and without Premium View.

Observing the images, we can clearly see that around the edges of a denser object, the LDA in this case, are displayed very differently (Figure 4.9). This is to be expected, as noted in the operator manual of the equipment: “Premium View is an image processing algorithm optimized for the structure of breast images. When imaging an object with thick or sharp borders (for example, a phantom), one might see an enhanced brightness at the border of the object. This enhancement is normal and is not expected to affect correct phantom scoring.” [39]

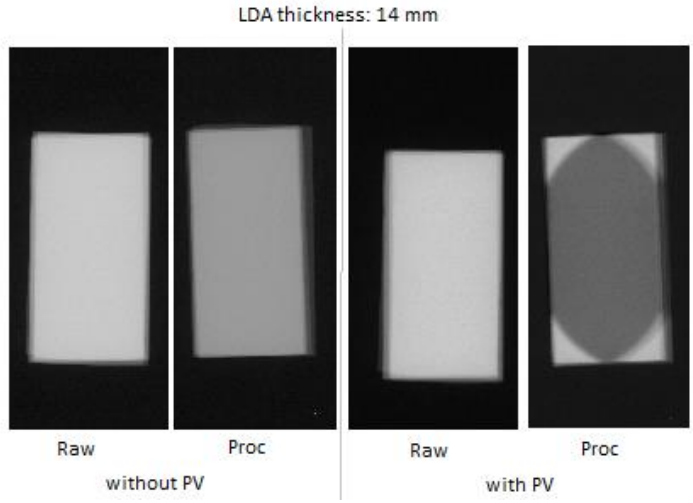


Figure 4.9 - Raw and processed (PROC) mages of a local dense area of 14mm of thickness, with and without Premium View.

4.1.4.2.1 *Pixel to Pixel comparison of raw and processed images with and without PV*

Once again, the raw images with the logarithmic adjustment found in section 4.1.3 (equation 4.1) were compared to the processed images in both sets of images with and without PV.

The MPV and SD of the whole images of subtraction and division were then measured and plotted against the corresponding LDA thickness being imaged (Figure 4.10).

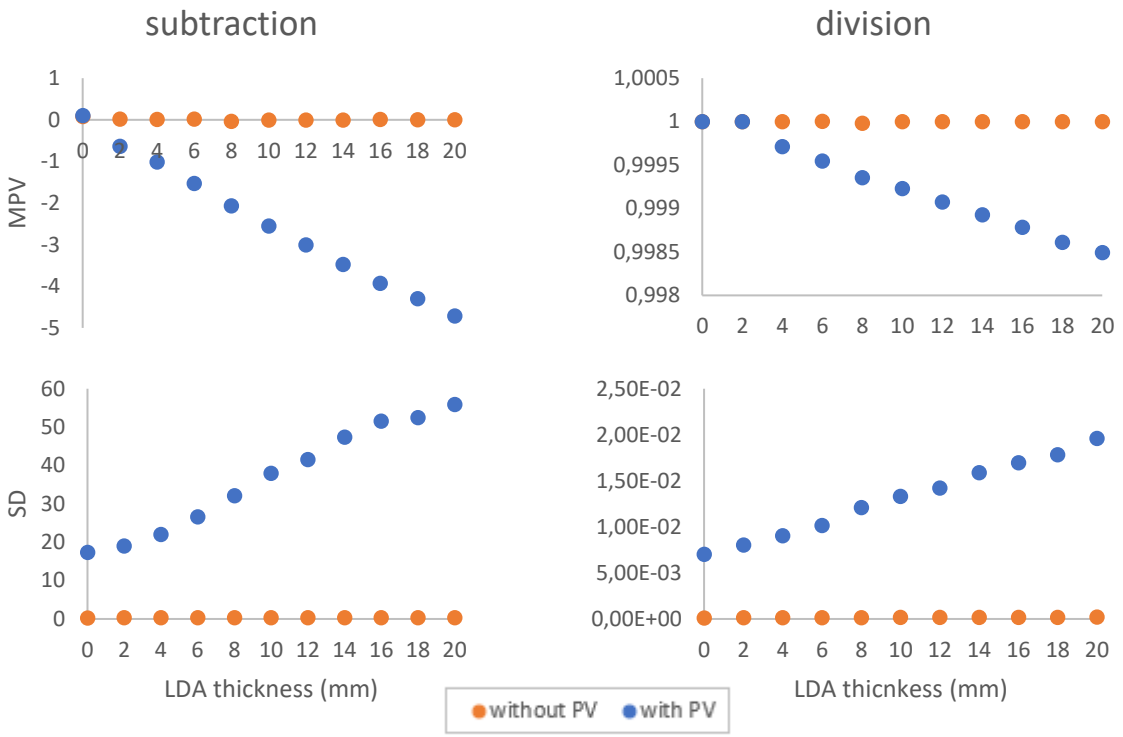


Figure 4.10 – Mean pixel value and standard deviation of the subtraction and division of the raw and processed images, with and without Premium View.

We can see that the difference in pixel values is null for the images without PV, as expected from the previous results. It is in the images with PV that we see a slight increase in the differences between processed and the adjusted raw images, when the density of the imaged object intensifies.

It is curious to measure a plot profile in these images that represent the differences. They are presented separately for the division (DIV) and subtraction (SUB) of the adjusted raw images and the processed images, with and without PV, for some LDA thicknesses in Figure 4.11.

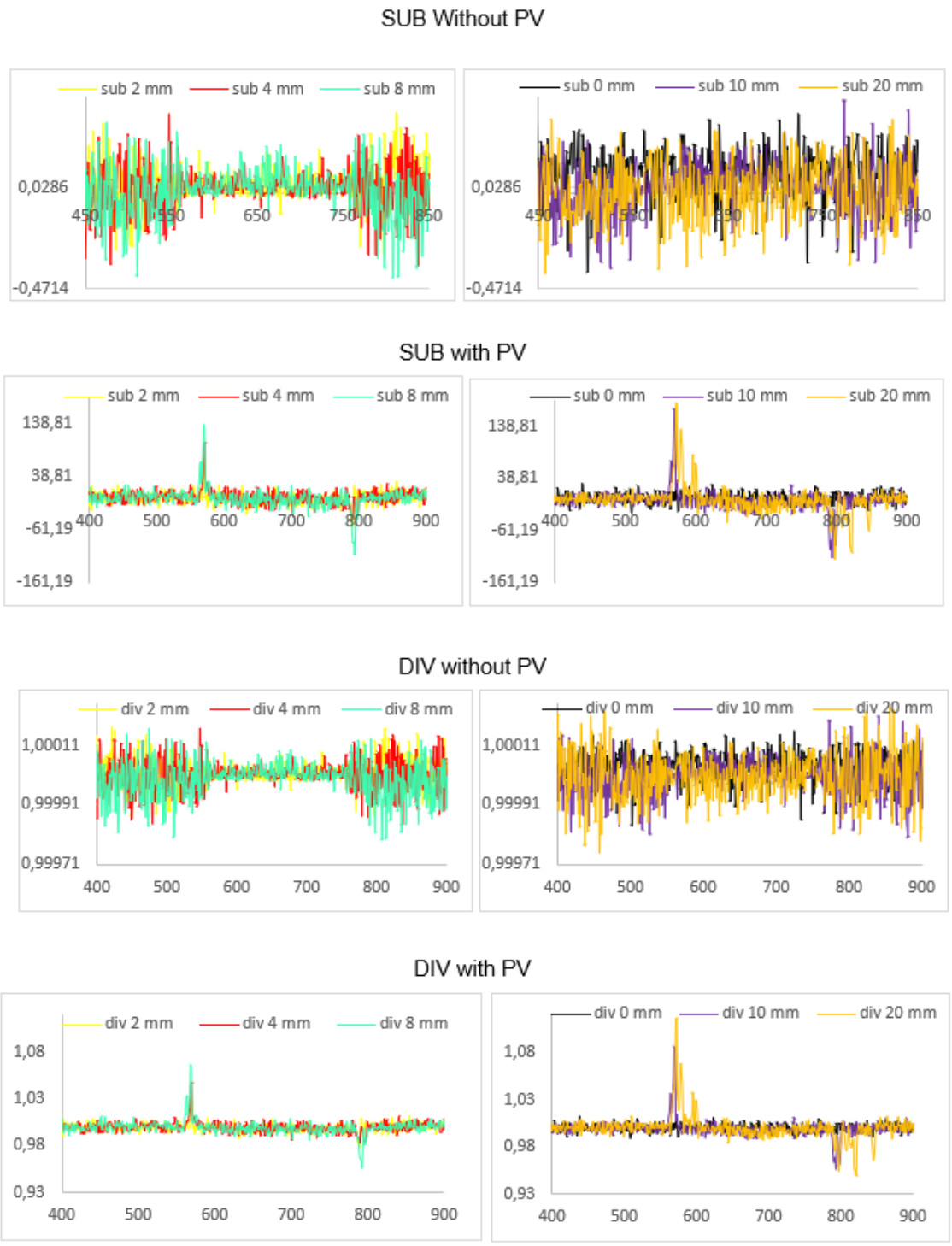


Figure 4.11 - Plot profiles (MPV vs. Position) of the subtraction (SUB) and division (DIV) images with and without PV.

These are not on the same scale for viewing purposes, but if they were we would see more noise in the differences of the images with PV. And as expected, the noise is more visible in the subtraction images then the division ones.

The results seem to agree with what GE described. The Premium View applies edge enhancement, since it is on the edges of the LDA that the differences are higher, and bigger contrasts are found.

According to Goldstraw et al. [46], when large variations of density are found, i.e. variations between fatty and glandular tissues, these are isolated with a low spatial frequency filter and used to form a mask. When little contrast variations are detected, within a tissue, these are isolated by subtracting the mask from the original image. The resultant frequency-enhanced image is further enhanced, and the final image should present reduced contrast between different tissues but enhanced contrast of smaller structures.

In conclusion the Premium View image-processing affects the contrast improving the visibility of breast structures.

4.1.5 Preliminary characterization conclusions

With this study of phantom images we could find that MPV and dose are proportional, on the raw images, and the processed images have a logarithmic relationship, instead of linear.

All the image processing that is applied to the raw images is to enhance the visualization of structures. One can see that the Premium View image processing practically does not affect the Mean Pixel Values, but only the noise. Based on these results, it was decided to use only processed clinical images in this study. The use of processed images instead of raw images considerably simplifies the data collection process, since processed images are archived in PACS and can be obtained retrospectively. The image processing effects studied in phantoms are sufficiently small that further computational processing of processed images, such as a contour delineation or detection of structures, should not be affected.

4.2 Study of Clinical Images

4.2.1 Computational processing

4.2.1.1 Contour finding and calculation of imaged breast area

A Li threshold was applied to all the mammography images and binary images were obtained. This served as a mask to apply to the original image and to delineate the contour of the breast edges. Figure 4.12 shows the steps in finding the breast contour on a random image.

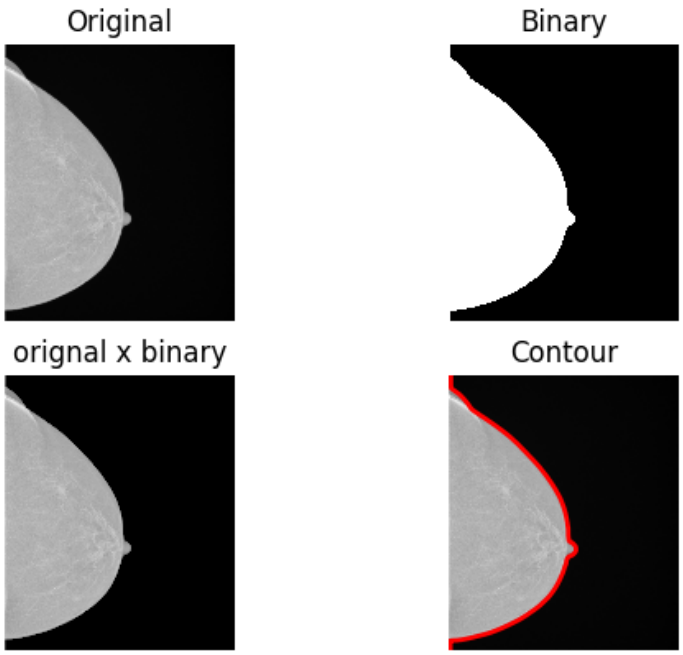


Figure 4.12 - Contour delineation method example.

The delineation of contours was successfully done and later used for comparison of contours of mammographic exams of the same person through the years and even to visually compare the left and right breasts.

As another example, Figure 4.13 shows the different contours detected for different connected regions that may appear in an image, including other tissue and some strange small structure outside the breast.

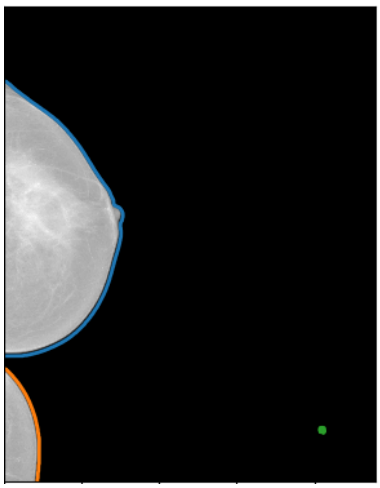


Figure 4.13 – Contours of the CC view of a left breast with separate detected regions. Lines of different colours delineate contours of different connected regions. Only the area within the blue contour was considered for analysis.

On another hand, structures like surgical clips and calcifications are not detected through this method, as seen in Figure 4.14. This is because the contours are found through a binary image, and all structures inside the breast have the same pixel value on that image.

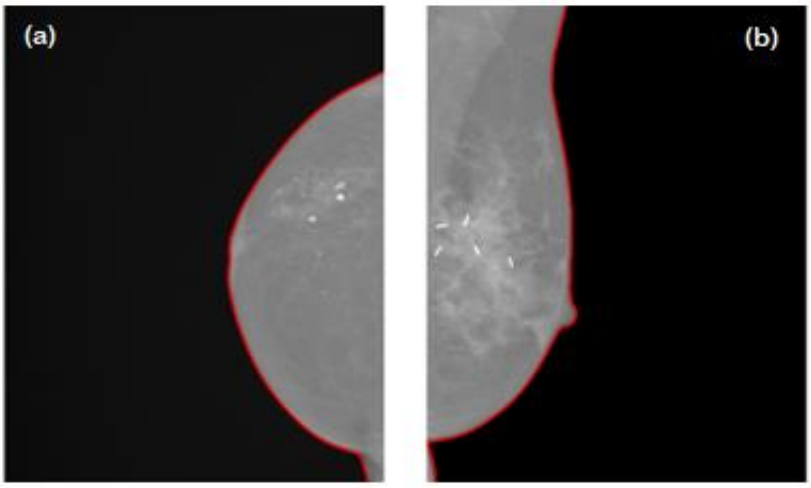


Figure 4.14 (a) - A CC view of a right breast with calcifications; (b) – a MLO view of a left breast with surgical clips.

The use of the mask obtained through the binary image allowed to calculate the area and the maximum width of the imaged breasts. Note that the area calculation would not consider separate structures like the ones in image Figure 4.13.

4.2.1.2 Surgical clips detection

Looking through the images of the first sample of the treated breast study, about 25% of patients categorized as treated did not present surgical clips in any view of the treated breast. Observing the clips present in some of the images they are always associated with the highest pixel values of the image, since they are made of metal.

Figure 4.15 shows an example of a detected surgical clip, and also a strange structure that we don't want to be considered, and the properties considered to classify the object.

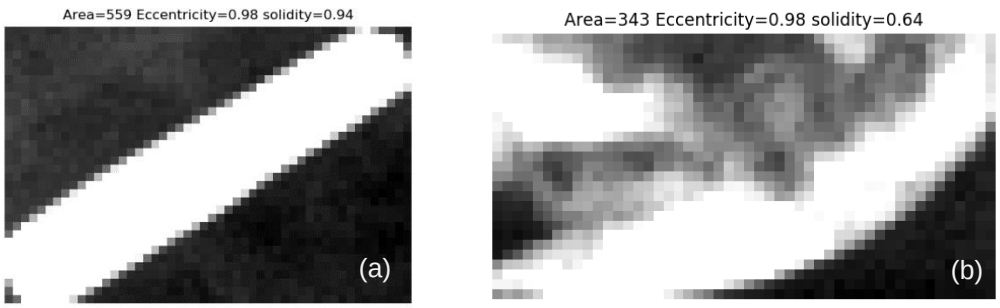


Figure 4.15 (a) Exemple of a detected surgical clip; (b) – Example of a non surgical clip.

After observing the images of the first sample the limiting values of area, eccentricity and solidity were adjusted. Then the script was run through all the images of the first sample. Correlating the visual observation of the clips and their automatic detection, 91% of the cases yield a positive result, i.e. at least a treated side was detected on treated patients with clips, and no treated side detected on patients with no clips.

Running the same script through images of the second sample presented a worse result. The images produced by the Senographe Essential are intrinsically different, starting with the fact that the pixel values of surgical clips are not as high as the ones on Senographe DS images. So, the first approach failed to detect any treated side in images of Essential, because a too high threshold was set. In the images of DS there was almost 100% of classifying correctly the treated side.

The detailed results are shown in Table 4.5. The samples from 2017-18 contain more patients than the numbers indicated in section 3.4.1.2.2, because this method was conducted before the final data selection for the study.

Table 4.5 – Results of the surgical clips detection method.

Sample	DS 2014-15	DS 2017-18	Essential
Number of patients	155	112	228
Patients with treatment	93	82	162
Patients with surgical clips	69	65	130
Treated patients identified (clips detected)	55	64	0
% related to number of patients with clips	80%	98%	-
Correct classification of the patient	141	111	-
% related to the total number of patients	91%	99%	-

The few times the results were not correct, on images from the DS mammography unit, it was because the clips sometimes appear open and the values of eccentricity and (mostly) solidity fall (see Figure 4.16). Other reason for the method to fail is if the clips are not fully imaged, i.e. when they are near the edge of the image.

This method needs further optimization, especially if we want to apply it to other mammography units. Still, we found good results in detecting a treated side for images in the GE Senographe DS unit.

Automatic surgical clips detection allowed identification of six misclassified patients: images of one treated patient had mistakenly been saved in the “no treatment” folder,

two treated patients on the left/right side that had been incorrectly classified as right/left due to an error and three cases where patients had treatment on both breasts.

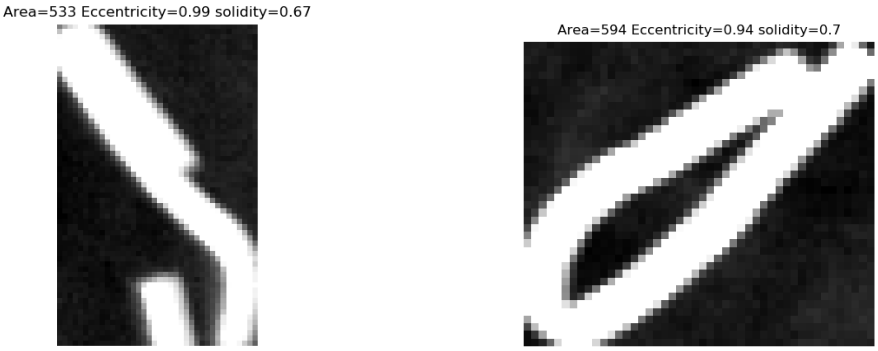


Figure 4.16 – Other forms the surgical clips can take.

4.2.2 MDG and ESAK estimation

The ESAK values extracted from DICOM agree with the calculated ones, within ±10%, as shows the box plot in Figure 4.17. The mean difference is less than 5% and this should be expected as a systematic error related to the different output values that are calculated for each unit each year and the output considered in the system's estimation of ESAK for each acquisition. The method the GE mammography units have to compute the ESAK uses "a calibrated model of the free-in-air-kerma, in the plane of the compression paddle in contact with the breast, with no backscatter contribution from the breast. It takes into account the attenuation of the X-ray beam by the compression paddle." [39]

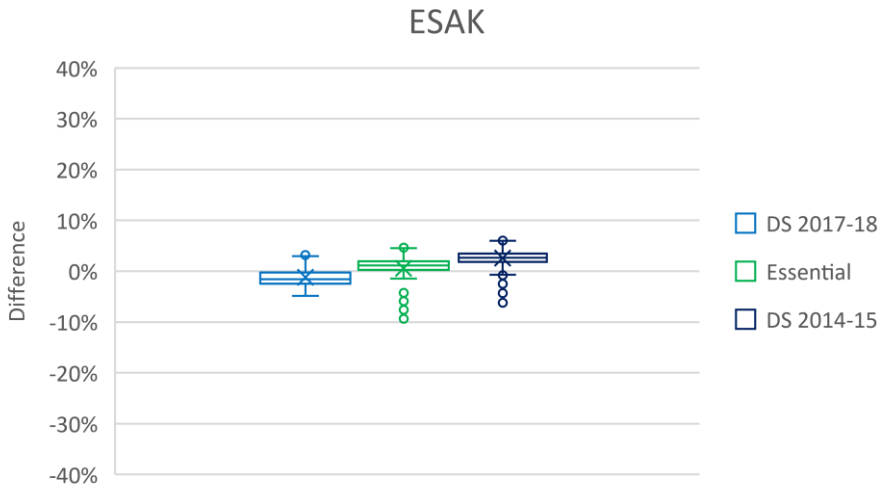


Figure 4.17 - Percentage differences between the estimated and the displayed ESAK values.

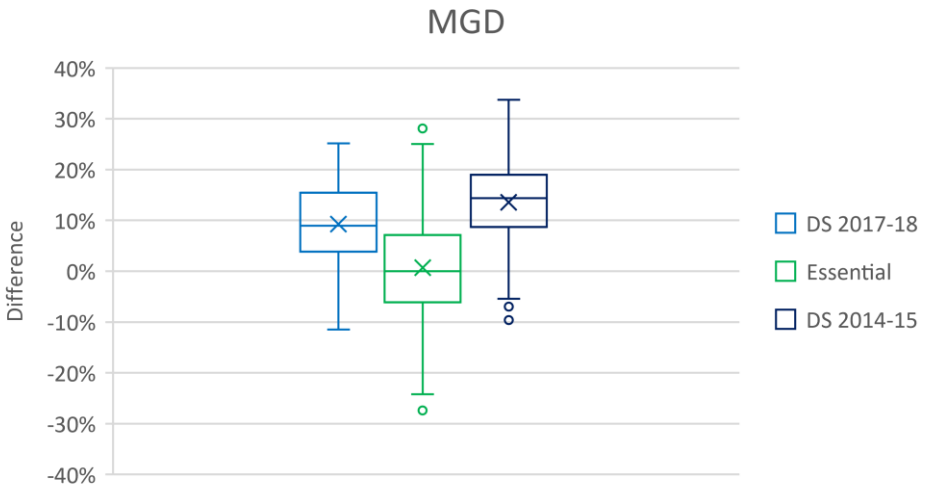


Figure 4.18 - Percentage differences between the estimated and the displayed MGD values.

Larger differences were found between the calculated MGD values and those indicated by the equipment, depending on breast compositions. These differences are reduced when calculating variations. As Figure 4.18 shows, the MGD calculated values are usually higher than the displayed ones. The differences go up until around 30%.

The difference from the values of Essential has a mean of 1% while the mean differences in the DS values are of 9% and 14%, for the 2017-18 and 2014-15 data, respectively. The trend of the mean values for the data only from Senogrpahe DS seems to correspond to the same trend in the mean difference of ESAK but amplified.

Supposedly, both units use the same method to estimate the MGD. It is computed considering the compressed breast thickness and a breast composition. The fact that the Essential unit has a larger detector and possibly images larger breasts could be the reason for this distinction.

These results agree with another study, by Tsalafoutas and Kanellopoulou [57], that estimated ESAK and MGD values for mammography exams performed in three different Senographe Essential units. They found lower differences on the MGD values when considering the glandularity values indicated in the DICOM headers instead of the Dance's method typical glandularity values.

The plot of the MGD difference vs. the glandularity values indicated in the DICOM headers suggests this difference increases with the percentage of glandularity, indicated by the equipment as seen in the next graphs for each data sample (Figure 4.19), except for glandularities of 30-40% and below where the calculated values are lower than the displayed ones.

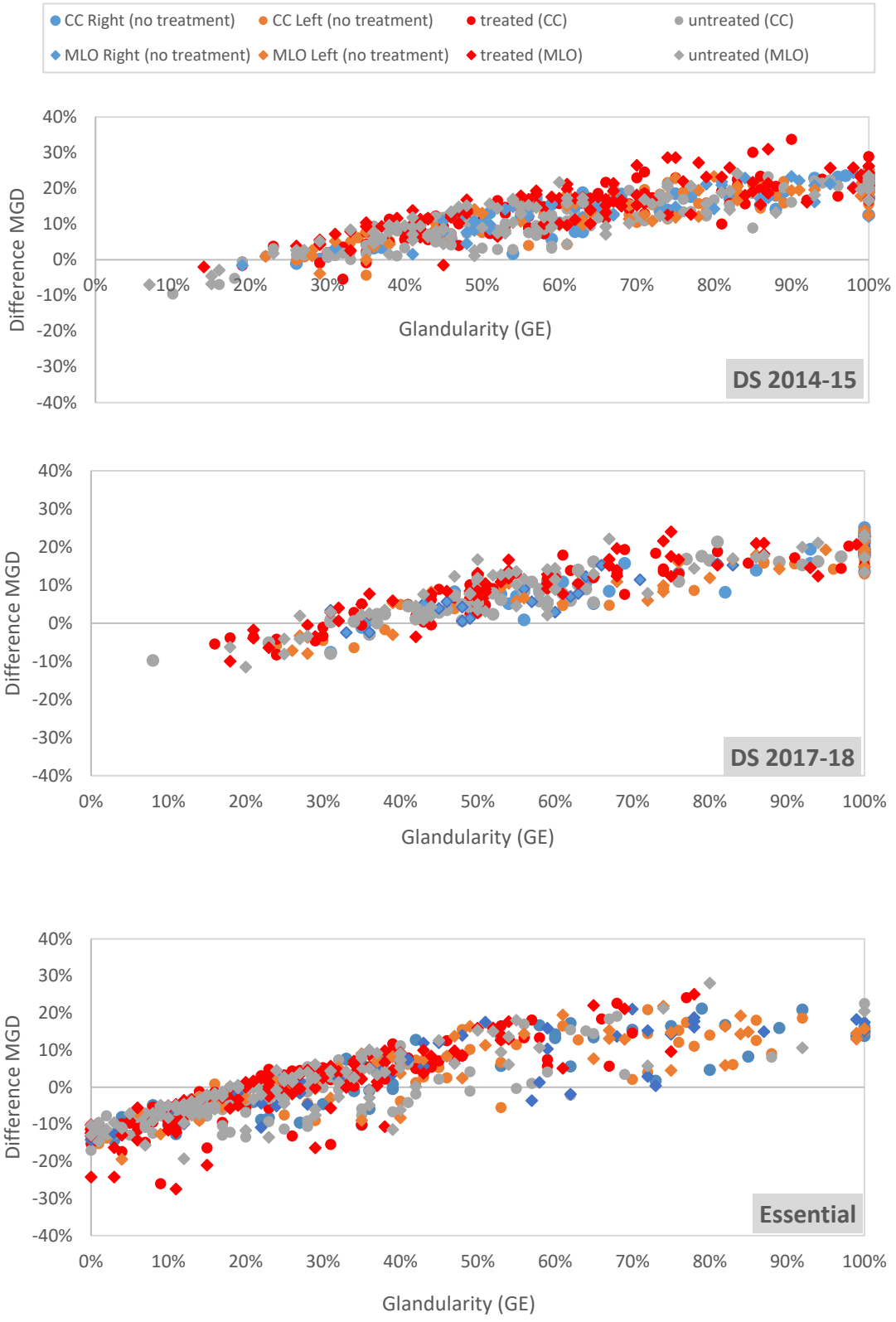


Figure 4.19 – Difference of the MGD values as a function of the glandularity values indicated in the DICOM header for the different clinical images samples.

The data from the Essential mammography unit indicates lower values of the glandularity. Curiously the data from treated breasts imaged on this unit does not seem to have any case of glandularities higher than 80%. The “threshold” between positive and negative differences appears to be slightly before, around 20-30%, when comparing with the data from the DS unit. The fact that the Essential equipment images higher compressed breast thicknesses might be contributing to the lower glandularities estimated. Still, this result suggests that, even though both units belong to the same manufacturer, their method to estimate glandularity might be different. This will be addressed again later on this study.

4.2.3 Normal variability between exams to the same patient

The mean values of the extracted parameters are presented in Table 4.6 for each view (MLO and CC), for all the exams of the 10 patients considered in this part of the study.

Table 4.6 - Mean values of the different parameters, obtained from the DICOM images. The non highlighted values don’t consider the exams taken on the Senographe Essential. (*****) MGD values in parenthesis were calculated by the EUREF method [1].

VIEW	Thickness (mm)	Compression Force (N)	Area (mm)	Maximum Width (mm)	MGD (mGy)*****	ESAK (mGy)	Glandularity
CC LEFT	52.0	92.8	15484.5	105.7	1.2 (1.4)	5.6	66%
	52.5	93.2	15690.8	106.3	1.2(1.4)	5.7	62%
CC RIGHT	52.1	97.9	14946.4	103.8	1.2 (1.4)	5.6	66%
	52.6	99.9	15068.8	104.1	1.2(1.4)	5.7	61%
MLO LEFT	54.0	104.3	19257.0	113.0	1.2 (1.4)	5.9	68%
	54.7	104.9	19772.8	113.1	1.3(1.4)	6.1	64%
MLO RIGHT	54.0	105.6	18967.8	109.2	1.2 (1.4)	6.0	66%
	54.7	109.4	19402.5	109.1	1.3(1.5)	6.1	63%
Mean	**53**	**100**	**17164**	**108**	**1.2 (1.4)**	**5.8**	**67%**
	54	**102**	**17484**	**108**	**1.3 (1.4)**	**5.9**	**63%**

As expected, all the parameters tend to have higher values for the MLO view.

For each patient, the mean deviation of each parameter was determined for different views, considering all the exams pertaining to that patient, in different years.

The mean values of these deviations were calculated for this patient group and are presented in Table 4.7. The most variable parameter proved to be the compression force, which for the same patient (same view, same side) had typical variability between exams around 18-23% of the mean value, reaching more than 30% in some patients. When including the data from Senographe Essential, the indicated glandularity shows an

average high variability as well. In some patients it even reached up to around 50%, while if we disregard this equipment the maximum variability reached was around 30%. This also suggests that both mammography units are computing the glandularity values differently. Graphs of the glandularity values through the years for each patient are presented in Appendix B.

Table 4.7 - Mean values of the mean deviations of each parameter for each view, considering all the exams pertaining to each patient, in different years. The first values don't consider the exams taken on the Senographe Essential. (*) MGD values in parenthesis were calculated by the EUREF method [1].

VIEW	Thickness	Compression Force	Area	Maximum Width	MGD*	ESAK	Glandularity
CC LEFT	6%	23%	4%	3%	7% (8%)	10%	14%
	7%	22%	5%	4%	8% (8%)	10%	22%
CC RIGHT	6%	22%	4%	3%	7% (7%)	9%	11%
	7%	22%	5%	4%	9% (8%)	11%	19%
MLO LEFT	6%	19%	4%	2%	7% (7%)	9%	14%
	7%	18%	7%	3%	9% (9%)	12%	21%
MLO RIGHT	6%	20%	4%	3%	6% (7%)	10%	11%
	6%	21%	7%	3%	7% (8%)	11%	17%
Mean	**6%**	**21%**	**4%**	**3%**	**7% (7%)**	**10%**	**13%**
	7%	**21%**	**6%**	**3%**	**8% (8%)**	**11%**	**20%**

The maximum width and the area of the imaged region had the least variability, when the data from the Senographe Essential is excluded. This difference might be due to the different detector sizes of the two imaging systems. Variability of MGD and ESAK was 7-8% and 9-10% respectively, for repeated views of the same untreated breast (without the data from Senographe Essential).

The mean values of the variations of the different parameters for all the patients (all views) are shown in Figure 4.20.

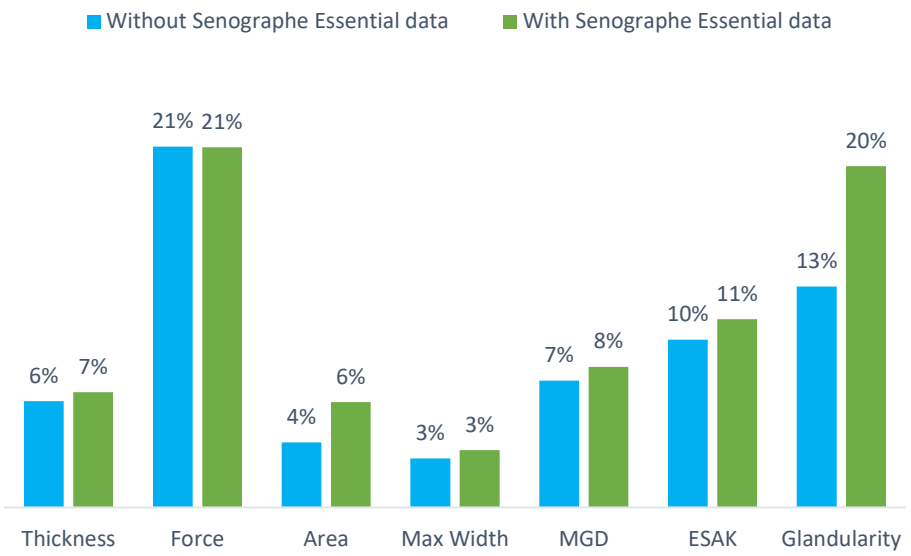


Figure 4.20 - Mean values of the variations for the different parameters considering all the exams (all views), with and without the Senographe Essential data

Visual comparison of the contours for different years confirms the good reproducibility of positioning, despite small translational variations. As examples, Figure 4.21 and Figure 4.23 show the contour of different exams for the patient E and G, respectively. Patient G is one who had performed two mammographic exams on GE Senographe Essential, in 2016 and 2017. The corresponding mean deviation of the parameters obtained for these patients are presented in Figure 4.22 and Figure 4.24.

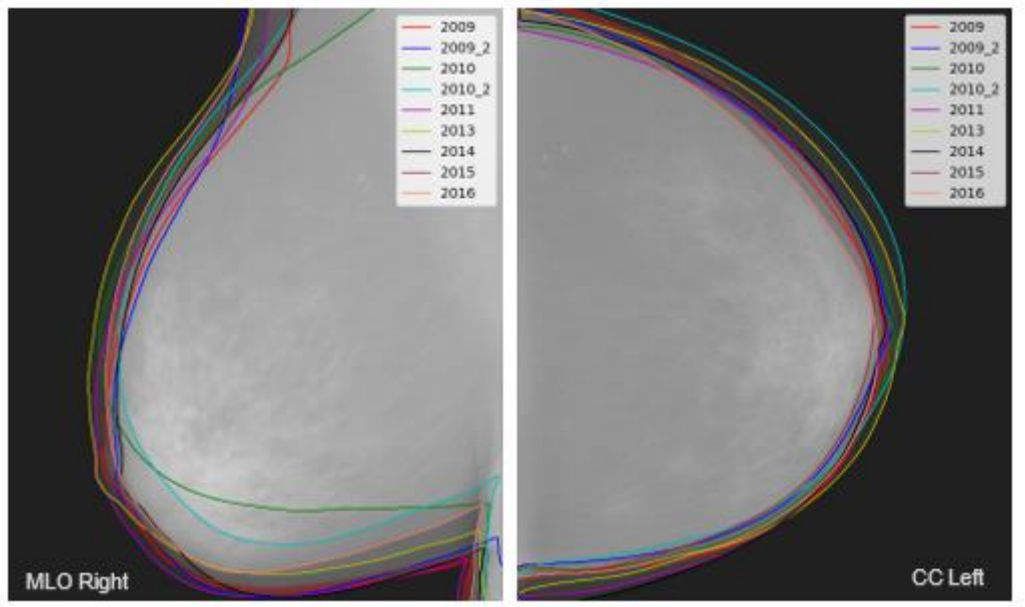


Figure 4.21 - Contours of exams for the patient E throughout the years: medio-lateral oblique (MLO) view of the right breast and cranio caudal (CC) view of the left breast.

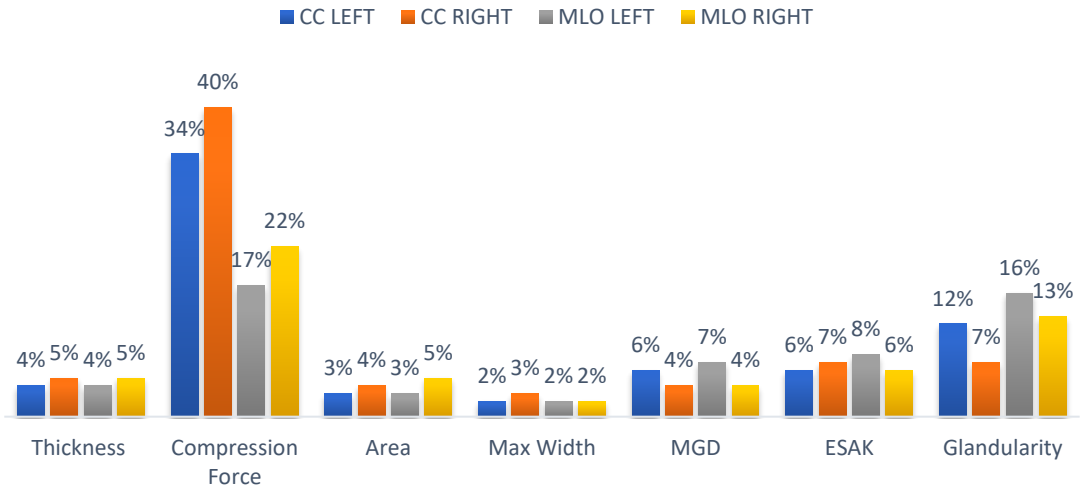


Figure 4.22 – Variations of the different parameters, corresponding to patient E.

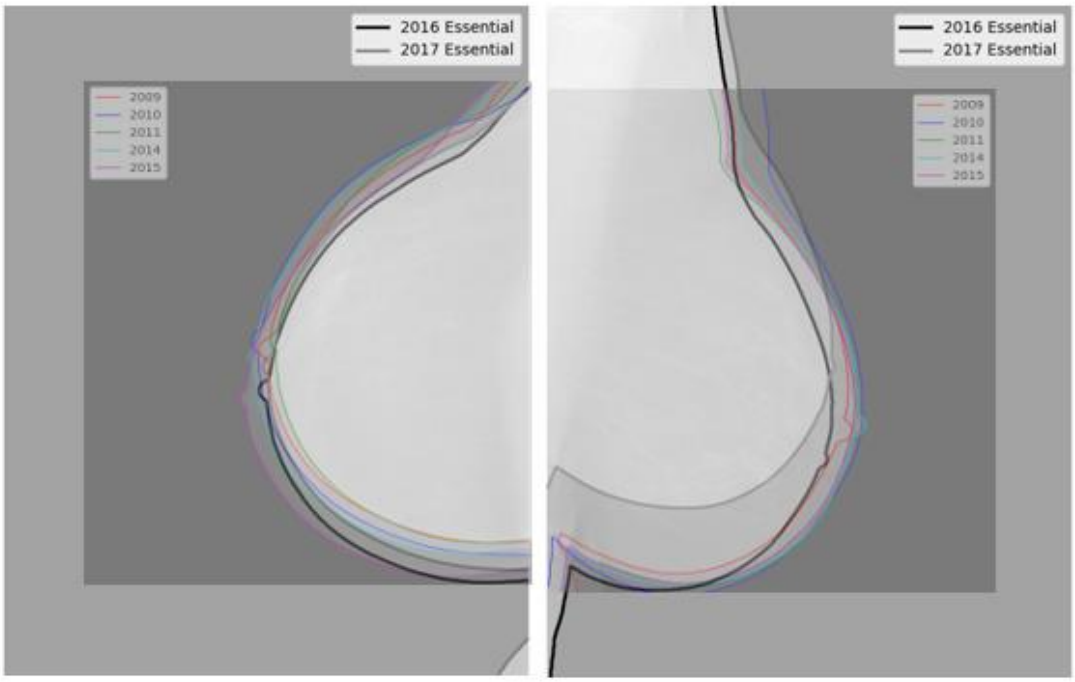


Figure 4.23 - Contours of exams for the patient G throughout the years: craniocaudal (CC) view of the right breast and medio-lateral oblique (MLO) view of the left breast.

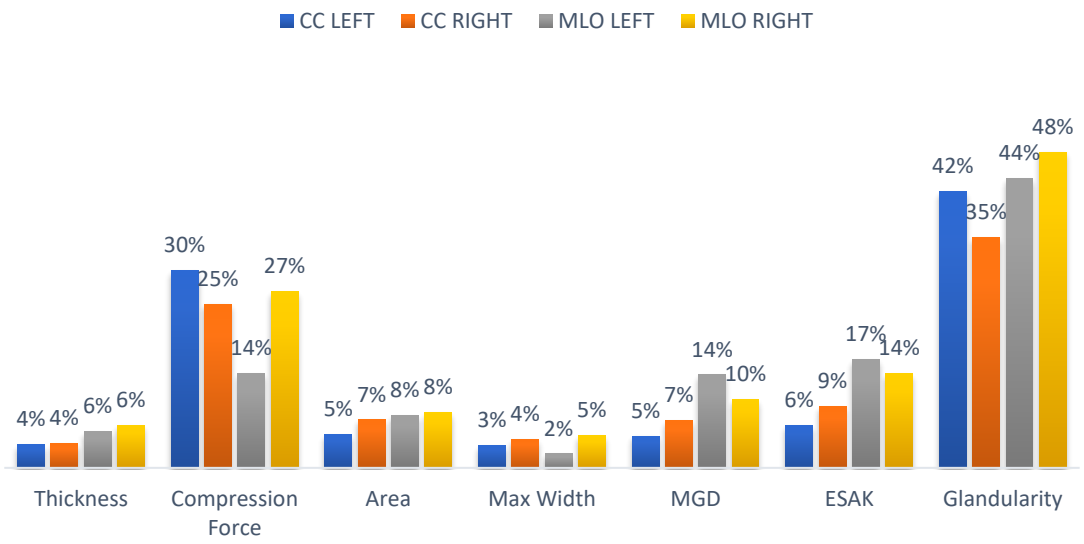


Figure 4.24 - Variations of the different parameters, corresponding to patient G.

We can conclude that there is good reproducibility in the positioning technique, and when using the same mammography unit, the compression force is the most variable parameter. This is consistent with studies that reported high variability in compression force between radiographers [14], [15] and probably the mammography unit operators were different through the years of the exams considered here.

4.2.4 Analysis of the samples DS 2014-15, DS 2017-18 and Essential

In this chapter the results from analysis of the first and second sample for the study of variability between left and right breasts of patients with no treatment and between treated and untreated breasts of treated patients are presented separately for each parameter.

In total 1596 images, of 399 different patients were considered, of which, 258 exams were of patients who mostly had undergone surgery and radiotherapy (123 on the right breast, 135 on the left breast).

The distribution of the ages of the all the patients, per sample is presented here.

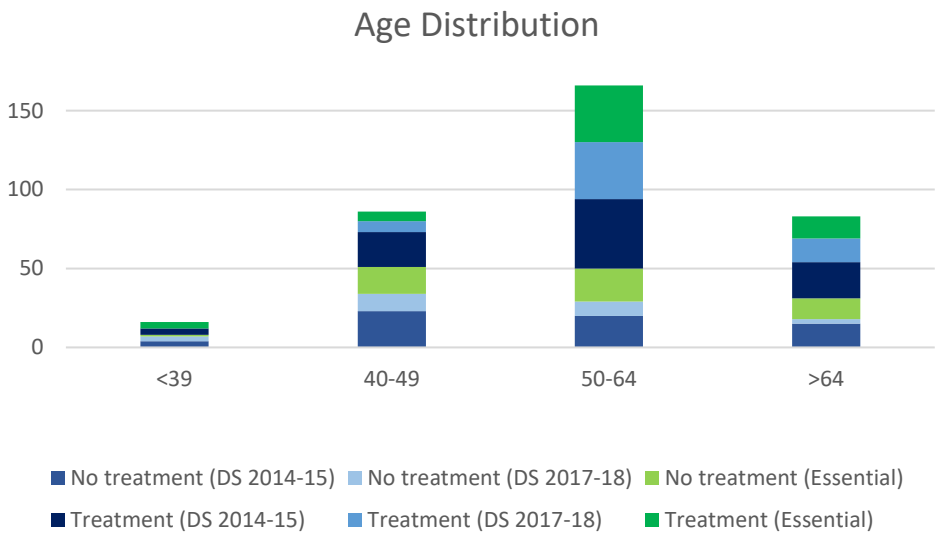


Figure 4.25 – Age distribution per sample.

4.2.4.1 Dose and Compressed Breast Thickness

The relationship between compressed breast thickness and dose is presented in the next graphs for the CC and MLO views separately, and for each sample. Here we present the values of MGD indicated by the manufacturer as the dose parameter. The versions of these graphs but with the calculated MGD values is presented in Appendix E.

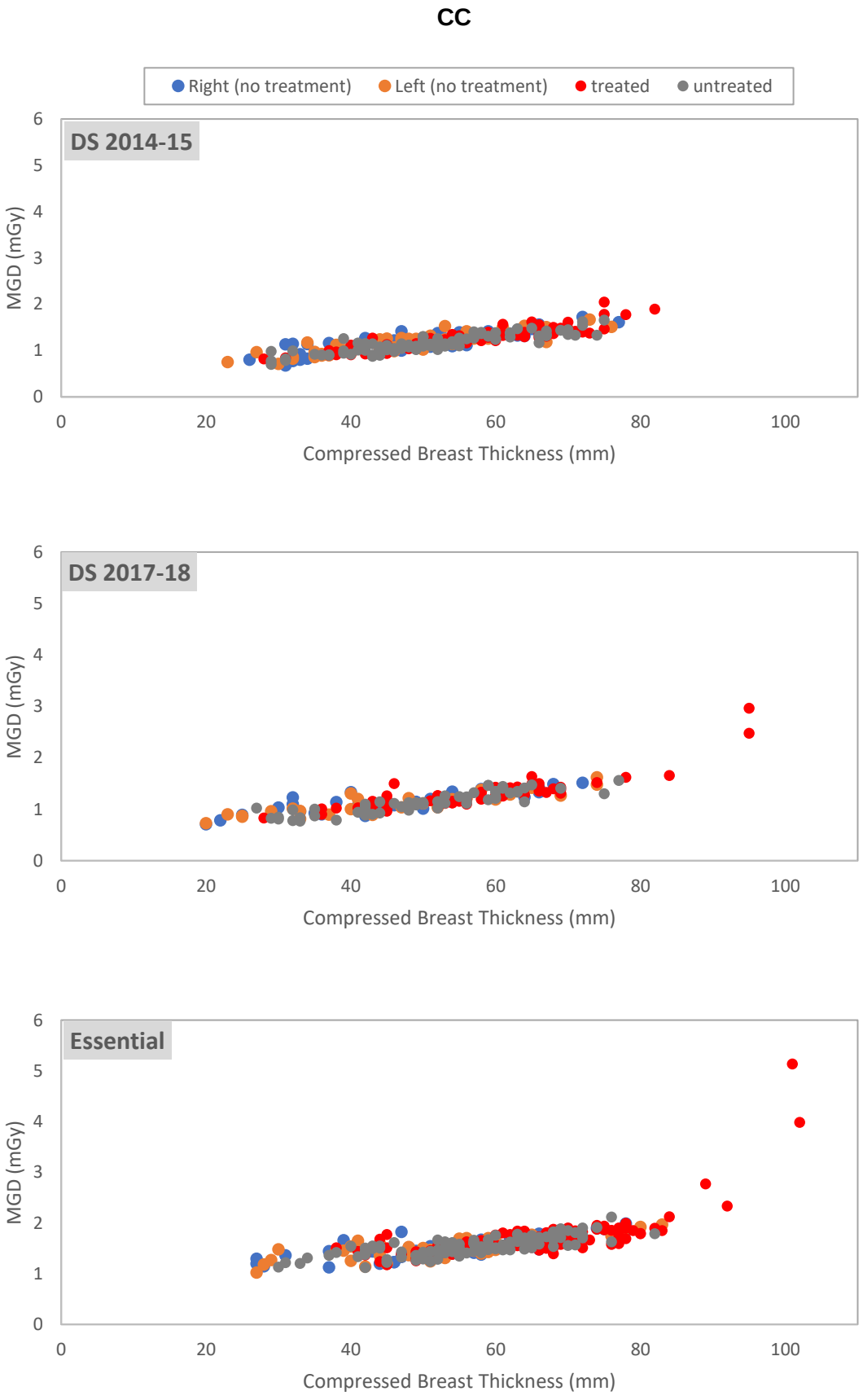


Figure 4.26 – Mean glandular dose (indicated by the system) vs. compressed breast thickness, for the CC view.

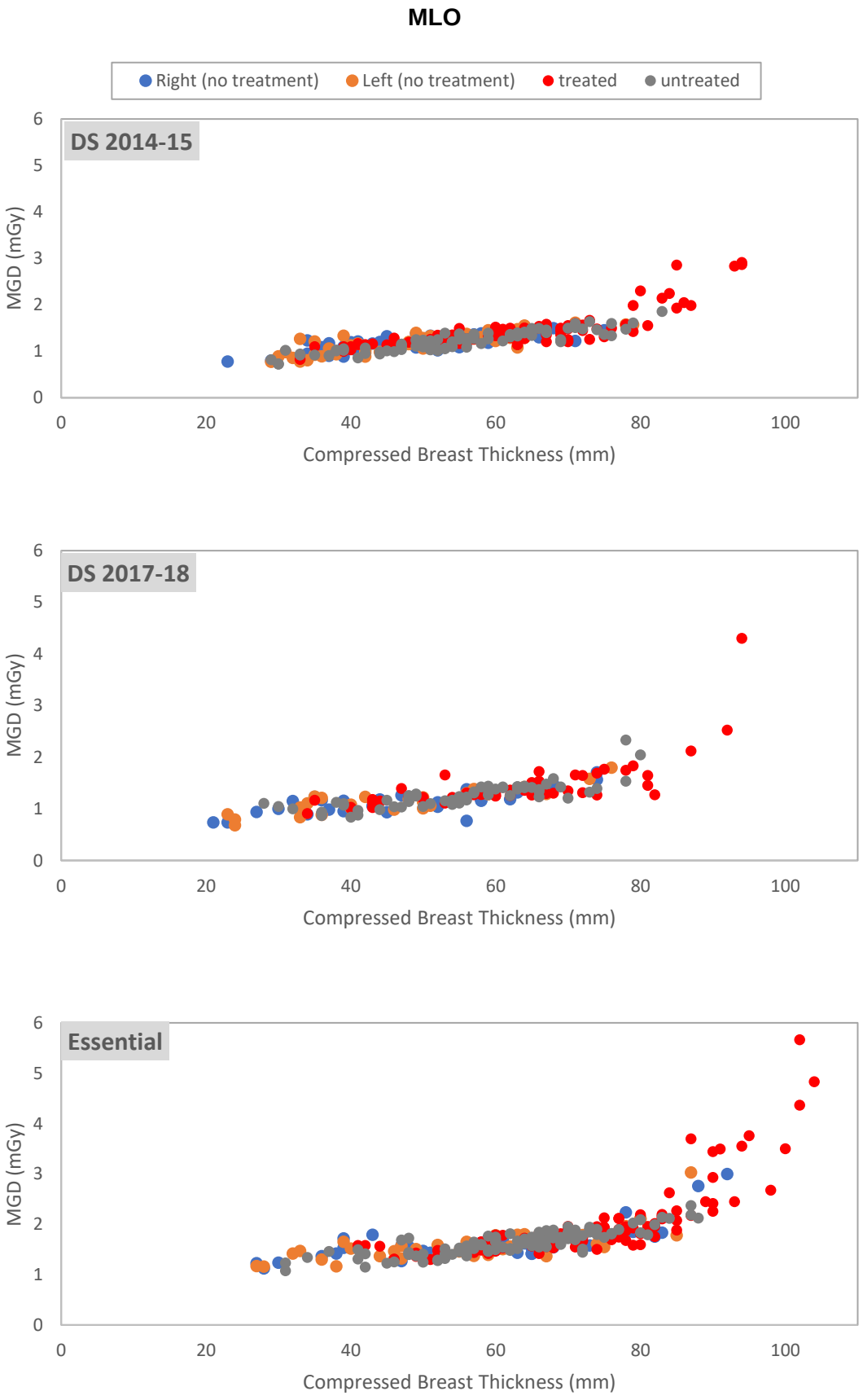


Figure 4.27 - Mean glandular dose (indicated by the system) vs. compressed breast thickness, for the MLO view.

The dose increases with the compressed breast thickness, at first slowly and then at over around 85mm more abruptly. The higher thicknesses seem to correspond mainly to treated breasts, and subsequently these have higher doses.
At the typical range of thicknesses (20-80mm) the distribution of MGD does not seem to differentiate treated and untreated cases.

4.2.4.2 ESAK and DRL

The plot of ESAK (indicated by the manufacturer) against the grossly estimated breast volume (compressed breast thickness x area) is shown in Figure 4.28. In this case, we grouped together both samples of the Senographe DS system. Untreated cases include both breasts of the “no treatment” patients and the untreated side of treated patients. The dashed line indicates the established European diagnostic reference level [27] for mammography of 10mGy of ESAK (for standard breasts).

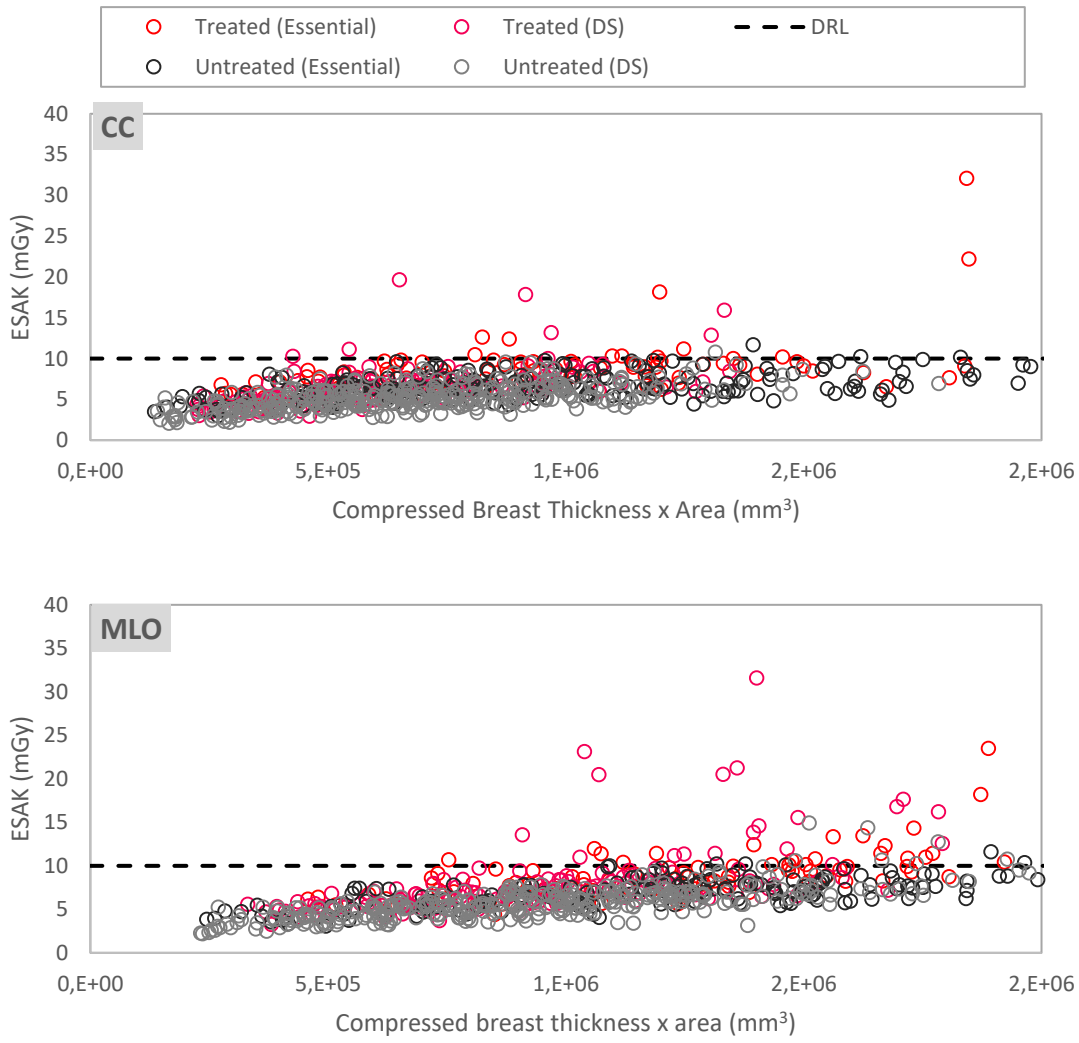


Figure 4.28 – ESAK vs. compressed breast thickness and imaged breast area.

This shows that the values above the DRL (for the standard breast) correspond mostly to treated patients. The values of the MLO view present more values above the reference level, including more of untreated patients.
The combination of the imaged breast area and the compressed breast thickness does not represent the breast volume, as this is considering as if the breast was a parallelepiped. But here we can see a comparison of these two parameters with a dose parameter. Because compressed breast thicknesses of different breast's sizes would present different imaged areas.
Most of the points above the DRL correspond to the same values of high thickness seen at the tail of the MGD vs. Compressed Breast Thickness graphs.
Higher thicknesses and higher doses seem to occur mainly for treated breasts with sometimes smaller imaged breast areas. Slightly high dose for smaller thicknesses are probably related to higher imaged breast areas.
DRLs are set as a function of compressed breast thickness, or for a set "standard" value of compressed breast thickness. These results start to suggest that there is no need to set DRLs differently for treated breasts or exclude treated breasts from the data collection when setting DRLs.

4.2.4.3 Variability between left and right untreated breasts and between treated and untreated breasts

4.2.4.3.1 Compressed Breast Thickness

The typical values of compressed breast thickness and its difference between breasts of the same patients are distributed in Figure 4.29 and Figure 4.30, respectively.

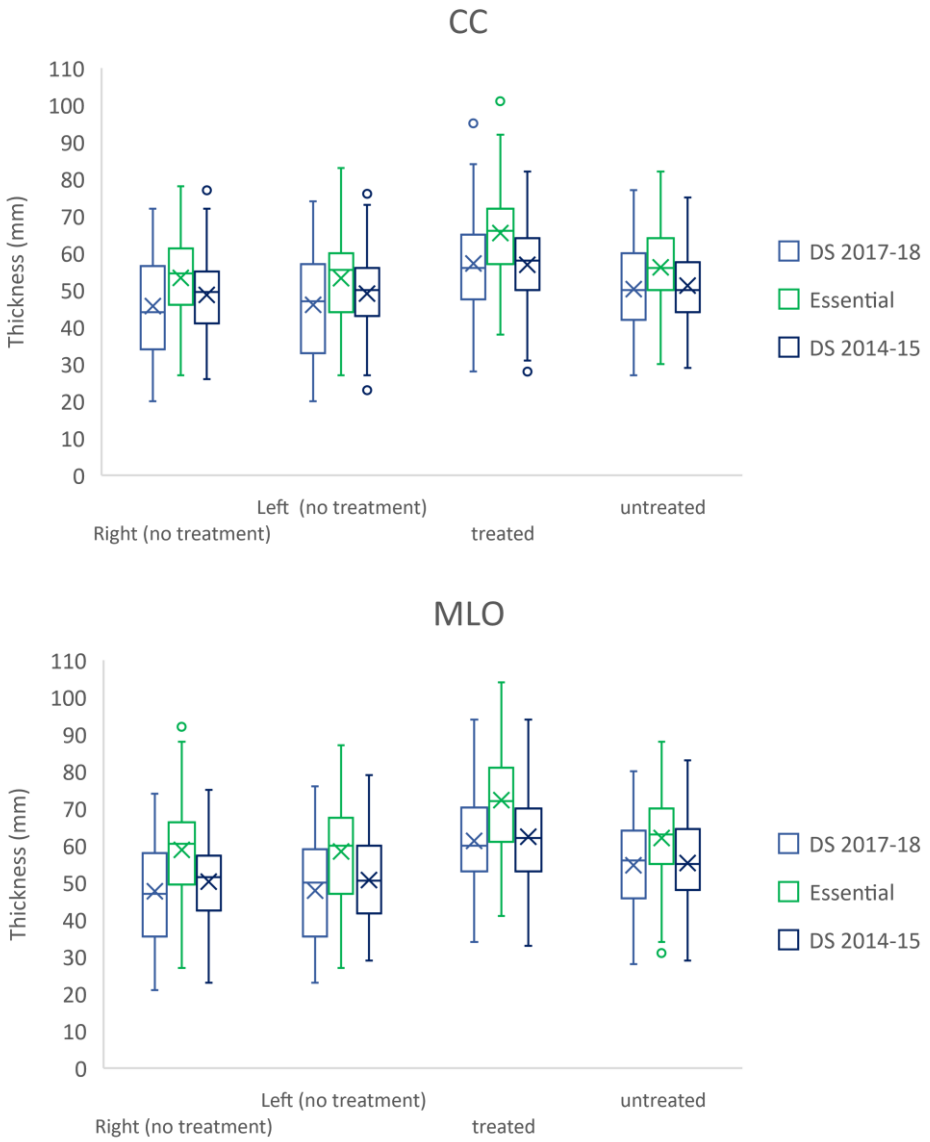


Figure 4.29 – Box plot of compressed breast thickness for each breast and each sample.

Breasts with no treatment have similar distributions of thickness. The Senographe Essential data has a higher mean thickness, as expected since it can image larger breasts. The mean values for the CC view for patients with no treatment are 50mm, 48mm and 54mm for the DS 2014-15, DS 2017-18 and the Essential data, respectively. For the MLO view, the values are a little bit higher, at 55mm, 54mm and 66mm, in the same order.

The patients who had undergone treatment prior to the mammography exam, have similar mean values for the compressed breast thickness on their untreated breast, but larger values on the treated side.

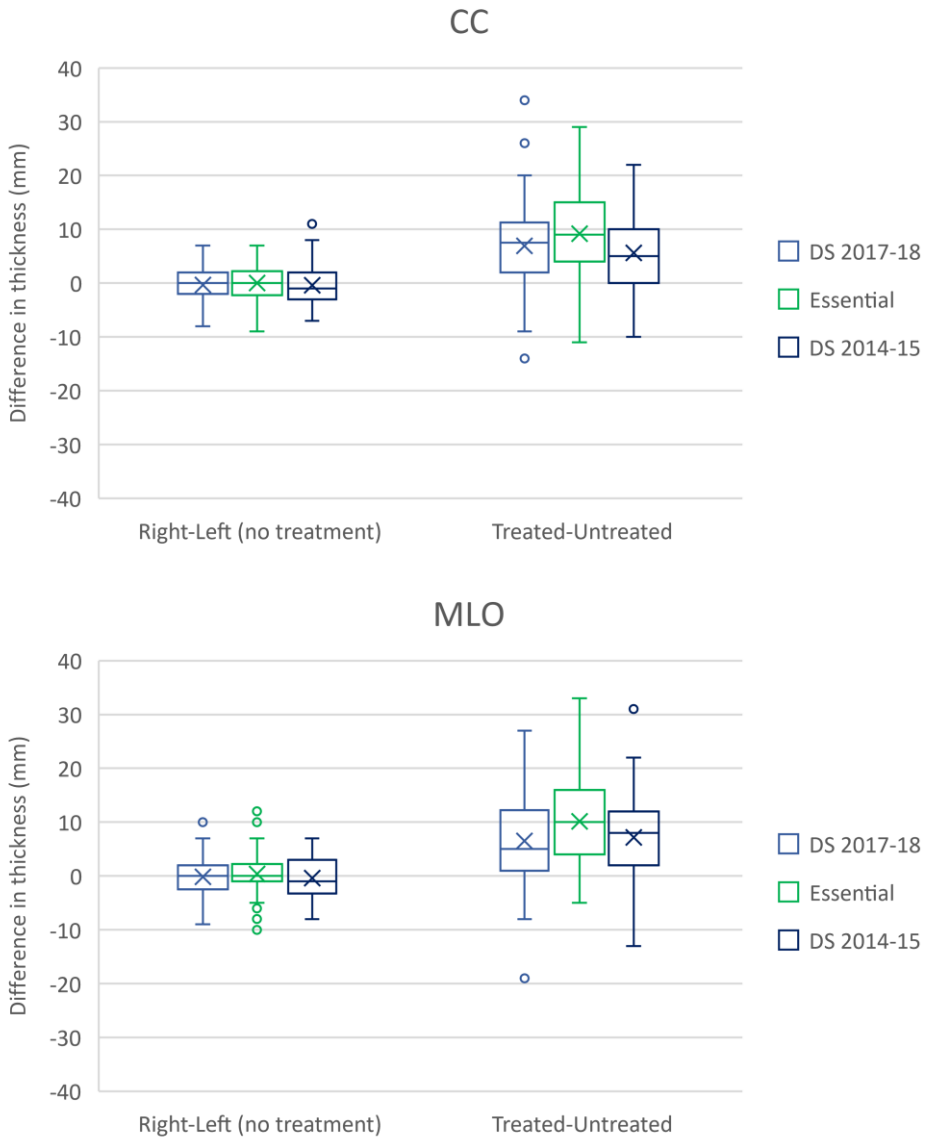


Figure 4.30 – Box plot of the differences in compressed breast thickness between the right and left breasts of patients with no treatment, and between the treated and untreated breasts of treated patients, for each sample.

The difference between the left and right untreated breasts has a mean close to zero. There are studies, on larger scales, that imply that the left breast is slightly larger than the right breast [58], but that is not evident here, and size is not precisely equal to the compressed breast thickness.

A treated breast's mammography indicates, on average, 6 to 10mm more of thickness than its untreated breast. One could think that a treated breast might be smaller due to tumour and tissue removal during the treatment, but we must remember that the indicated thickness depends on the size of the breast and the pressure applied on it.

The compression force and imaged area are studied next.

4.2.4.3.2 Compression force

The typical values of compression force and its difference between breasts of the same patients are distributed in Figure 4.31 and Figure 4.32, respectively.

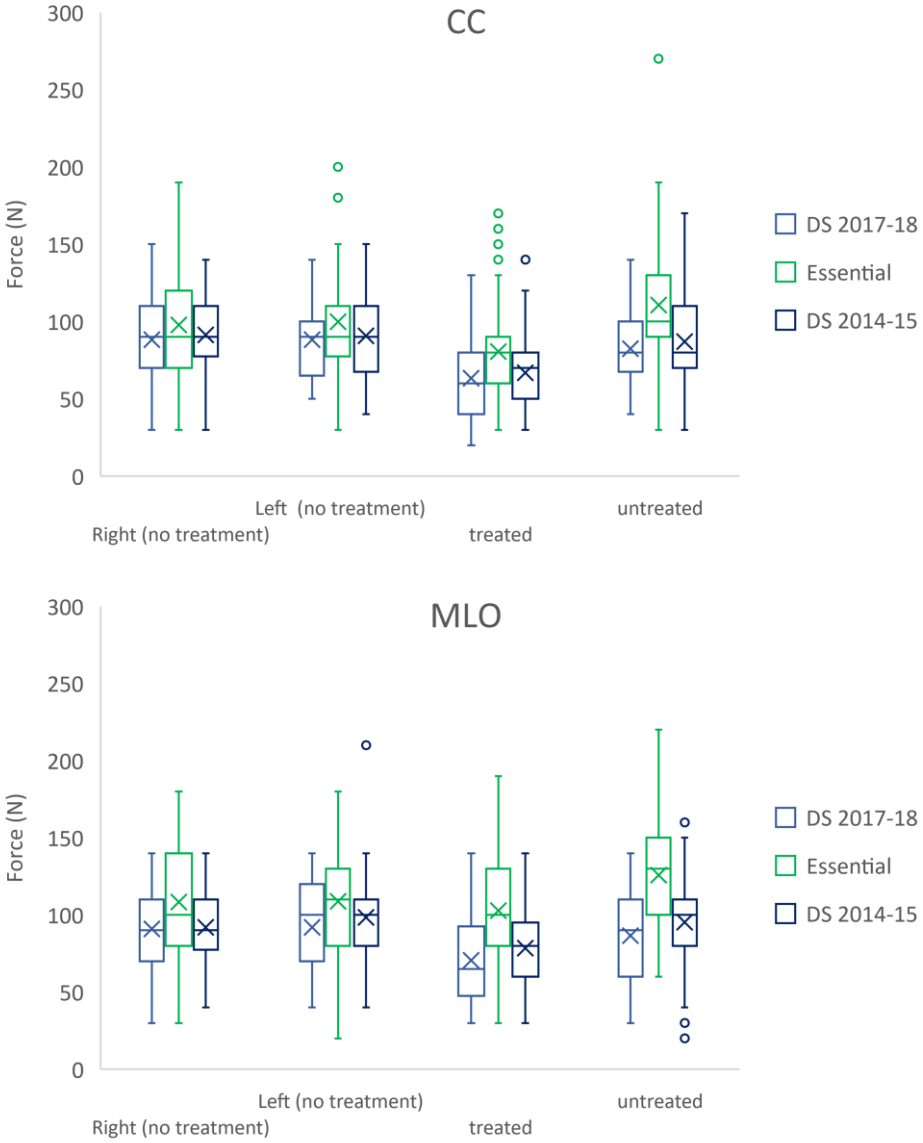


Figure 4.31 - Box plot of compression force for each breast and each sample.

The compression force is subject to larger variations as seen in the study of the normal variability (section 4.2.3).

The median force for the CC view, for patients with no treatment, proved to be 90N for all samples, while on the MLO view a larger median value was found for the samples of more recent exams. Namely, the DS 2017-18 had a median force of 110N and the Essential data had a median force of 100N.

There are some cases where it seems too much compression force was applied (above 200N [1]), two cases on the Senographe Essential, where the CC view had a force that reached 270N on an untreated breast. Another case of compression force higher than 200N appears in an MLO view of a patient with no treatment imaged by the Senographe DS. These values exceed the maximum compression achievable with these mammography units, and must be an error. But in general, the force values indicated on Essential's images are higher on average.

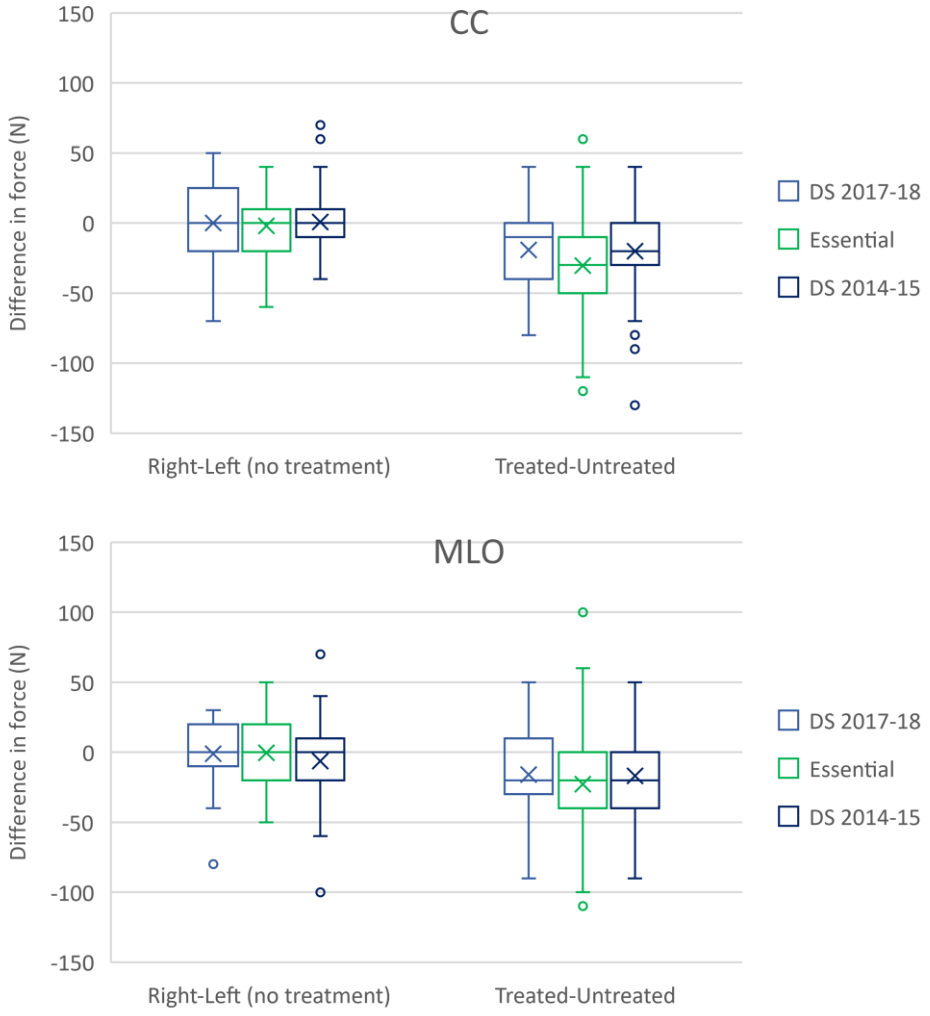


Figure 4.32 - Box plot of the differences in compression force between the right and left breasts of patients with no treatment, and between the treated and untreated breasts of treated patients, for each sample.

The mean difference between breasts with no treatment is around null, while on average a treated breast received around 15-30N less than the same patient's untreated breast. The patients considered on the first sample (DS 2014-15) had mostly the exams preformed up to three years after the treatment, and the patients from the other two groups had a time interval of four to five years. So, the patient could be psychologically

more sensitive towards the treated breast and the radiographer could be influenced by that knowledge and generally apply less compression. It could also be that the tissues were affected by treatment in that the breast tissues became denser, more rigid and/or more sensitive to the touch.

This result may be related with the higher thicknesses indicated for treated breasts, because with more compression the tissues spread out more and the thickness reduces.

4.2.4.3.3 Area

The typical values of the imaged breast area and its difference between breasts of the same patients are distributed in Figure 4.33 and Figure 4.34, respectively.

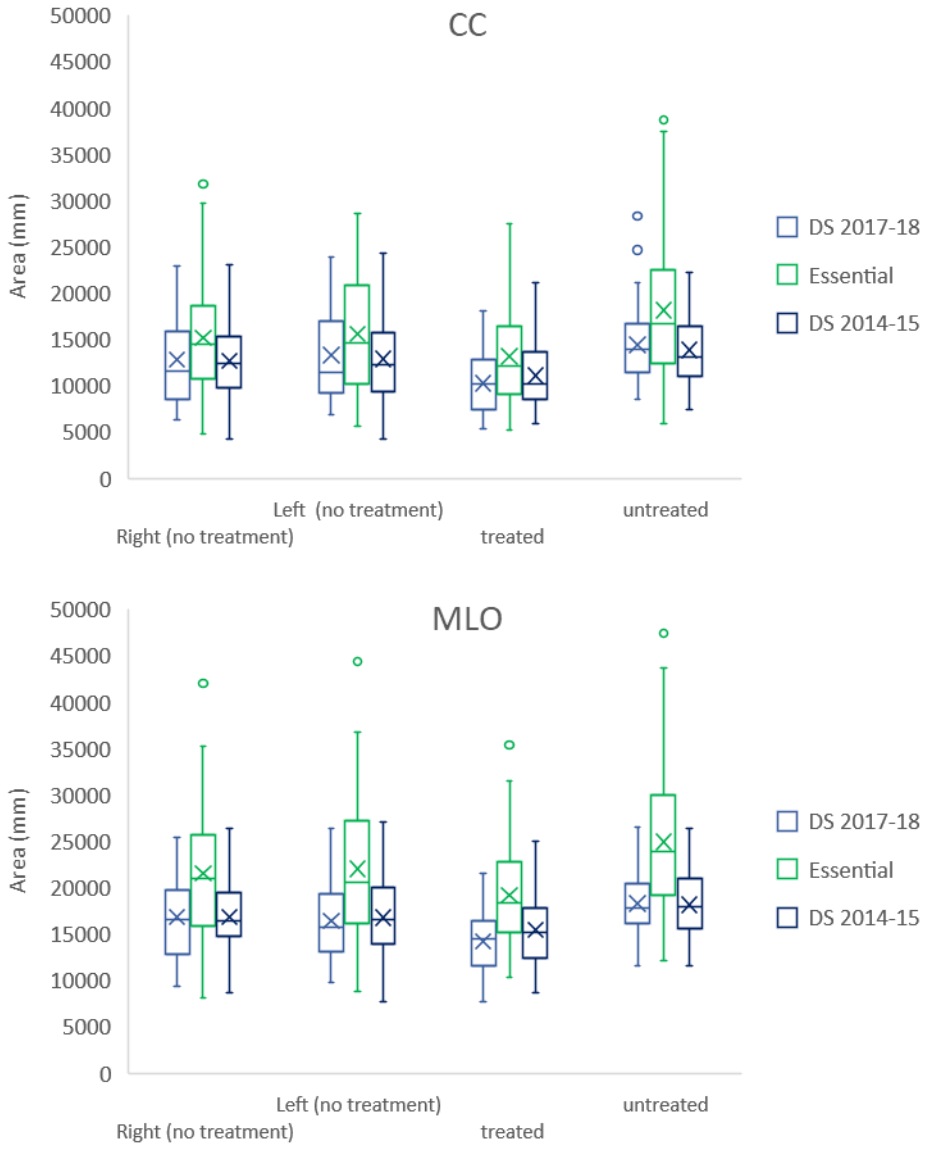


Figure 4.33 – Box plot of imaged breast area for each breast and each sample.

There is more imaged breast area for the Essential images, and for the MLO, as expected.

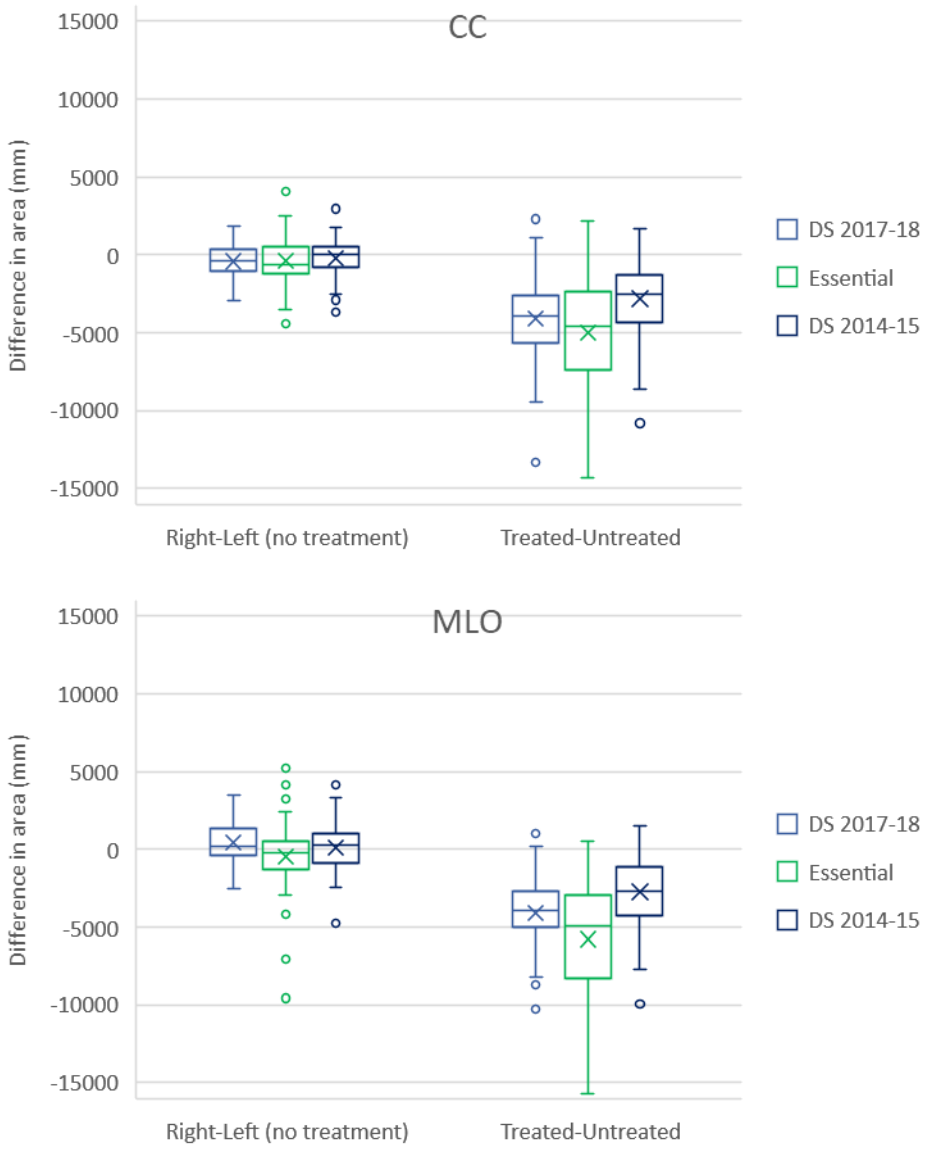


Figure 4.34 - Box plot of the differences in imaged breast area between the right and left breasts of patients with no treatment, and between the treated and untreated breasts of treated patients, for each sample.

Between the left and right breast of untreated patients, the difference of areas does not appear significant. As for the treated patients, treated breasts have almost always a smaller imaged area than the untreated ones. This, again, agrees with the findings of the variability of compressed breast thickness and compression force. Lower compression leads to less area being imaged and bigger thicknesses. There is also less tissue in a lot of treated breasts, caused by its treatment and it reflects on the imaged area which is

sometimes obvious when visualizing the images side to side, as shows Figure 4.35 as an example. If the treated breast actually becomes more rigid and/or denser, the imaged area may be smaller as well, because the tissues spread less.

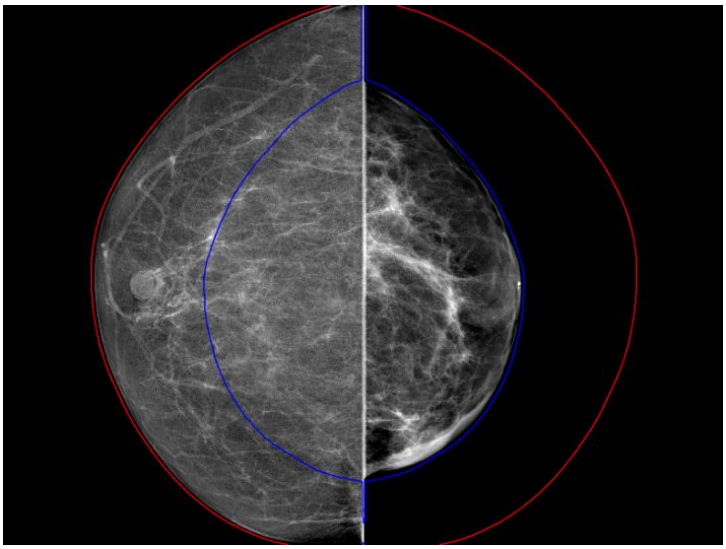


Figure 4.35 – Example of the CC view of the right untreated breast (red) and of the left treated breast (blue) of the same patient. The contours are overlaid on each side for better observation of the differences of the imaged areas.

4.2.4.3.4 ESAK

The distribution of the values of the Entrance Surface Air Kerma (ESAK), estimated by the EUREF method [1], and its difference between breasts of the same patients are distributed in Figure 4.36 and Figure 4.37, respectively.

The average values of ESAK are higher for the Essential images.
As it was seen in the section 4.2.4.1 (page 58), the treated patients receive, on average, more dose than the untreated patients, specifically on their treated breast. The outsider values presented in the distribution of the treated breast correspond mostly to the values above the DRL (for standard breasts). This could be related to the lower mean compression and the higher mean thicknesses observed.

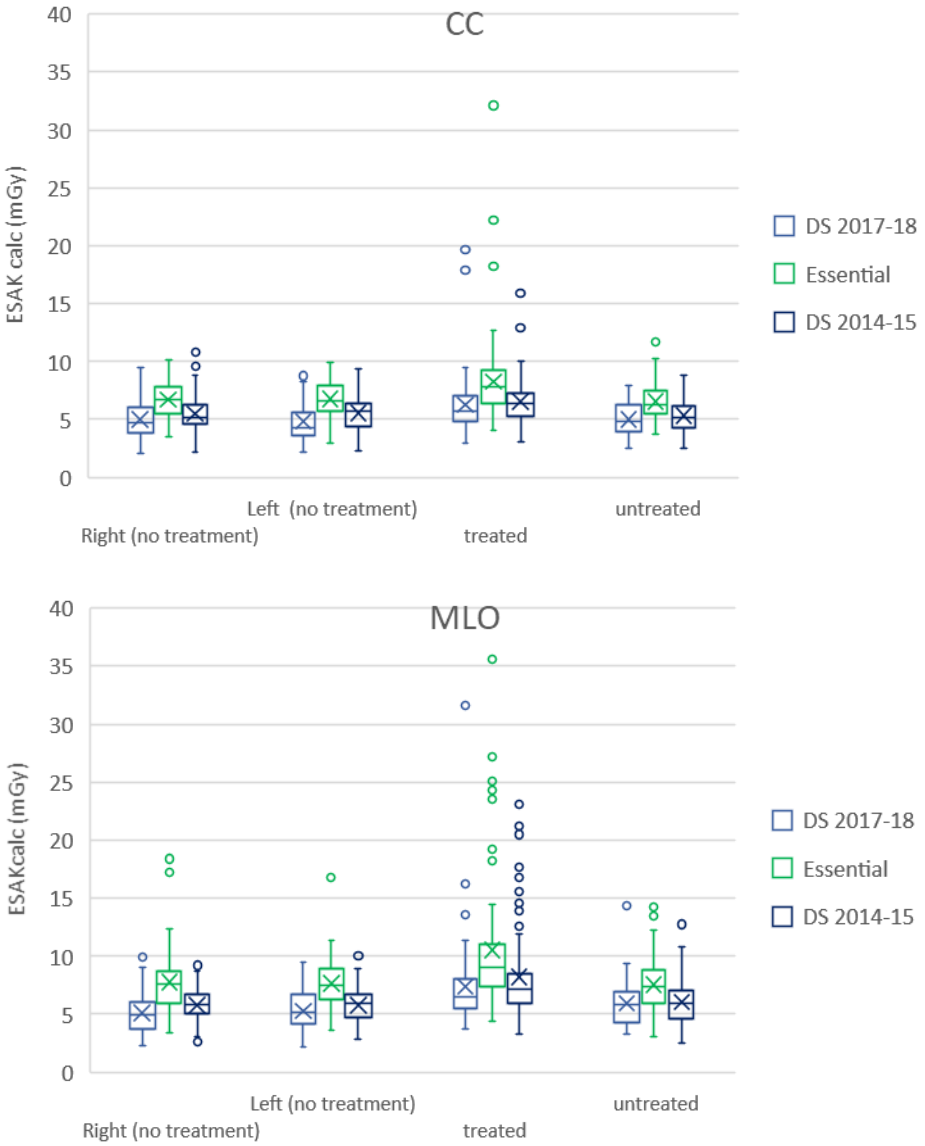


Figure 4.36 – Box plot of the calculated ESAK for each breast and each sample.

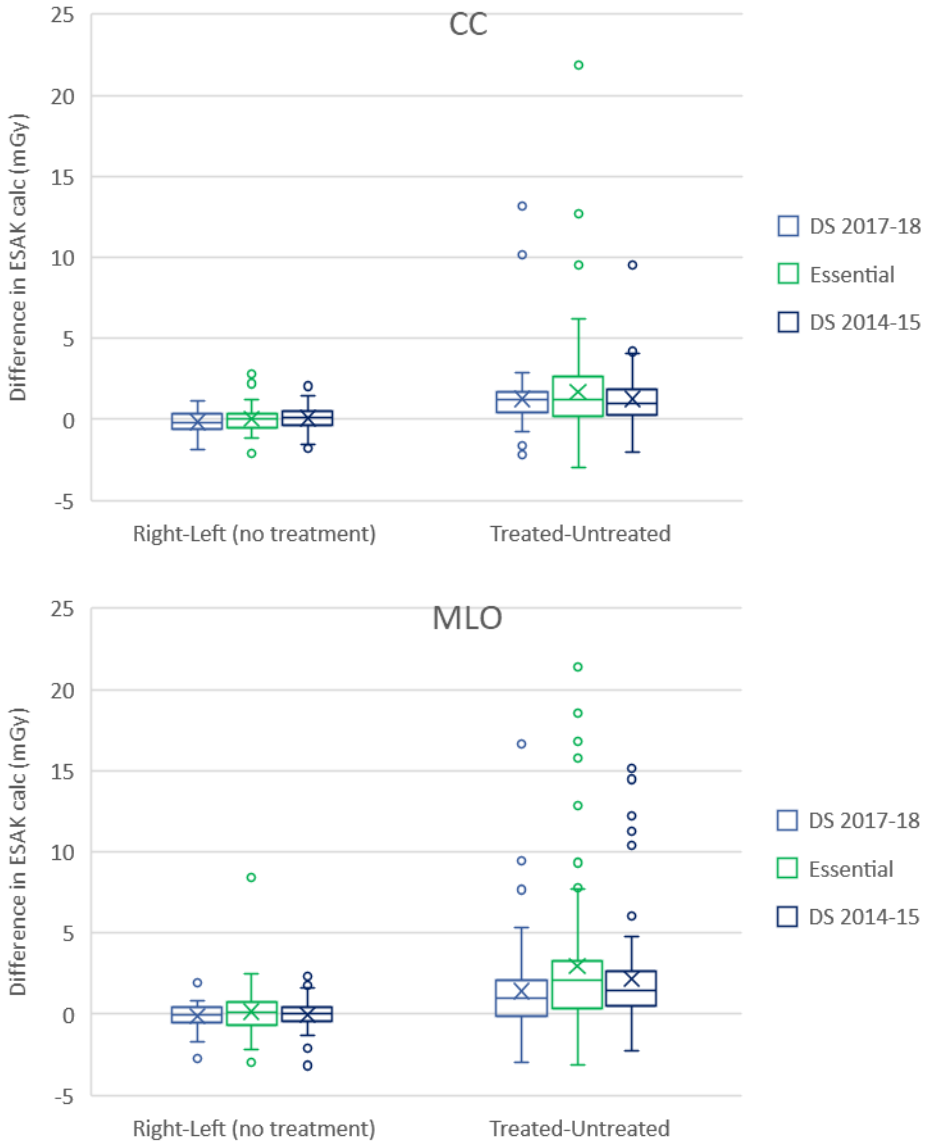


Figure 4.37 - Box plot of the differences in the calculated ESAK between the right and left breasts of patients with no treatment, and between the treated and untreated breasts of treated patients, for each sample.

The difference of ESAK between the right and left breast is pratically null. Treated breasts receive on average 2 mGy more than the untreated breast.

4.2.4.3.5 MGD

The typical values of the Mean Glandular Dose, estimated by the EUREF method [1], and its difference between breasts of the same patients are distributed in Figure 4.38 and Figure 4.39, respectively.

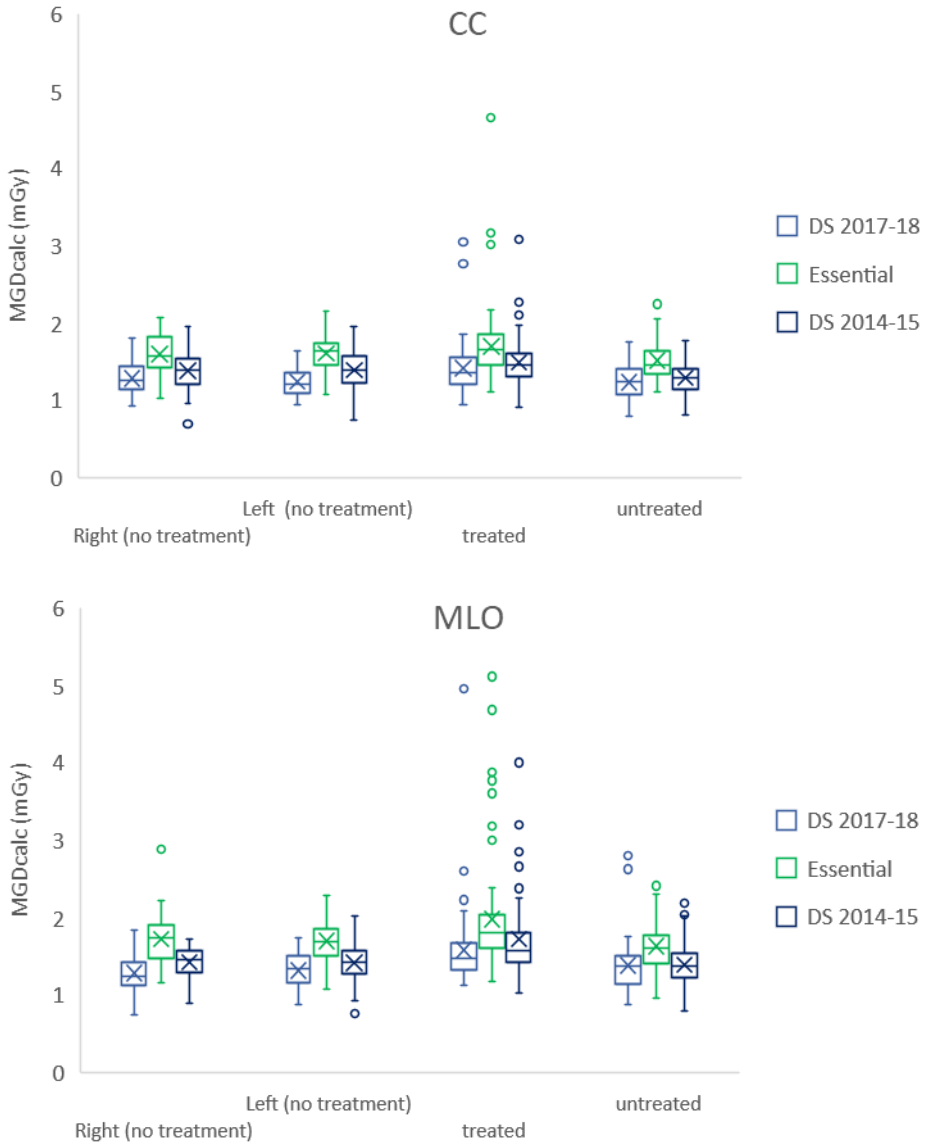


Figure 4.38 – Box plot of the calculated MGD values for each breast and each sample.

The mean MGD calculated values for the patients with no treatment (for both breasts) were 1.38mGy, 1.26mGy and 1.38mGy in the CC view for the samples DS 2014-15, DS 2017-18 and Essential, respectively. As for the MLO view, the values were 1.42mGy, 1.3mGy and 1.71mGy, in the same order. The indicated MGD values, shown on the relationship of MGD and compressed breast thickness, had a higher maximum MGD values.

The untreated breast of the set of treated patients had similar mean MGD values. Treated breasts reach higher maximum values for the mean glandular dose, up to 5.11mGy on the MLO views of the Essential data. The “outsider” values correspond to

the points in the tail of the MGD and compressed breast thickness relationship, as seen in section 4.2.4.1.

The mean values for the treated breasts were 1.49mGy, 1.41mGy and 1.70mGy in the CC view for the samples DS 2014-15, DS 2017-18 and Essential, respectively. As for the MLO view, the values were slightly higher, 1.72mGy, 1.58mGy and 1.98mGy, in the same order.

A study conducted by Baek et al. [59], with different units (Mammomat Inspiration, Siemens and Selenia, Hologic) found higher MGD values for patients who had prior breast conserving surgery, as well.

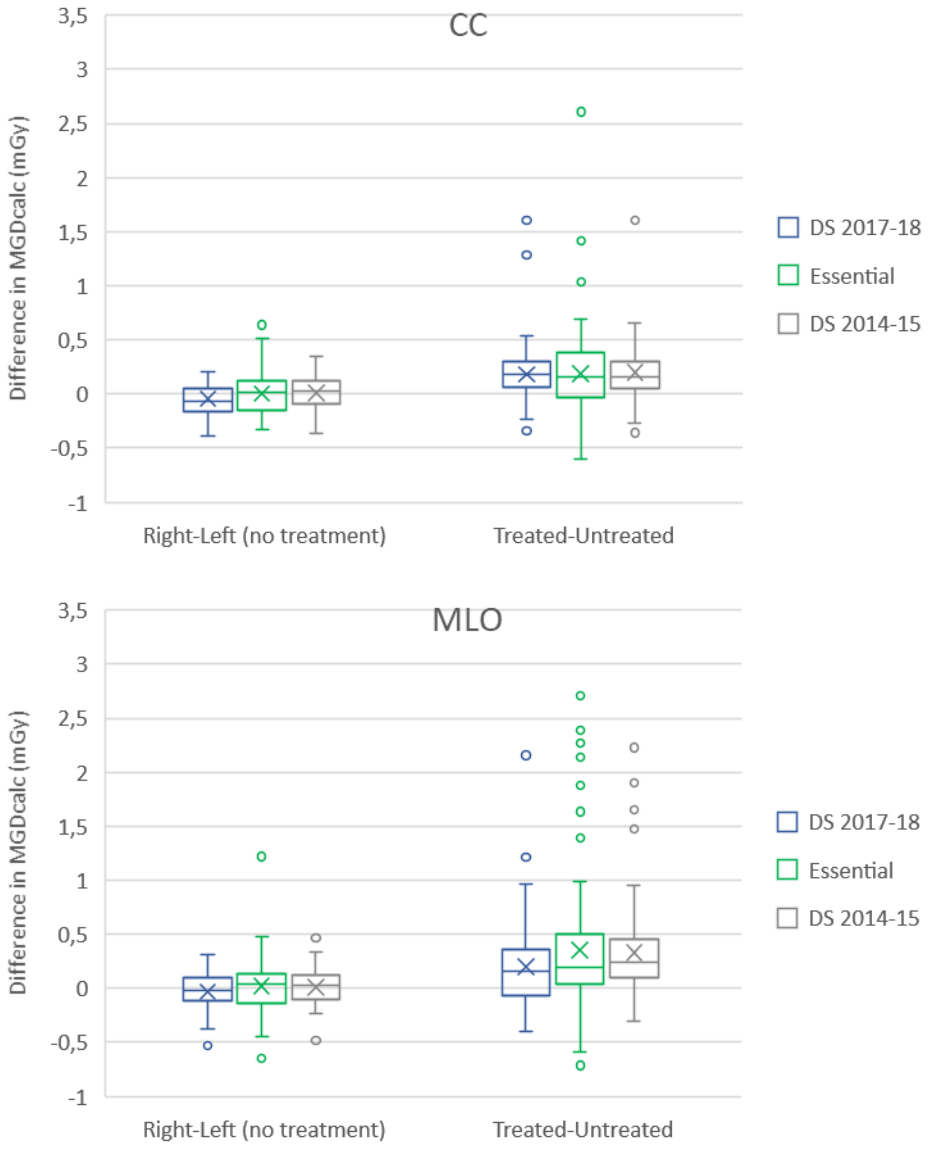


Figure 4.39 - Box plot of the differences in the calculated MGD between the right and left breasts of patients with no treatment, and between the treated and untreated breasts of treated patients, for each sample.

Again, the difference in mean glandular dose between the breasts of patients with no treatment is very low. Patients with treatment have a mean of around 0.2mGy more than the untreated breast, on the CC view, and of 0.2-0.3mGy for the MLO view.

4.2.4.3.6 Glandularity

The distribution of the indicated values of glandularity and its difference between breasts of the same patients are distributed in Figure 4.40 and Figure 4.41, respectively.

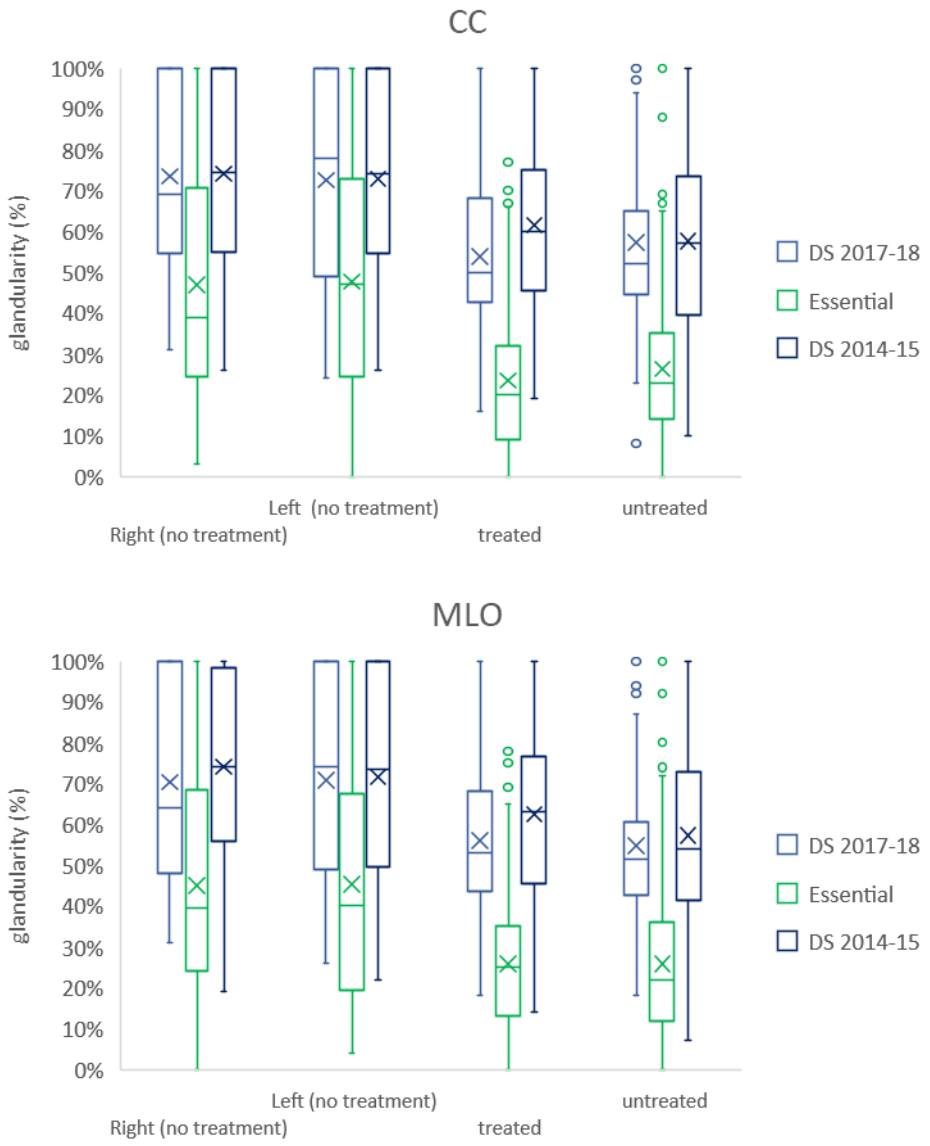


Figure 4.40 – Box plot of glandularity for each breast and each sample.

Glandularity of the breast is a property that varies much from patient to patient, as can be seen from its broad range of distribution. There does not seem to be a significant distinction between the patients with and without treatment. With other parameters

studied above, normally the treated breast values had a distinct distribution to all untreated breasts. Here, the distinction in glandularity between the two breasts of the same patient, whether with treatment or not, does not seem to be significant. We can see, though, an apparent difference between the breasts with no treatment (right and left) vs the untreated breasts of patients with treatment, in terms of the mean glandularity. Untreated breasts in the treated set of patients had lower mean glandularity values (~58% for the DS unit and ~26% for Essential) than the no treatment set of patients, for both mammography units (~73% for the DS unit and ~46% for Essential). Remembering that none of the patients considered in this study are “standard” patients, maybe this could be related to risk factors, age groups or symptoms of benign conditions. This needs to be looked into further.

A distinction can also be found between exams of different mammography units. As hinted before, in the section 4.2.2 (page 50), the glandularity indicated by the Essential unit reaches lower values, down to 0%, compared to that of the DS unit which never indicated glandularity values lower than 7%. This will be investigated further, when comparing the same patient imaged on both units.

It was also pointed out that Essential’s images of treated breasts never indicated glandularity values over 80%.

As for the difference between the glandularity of both breasts of the same patients, there does not seem to be a significant distinction from patients with and without treatment. It can only be seen a slightly larger range of differences when comparing treated and untreated breasts of the same patients (see Figure 4.41).

We can see, mainly in the MLO view, that the difference in the glandularity values for the older sample, of 2014-15, seems to be higher, on average. The 2014-15 and 2017-18 samples were differently acquired, and the latter had a fixed time gap between the treatment and the mammographic exam, contrary to the former case.

It seems important to point out that some treated breast have surgical clips, and according to Senographe’s Operator Manual [39], when imaging dense objects with areas greater than $2mm^2$ like radio-opaque markers within the AOP ROI, it can affect the estimation of tissue density and even produce a degraded image. Since surgical clips are literally metallic clips, that are subject of study later on, one could expect them to affect the indicated value of glandularity and the AOP parameters which would consequently affect the dose. But in general, the glandularity values indicated by the

mammography unit does not seem to reflect that the automatic exposure control detects a treated and untreated breast composition.

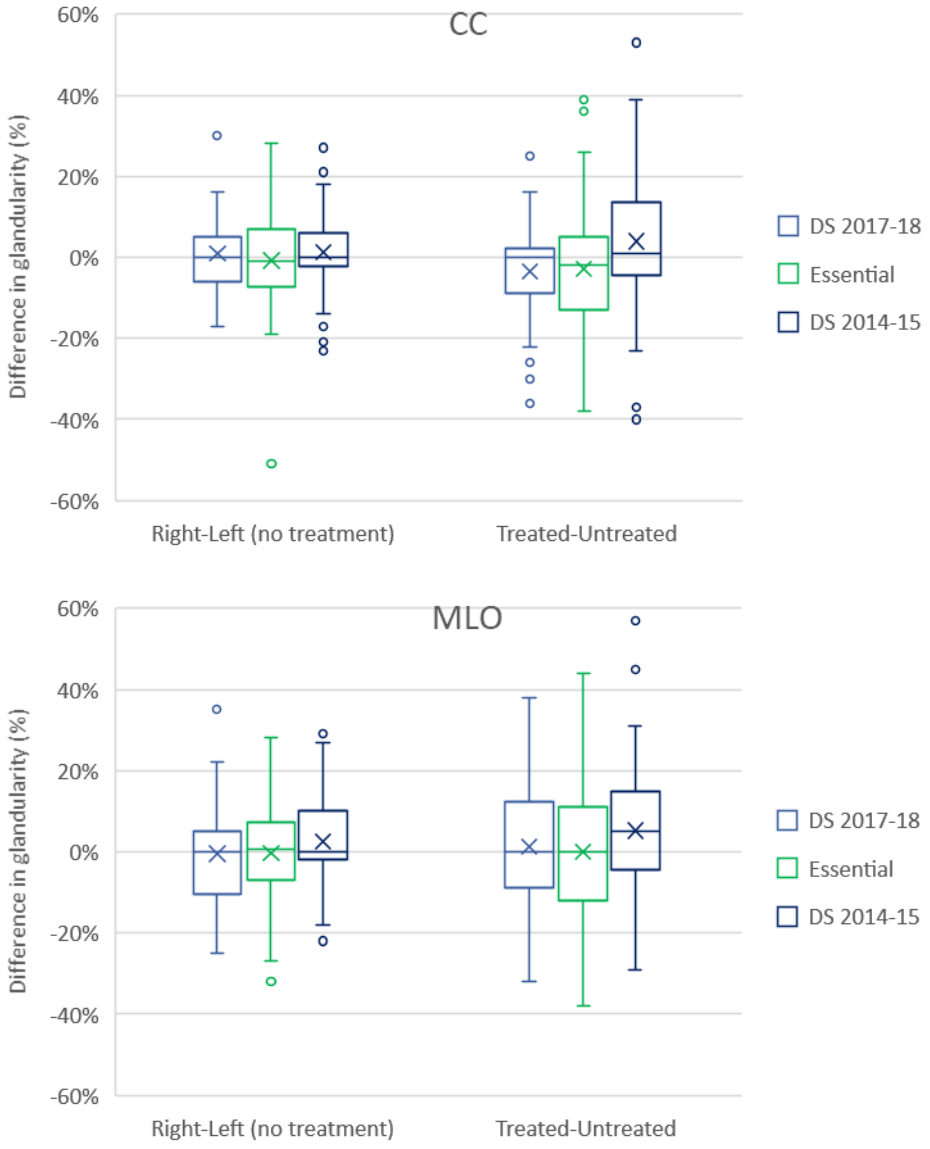


Figure 4.41 - Box plot of the differences in the glandualrity (indicated by the system) between the right and left breasts of patients with no treatment, and between the treated and untreated breasts of treated patients, for each sample.

Glandularity vs. Compressed Breast Thickness:

Dance's method to estimate MGD considers the same dependence of glandularity on compressed thickness for all exams within an age group [1], whereas the equipments estimate glandularity for each case separately. The observed differences in MGD in section 4.2.2 (page 50) reflected the difference in glandularity.

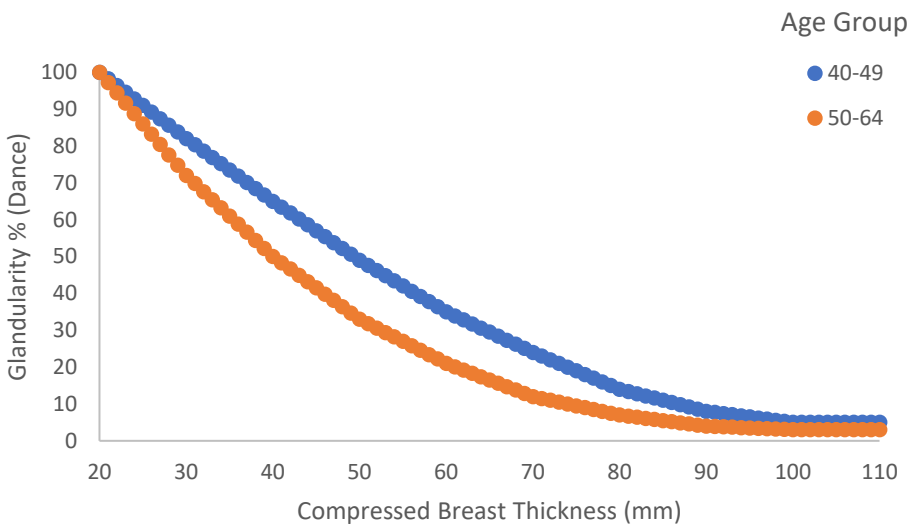


Figure 4.42 – Glandularity as indicated by Dance's method [1] vs. compressed breast thickness, for the age groups 40-49 and 50-64

In Figure 4.42 can observe the typical glandularity values for different compressed breast thicknesses for the two age groups considered by Dance's method. For the older age group, the glandularity values are lower, for the same thickness. Grouping all the samples by the two age groups (40-49 and 50-64), we can see that the glandularity indicated by the manufacturer does not take into account the age of the patient (Figure 4.43) as expected. We should also remember that our samples include patients with ages outside those groups as well. We can see that the values estimated by the system disperse around the Dance typical values, as expected. The indicated glandularity still decreases with thickness.

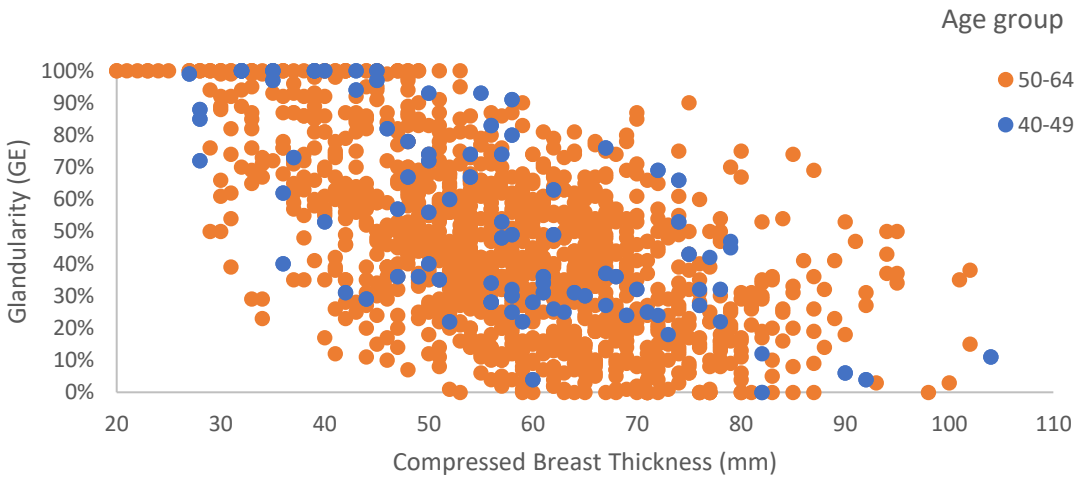


Figure 4.43 – Glandularity indicated by the equipments vs. compressed breast thickness, per age group.

The EUREF method to estimate the dose parameters was included in this study as a standard, precisely to avoid the different methods of detecting glandularity.
An independent estimation of the glandularity should be conducted for, but it was not included in this work.

4.2.4.3.7 Summary

To understand the whole scope of the variations found, the variabilities are presented next for all the parameters as a percentage difference.
These values were calculated with the next equations, depending on the patient's case.

$$Dif\ \% = \frac{Right - Left}{Mean\ (Right\ and\ Left)} \times 100\% \quad (4.2)$$

$$Dif\ \% = \frac{Treated - Untreated}{Untreated} \times 100\% \quad (4.3)$$

For the patients with no treatment, the differences to the mean value of both breasts (of each patient) were compared, since the typical differences between the two untreated breasts were found to be small, and it was impossible to decide which side was the reference.
As for the patients who had treatment, the reference was considered to be the untreated breast.
The mean percentage differences were then calculated for each sample and are presented in Table 4.8.
The mean absolute values for each sample can be found in Appendix F.
An exception was made for the glandularity case. The glandularity value is already in percentage and it is the parameter with lowest values (reaching 0% in some cases). Using equation 4.3 to determine a difference with a null value as a reference cannot be done. Therefore, the absolute differences in glandularity are included.

Table 4.8 - Mean percentage difference between the right and left breasts of patients with no treatment and between the treated and untreated breasts of treated patients. (*) refer to absolute differences/deviations.

	Right - Left			Treated-Untreated			Normal variability
	DS 14-15	DS-17-18	Essential	DS 14-15	DS-17-18	Essential	
Thickness	-1%	0%	0%	12%	16%	18%	**6%**
	0%	-1%	1%	14%	13%	18%	**6%**
Compression Force	1%	-2%	-2%	-20%	-21%	-22%	**22%**
	-7%	-3%	0%	-10%	-14%	-15%	**19%**
Area	-1%	-4%	-3%	-20%	-28%	-26%	**4%**
	1%	2%	-2%	-15%	-22%	-22%	**4%**
Glandularity	-1%*	1%*	-1%*	4%*	-3%*	-3%*	**7%***
	3%*	0%*	0%*	5%*	1%*	0%*	**7%***
MGD	-1%	3%	0%	11%	16%	13%	**7%**
	0%	-3%	1%	18%	14%	21%	**6%**
MGDcalc	-1%	3%	0%	17%	16%	13%	**7%**
	1%	-3%	1%	25%	16%	22%	**7%**
ESAK	-1%	3%	-1%	25%	28%	26%	**10%**
	0%	-3%	1%	36%	26%	37%	**10%**
ESAKcalc	-1%	4%	-1%	25%	28%	27%	**10%**
	-4%	-3%	1%	37%	26%	39%	**10%**
			CC	**MLO**			

Normal differences between the right and left breasts of patients with no treatment fall below 10% of the mean value, which we can consider to be insignificant when compared to the normal variability between exams. Compression Force has slightly higher mean difference values (>5%) in the case of the MLO view of the 2014-15 sample.

As for the differences between treated and untreated breasts, it reflects the differences between a treated and untreated breast in terms of radiation doses and positioning technique parameters. For comparison, the normal variability obtained earlier in this study is included in the table. Only the compression force presents similar variability for both cases, which makes sense because it is an independent parameter that varies much between exams.

Glandularity between treated and untreated breasts, of patients who had treatment, proved to have lower or similar mean (absolute) differences than the found mean deviation (of the exams from Senographe DS only, from the normal variability study), and it is the only parameter to do so.

These results suggest that the AEC system does not notice different breast compositions between a treated and untreated breast of a patient.

4.2.5 Comparison of the mammography units

The 40 patients set aside for this comparison, had an average time span of a 12.35 months between their exams performed in each unit, with some patients having their second exam up until 18 months after the first one. Even though breast density, and intrinsically the glandularity, decreases with age, it has been reported that this change is "likely to be imperceptible over 1 year" [60].

In general, the previous results of the Senographe Essential data, when compared to the DS model's data, included higher imaged areas, compressed breast thicknesses, compression force and doses. As of the indicated values of glandularity, the Essential data presented lower values, on average. With this smaller subsample, we compared directly those differences between the mammography unit for the same patients.

Table 4.9 presents the mean, median, minimum and maximum percentage differences after calculating them with equation 3.8 (page 36) for some of the image parameters, including the maximum and mean pixel value and the standard deviation (SD) of the imaged breast area. In this case the MGD and ESAK percentage differences referred are the values indicated by the manufacturer, for obvious reasons. The distribution of the difference in glandularity between the Essential and the DS units is presented in the box plot in Figure 4.44, because greater differences were found.

Table 4.9 – Mean percentage difference between values indicated by the Senographe DS system and the Senographe Essential system.

	Difference (Essential-DS)/Mean							
	Mean		**Median**		**Minimum**		**Maximum**	
	CC	MLO	CC	MLO	CC	MLO	CC	MLO
Thickness	7%	11%	8%	11%	-50%	-39%	38%	62%
Force	17%	22%	18%	19%	-120%	-100%	113%	120%
Area	6%	19%	7%	18%	-16%	-14%	51%	42%
MGD	23%	25%	23%	25%	-17%	-46%	72%	69%
ESAK	17%	22%	18%	22%	-56%	-60%	65%	89%
Maximum Pixel Value	-10%	-7%	-9%	-7%	-34%	-30%	26%	28%
Mean Pixel Value	-1%	-1%	-1%	1%	-12%	-16%	17%	11%
SD	-6%	4%	-2%	6%	-67%	-56%	33%	50%

Positive differences implicate that the Essential's values are higher, and that was true for most of the mean differences. The compressed breast thickness had a mean

difference of around 10%. The imaged breast area had a mean and median difference below 10% as well but only for the CC view. The dose parameters, namely the MGD and ESAK indicated in the DICOM header, presented mean and median differences around 20%.

As for the parameters related to pixel values of the imaged breast area, there is a small tendency of the Essential's image to have lower standard deviation and maximum pixel values. If the difference is intensified in the presence of more attenuating regions like surgical clips, then it should be noted that about half of the patients considered here had surgical clips present on their treated breast, making up a fourth of all the images considered. Further investigation is needed by comparing phantom images of the two units.

The distribution of the absolute differences in glandularity is presented in Figure 4.44. The Essential's unit seem to indicate, on average, less 25% of glandularity than that indicated by the DS's unit.

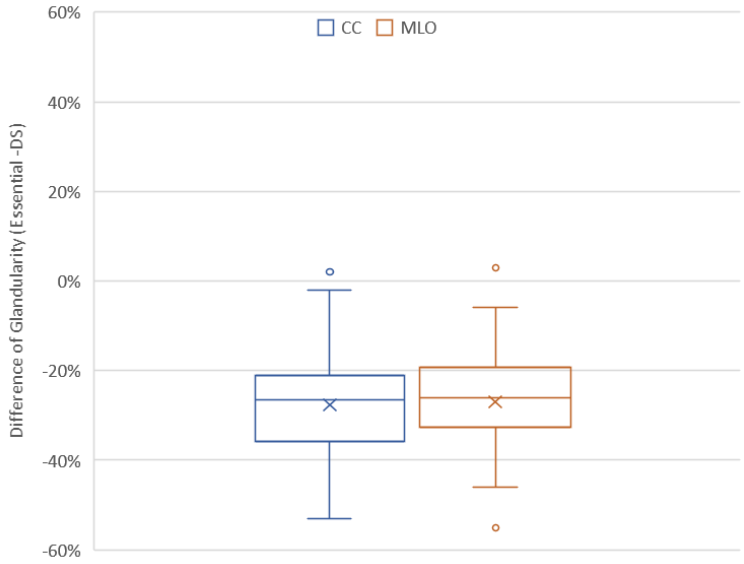


Figure 4.44 – Absolute difference in the estimated glandularity values between the two systems.

5 Conclusions

The initial assessment of a digital mammography unit through phantom images confirmed that the mean pixel value in the raw images is proportional to the dose, while in the processed images this relationship is logarithmic. This initial assessment allowed a better understanding of how the equipment and the automatic exposure control works. Clinical images are always processed images, stored after a proprietary image processing is applied to the initial raw image, to enhance the visualization of breast structures. The proprietary Premium View feature practically does not affect the Mean Pixel Values, but only the noise. Therefore, further computational processing done on clinical (processed) images, such as a contour delineation or detection of structures, should not be affected by these intrinsic image processing methods.

When clinical images were studied, the MDG and ESAK were estimated using the method suggested by EUREF, and differences to values indicated by the equipment were within around ±10% and ±30%, respectively. The different ESAK values are probably related to minor differences between the measured outputs considered for calculations in this work, and the output values indicated by the system. The larger differences in MGD values are probably related to the glandularities measured by the mammography unit, which are different and more varied than the standard dependence on thickness and age group considered for calculating MGD with Dance's method [1]. The results also suggest that two mammography units measure glandularity in a different way, despite being from the same manufacturer. The Essential unit seems to indicate, on average, less 25% of glandularity than that indicated by the DS unit.
The differences between the DS and Essential units for positioning techniques were evaluated, with the newer system yielding, on average, higher compressed breast thickness and radiation dose, which should be related to its larger detector. As for the difference in the glandularity values, further investigation is needed.

A preliminary study of ten patients with repeated mammography exams between 2009 and 2017 was used to establish a baseline of normal variability of parameters in mammographic exams. In total 276 images of 69 exams were considered in this preliminary analysis.
Variability of exam doses for repeated exams of the same patient, through the years, was found to be small (less than 10% for MGD and around 10% for ESAK). This is

probably related to the very consistent positioning found. The variability of the indicated glandularity by the GE Senographe DS system was about 13% while variability in compressed breast thickness and area were of about 5%. The most variable parameter proved to be the compression force This finding is consistent with Mercer et al. [14][15] studies of variability in compression force.

A bigger sample, comprising in total 1596 images, of 399 different patients was considered for the study of oncological treatment effects. This included 258 exams of patients who had undergone surgery and radiotherapy (123 on the right breast, 135 on the left breast), and 141 exams of patients who were being monitored for benign conditions and had no treatment until the date of the exam. The exams were retrieved from PACS, from two different GE mammography units. The exams were taken either between 2014-15 or 2017-18.

In the set of patients with no treatment, differences between the left and right breasts were practically null, (within ±3% in most cases), proving a good reproducibility within the same exams.

Breast cancer treatments effects seem to influence some mammography technique factors, especially the compression force and the breast thickness. The treatments probably cause the breast tissue to become more rigid and sensitive, which in turn influences the compression force applied. Lower compression forces may be related to higher compressed breast thicknesses, which in turn yield higher doses when imaged by a digital mammography unit. In fact, patients with doses above the established European diagnostic reference level (DRL) were mainly patients who had undergone treatment. It is important to remember that the DRLs are established for standard patients, and at IPO-Porto that is not the usual case.

Mean exam doses are higher for treated breasts (CC: 1.49mGy, 1.41mGy and 1.70mGy; MLO: 1.72mGy, 1.58mGy and 1.98mGy for DS 2014-15, DS 2017-18 and Essential, respectively), but this seems to be mostly the result of increased breast thickness. Thedependence of exam doses on compressed breast thickness appears to be very similar for treated and untreated breasts, with the exception of a few outliers at extremely high compressed thicknesses, which occurred only for treated breasts. In these cases, the volume of the treated breast is actually smaller than might be expected, e.g. these treated breasts with extremely high thickness tend to have relatively small areas.

If DRLs are set as a function of compressed breast thickness, or set for a “standard” value of compressed breast thickness, then the results presented in this study strongly

suggest that there is no need to set DRLs differently for treated breasts or exclude treated breasts from the data collection when setting DRLs.

Of all the parameters studied, breast glandularity was found to have the larger range of values. It was also the one parameter where the indications of the two mammography units considered in this study differed most. Interestingly, indicated values of glandularity are similar for treated and untreated breasts of the same patient, in both mammography units (mean diferences were around 3% for the DS unit and 2% for Essential). Untreated breasts in the treated set of patients had lower mean glandularity values (~58% for the DS unit and ~26% for Essential) than the no treatment set of patients, for both mammography units (~73% for the DS unit and ~46% for Essential). This should be investigated further, using an automated method for independent estimation of glandularity for comparison with other studies.

The main conclusion of this study is that the automatic exposure control systems of the two GE digital mammography units do not appear to detect changes in breast composition when imaging treated breasts.

5.1 Future work

The results obtained are very interesting, but unfortunately restricted to one institution with two mammography units of the same manufacturer. However, the methodology developed here can easily be adapted for a large-scale study in different hospitals, with a more representative variety of digital mammography units. This study could be conducted more thoroughly, with statistical analysis of results for different populations. It would also be interesting to produce a prospective study to compare mammography exams of the same patient, before and after the treatment.

To optimize the treated side detection method, machine learning should be used next. Considering the treated patients with no observable surgical clips on their exams, further parameters should be included in the machine learning process, like the difference in area and thickness between images of the same view of a treated and untreated breast.

An automated method for independent estimation of the glandularity would have been a valuable tool in this study, given the differences in glandularity values indicated by these two mammography units. There was no time to implement it in this work, but an independent estimate of glandularity would certainly be necessary for a large-scale study involving many mammography units of different manufacturers.

Bibliography

[1] N. Perry, M. Broeders, C. de Wolf, S. Törnberg, R. Holland, and L. von Karsa, "European guidelines for quality assurance in breast cancer screening and diagnosis*"*., *Annals of oncology: official journal of the European Society for Medical Oncology / ESMO*, vol. 19, no. 4. 2006.

[2] IARC, "Globocan," 2012. [Online]. Available: http://gco.iarc.fr/today/fact-sheets-cancers?cancer=15&type=0&sex=2. [Accessed: 31-Jul-2018].

[3] Serviço Nacional de Saúde, "Cancro da Mama," 2017. [Online]. Available: https://www.sns.gov.pt/noticias/2017/12/07/cancro-da-mama/. [Accessed: 31-Jul-2018].

[4] M. Ghoncheh, Z. Pournamdar, and H. Salehiniya, "Incidence and Mortality and Epidemiology of Breast Cancer in the World," *Asian Pacific J. Cancer Prev.*, vol. 17, no. sup3, pp. 43–46, 2016.

[5] Liga Portuguesa Contra o Cancro, "Programa de rastreio do cancro da mama." [Online]. Available: www.ligacontraocancro.pt/serviços/detalhe/url/programa-de-rastreio-de-cancro-da-mama/. [Accessed: 31-Jul-2018].

[6] L. M. Warren, D. R. Dance, and K. C. Young, "Radiation risk of breast screening in England with digital mammography," *Br. J. Radiol.*, vol. 89, no. 1067, p. 20150897, 2016.

[7] R. E. Hendrick, "Radiation Doses and Cancer Risks from Breast Imaging Studies," *Radiology*, vol. 257, no. 1, pp. 246–253, 2010.

[8] M. B. Pereira, J. Oliveira, D. P. Ribeiro, B. Castro, J. Yaphe, and J. C. de Sousa, "Grupo etário e periodicidade recomendados para a mamografia de rastreio: uma revisão sistemática," *Cien. Saude Colet.*, vol. 19, no. 4, pp. 1135–1140, 2014.

[9] I. M. D. Dance, S. Christofides, M. Maidment, "Diagnostic Radiology Physics: A handbook for teachers and students," *Iaea*, p. 710, 2014.

[10] C. S. dos Reis, M. J. Fartaria, J. H. G. Alves, and A. Pascoal, "Portuguese study of mean glandular dose in mammography and comparison with European references," *Radiat. Prot. Dosimetry*, vol. 179, no. 4, pp. 391–399, 2018.

[11] J. I. Ten, E. Vano, R. Sánchez, and J. M. Fernandez-Soto, "Automatic Patient Dose Registry and Clinical Audit on Line for Mammography.," *Radiat. Prot. Dosimetry*, vol. 165, no. 1, pp. 346–349, 2015.

[12] K. C. Young and J. M. Oduko, "Radiation doses received in the United Kingdom

breast screening programme in 2010 to 2012," *Br. J. Radiol.*, vol. 89, no. 1058, 2016.

[13] J. V. Nguyen, M. B. Williams, J. T. Patrie, and J. A. Harvey, "Do women with dense breasts have higher radiation dose during screening mammography?," *Breast J.*, vol. 24, no. 1, pp. 35–40, 2018.

[14] C. E. Mercer *et al.*, "A 6-year study of mammographic compression force: Practitioner variability within and between screening sites," *Radiography*, vol. 21, no. 1, pp. 68–73, 2015.

[15] C. E. Mercer, P. Hogg, R. Lawson, J. Diffey, and E. R. E. Denton, "Practitioner compression force variability in mammography: A preliminary study," *Br. J. Radiol.*, vol. 86, no. 1022, 2013.

[16] K. Holland, I. Sechopoulos, R. M. Mann, G. J. den Heeten, C. H. van Gils, and N. Karssemeijer, "Influence of breast compression pressure on the performance of population-based mammography screening," *Breast Cancer Res.*, vol. 19, no. 1, pp. 1–8, 2017.

[17] J. T. Bushberg, J. A. Seibert, E. M. Leidholdt, J. M. Boone, and E. J. Goldschmidt, *The Essential Physics of Medical Imaging*, vol. 30, no. 7. 2003.

[18] R. Klausz and N. Shramchenko, "Dose to population as a metric in the design of optimised exposure control in digital mammography," *Radiat. Prot. Dosimetry*, vol. 114, no. 1–3, pp. 369–374, 2005.

[19] I. A. E. Agency, "Quality assurance programme for digital mammography," *Young*, pp. 1–195, 2011.

[20] M. J. Yaffe *et al.*, "Comparative performance of modern digital mammography systems in a large breast screening program," *Med. Phys.*, vol. 40, no. 12, 2013.

[21] A. M. Chiarelli *et al.*, "Digital Compared with Screen-Film Mammography: Performance Measures in Concurrent Cohorts within an Organized Breast Screening Program," *Radiology*, vol. 268, no. 3, pp. 684–693, 2013.

[22] "Dicom Standard." [Online]. Available: www.dicomstandard.org/history/. [Accessed: 12-Mar-2018].

[23] NEMA, "DICOM PS3.5 2018d - Data Structures and Encoding." [Online]. Available: http://dicom.nema.org/medical/dicom/current/output/html/part05.html. [Accessed: 23-Oct-2018].

[24] NEMA, "DICOM PS3.6 2018d - Data Dictionary." [Online]. Available: http://dicom.nema.org/medical/dicom/current/output/html/part06.html#PS3.6. [Accessed: 23-Oct-2018].

[25] M. E. Suleiman, P. C. Brennan, and M. F. McEntee, “Mean glandular dose in digital mammography: a dose calculation method comparison,” *J. Med. Imaging*, vol. 4, no. 1, p. 013502, 2017.

[26] D. R. Dance, C. L. Skinner, K. C. Young, J. R. Beckett, and C. J. Kotre, “Additional factors for the estimation of mean glandular breast dose using the UK mammography dosimetry protocol,” *Phys. Med. Biol.*, vol. 45, no. 11, pp. 3225–3240, 2000.

[27] S. Vogiatzi *et al.*, “Guidance on diagnostic reference levels (DRLs) for medical exposures,” *Radiat. Prot. Dosimetry*, vol. 80, no. 1, pp. 1–26, 1998.

[28] EUREF, “Update Digital Mammography Protocol 01-2017.”

[29] S. Exhibit, G. Gennaro, I. Sechopoulos, L. Gallo, and V. Rossetti, “Impact of objective volumetric breast density estimates on mean glandular dose calculations in digital mammography,” 2015.

[30] P. R. Bakic, A.-K. Carton, D. Kontos, C. Zhang, A. B. Troxel, and A. D. A. Maidment, “Breast Percent Density: Estimation on Digital Mammograms and Central Tomosynthesis Projections,” *Radiology*, vol. 252, no. 1, pp. 40–49, 2009.

[31] E. U. Ekpo and M. F. McEntee, “Measurement of breast density with digital breast tomosynthesis-a systematic review,” *Br. J. Radiol.*, vol. 87, no. 1043, pp. 1–9, 2014.

[32] I. Thierry-Chef, S. L. Simon, R. M. Weinstock, D. Kwon, and M. S. Linet, “Reconstruction of Absorbed Doses to Fibroglandular Tissue of the Breast of Women Undergoing Mammography (1960 to the Present),” *Radiat. Res.*, vol. 177, no. 1, pp. 92–108, 2012.

[33] American College of Radiology, “ACR BI-RADS® Mammography,” *ACR BI-RADS® Atlas, Breast Imaging Report. Data Syst.*, pp. 121–140, 2013.

[34] A. Oliver *et al.*, “A statistical approach for breast density segmentation,” *J. Digit. Imaging*, vol. 23, no. 5, pp. 527–537, 2010.

[35] C. Nickson *et al.*, “AutoDensity: an automated method to measure mammographic breast density that predicts breast cancer risk and screening outcomes.,” *Breast Cancer Res.*, vol. 15, no. 5, p. R80, 2013.

[36] S. Ciatto *et al.*, “A first evaluation of breast radiological density assessment by QUANTRA software as compared to visual classification,” *Breast*, vol. 21, no. 4, pp. 503–506, 2012.

[37] R. Highnam, “Patient-Specific Radiation Dose Estimation in Breast Cancer Screening Keeping Patients Safe and Informed,” no. 1, 2017.

[38] Volpara® Solutions, “Volpara® density.” [Online]. Available: https://volparasolutions.com/our-products/volparadensity/.

[39] GEHealthcare, *Senographe Essential Operator Manual*. 2010.

[40] “IPO-Porto: Clínica da mama.” [Online]. Available: http://www.ipoporto.pt/clinica/mama/. [Accessed: 11-Oct-2018].

[41] H. G. Kaplan, J. A. Malmgren, M. K. Atwood, and G. S. Calip, “Effect of treatment and mammography detection on breast cancer survival over time: 1990-2007,” *Cancer*, vol. 121, no. 15, pp. 2553–2561, 2015.

[42] M. Çakir *et al.*, “Comparison of mammography sensitivity after reduction mammoplasty targeting the glandular and fat tissue,” *Turkish J. Surg.*, vol. 31, no. 2, pp. 68–71, 2015.

[43] C. T. Rueden *et al.*, “ImageJ2: ImageJ for the next generation of scientific image data,” *BMC Bioinformatics*, vol. 18, no. 1, pp. 1–26, 2017.

[44] “Python.” [Online]. Available: https://python.org/. [Accessed: 08-Sep-2017].

[45] *Supplement to the European Guidelines fourth edition*, no. 1.0. 2011.

[46] E. J. Goldstraw, I. Castellano, S. Ashley, and S. Allen, “The effect of Premium View post-processing software on digital mammographic reporting,” *Br. J. Radiol.*, vol. 83, no. 986, pp. 122–128, 2010.

[47] “Pydicom.” [Online]. Available: https://pydicom.github.io/pydicom/stable/index.html. [Accessed: 08-Sep-2017].

[48] “Scikit-image: Image Processing in Python.” [Online]. Available: http://scikit-image.org/. [Accessed: 11-Sep-2017].

[49] “Edge Operators.” [Online]. Available: http://scikit-image.org/docs/dev/auto_examples/edges/plot_edge_filter.html#sphx-glr-auto-examples-edges-plot-edge-filter-py. [Accessed: 11-Sep-2017].

[50] “Canny Edge Detector.” [Online]. Available: http://scikit-image.org/docs/dev/auto_examples/edges/plot_canny.html. [Accessed: 11-Sep-2017].

[51] “Frangi Filter.” [Online]. Available: http://scikit-image.org/docs/dev/auto_examples/filters/plot_frangi.html#sphx-glr-auto-examples-filters-plot-frangi-py. [Accessed: 11-Sep-2017].

[52] “Find Contours.” [Online]. Available: http://scikit-image.org/docs/dev/api/skimage.measure.html#skimage.measure.find_contours. [Accessed: 12-Sep-2017].

[53] “Scikit-image (image processing in python); Module: filters.” [Online]. Available:

http://scikit-image.org/docs/dev/api/skimage.filters.html#skimage.filters.threshold_li. [Accessed: 08-Feb-2018].

[54] "Regionprops." [Online]. Available: http://scikit-image.org/docs/dev/api/skimage.measure.html#skimage.measure.regionprops. [Accessed: 18-Jul-2018].

[55] GE Healthcare, *Senographe DS: DICOM Conformance Statement V3.0*. 2010.

[56] GE Healthcare, *Senographe Essential Version ADS 55 . 30 and ADS 55 . 40 DICOM CONFORMANCE STATEMENT*. 2014.

[57] I. A. Tsalafoutas and S. Kanellopoulou, "Average glandular dose in digital mammography: comparison of real time calculations with retrospective calculations using the dicom header data," *Phys. Medica*, vol. 30, p. e41, 2014.

[58] R. T. Senie, P. P. Rosen, M. L. Lesser, R. E. Snyder, D. Schottenfeld, and K. Duthie, "Epidemiology of breast carcinoma II: Factors related to the predominance of left-sided disease," *Cancer*, vol. 46, no. 7, pp. 1705–1713, 1980.

[59] J. E. Baek, B. J. Kang, S. H. Kim, and H. S. Lee, "Radiation dose affected by mammographic composition and breast size: First application of a radiation dose management system for full-field digital mammography in Korean women," *World J. Surg. Oncol.*, vol. 15, no. 1, pp. 1–9, 2017.

[60] C. N. Damases, P. C. Brennan, and M. F. McEntee, "Mammographic density measurements are not affected by mammography system," *J. Med. Imaging*, vol. 2, no. 1, p. 015501, 2015.

[61] R. Heidsieck, "Method of calibrating a radiological system and of measuring the equivalent thickness of an object.," 5,528,649, 1996.

Appendices

A. Breast thickness and composition compensation

Exposure factors chosen by the AEC system and the dose indexes for images of PMMA plates of 20mm thickness are considered. This is repeated for other thicknesses of PMMA, namely 30, 40, 45, 50, 60 and 70mm.

An aluminium object positioned between the first two PMMA plates and with different amounts of PMMA plates and spacers above it, as exemplified in Figure A.1, was considered.

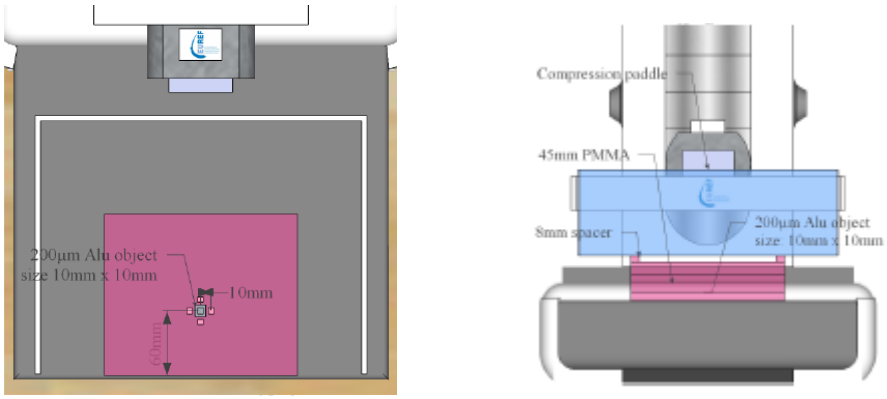


Figure A.1 – Setup for the dosimetry measurements, top view and front view. [45]

A small ROI is drawn on the imaged aluminium object and other 4 ROIs are drawn on the background as the scheme in Figure A.2 shows.

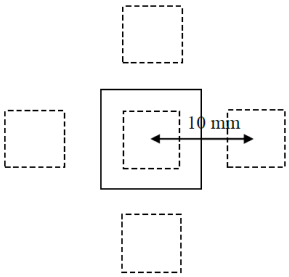


Figure A.2 – Scheme of regions of interest (ROI) considered to calculate the signal difference to noise ratio (SDNR).[45]

The MPV and SD of the background are estimated as:

$$SD_{background} = \frac{\sum_1^4 SD(ROI_n)}{4} \quad \text{(A.1)}$$

$$PV_{background} = \frac{\sum_1^4 PV(ROI_n)}{4} \quad \text{(A.2)}$$

The signal difference to noise ratio (SDNR) of the aluminium object can be calculated as:

$$SDNR = \frac{PV_{signal} - PV_{background}}{\sqrt{\frac{SD_{signal}{}^2 + SD_{background}{}^2}{2}}} \quad (A.3)$$

The determination of the limiting values of SDNR is done to apply the standards of the European protocol. One of the methods described in the EUREF supplement [45] determines the limiting values of SDNR for an attenuation equivalent to 50mm PMMA. Table A.1 presents the SDNR for the raw and processed images, related to each PMMA thickness.

Table A.1 - SDNR for different thicknesses of PMMA

Thickness of PMMA (mm)	SDNR	
	RAW	**PROC**
20	-18,5	18,3
30	-14,3	14,3
40	-11,6	11,6
45	-10,3	10,3
50	-10,8	10,8
60	-8,4	8,4
70	-6,6	6,6

A.1 Dosimetry

Using the same images and considering the organ dose value displayed and presented on the DICOM header, it was evaluated if these values did not exceed the achievable and acceptable levels established by EUREF.

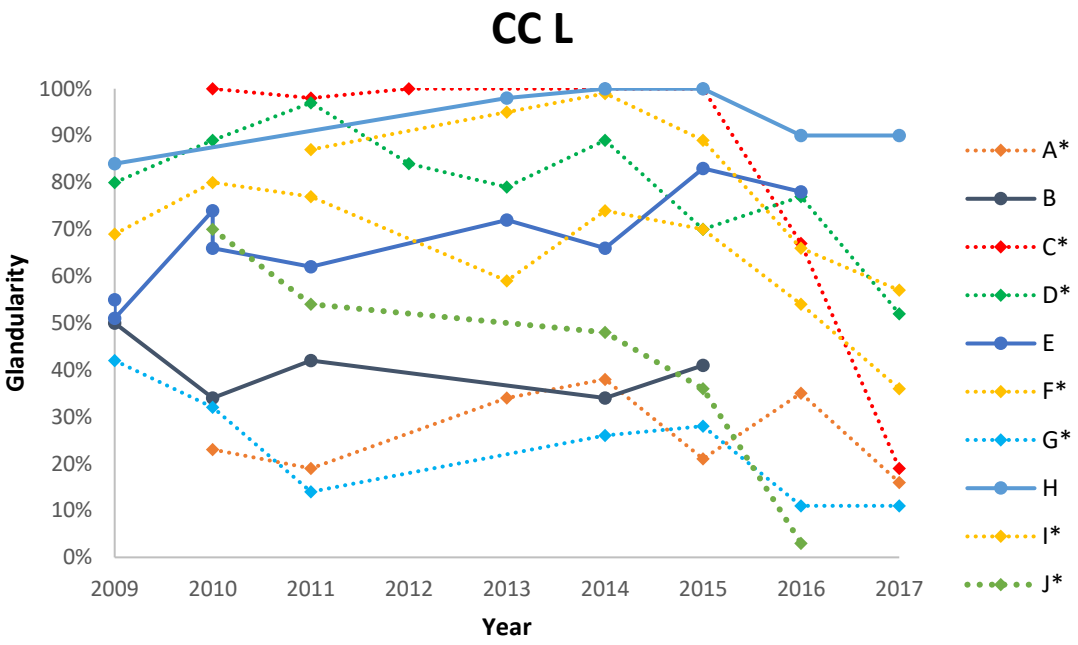
Exposure factors chosen by the AEC and the dose indexes for the dosimetry images are presented in Table A.2. All images were taken in the STD AOP mode, with 50N of compression force, and the aluminium object positioned as indicated in C.1. It can be seen that the dose levels are below the acceptable and achievable levels established by EUREF [28].

Table A.2 – Exposure and dose parameters of the dosimetry images. Highlighted are the images's information and the corresponding acceptable and achievable levels established by EUREF [1].

PMMA thickness (mm)	Spacers' thickness (mm)	mAs	kVp	Target/ Filter	Thickness (mm)	Organ Dose (mGy)	Acceptable level	Achievable level
20	1	23	26	Mo/Mo	21	0,565	1,2	0,8
30	2	36	26	Mo/Rh	33	0,738	1,5	1,0
40	5	39	29	Rh/Rh	45	0,968	2,0	1,6
45	9	45	29	Rh/Rh	52	1,02	2,5	2,0
50	10	68	29	Rh/Rh	59	1,448	3,0	2,4
60	15	73	30	Rh/Rh	73	1,594	4,5	3,6
70	20	91	30	Rh/Rh	88	1,842	6,5	5,1

B. Glandularity through the years

The glandularity estimated by the systems are plotted with the year of each exam, in the next figures (for each view and all patients). Dashed lines refer to patients who had exams performed in Senographe Essential in 2016 and/or 2017.



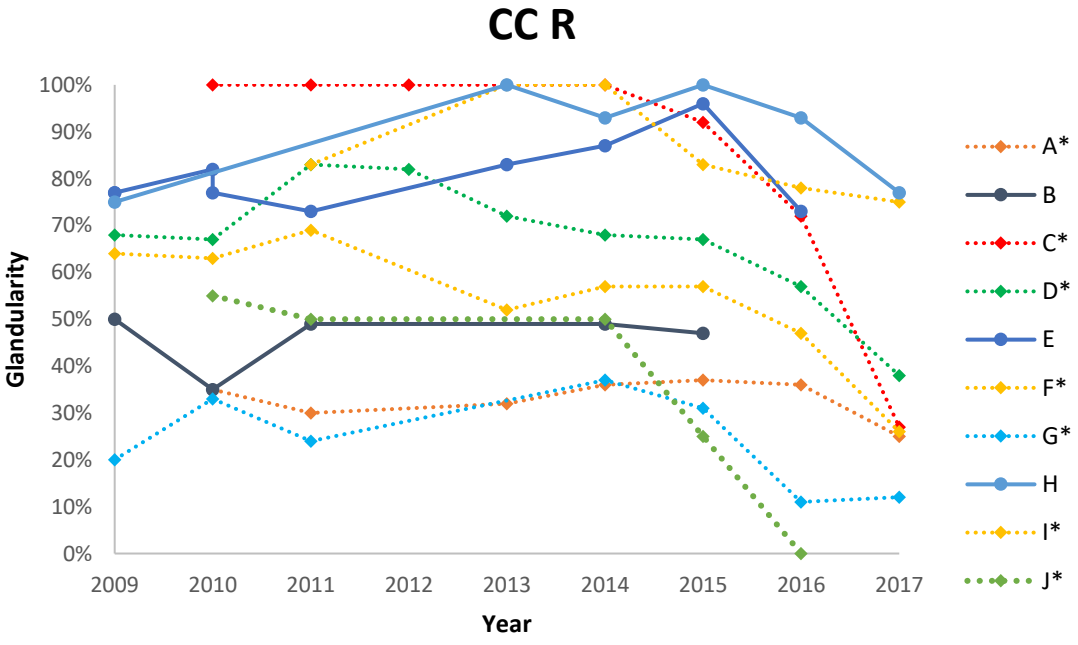


Figure B.1 – glandularity estimated by the systems through the years (CC views).

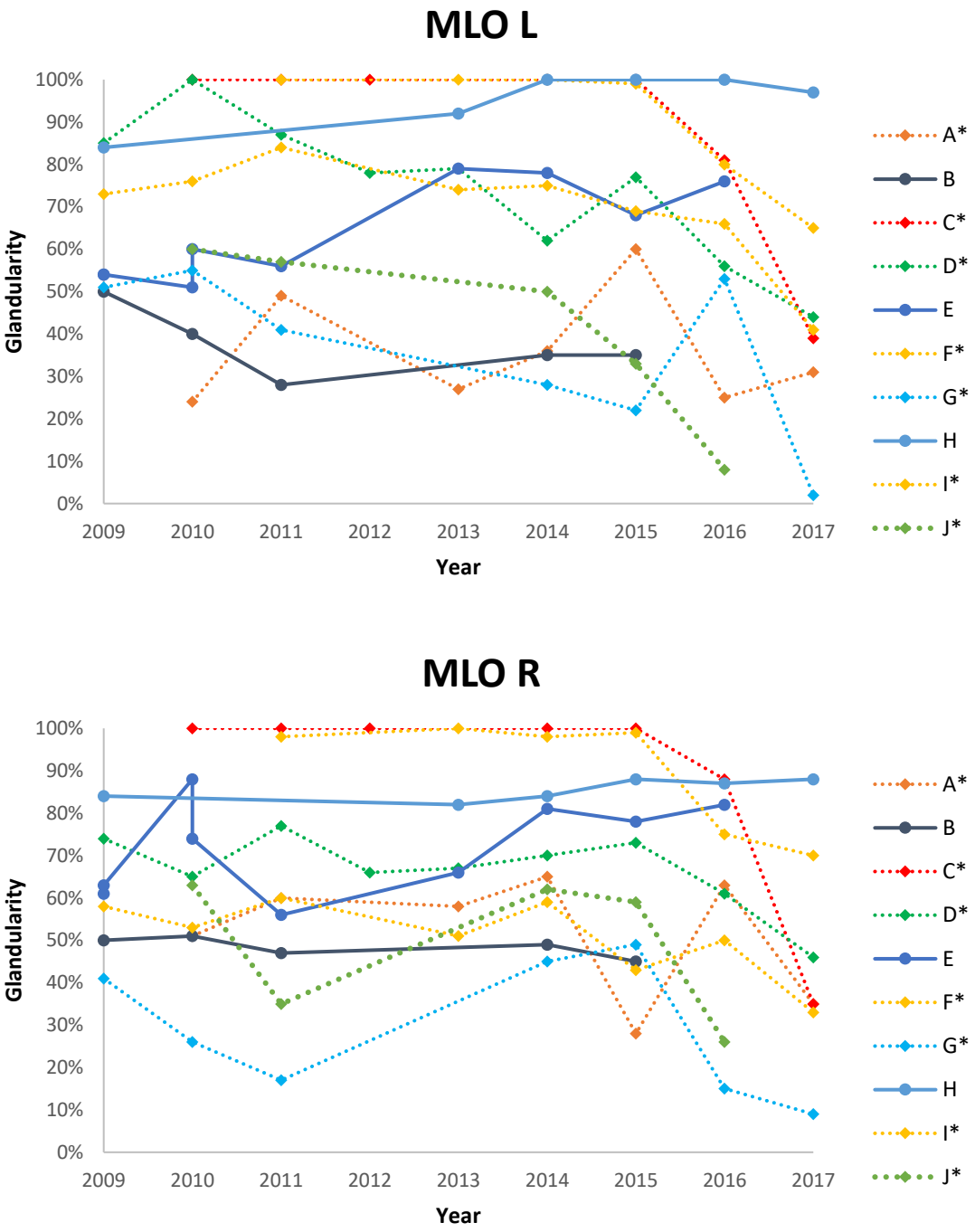


Figure B.2 - glandularity estimated by the systems through the years (MLO views).

It is here clearly visible the bigger variation if we consider the two mammography units, as we found the two systems yield different glandulairty values.

C. Dance's factors for MGD estimation

In this appendix are included the tables for the s, g and c factors for the estimation of the mean glandular dose, as established in the European guidelines [1][45].

Table C.1 – s factors for the typical target and filter materials of the GE mammography systems.

Target material	Filter material	Filter thickness	s-factors
Mo	Mo	30	1.000
Mo	Rh	25	1.017
Rh	Rh	25	1.061

Table C.2 – Typical HVL values ((±0.02mm, of Aluminimum)) for the different tube voltages and target filter combinations, used by the GE systems. Highlighted values were obtained by linear interpolation.

kV	**MoMo**	**MoRh**	**RhRh**
25	0,32	0,38	0,37
26	0,33	0,39	0,39
27	0,34	0,41	0,40
28	0,35	0,42	0,42
29	0,36	0,43	0,43
30	0,37	0,44	0,44
31	0,38	0,45	0,45
34	0,40	0,47	0,47

Table C.3 - g-factors for different breast thickness and HVL. Highlighted in yellow are the additional values included in the 2011 EUREF supplement.

HVL	Breast thickness (cm)									
mm Al	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*
0.30	0.390	0.274	0.207	0.164	0.135	0.114	0.098	0.0859	0.0763	0.0687
0.35	0.433	0.309	0.235	0.187	0.154	0.13	0.112	0.0981	0.0873	0.0786
0.40	0.473	0.342	0.261	0.209	0.172	0.145	0.126	0.1106	0.0986	0.0887
0.45	0.509	0.374	0.289	0.232	0.192	0.163	0.14	0.1233	0.1096	0.0988
0.50	0.543	0.406	0.318	0.258	0.214	0.177	0.154	0.1357	0.1207	0.1088
0.55	0.573	0.437	0.346	0.287	0.236	0.202	0.175	0.1543	0.1375	0.1240
0.60	0.587	0.466	0.374	0.31	0.261	0.224	0.195	0.1723	0.1540	0.1385
0.65	0.622	0.491	0.399	0.332	0.282	0.244	0.212	0.1879	0.1682	0.1520
0.70	0.644	0.514	0.421	0.352	0.300	0.259	0.227	0.2017	0.1809	0.1638
0.75	0.663	0.535	0.441	0.371	0.317	0.274	0.241	0.2143	0.1926	0.1746
0.80	0.682	0.555	0.460	0.389	0.333	0.289	0.254	0.2270	0.2044	0.1856

Table C.4- c-factors for average breasts in the 50-64 age group. Highlighted in yellow are the additional values included in the 2011 EUREF supplement.

Breast thickn	Gland	HVL (mm Al)										
(cm)	%	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80
2	100	0.885	0.891	0.9	0.905	0.91	0.914	0.919	0.923	0.928	0.932	0.936
3	72	0.925	0.929	0.931	0.933	0.937	0.94	0.941	0.947	0.950	0.953	0.956
4	50	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	33	1.086	1.082	1.081	1.078	1.075	1.071	1.069	1.064	1.060	1.057	1.053
6	21	1.164	1.160	1.151	1.15	1.144	1.139	1.134	1.124	1.117	1.111	1.103
7	12	1.232	1.225	1.214	1.208	1.204	1.196	1.188	1.176	1.167	1.157	1.147
8	7	1.275	1.265	1.257	1.254	1.247	1.237	1.227	1.213	1.202	1.191	1.179
9	4	1.299	1.292	1.282	1.275	1.27	1.26	1.249	1.236	1.225	1.213	1.200
10	3	1.307	1.298	1.29	1.286	1.283	1.272	1.261	1.248	1.236	1.224	1.211
11	3	1.306	1.301	1.294	1.291	1.283	1.274	1.266	1.251	1.240	1.228	1.215

Table C.5- c-factors for average breasts in the 40-49 age group. Highlighted in yellow are the additional values included in the 2011 EUREF supplement.

Breast thickn	Gland	HVL (mm Al)										
(cm)	%	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80
2	100	0.885	0.891	0.9	0.905	0.91	0.914	0.919	0.923	0.928	0.932	0.936
3	82	0.894	0.898	0.903	0.906	0.911	0.915	0.918	0.924	0.928	0.933	0.937
4	65	0.940	0.943	0.945	0.947	0.948	0.952	0.955	0.956	0.959	0.961	0.964
5	49	1.005	1.005	1.005	1.004	1.004	1.004	1.004	1.004	1.003	1.003	1.003
6	35	1.080	1.078	1.074	1.074	1.071	1.068	1.066	1.061	1.058	1.055	1.051
7	24	1.152	1.147	1.141	1.138	1.135	1.130	1.127	1.117	1.111	1.105	1.098
8	14	1.220	1.213	1.206	1.205	1.199	1.190	1.183	1.172	1.163	1.154	1.145
9	8	1.270	1.264	1.254	1.248	1.244	1.235	1.225	1.214	1.204	1.193	1.181
10	5	1.295	1.287	1.279	1.275	1.272	1.262	1.251	1.238	1.227	1.215	1.203
11	5	1.294	1.290	1.283	1.281	1.273	1.264	1.256	1.242	1.232	1.220	1.208

D. Output

The measured output values (mGy/mAs) are presented in this appendix for each mammography system and through the years.

D.1 Senographe DS

Table D.1– Output values for the Senographe DS system. Highlighed in yellow are values calculated through interpolation.

	MoMo	MoRh	RhRh	MoMo	MoRh	RhRh	MoMo	MoRh	RhRh
kV	**2009**	**values of 2008**		**2010**			**2011**		
25,0	23,8	----	----	24,2	----	----	24,6		
26,0	27,2	22,6	----	27,8	23,3	----	28,2	23,6	
27,0	30,7	25,9	----	31,6	26,6	----	32,2	27,0	
28,0	34,5	29,2	27,0	35,6	30,1	29,1	35,9	30,4	29,2
29,0	38,4	32,7	30,0	39,7	33,8	32,3			32,4
30,0	----	----	33,3	----	----	35,9			36,0
31,0	----	----	36,6	----	----	39,7			39,7
kV	**2012**			**2013**			**2014**		
25,0	25,0	----	----	25,0	----	----	26,8	----	----
26,0	28,5	23,9	----	29,5	24,0	----	30,4	25,5	----
27,0	32,5	27,3	----	32,1	27,0	----	34,3	29,0	----
28,0	36,3	30,7	29,2	35,8	30,4	27,7	38,3	32,6	29,6
29,0	----	----	32,5	----	----	30,9	----	36,4	32,8
30,0	----	----	36,1	----	----	34,1	----	----	36,3
31,0	----	----	39,7	----	----	37,5	----	----	39,9
kV	**2015**			**2016**			**2017**		
25,0	26,2	----	----	25,6	----	----	26,1	----	----
26,0	29,8	24,0	----	29,2	24,4	----	29,7	24,8	----
27,0	32,3	27,3	----	32,9	27,8	----	33,4	28,1	----
28,0	36,4	30,7	28,7	36,8	31,3	27,9	37,4	31,7	28,3
29,0	40,1	30,6	32,0	40,9	34,9	31,1	41,5	35,3	31,5
30,0	44,2	----	33,7	----	----	34,4	45,7	39,1	34,8
31,0	----	----	37,0	----	----	37,8	----	----	38,3

D.2 Senographe Essential

Table D.2 - Output values for the Senographe Essential system.

	MoMo	MoRh	RhRh	MoMo	MoRh	RhRh
kV	**2017**			**2016**		
25	----	----	----	----	----	----
26	29,4	22,2	----	30,9	23,3	----
27	33,3	25,4	----	34,9	26,6	----
28	37,4	28,8	26,8	39,3	30,2	27,8
29	41,7	32,3	29,9	43,7	33,9	31,1
30	----	35,9	33,3	48,4	37,7	34,6
31	----	----	36,8	----	----	38,1

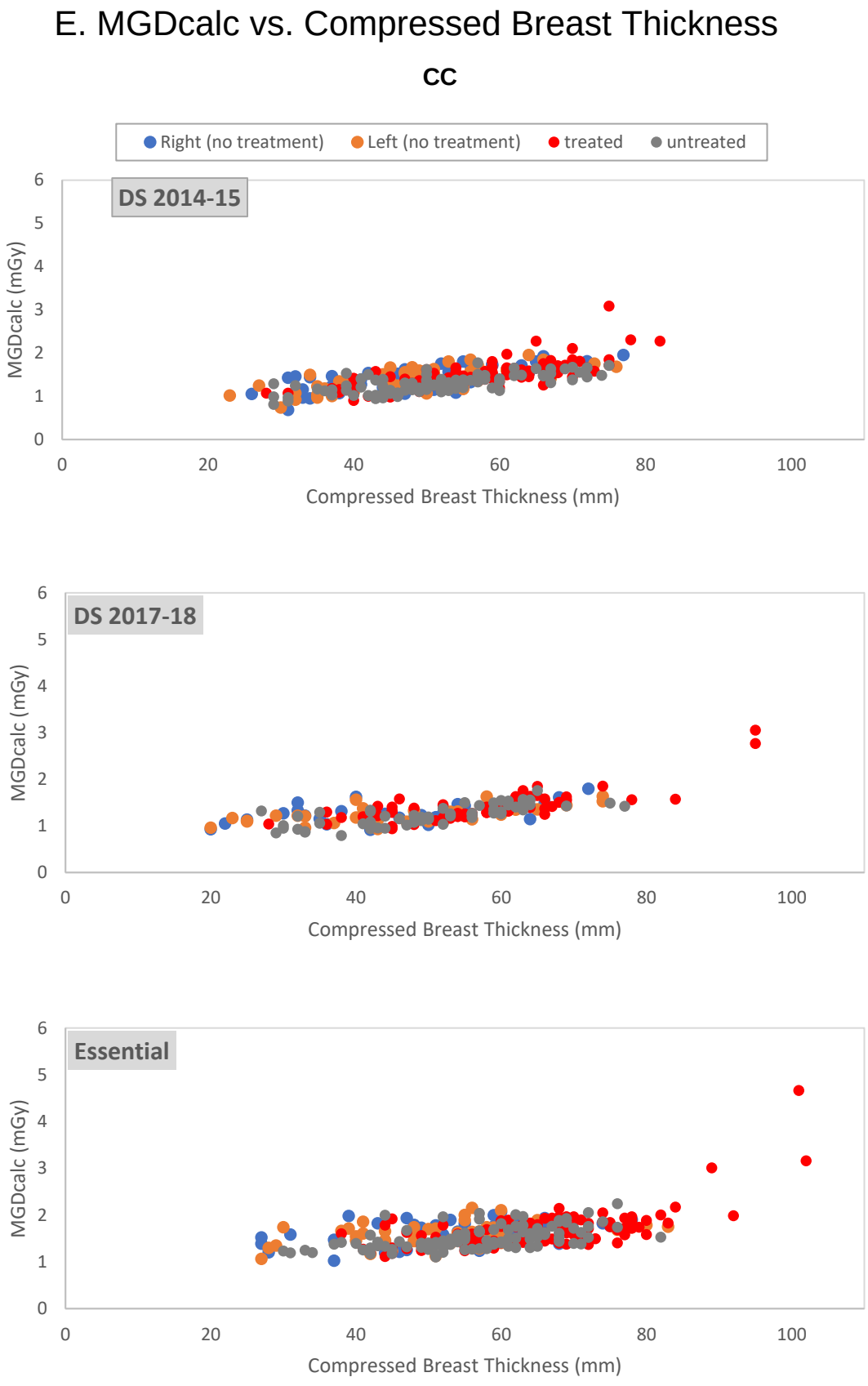


Figure E.1 - MGD, estimated by the EUREF [1] method, vs compressed breast thickness (CC).

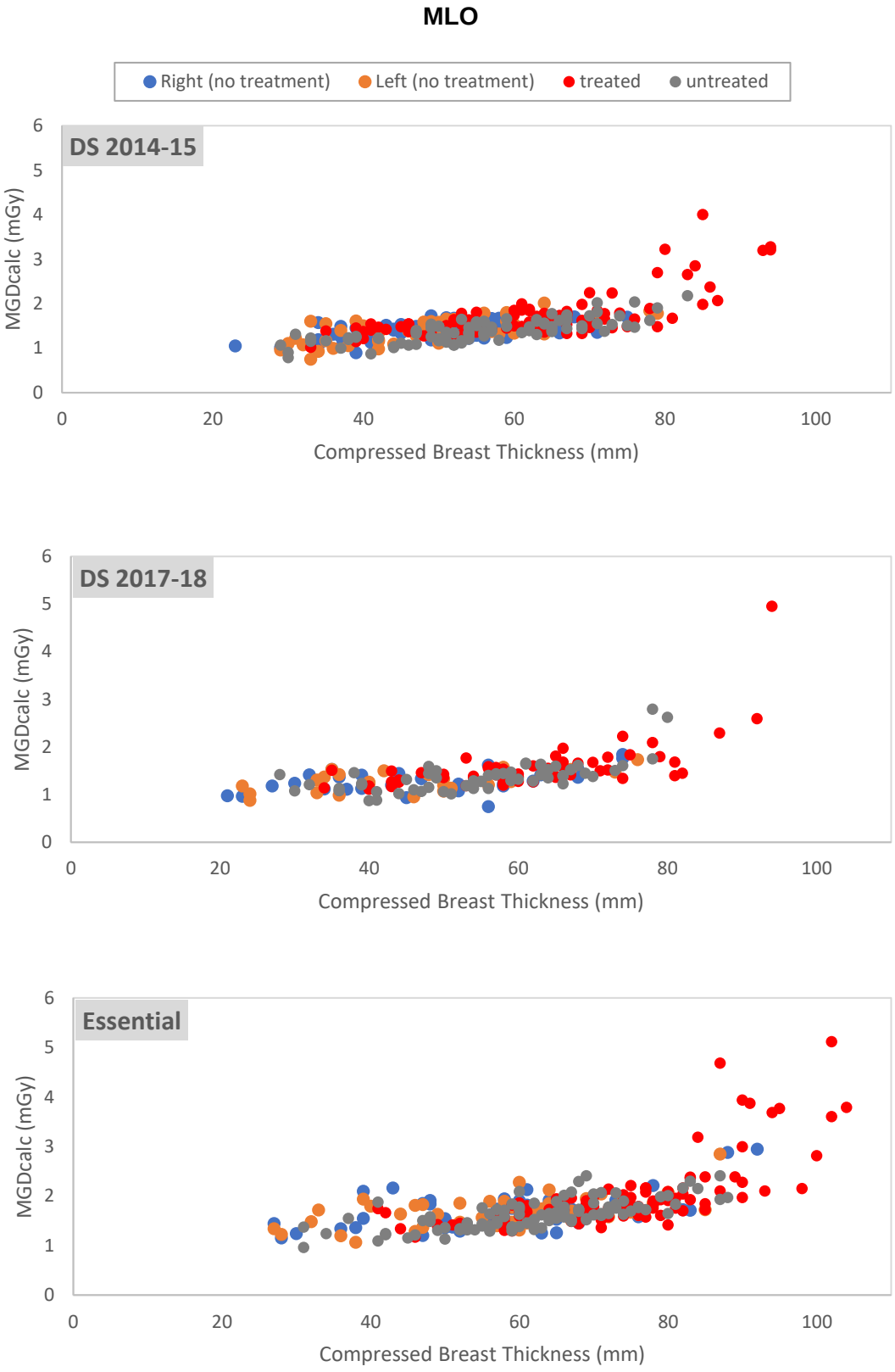


Figure E.2 - MGD, estimated by the EUREF [1] method, vs compressed breast thickness (MLO).

F. Mean values

Table F.1– Mean values of the diferent parameters for the study of oncological effects. (CC) G% refers to the glandularity values indicated by the system.

			CC							
			Thickness (mm)	Force (N)	Area (mm)	G%	MGD (mGy)	MGDcalc (mGy)	ESAK (mGy)	ESAKcalc (mGy)
DS 2014-15	No treatment	Right	49	91	12709	74%	1,16	1,38	5,35	5,51
		Left	49	91	12912	73%	1,18	1,39	5,43	5,58
		Both	49	91	12810	73%	1,17	1,38	5,39	5,54
	Treatement	Treated	57	67	11070	62%	1,28	1,49	6,40	6,57
		Untreated	51	87	13891	58%	1,16	1,29	5,20	5,32
DS 2017-18	No treatment	Right	46	88	12847	73%	1,13	1,28	5,02	4,98
		Left	46	88	13285	73%	1,11	1,24	4,89	4,82
		Both	46	88	13066	73%	1,12	1,26	4,96	4,90
	Treatement	Treated	57	63	10277	54%	1,30	1,41	6,34	6,26
		Untreated	50	82	14382	57%	1,13	1,23	5,07	4,99
Essential	No treatment	Right	53	98	15205	47%	1,52	1,60	6,69	6,73
		Left	53	100	15589	48%	1,52	1,60	6,72	6,75
		Both	53	99	15397	47%	1,52	1,60	6,70	6,74
	Treatement	Treated	65	80	13153	23%	1,72	1,70	8,16	8,25
		Untreated	56	111	18145	26%	1,53	1,51	6,52	6,55

Table F.2- Mean values of the diferent parameters for the study of oncological effects. (MLO) G% refers to the glandularity values indicated by the system.

			MLO							
			Thickness (mm)	Force (N)	Area (mm)	G%	MGD (mGy)	MGDcalc (mGy)	ESAK (mGy)	ESAKcalc (mGy)
DS 2014-15	No treatment	Right	50	92	16859	74%	1,19	1,42	5,60	5,76
		Left	51	98	16739	72%	1,20	1,41	5,65	5,80
		Both	50	95	16799	73%	1,20	1,42	5,63	5,78
	Treatement	Treated	62	78	15424	62%	1,45	1,72	7,92	8,19
		Untreated	55	95	18173	57%	1,23	1,39	5,89	6,03
DS 2017-18	No treatment	Right	48	91	16835	70%	1,14	1,28	5,17	5,13
		Left	48	92	16401	71%	1,18	1,31	5,32	5,26
		Both	48	92	16618	71%	1,16	1,30	5,24	5,19
	Treatement	Treated	61	71	14204	56%	1,42	1,58	7,41	7,35
		Untreated	55	87	18289	55%	1,26	1,38	6,01	5,93
Essential	No treatment	Right	59	108	21521	45%	1,62	1,72	7,70	7,79
		Left	58	109	21989	45%	1,60	1,70	7,56	7,63
		Both	59	109	21755	45%	1,61	1,71	7,63	7,71
	Treatement	Treated	72	103	19164	26%	1,98	1,98	10,30	10,52
		Untreated	62	126	24928	26%	1,63	1,63	7,50	7,57