Master of
Science Thesis
VT2020

Assessment of mammography screening using clinical and virtual data

Hanna Tomic

Supervisors

Predrag Bakic, Anna Bejnö, Magnus Dustler,
Gustav Hellgren, Anders Tingberg

The work has been performed at the
Department of Radiation Physics
Skåne University Hospital, Malmö

Department of Medical Radiation Physics,
Clinical Sciences, Lund
Lund University

Populärvetenskaplig sammanfattning

Bröstcancer är idag en av de vanligaste cancerformerna hos svenska kvinnor; mer än 1400 avlider varje år till följd av sjukdomen. Ett sätt att reducera dödligheten är genom screeningprogram, där intentionen är att detektera brösttumörer i ett tidigt stadie med hjälp av en röntgenundersökning som kallas mammografi. Alla kvinnor i åldrarna 40–74 erbjuds därför mammografi i förebyggande syfte, trots att misstanke om sjukdom inte finns.

I Malmö har det utförts stora studier på att vidareutveckla screening till att även innefatta tre-dimensionell bildtagning av bröstet, så kallad brösttomosyntes. Det krävs dock stora resurser för att utföra kliniska studier för att bevisa att bildtagningsmetoden gör nytta. De faktorer som främst spelar roll är de ekonomiska och etiska aspekterna då det är en majoritet friska kvinnor som bestrålas. Därför utvecklas nu idén om att utföra virtuella kliniska studier (VCT) där bröstanatomin och bildtagningstekniken datorsimuleras. Detta möjliggör virtuella mammografibilder som efterliknar de verkliga och som det går att utföra kliniska studier på. VCT är ett både snabbt och effektivt komplement till kliniska studier, som inte är beroende av riktiga patienter. Det har dock ännu inte implementerats en modell som innefattar tumörers tillväxt. Detta innebär att det i dagsläget inte är möjligt att simulera kliniska studier över tid, då det saknas information om hur de virtuella tumörerna utvecklas.

I denna studie har en initial modell av tumörtillväxten därför tagits fram och utvärderats i en virtuell klinisk miljö. Modellen baseras på tidigare publicerad klinisk data från Malmö gällande tillväxt hos brösttumörer. I studien simulerades 30 bröst med tumörer, vid två olika tidpunkter som motsvarar två screeningtillfällen. Utifrån de virtuella mammografibilderna kunde den tid som det tar för tumören att dubblera sin storlek (dubbleringstid) att mätas.

Resultaten visade att tumörtillväxtmodellen hade god överenstämmelse med tidigare publicerade data. Vidare visade det sig att simuleringen av tumörtillväxten blev lyckad och att det gick att visuellt följa tumörens utveckling i en simulerad röntgenbild. Modellen hade även god noggrannhet när den testades i en virtuell miljö, då det var små skillnader vad beträffar de ursprungliga och de uppmätta värdena på dubbleringstiden. Även radiologens uppskattning av tumörernas storlek sammanföll väl med den sanna storleken.

I framtiden bör fler tumörtyper utvärderas i liknande studier och modellen vidareutvecklas med grund i mer omfattande klinisk data gällande tumörtillväxt. Studiens resultat lägger grunden för ett nytt och mer realistiskt sätt att modellera cancerfall inom virtuella kliniska studier på mammografi och brösttomosyntes.

Abstract

Aim: Virtual clinical trials (VCT) in medical imaging can be used to predict the outcome of clinical trials, by simulating anatomy, imaging methods and image interpretation. VCTs could reduce the cost and duration of clinical trials and their dependence on available patients. This thesis is motivated by the limitations in the field of breast cancer screening. The aim of the project is to expand a specific VCT-software, OpenVCT, to support future VCTs and complement the results from the Malmö Breast Tomosynthesis Screening Trial (MBTST). The virtual patients and tumours currently simulated in OpenVCT need to be further improved and assessed, especially when it comes to their progression over time. The foremost objective of this study was to initiate a simulation of breast tumour growth and to implement growing lesions into virtual breast phantoms and thus allow for the simulation of multiple examinations over time. More specifically, a tumour growth model that is based on the characteristics of the Malmö screening population. A secondary objective was to evaluate the tumour growth model in a virtual clinical environment by estimating tumour volume doubling times (TVDT) from virtual mammograms and comparing with the theoretical values of the model.

Material and Methods: The tumour growth model was based on previous studies of TVDT in breast cancer patients in Malmö, Sweden. A gamma probability distribution was fitted to the existing data and a program was developed that randomly samples a TVDT for a virtual breast cancer patient. Based on this, 30 virtual breasts were simulated using simplified tumour characteristics such as spherical lesions and exponential growth functions. The patient age and TVDT was specific for the Malmö population. Two mammograms, at different time points, were simulated per patient in order to display the tumour growth. TVDTs were then estimated from the mammograms by having a radiologist measure the lesion size. The estimated TVDTs were compared with their corresponding nominal values.

Results: The initial tumour growth model was successfully implemented, and virtual mammograms were simulated for the 30 patients, depicting tumour growth. The model was estimated to have a mean TVDT of 297 $\pm$ 169 days, whereas the sampled virtual patient cohort had 322 $\pm$ 217 days. The estimated TVDT from the simulated mammograms had a mean of 306 $\pm$ 209 days. The data sets were found to originate from the same distribution as no significant difference was found between them ($p>0.54$). However, it was observed that the median difference between the sampled and estimated TVDTs was 12 days (IQR = 20.75) and significantly larger than zero ($p<0.01$). The mean difference between the sampled and estimated TVDTs was 16 $\pm$ 57 days. Median differences between the other data sets showed no significant distinction from zero ($p>0.64$).

Conclusion: The initial tumour growth model displayed high accuracy and reliability when used in a possible virtual clinical trial and showed potential for further development.

Keywords: tumour growth, virtual clinical trials, mammography, tumour volume doubling time, simulations

Abbreviations

3D – Three-dimensional
AEC – Automatic Exposure Control
AUC – Area Under the Curve
CC – Craniocaudal
DBT – Digital Breast Tomosynthesis
DCIS – Ductal Carcinoma In Situ
DM – Digital Mammography
ECDF – Empirical Cumulative Distribution Function
FE – Finite Element
FOV – Field of View
GPU – Graphics Processing Unit
IDC – Invasive Ductal Carcinoma
ILC – Invasive Lobular Carcinoma
kV – Kilovoltage
LCIS – Lobular Carcinoma In Situ
mA – Milliampere
MBTST – Malmö Breast Tomosynthesis Screening Trial
MLO – Mediolateral Oblique
SDD – Source to Detector Distance
TVDT – Tumour Volume Doubling Time
UOQ – Upper Outer Quadrant
VCT – Virtual Clinical Trial
XML – Extensible Markup Language

Contents

1 Introduction

Breast cancer is the most common type of cancer in Swedish women, accounting for almost 9000 new cases [1] and 1400 deaths each year [2]. A way of reducing mortality is the implementation of adequate screening programs with the use of Digital Mammography, DM. It is already known that mortality has dropped by 16-25% since the introduction of breast cancer screening in Sweden [3], as tumours can be found at early and more treatable stages. However, there are limitations with DM such as false-positives or missed cancers due to overlapped anatomical structures and anatomical noise, which can either hide or mimic cancer. Recent studies such as the Malmö Breast Tomosynthesis Screening Trial MBTST, one of the largest prospective breast tomosynthesis trials, have shown that Digital Breast Tomosynthesis DBT has an increased cancer detection rate, with 34% more cancers found compared to DM [4]. Due to economics, ethics and time aspects large population-based studies regarding DBT are difficult to conduct. This complicates further evaluation of DBT as a supplement to DM in screening programs, which could for example be beneficial for women with mammographically dense breasts where DM might be insufficient [4].

This has led to the motivation of developing Virtual Clinical Trials, VCT, as a method to evaluate breast screening modalities. VCTs are based on simulating anthropomorphic breasts and X-ray modalities to generate virtual mammograms. Components from the now free online source OpenVCT have been used in several studies in an effort to both evaluate and optimize DBT [5–10] and to further develop the software [11–13]. However, software such as OpenVCT do not currently model tumour growth which prevents running VCTs of multiple screening exams spanning a period of several years, which is a crucial factor for these research purposes.

This study aims therefore at initiating a first tumour growth model, specifically one based on the Malmö screening population, and to analyze it in a clinical environment. It will be possible to further assess the use of DBT in Sweden, as well as to assess the use of VCT in comparison with real clinical trials.

Aim

Create a tumour growth model for breast cancer

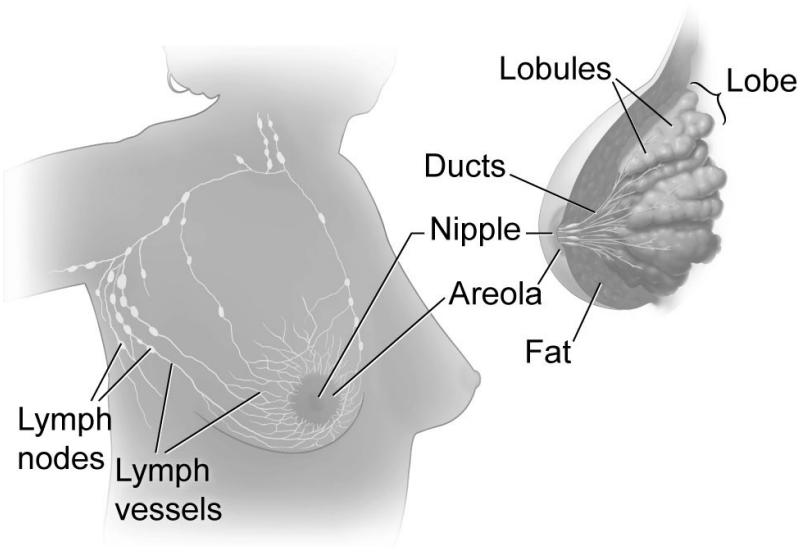
- Generate a pipeline for the model
  - find relevant published data on tumour growth
  - find a way how to simulate growing lesions
- Evaluate the growth model compared to published clinical data
- Evaluate the model in a clinical situation

2 Theory

To design a breast tumour growth model, it is essential to be familiar with the human breast and its corresponding radiographic properties, as well as those of breast tumours. It is also important to understand the structure of VCTs and how such a trial is applied in order to reach high correspondence with real clinical trials and specifically with mammography screening.

2.1 Radiographic appearances of the breast tissue and breast lesions

From a radiological standpoint, breast tissue consists of dense and none-dense areas. The dense compartments comprise glandular tissue, meaning the milk producing lobules and fibrous tissue, a sort of connective tissue [14,15]. The non-dense areas are adipose tissue, another form of connective tissue [14]. Moreover, the breast consists of lactiferous ducts that are branching from the nipple towards the mammary glands as seen in Figure 1 [15,16].



Breast cancer can be divided into two larger subgroups, ductal and lobular breast cancer [16,17].

Figure 1: The basic anatomy of the female breast [18].

Most commonly, breast cancer arises from the lactiferous ducts, where the ductal cancer is either in situ, meaning that it has not spread beyond its origin and is then named ductal carcinoma in situ, DCIS [16,17]. If the ductal cancer has spread into surrounding breast tissue it is classified as an invasive ductal carcinoma (IDC) [17]. Similarly, there are invasive lobular cancers (ILC) and lobular cancers in situ (LCIS) [17]. These pathologies are classified as malignant breast neoplasms and occur on mammograms as masses, architectural distortions, asymmetries or calcifications [19,20].

A selection of mammographic images can be seen in Figure 2 depicting the mentioned pathologies.

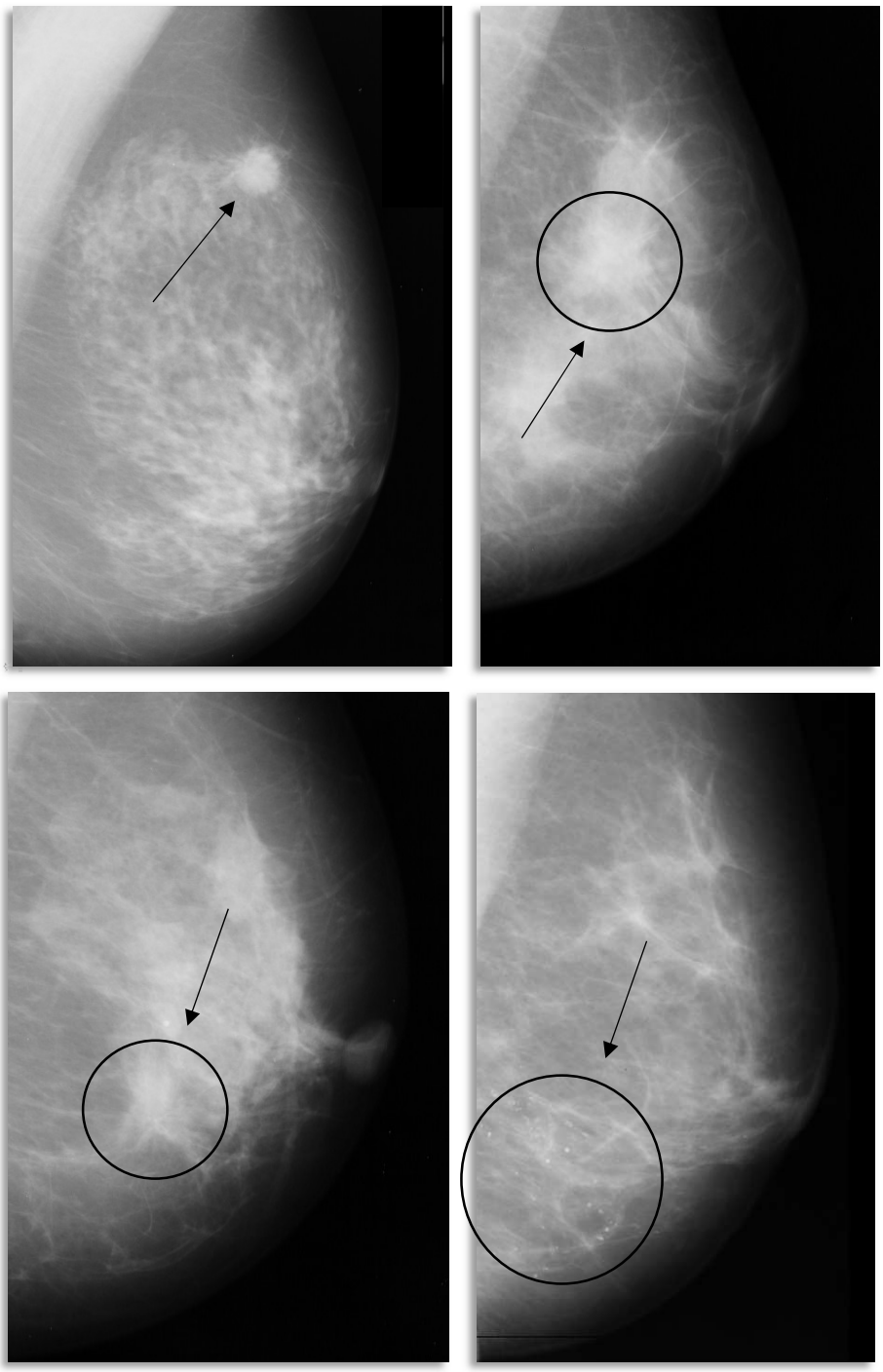


Figure 2: Breast pathologies can appear in mammograms in different ways. The upper left image shows a circumscribed mass that is slightly spiculated, whilst the upper right and lower left show asymmetry and architectural distortion. Calcifications are seen as bright spots in the lower right image. The asymmetric and architectural characteristic are not as obvious for the untrained eye and have a more complex visual variation compared to masses or calcifications. [21]

Masses

Radiographic masses are characterized, amongst other features, by their density and the appearance of their margin. This is related to the risk of malignancy, as high-density lesions can be seen as a sign of malignancy [19,22]. The margin provides information on how well defined the lesion circumference is and is divided into subgroups, in order of risk of malignancy (from low to high): circumscribed (well defined), obscured (partially hidden by fibroglandular tissue), microlobulated (undulated), indistinct (not well defined) and spiculated (radiating spikes) [19,20]. Besides this, masses can be described either as round, oval or irregular and by its location and size [23]. Most commonly lesions are found in the upper outer quadrant (UOQ) close to the axilla as seen in Figure 3, since more fibroglandular tissue is found there [24].

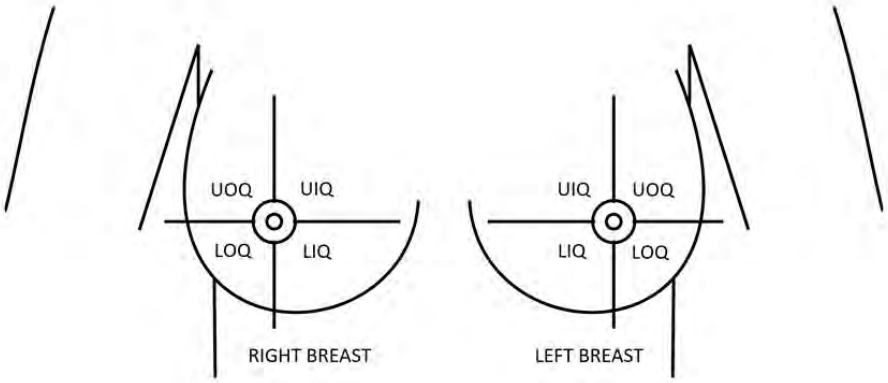


Figure 3: Schematic image of the breast divided into compartments [25]. Lesions are most found in the UOQ [24].

Calcifications

Calcifications can be classified as either macro- or microcalcifications; the latter (typically < 0.5 mm) can be an indication of malignancy [19]. Macrocalcifications however are normally benign and more commonly found in older women [19]. Both calcifications appear as bright dots in mammograms, corresponding to high x-ray attenuation. They are either randomly distributed or found as clusters or linear formations (Figure 2) [19].

Architectural distortions and asymmetries

Architectural distortions and asymmetries are alterations in the breast anatomy that are not related to a distinguishable mass [16,26]. They can be spiculations or asymmetries in the dense tissue when comparing with the opposite breast or the previous mammogram [16,26]. The upper right and lower left image in Figure 2 shows this type of pathology.

2.2 Breast tumour growth rates and current mathematical models

Tumour growth is controlled by complex biological processes that can be manifested in different aspects. Clinically, histopathological findings can be classified in molecular subgroups of breast cancer. They are often related to the status of hormone receptors [27] and growth-promoting proteins such as HER2 and Ki-67 [28,29], as well as with age and histological grade [28,30]. However, mathematical models of tumour growth use simplified tumour characteristics since we do not yet know enough to accurately model these detailed processes.

Available approaches

There are many diverse ways of modelling tumour growth through mathematical principles. Previous studies have demonstrated growth models in terms of proliferating cells [31,32], including both dividing and necrotic cells. These models follow continuous growth models due to the nature of cell proliferation and are described either by Gompertz Law or by logistic powers [32,33]. They are jointly known as continuous functions and are a recurrent factor with regards to modelling tumour growth [28,31–35]. Other approaches of estimating tumour growths include assuming an expanding geometrical model under free boundaries, known as Hele-Shaw models [33,36]. The geometrical shape is often spherical due to the usually well-defined tumour delineation [33]. This model is useful when determining tumour growth rates from looking at radiographically detectable tumours [28,29,37]. This is described in detail in the following subsection.

Many studies advocate other theories consisting of likelihood-based estimating procedures such as Markov Models [38,39]. However, they are only easily executable under rough assumptions that all women are equal regarding their disease [36], making continuous functions more attractive as they are easy and flexible compared to their complex counterparts.

According to a study by Förnvik et al. [28], there was not a significant difference between the exponential growth curve and the Gompertzian model when it came to estimating tumour growth. This indicates that both models are adequate for the task [28]. As a first step this project aims therefore on further examining how exponential functions can be used to model tumour growth rates that can later be applied within the OpenVCT software.

Tumour volume doubling time and exponential functions for describing tumour growth rates

The geometrical model mentioned earlier, Hele-Shaw, can be used when estimating tumour growth rates from radiographic or ultrasonic imaging [28,29,37]. The model is characterized by tumour volume doubling times, TVDT, which is the time it takes for a tumour to double its size and is often assumed to be constant over the course of tumour progression [28].

According to Förnvik et al. [28] tumour growth rates and TVDT can be derived under the assumption that each tumour is spherical. The volume V and diameter d changes over time t, given by

$$V(t) = \frac{4\pi}{3}\left(\frac{d(t)}{2}\right)^3 \quad (1)$$

Assuming exponential tumour growth, the volume at time t with a given growth rate k is as follows

$$V(t) = V_{start} \cdot e^{k \cdot t} \quad (2)$$

Where V_{start} is the initial tumour volume at time t=0. The growth factor k can then be derived by assuming that at the tumour volume doubling time $t = t_d$, the volume $V(t_d)$ is twice the size of the initial volume, which results in the growth factor being represented in terms of tumour volume doubling time

$$2V_{start} = V_{start} \cdot e^{k \cdot t_d} \quad (3)$$

$$k = \frac{\ln 2}{t_d} \quad (4)$$

TVDT (t_d) is derived by combining Equation 1 and 2, resulting in

$$k = \left(3 \cdot \ln\left(\frac{d_1}{d_2}\right)\right)/\Delta t \quad (5)$$

where Δt is the time that has passed between the tumour having diameter d_1 and later d_2. At time $t = t_d$ Equation 4 applies. Combining this with Equation 5 results in TVDT (t_d) being expressed as

$$t_d = \frac{\ln 2 \cdot \Delta t}{3 \cdot (\ln d_1 - \ln d_2)} \quad (6)$$

TVDT can easily be calculated from existing mammograms for those women who have participated in two or more screenings, simply by measuring tumour diameter at each mammogram. Past studies have estimated TVDT by measuring the tumour in the projection where it is most distinguishable and then tracing back to find the same abnormality in previous mammograms [28]. If no tumour could be outlined in the

previous mammogram an initial lesion size was set to 5mm if the tumour was located in non-dense tissue or 10 mm in dense tissue [28].

2.3 Breast cancer screening

In Sweden women aged 40-74 are offered to participate in an optional, free of charge breast screening program, designed to detect cancer in early stages of progression [40]. Those of age 55 years and older are called for screening every 2 years, whilst younger women have a narrower gap of 1.5 years [40,41].

For the routine screening examinations, two digital mammograms are collected for each breast, in craniocaudal (CC) and mediolateral oblique (MLO) view, see Figure 4 [16,41,42]. Before imaging, the breast is compressed in order to reduce the scattered radiation and achieve lower radiation dose to the woman and better image quality [16,42,43]. Another benefit of compressing the breast is that it diminishes the risk for motion artefacts and further decreases the issue of overlapping breast structures [42,43].

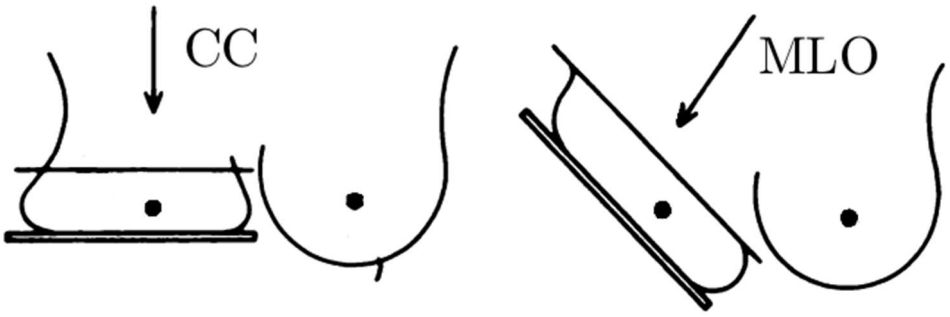


Figure 4: Schematic image of the breast compressed to achieve CC and MLO views in the mammogram. Arrows indicate the X-ray tube angle [44].

Because of the low attenuating soft breast tissue, the X-ray tube voltage (kV) and tube loading (mAs) must be adequately chosen to achieve appropriate image contrast. A high mA and low kV technique, with relatively long exposure times of seconds, was proposed early [45] and is still the gold standard for mammography [42,43,46]. Today however most modalities have an automatic exposure control (AEC) system that self-adjusts the kV and mAs needed to reach a desired detector signal and image quality [42,46,47]. AEC estimates desired parameters by a test exposure that determines the X-ray transmission rate [46]. Generally larger breasts, in terms of compressed breast thickness, and mammographically dense breast require a higher tube voltage and mAs to reach a desired signal level. This in turn results in higher absorbed dose compared to smaller or less dense breasts [46]. The function of the AEC system varies between different manufacturers, mainly whether the desired detector output is reliant on tube current or image noise level [48].

Future prospects of screening with tomosynthesis – and the benefit of VCTs

Another technique on the rise, but not yet adopted in the Swedish screening program, is digital breast tomosynthesis (DBT). A DBT system resembles a conventional mammography machine, using a very similar arrangement of X-ray tube and detector. The difference is that tomosynthesis acquires multiple low dose projections while the tube moves along a limited arc around the breast [46,49].

From a set of DBT projection images, quasi-three-dimensional images of the breast can be reconstructed. Reconstructed images show the breast slices at different depth (distance from the detector), separated by a selected slice distance [49]. The limitation of DBT is the limited angular range compared to computed tomography (CT), which results in anisotropic resolution and artifacts in the direction orthogonal to the detector plane [50].

The advantage of DBT is that it reduces the impact of anatomical noise due to overlapping tissue, which may mask or mimic malignant findings [46,49], and it has also been shown to be beneficial for women with radiographically dense breasts [51,52]. Several studies have shown that including tomosynthesis (DBT together with DM) increases the number of detected cancer cases [53–56].

Combining DM and DBT does however increase the absorbed dose to the breast tissue. Studies have therefore investigated imaging with one-view-only-tomosynthesis [4,57]. Recently the Malmö Breast Tomosynthesis Screening Trial (MBTST) [4] was published, showing that one-view breast tomosynthesis has increased sensitivity compared to two-view digital mammography, at a slightly reduced dose.

There are still challenges that impede on full implementation of DBT as a stand-alone imaging system. Generally, time- and cost effectiveness, health risks from radiation exposure to an overall healthy population and the logistics in producing clinical trials are barriers for further clinical trials [12]. These obstacles may be overcome by the use of VCTs, which enable performing clinical studies using simulated patient populations based on available clinical data assets [12]. However, today's VCT software lacks the ability to simulate tumour growth, which would be a beneficial supplement to the current pipelines and crucial for simulating trials that run over time, with multiple screening rounds of the same virtual patient.

2.4 Simulating breast cancer screening

VCTs aimed at the evaluation of DBT have been carried out previously [5–10]. There are also studies that have focused on creating an adequate pipeline (see Figure 5 below) for VCTs related to breast imaging [11–13] in the OpenVCT software. Common for the studies is that the framework is based on the open source software OpenVCT (OpenVCT), where the concept of the program has been to provide a realistic simulation of breast tissue related to the population of interest. More precisely, to create a model of a deformable breast, with the ability of inserting relevant pathology such as calcifications or spherical lesions [11,12]. The pipeline also consists of simulating the image acquisition, with ray tracing based on given system parameters and post processing systems related to reconstruction models of the projected images [12]. It also accounts for simulating the human observer in terms of the radiologist [12]. The performance of OpenVCT has been compared to clinical outcomes, to verify its credibility to mimic realistic events [10]. The simulated results from OpenVCT showed close agreement with clinical data when comparing the area under the curve (AUC) for DBT against DM [10].

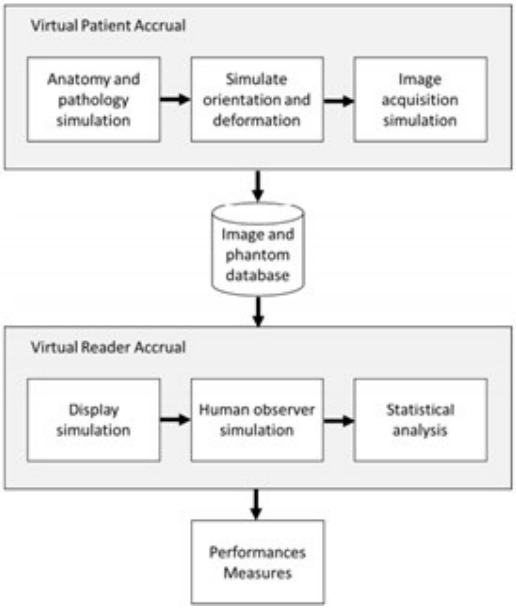


Figure 5: An extract of the pipeline principle for the OpenVCT software [12].

Adding growth models would enable a broader range of VCTs, including studies that span a longer period with multiple screening intervals. Interval cancers, cancers that are found in between screening rounds, can therefore not be included in the trials with the current pipeline. This likely affects the sensitivity, specificity and number of detected cancers compared to a real clinical trial since interval cancers can be regarded as a false-negative report [4,58]. Previously mentioned ongoing research on DBT also gives another highly motivated reason to continue developing the OpenVCT source to include tumour growth.

2.5 Pipeline for OpenVCT

Breast phantom

Creating an anthropomorphic breast phantom implies that the phantom must be customisable regarding breast density, thickness, volume, as well as skin thickness and ligament size in order to represent patient variations seen in clinical studies. The breast phantoms used in the OpenVCT software are created through a quick and efficient method based on recursive partitioning of the phantom volume using octrees, which divides the volume in eight parts through a tree data structure [11,13]. OpenVCT can simulate skin, adipose tissue, fibrous ligaments and dense fibro glandular tissue [11]. The adipose or dense fibro glandular compartments are separated by simulated Cooper's ligaments, that can vary in size and shape [11]. Most structures are simulated using octree-based recursive partitioning, based upon a region growing algorithm [13], as well as beta-functions that can be set to match the parenchymal patterns that are present in clinical cases [11]. Phantoms are later stored in voxel arrays, where each breast compartment is given a certain pixel value depending on the tissue it mimics, which is directly linked to its attenuating properties [11].

In the user interface of OpenVCT, each breast phantom can be designated an individual cup size, skin thickness, ligament thickness and percent dense compartments. Height, weight, birthdate and several other options are also selected by the user. These properties are stored in an XML-file for each phantom and can be changed to suit specific requests. The XML-files can be run through the program in the Windows command line in order to generate the phantom of interest.

Breast phantom deformation

The breast deformation during DM or DBT exam in OpenVCT is simulated by a finite element model (FE), using the open-source software FEBio (University of Utah). FE simulate deformation by dividing the phantom volume into a collection of elements with small volumes. Each element is then deformed using appropriately linearized mechanical equations [59,60].

Simulated breast lesions

The lesions that are simulated need to represent clinically relevant features. For breast cancer, based on what is mentioned in Section 2.1, this can be reduced to clinically significant masses and microcalcifications. In the OpenVCT software the microcalcifications come as either single ones or multiples arranged in clusters with a given radius [11]. In the user interface, the masses can be chosen to be circular, lobular or irregular, which defines the shape of the circumscribed lesion [11].

However, version 0.1 of OpenVCT that was used in this project was limited and did not allow for lobular or irregular masses, hence only circular masses could be used.

The circular masses can differ in radius, position and in the number of shell layers [12]. The position and radius are chosen in the user interface but can simply be altered by editing the XML-file of each phantom. The layers can also be changed by rendering the XML files. The default lesion consists of three concentric shell layers, meaning three spheres with different sizes relative to the nominal input size, stacked upon each other [11]. The layers can be assigned different attenuation properties [11], which helps to blend the lesion into the surrounding simulated tissue and improves the appearance and contrast of synthetic x-ray images [11]. This type of lesion model has been used in previous VCTs of breast lesion detectability in DM and DBT [10] and in human observer studies regarding lesion detectability [61].

The lesions are inserted into the phantom after it has been deformed and the lesion shape is hence not affected by the compression of the breast.

Ray tracing and post processing

In order to create accurate radiation simulations, it is essential to include all the relevant parameters present in the X-ray modality. Previous studies [12] have included essential parameters for X-ray image acquisitioning, a selection of these are presented below in Table 1.

Table 1: Selection of important parameters for X-ray simulations [12].

Image Acquisition
Modality
Anode material
Filter material
Filter thickness (mm)
Detector element size (mm)
Detector element count
Number of acquisitions

The ray tracing algorithm in the OpenVCT software emanates from Siddon's method [12,62] and has been used for determining the ray path in the simulated breast phantoms, but with the modification of using graphic processing units, GPU's, for an accelerated calculation [63]. The choice of Siddon's method compared to more realistic alternatives, such as Monte Carlo, is mainly due to the speed and

effectiveness of Siddon. These synthetic images have been used in previous studies, including reader studies with human readers [61].

The current default modality is Hologic Selenia Dimensions digital breast tomosynthesis system (Hologic Inc., Bedford, MA, USA) [12], which was used for the Pennsylvanian trials [10]. The simulated AEC settings are based on a method by Feng and Sechopoulos [64] and are defined to mimic the mentioned Hologic modality. Scattered radiation is currently not being simulated, however it is possible to incorporate it using appropriate methods [7].

2.6 Statistical tests

The tumour growth model has been evaluated by calculating various statistical tests of the agreement between parameters estimated from clinical and simulated data. The statistical tests that were used are briefly described below.

Pearson coefficient correlation and Student's T-test

Pearson correlation is commonly used when comparing the linear correlation between metric data sets [65]. The correlation can be evaluated through Student's T-test, where the null hypothesis states that the difference between the variables is due randomness [66]. Normally distributed populations are not a requirement for sample sizes N ≥ 20 and independent observations [65]. The correlation coefficient R for two variables X and Y and the sample size n is given by

$$R_{xy} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{\bar{Y}})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \tag{7}$$

where X_i and Y_i is a specific data points and $\bar{X}$ and $\bar{Y}$ are mean values. As R approaches 1 or -1 the variables collectively move towards the linear regression line. $R = 0$ means there is no linear correlation found between the variables.

Shapiro-Wilks test

The null hypothesis of Shapiro-Wilks test is that the data set follows a normal distribution, which is necessary information when deciding which statistical test to apply [67]. Shapiro-Wilks is recommended because of its higher statistical power compared to other tests [68].

Two-sampled Kolmogorov-Smirnov test

The two-sampled Kolmogorov-Smirnov is useful for determining if data sets are from a specific distribution [69] and is appropriate for small samples and non-normal distributions due to being non-metric [70]. It orders data points in an empirical

cumulative distribution function (ECDF) and estimates the distance between the ECDF's that are being compared [69].

Wilcoxon signed-rank test

Wilcoxon is another rank-based test applicable for non-normal distributions [71]. It is a more sensitive test compared to the Wilcoxon rank-sum test (also known was Mann-Whitney U-test) and is the non-parametric equivalent to a paired t-test [71]. Wilcoxon tests the similarity between two data sets, more specifically the null hypothesis states that the median of the difference between two paired data points is zero [71].

3 Material and methods

The intention of this project is to provide an initial tumour growth model that is suitable for women from the Malmö population. Therefore, the basis for the model comes from a study conducted by Förnvik et al., who estimated tumour growth rates from clinical data at the Malmö University hospital [28]. The underlying data for the virtual population comes from the information provided by the MBTST [4]. Both studies were approved by The Regional Ethical Review Board at Lund University (Dnr 2015/105 and Dnr 2009/770).

3.1 Simulation of the tumour growth model

In Förnvik et al. [28], TVDT was calculated for 31 women in who had their screening mammography and follow up cancer treatment at Malmö Hospital. The calculation was performed by using available mammograms, based on the exponential tumour growth model proposed in Section 2.2, see specifically Equation 6. The clinical data regarding TVDT from this study was used for the tumour growth model in the current project.

All simulations regarding the tumour growth model were done in MATLAB (R2019b, The MathWorks, Inc., Natick, MA, USA). A gamma probability distribution was fitted to the clinical TVDT histogram to be able to randomly sample a TVDT from the Malmö population distribution for each virtual patient. The tumour growth factor k was calculated using Equation 4. Initial lesion sizes were randomly selected based on previous findings of clinically detectable sizes in mammograms [28,37,72,73], the range was set between 2-11 mm. The initial tumour volumes were calculated as spheroids (see Equation 1). As a result, tumour diameters could be derived at different times t, by assuming exponential tumour growth (applying Equation 2). A curve was fitted for each tumour growth function that was simulated. A selection of virtual lesions was visualized as growing spheroids through surface plots in MATLAB.

The simulated lesions have no upper limit in size, because of the assumption made regarding sufficient nutrients to the tumour and that no intervention is being done after any inevitable self-detection. Following the example of Förnvik et al. [28], TVDT was assumed constant throughout the tumour progression.

3.2 Evaluating the model in a virtual clinical environment

One of the crucial aims of the project was to assess the model in a virtual clinical situation. This was done by simulating mammograms for a cohort of 30 virtual patients, at two different time points, based on the tumour growth model. The model was evaluated by a clinical radiologist, who measured the size of simulated tumours seen in synthetic mammograms. Based on this, TVDT could be estimated in the same manner as mentioned in Section 2.2, see Equation 6.

To ensure the model plausibility the estimated TVDT from the simulated images was compared with the theoretical model and the clinical data from Förnvik et al. [28]. In a similar approach, the TVDT selected from the theoretical tumour growth model was compared to the clinical TVDT.

3.3 Set up of the virtual clinical assessment

Population simulation

As previously mentioned, the statistics provided for the virtual population is based on the MBTST [4]. Extensive work has already been made, outlining important features of the Malmö population [74]. However, for the intention of simulating a tumour growth model in a virtual screening environment through the OpenVCT source, this study mainly focused on setting the age of the virtual population to mimic the Malmö population. This was done in order to produce virtual mammograms in screening intervals that follow the Swedish standards of 1.5 or 2 years depending on the age of the woman. The MBTST study reported age at diagnosis, divided in age groups of 10 years, for 139 cancer cases [4]. For this study, the age at examination was grouped instead in spans of 5 years.

The breast volume, represented by bra cup size, of virtual patients was selected partially referring to MBTST. In total, 22 virtual patients were assigned a C-cup, five were assigned with B-cup and three with D-cup. However, this was not thoroughly examined in the first assessment of this study and should not be regarded as concluding or absolute matter.

A cohort of 30 virtual patients were selected for the study, comparable to the number from the clinical study by Förnvik et al. [28] with 31 women.

X-ray modality simulation

The default modality built in the OpenVCT software is a simulation of the Selenia Dimensions DBT system (Hologic Inc. Bedford, MA, USA) [12]. Currently Malmö uses Siemens or GE DBT systems in the screening program. For the evaluation of the tumour growth model this was not a crucial aspect. There are plans to implement the simulation of the Siemens DBT system in the software, it was however beyond the scope of this thesis.

The simulated modality assumes poly-energetic X-rays and no scattered radiation. Geometrical parameters that are used in the OpenVCT software to simulate Hologic, such as element size, field of view (FOV), focal point and distance from source to detector (SDD) are presented in Table 2 and can be found in XML-files of the OpenVCT. Because of time limits, all mammograms were simulated in MLO-view only.

Table 2: The applied geometry in the OpenVCT source emulates that of Selenia Dimensions DBT system (Hologic Inc. Nedford MA, USA). The parameters are presented in their corresponding XML-files of the OpenVCT software.

Element size [mm]	FOV [mm^2]	Focal point [mm]	SDD [mm]	Target material	Added filter
0.07	250.88 × 286.72	700	725	Tungsten	Rhodium

Breast phantom generation

Breast phantoms were generated in OpenVCT with default lesions, which were later modified in order to receive the desired lesion size and shape. All breasts were simulated with spherical lesions consisting of three layers, an inner layer being 17% smaller than the nominal size (based on default lesions in the OpenVCT software) and an outer layer with the same diameter as the nominal lesion size. Because of the current limitations in the OpenVCT source, percent dense compartments, ligament thickness and skin thickness were selected arbitrarily from the options that were available. Cup sizes were mainly set to be C-cup (22 patients), with a small fraction of B- (5 patients) and D-cups (3 patients).

Tumour growth and mammography simulation

Each patient was associated with a specific TVDT that was selected from the probability distribution in the tumour growth model and a certain initial tumour size. Two mammograms at different time points were simulated for each patient, resulting in altogether 60 virtual mammograms. The time points of the mammograms were based on the age of the virtual patient. All patients within the screening age but older than 55 had a 2-year interval between their mammograms. Patients aged 40-54 had a 1.5-year interval between their mammograms.

Lesion measurements by radiologist and TVDT estimations

A radiologist from Skåne University Hospital SUS estimated lesion size for the 60 virtual mammograms. Two measurements were made for each lesion and a mean value was calculated. Figure 6 presents an example of how the cursor was placed for each measurement. All measurements were done with ImageJ (Research Services Branch, National Institute of Mental Health, Bethesda, MD, USA), using the built-in measurement tool. Based on the measurements TVDT was calculated using Equation 6.

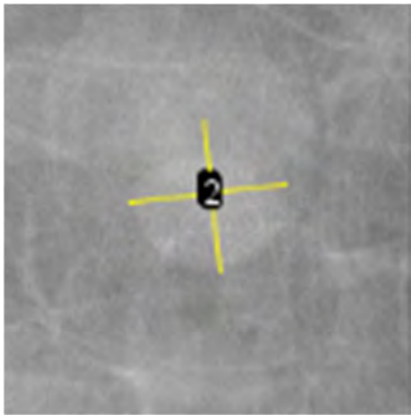


Figure 6: Placement of cursors on the lesion before measuring. Each lesion was measured using two cursors to ensure higher accuracy.

Evaluations

The method can be summarized as resulting in four data sets, see Figure 7 for clarification. The model was compared to the clinical values of Förnvik et al. [28]. The data set for the estimated TVDTs was compared with the TVDTs that were sampled from the model for the 30 virtual patients, as it represents the ground truth. The estimated data was also compared to the clinical data, to sum up the overall uncertainties of this study. Lastly, the modelled and sampled TVDTs were compared with each other to effectively see the error caused by having a smaller sample of 30 virtual patients.

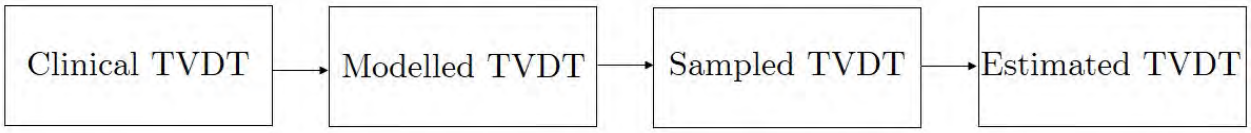


Figure 7: The four data sets of TVDT in chronological order. The modelled TVDT, the sampled TVDT and the estimated TVDT trace back to the clinical TVDT from Förnvik et al [28]. They were therefore compared with the clinical values. The estimated TVDT was also compared with the sampled TVDT since it represents its ground truth.

Tumour growth functions were estimated based on the tumour growth model and based on estimates from measuring lesion sizes in the simulated images, which were compared as well.

Measured lesion size was compared with the nominal lesion size to estimate the performance of the radiologist when measuring virtual mammograms. Comparisons were also made between the two measurements that were made on each lesion (Figure 6), to see if any systematic errors could be found.

3.4 Statistical analysis

To compare the TVDT-data sets seen in Figure 7, it was necessary to establish the nature of their distribution. Hence a Shapiro-Wilks test was used to determine whether the data follows a normal distribution. Based on the results from the Shapiro-Wilks test, both Wilcoxon signed-rank test and the two-sampled Kolmogorov-Smirnov tests were selected for comparing the data sets with each other. The tests were chosen since they are non-parametric and do not assume a certain distribution or shape. All tests were performed in MATLAB (R2019b, The MathWorks, Inc., Natick, MA, USA) and the results are thoroughly presented in Section 4.6.

The correlation between the compared data sets were also determined by using Pearson correlation. It was assumed that it would yield high correlation coefficients. These calculations were done in Excel for Office 365.

The mean of the TVDTs for the four data sets and their standard deviation were also calculated and are presented in the beginning of Section 4.2. This was mainly done to ensure that the mean of the modelled, sampled and estimated TVDT-data sets coincide within the range of the clinical TVDTs.

Moreover, the mean difference between the estimated and sampled TVDTs and the corresponding standard deviation was calculated.

4 Results

4.1 Patient age

For the cohort of the 30 virtual patients, age at examination was selected based on available data regarding cancer incidence from MBTST [4] (Table 3). The table shows the distribution of cancer cases by age at examination and the corresponding number of virtual patients (30 in total in this study) assigned with each age group. Screening intervals for the virtual patients are also presented in Table 3, being 1.5 years for women younger than 55 years and otherwise 2 years.

Table 3: Data based on 139 cancer cases from MBTST and the corresponding breast cancer distribution between age groups. The age of the virtual patients was selected based upon on the distribution. The allocated age group and number of virtual patients are presented together with the corresponding screening interval associated with the given age group, either 1.5 years for women under 55 years or 2 for women older than 55.

Age group at examination (MBTST) [years]	**Cancer distribution [%]**	**Number of virtual patients (rounded to sum up to 30)**	**Screening interval [years]**
70-74	19.4	6	2
65-69	20.9	6	2
60-64	20.1	6	2
55-59	15.1	5	2
50-54	6.5	2	1.5
45-49	12.9	4	1.5
40-44	5.0	1	1.5

Results show that the highest cancer incidence in the MBTST cohort was associated with women between the age of 60 to 74 years, with a proportion of 20% for each age group. Least cancer cases were linked to the youngest age span, 40-44 years, with only 7 cancer incidences out of 139 in total corresponding to 5% of all cancer cases.

4.2 Tumour Volume Doubling Time

General results

The TVDT-model can be seen in Figure 8 as the red gamma distribution function fitted to the clinical histogram from Förnvik et al. [28]. The gamma probability distribution was automatically assigned scale and shape coefficients with 95% confidence interval of 2.88 (1.78, 4.65) and 102.97 (60.98, 173.88). The mean TVDT for the TVDT-model was 297 ± 169 days, compared to published clinical value of 282 ± 164 days from Förnvik et al. [28]. This gives a relative difference of 5% between the mean values of the data sets.

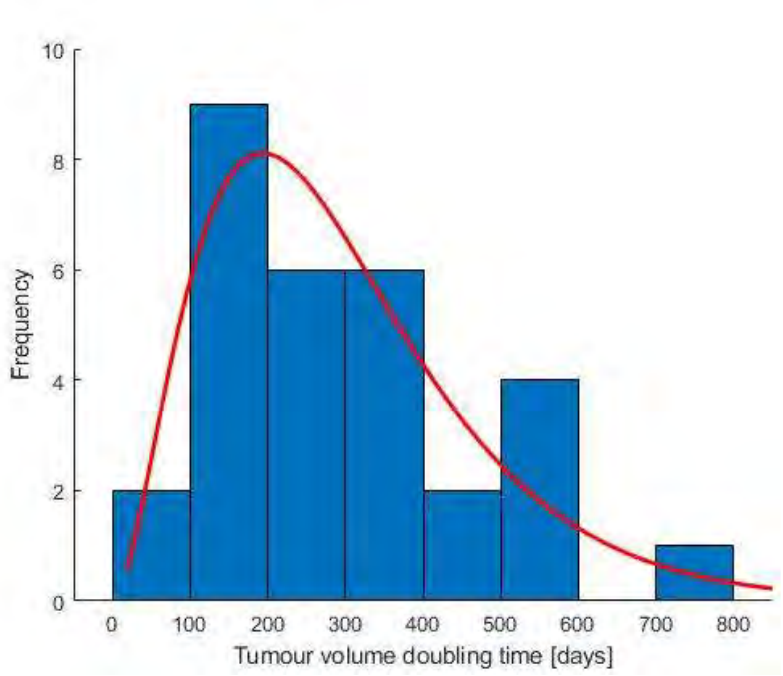


Figure 8: The red function indicates the TVDT-model based on the histogram seen in Förnvik et al. [28] on Tumour Volume Doubling Times for 31 women participating in screening programs in Malmö, Sweden.

Each virtual patient was assigned a theoretical TVDT after randomly sampling values from the TVDT-model. The sampled TVDT-distribution is presented in Figure 9. The TVDT-distribution later estimated from the radiologist's measurements of lesion sizes is presented in Figure 10.

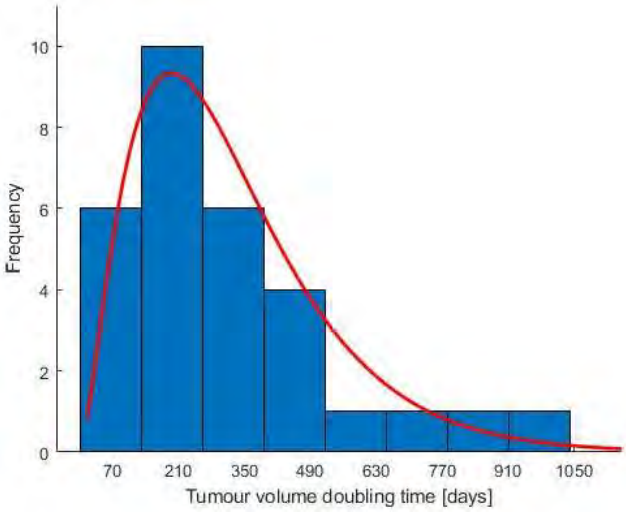


Figure 9: TVDT for the 30 virtual patients, sampled from the gamma distribution model in Figure 3. Mean value was determined to be 322 days.

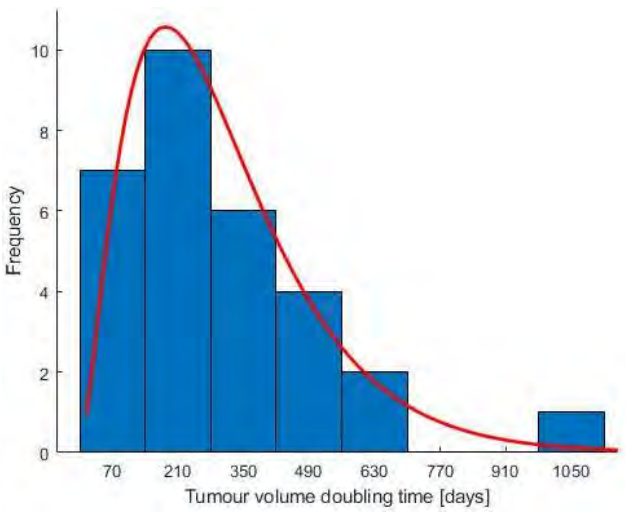


Figure 10: TVDT calculated for 30 virtual patients, based on the radiologist's measurements of simulated lesions size Mean TVDT was estimated to be 306 days.

The mean of the sampled TVDT was 322 ± 217 days. The mean of the estimated TVDT, that was calculated from the lesion size measurements, was 306 ± 209 days. The mean of the estimated TVDT had a relative difference of -5% from the sampled TVDT mean (322 ± 217 days) and 8% from the clinical mean from Förnvik et al. (282 ± 164 days) [28]. From these results it may be concluded that the standard deviations of the data sets overlap.

The mean of the sampled TVDTs (322 ± 217 days) had a relative difference of 9% from the TVDT-model (297 ± 169 days).

Differences between the data sets

The TVDT data sets (see Figure 7 for clarification) were tested for normality using the Shapiro-Wilks test. The clinical data followed a normal distribution ($p=0.053$). However, it was determined that none of the other data sets had normal distributions ($p<0.05$). P-values are represented in Table 4.

Table 4:. Apart from the clinical data set, the null hypothesis of the Shapiro-Wilks was rejected at the 0.05 level for the remaining data sets. The corresponding p-values were well below 0.05.

P-value for the clinical data set	P-value for the modelled data set	P-value for the sampled data set	P-value for the estimated data set
0.053	0.016	0.0016	0.00099

Moreover, no significant difference was found between any of the four data sets according to the two-sampled Kolmogorov-Smirnov test ($p>0.5$).

The Wilcoxon signed-rank test demonstrated that both the modelled, the sampled and the estimated TVDT-data are very similar when compared to the clinical data ($p>0.6$). The sampled and the estimated data did show a significant difference ($p=0.0043$), in the sense that the difference between the data sets was consistent. The median difference between the sampled and estimated data was 12 days (IQR=20.75).

The mean difference between the sampled and estimated data was 16 ± 57 days.

All p-values related to the Kolmogorov-Smirnov and Wilcoxon tests are presented in Table 5.

Table 5: The TVDT data sets were compared through two statistical tests. P-values are presented for the Kolmogorov-Smirnov and the Wilcoxon signed-rank test.

Type of statistical test	**Clinical and modelled data**	**Clinical and estimated data**	**Sampled and estimated data**	**Modelled and sampled data**
Kolmogorov-Smirnov p-value	0.54	0.94	0.54	0.76
Wilcoxon p-value	0.94	0.69	0.0043	0.64

Difference in gamma distribution characteristics between the data sets

The fitted gamma distribution parameters shape and scale were compared between the TVDT-model and the sampled and estimated distributions and are presented in Table 6. The values differed with -13% for the sampled data and with -14% for the estimated data. However, the values fitted within the 95% confidence interval of the nominal data.

Table 6: Gamma distribution parameters for the TVDT-model, the sampled theoretical TVDT distribution and the estimated TVDT distribution. The deviation from the nominal model is given in percent and each parameter is presented with 95% confidence intervals seen in brackets.

Gamma distribution parameters	Nominal data from TVDT-model	Sampled TVDT data	TVDT data estimated from simulations
Shape PD	2.88 (1.78, 4.65)	2.51 -13% (1.56, 4.05)	2.49 -14% (1.55, 4.01)
Scale PD	102.97 (60.98, 173.88)	128.35 +25% (75.77, 217.43)	122.80 +20% (72.48, 208.07)

Correlations between data sets

As assumed, it was found that the estimated TVDT had a positive correlation (R=0.97) to the sampled TVDT that was assigned for the virtual patients, see Figure 11.

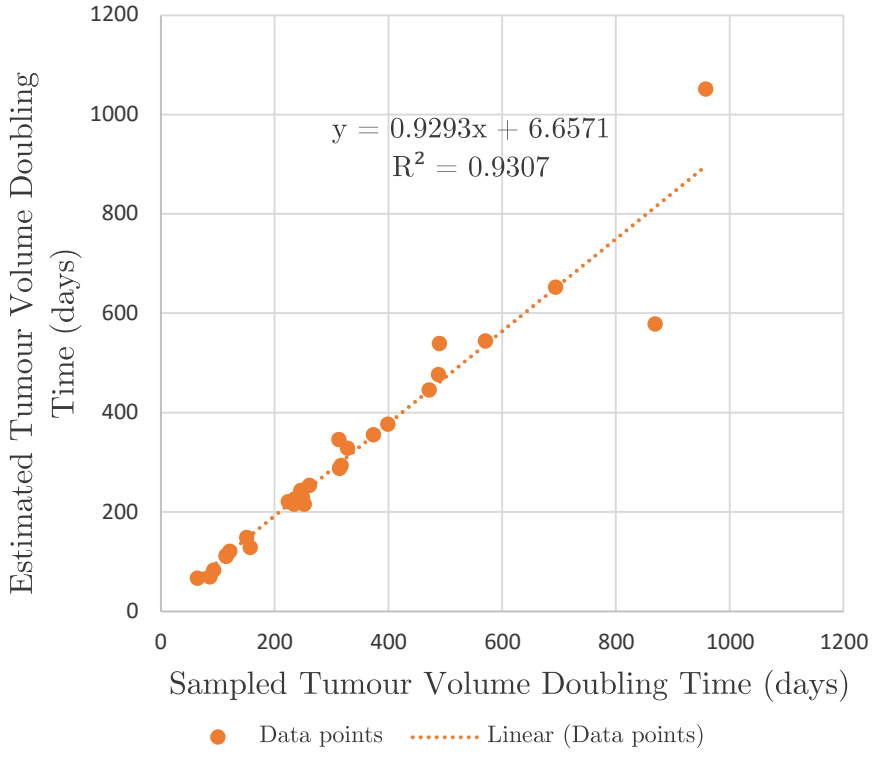


Figure 11: Positive correlation was found between he estimated TVDT and the sampled TVDT. The largest deviation was found for larger TVDTs.

4.3 Growth factors and growth functions

The estimated growth factor k_{est}, based on the radiologist's measurement of lesion size and calculated with Equation 4 or 5, showed strong correlation to the theoretical values k_{theo} (R=0.99) that were derived from the sampled TVDT. A more considerable deviation was found for larger k-values as seen in Figure 12.

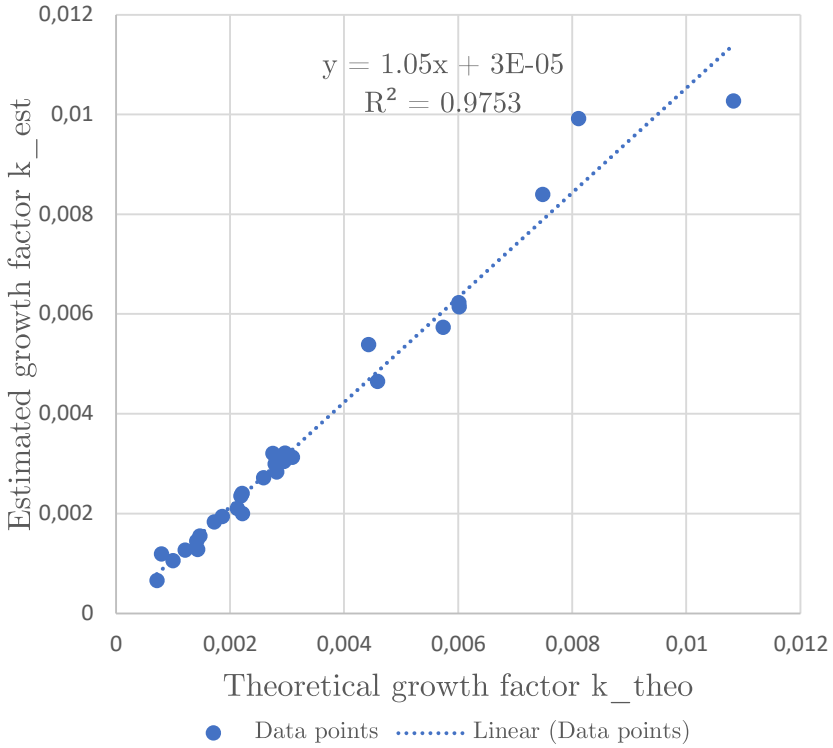


Figure 12: The estimated k-factor compared with the theoretical values that were based on the model. Statistical analysis using Pearson coefficient showed strong correlation between the data sets.

The growth curves, both estimated and those sampled from the model, can be seen for two virtual patients in Figure 13 and 14. It was noted that the divergence between the curves grew larger over time.

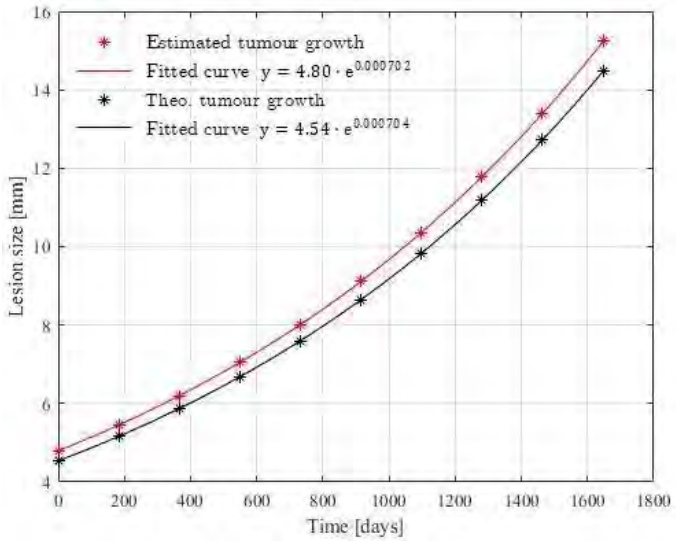


Figure 13: Tumour growth function for the first virtual patient. Red curve represents the estimated growth curve, that was based on lesion sizes measured by the radiologist in order to estimate TVDT. The darker curve shows the theoretical tumour growth that was sampled from the model and acts as the ground truth for the estimated data.

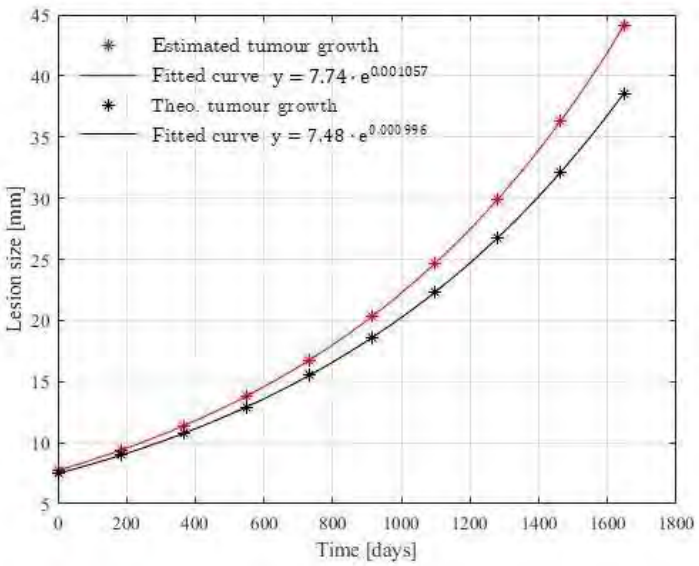


Figure 14: Tumour growth function for the second virtual patient. In this Figure it is clearer that the curves deviate more over time.

4.4 Lesion size measurements

The mean difference between the measured lesion size and the theoretical size of the simulated lesions was 0.49 mm, the absolute difference ranging between 0.05-5 mm. The largest difference was linked to the largest tumour size. As assumed, there was a strong correlation (R=0.99) found between data sets (Figure 15).

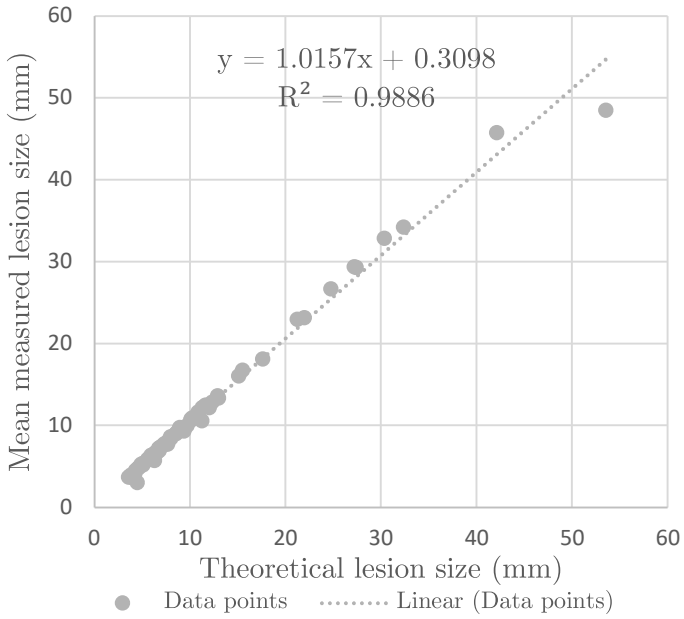


Figure 15: The theoretical lesion size compared to the mean measured size, estimated by the radiologist. A strong correlation was found between the data sets.

All size measurements by the radiologist were done with two cursors for each lesion, in total 120 measurements (see Figure 6). The accuracy of the two measurements was compared and can be seen in Figure 16. As assumed, there was a strong correlation between the measurements from the two cursors (R=0.99). The mean value of the differences between the measurements was 0.075 mm, the absolute difference ranging between 0.00-1.54 mm.

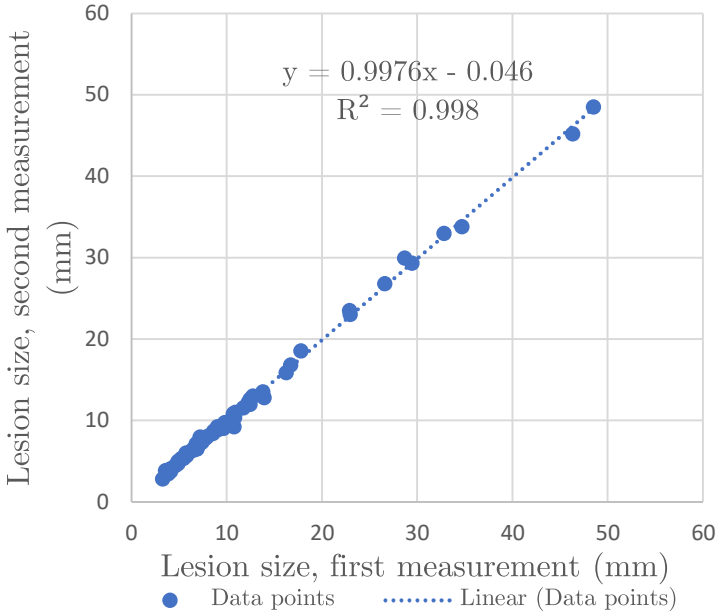


Figure 16: First measurement of lesion size, compared to second measurement of the same lesion for 60 mammograms and in total 30 lesions.

4.5 Simulated images and tumour visualization

Three of the 60 simulated mammograms presented to the radiologist can be seen below (Figure 17), representing the difference in breast size between B-, C- and D-cup.

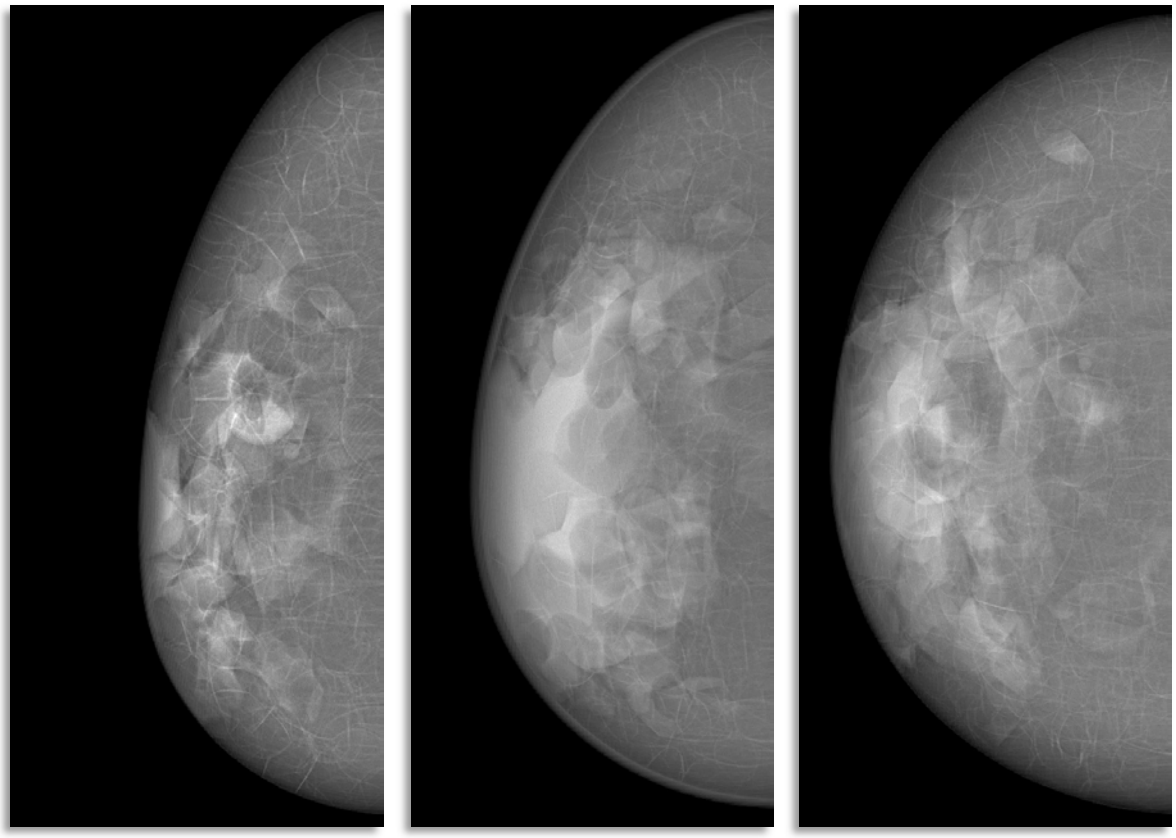


Figure 17: Three breast phantoms created with OpenVCT displayed in MLO-view. The left phantom was assigned a B-cup, whilst the central image shows a phantom with C-cup. As seen, the central phantom consists of slightly denser compartments. The right image presents a D-cup breast phantom.

Figures 18 and 19 show two virtual mammograms at different time points for two separate patients. The breast phantom depicted in Figure 19 is less dense compared to the one seen in Figure 18, this results in a more visible lesion.

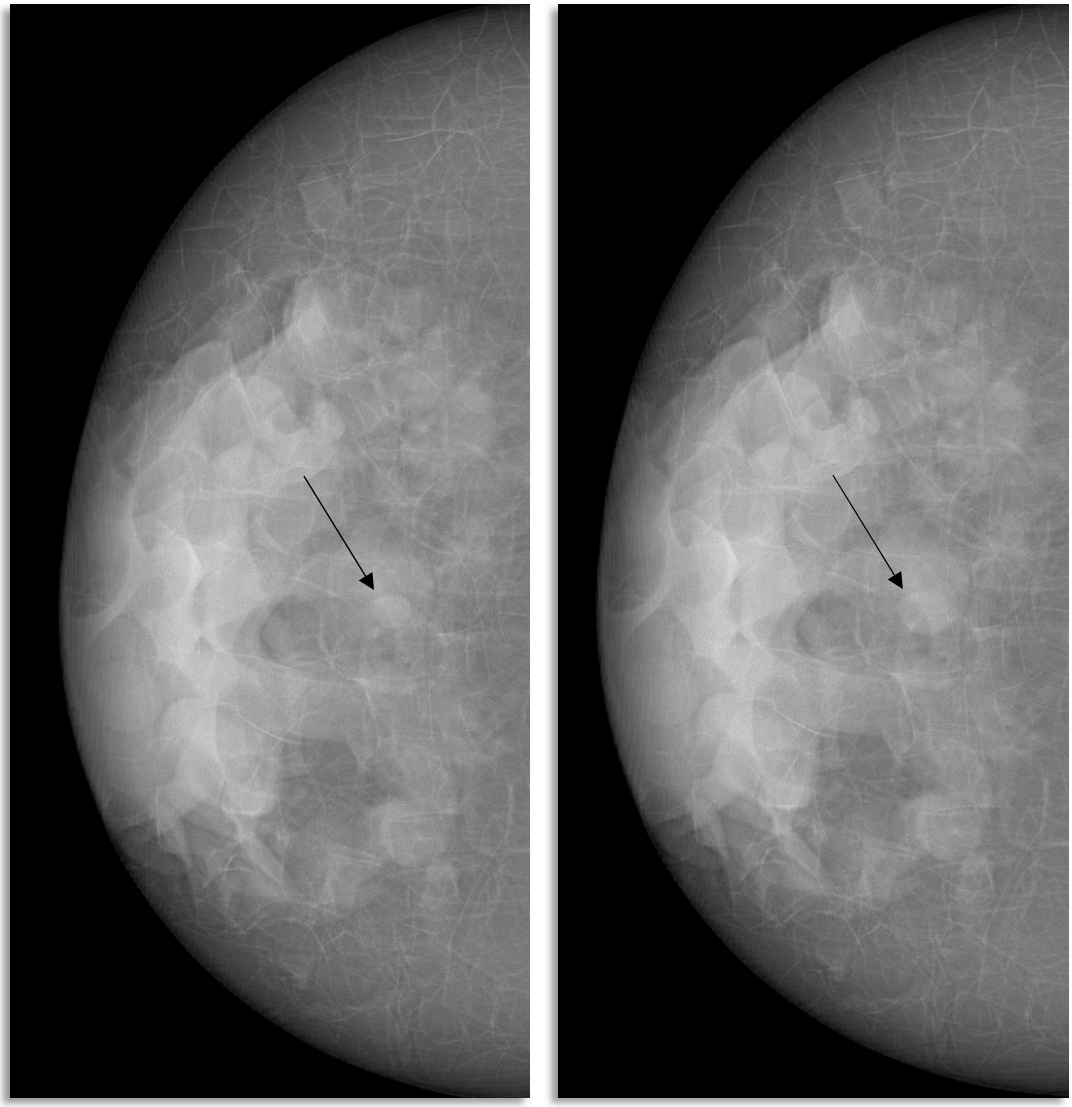


Figure 18: Two simulated mammograms of the left breast belonging to a virtual patient born 1956. The breast phantom is a D-cup consisting of 40% dense compartments and has a ligament thickness of 0.18 cm and a skin thickness of 0.05 cm. The patient was assigned an initial lesion diameter of 7.18 mm and a TVDT of 374 days. The left image shows the lesion at the first virtual screening occasion, whereas the right image depicts the lesion after two years, having grown to 11.3 mm. The radiologist estimated the sizes to 7.6 and 12.2 mm respectively. The estimated TVDT became 356 days.

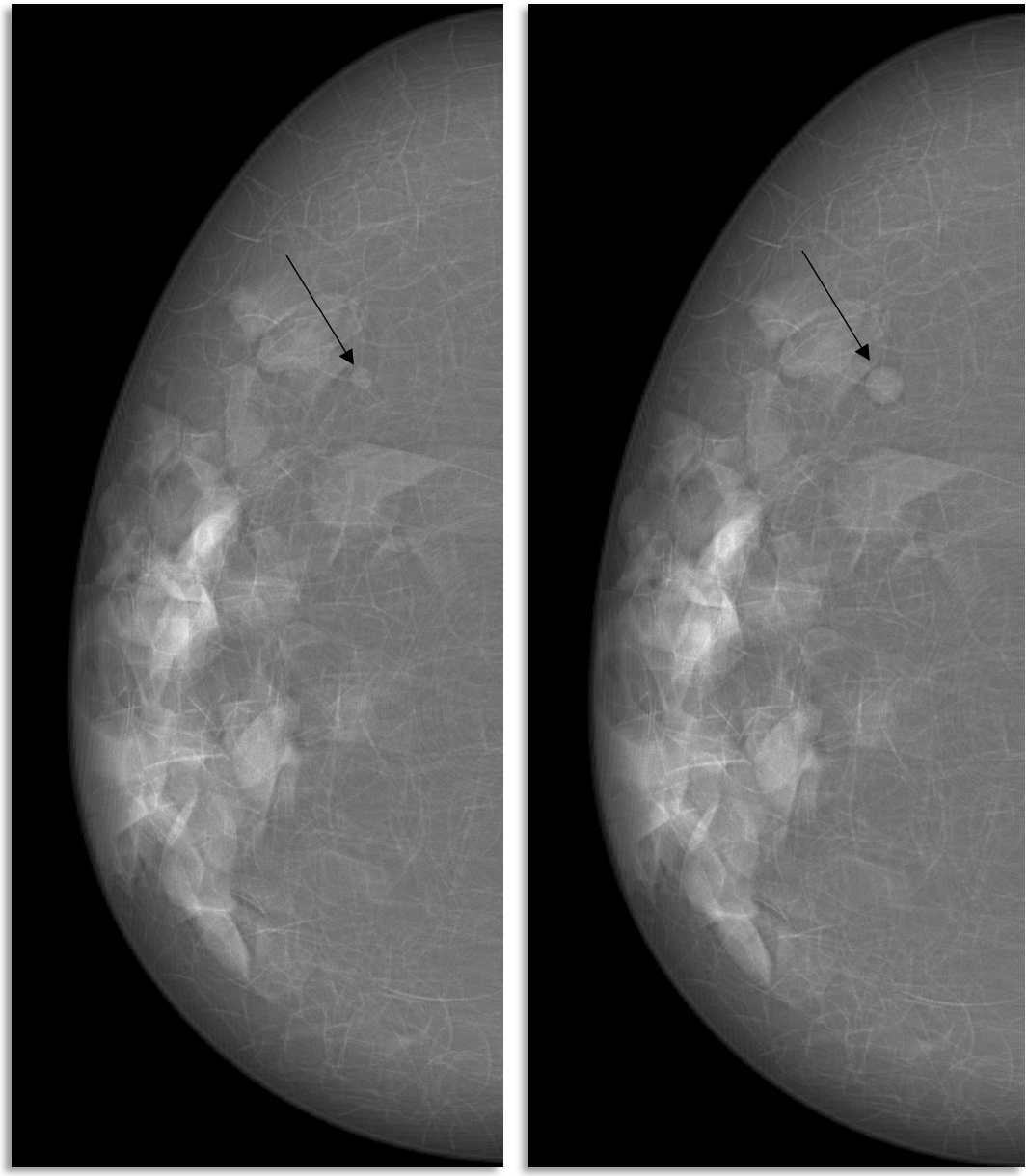


Figure 19: Mammogram showing left breast of virtual patient born in 1958. Left image shows the initial lesion size of 4.00 mm and the right shows the lesion after 2 years having grown to 6.82 mm. The sampled TVDT was 317 days. The breast phantom was assigned a C-cup with 0.15 cm skin thickness, 0.02 cm ligament thickness and 15% dense compartments. The radiologist estimated the sizes to be 4.1 mm and 7.3 mm, leading to an estimated TVDT of 294 days.

The lesions were also visualized as spheroids for a few virtual patients. Two examples can be seen in Figure 20 and 21.

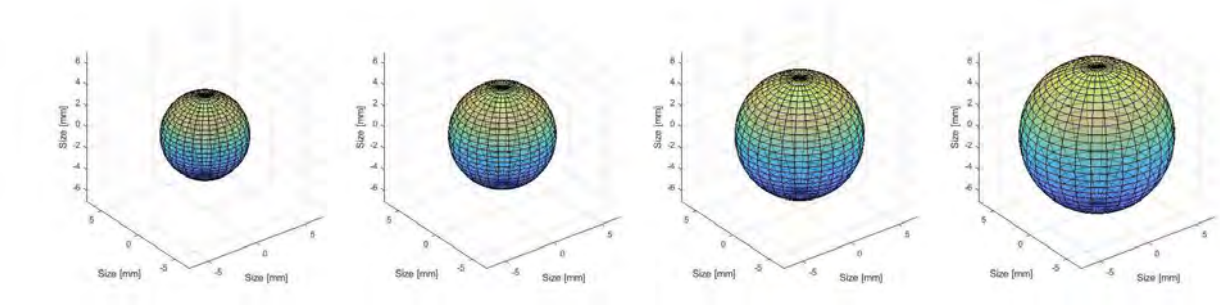


Figure 20: Spherical lesion visualized with a half-year interval for a virtual patient with initial lesion size of 7.48 mm. The sampled TVDT was 232 days.

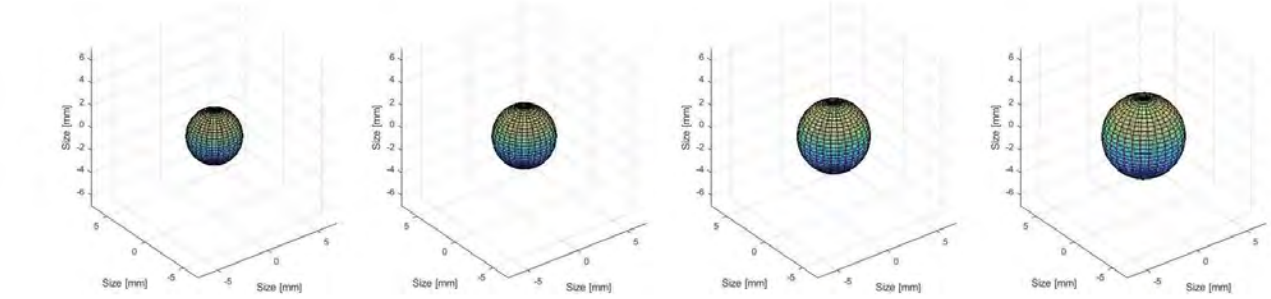


Figure 21: Lesion displayed with a half-year interval for a virtual patient with TVDT of 328 days. Initial tumour size was 4.54 mm.

5 Discussion

The initial tumour growth model was successfully established and assessed in a virtual clinical situation as desired. The model showed to accurately mimic the clinical TVDTs that were based on breast cancer cases of the Malmö screening population. It was concluded that no significant difference could be found between the modelled and the clinical data sets. This is considering the given limitations, such as simulated tumour geometry and a small sample size.

There were relative differences found between the mean values of the TVDT data sets. The largest was 9%, between the sampled TVDT and the TVDT-model. This is most likely due to the sampled patient cohort being small (30 virtual patients). However, compared to the large range of TVDTs, these can be considered as minor relative differences. Especially since it was found that there is no significant difference between the TVDT data sets (Kolmogorov-Smirnov test yielding a $p>0.54$, see Table 5).

The estimated TVDT did however have a median difference that was significantly ($p=0.0043$) larger than zero when compared with the sampled TVDT. However, since the Wilcoxon signed-rank test is a non-parametric test it only takes into account that there is a systematic difference. It does not consider the magnitude of the difference. Therefore, the median difference was estimated as well. It showed to be low (12 days with IQR = 20.75 days), which can be put in relation to TVDT that is in the range of a year. Hence the conclusion can still be drawn that there are small differences between any of the data sets. Furthermore, the Wilcoxon test showed that the estimated TVDT-data is similar to the clinical data ($p>0.64$, see Table 5).

The correlations that were examined regarding TVDT (Figure 11), growth factors (Figure 12) and lesion sizes (Figure 15 and 16), were assumed to be good. This was confirmed through the Pearson correlation that produced R-values above 0.97. The growth functions that were simulated (Figure 13 and 14) showed that there was a deviation between the estimated growth curve and the theoretical one. This is a result of the existing parametric differences between the data sets, that becomes more pronounced over time. However, it was not fully examined how much the divergence over time between the curves matters, as any lesion would inevitably be self-discovered by the woman before reaching a certain size. An upper limit of the lesion size was not implemented in the model and was one of the limitations with this study. Large tumour sizes are usually associated with fast growing tumours with short TVDT and would in reality probably comprise the cases with interval cancers that are detected in between screenings.

One of the most important limitations of the study was the small sample size, both regarding the clinical data that was available from Förnvik's publication [28] as well as the sampled virtual patient population. The publication was chosen as the foundation of the model since it was specifically examining TVDT for the Malmö population and since there was detailed data available to use. Since the data was based on an already limited sample size, having a larger virtual population would not yield a much lower statistical uncertainty. Apart from providing larger uncertainty when it comes to TVDTs, the small sample sizes also affected the fitted gamma distributions (Figure 8-10 and Table 6). This in turn leading to a more considerable difference between the data distributions, that would perhaps not be reproduced if a larger patient cohort was used. Moreover, it was noted that apart from the clinical data, the distribution fits depicting TVDT did not follow a normal distribution (Table 4). This might also be due to the small sample size. A larger cohort may have resulted in a normal distribution, which would affect the statistical power as the non-parametrical statistical tests could be disregarded.

Another essential limitation was that the simulated lesions were assumed to have a spherical geometry as seen in Figures 17-21. The uniformity of the lesion shape may have supported relatively high agreement between the nominal lesion size and the one estimated by the radiologist (mean difference was 0.49 mm, see also Figure 15). A spiculated or obscured mass would be more difficult to outline and may result in larger differences between the estimated TVDT compared to the other data sets. However, studies would need to confirm this, as there is a possibility that the accuracy would not be significantly impaired when using a more detailed tumour model. Moreover, simulating other lesion shapes would increase the similarity to a clinical outcome. The fact that scattered radiation is not currently simulated might also affect the lesion measuring certainty. Figure 15 and 16 also provide information about any systematic errors that the radiologist might have caused while preforming measurements, which showed to be low. The errors seem to grow larger for larger tumours (Figure 15), however the relative difference did not vary systematically.

The characterizations of the Malmö population also included accurately depicting patient age and breast size in the virtual cohort. The data regarding patient age was conducted from a larger cohort of 139 patients from the MBTST [4]. The sample size was considered to be sufficient for the purpose of assigning a fair distribution of screening intervals within the virtual patient cohort. However, research regarding the distribution of breast sizes can be more comprehensively regarded in future studies. Another highly motivating parameter that would increase the individualization of the virtual populations is taking into account the breast density distribution in the Malmö population. It was not examined in this study, but would likely improve the resemblance between VCTs and clinical trials. Above all, the breast density sets the attenuation properties and simulating this according to real clinical cases would considerably affect the number of lesions that are masked due to anatomical features. Hence it would probably affect the specificity of potential VCTs. However, variation in breast density could still be simulated for the virtual patient cohort and was demonstrated as the difference between Figure 18 and 19.

The aim of this study was to pursue an exponential tumour growth model. As mentioned previously, Förnvik et al. [28] found that there was not a significant difference between the exponential growth curve and the Gompertzian model. However, according to Ryu et al. [29] TVDT changes over time for some breast cancer subtypes and the Gompertzian model does not assume that TVDT remains constant. This suggests Gompertzian models should be considered in future studies.

5.1 Future work

In a next step of the study it would be necessary to simulate other lesion shapes and above all include spiculated masses. Moreover, to evaluate how to accurately simulate the breast density distribution of the Malmö population. It would also be of interest to simulate temporal changes of the breast tissue and address local tissue variation around the growing tumour.

Another crucial aspect is to include larger sample sizes. This would result in estimating TVDT in a larger number of clinical cases, to later base the model on a greater sample. In this study it would be also useful to examine how irregular clinical lesions vary in size, which can be done by estimating lesion sizes in clinical DBT slices. This way future VCTs can accurately represent lesions in DBT simulations. It would also be important to extend the information about TVDTs that vary over time and include Gompertzian models. Moreover, it would be worth considering including an upper limit of tumour size. This size would stand into relation with the lesion size at which the women self-detect the tumour and might affect the number of interval cancers seen in the VCT.

Apart from the clinical work, it would also be essential to simulate the Siemens and GE X-ray modalities to represent the mammography equipment most commonly used in clinics today. Moreover, to include simulations of scattered radiation and learn more about how the AEC is being simulated.

The study should be repeated as well simulating lesions in DBT instead of mammography. The results could be compared to this study, to see if the same results regarding TVDT would be received when measuring lesion sizes in DBT-images instead.

More work should be done to establish a virtual clinical trial where the MBTST can be continued in a second screening round.

6 Conclusion

The tumour growth model was shown to accurately mimic clinical TVDTs from a previous clinical study by Förnvik et al. [28] in Malmö, both when simulated and when estimated in a virtual clinical situation. Results showed no significant difference between the clinical TVDTs and the TVDTs estimated by the radiologist on the virtual mammograms. High accuracy was also found between theoretical and measured lesion sizes. However, larger sample sizes and a more realistic tumour geometry should be included in future work to strengthen the clinical resemblance.

The study lays the foundation for adding an additional dimension of realism when modelling cancer cases in virtual clinical trials on digital mammography and DBT.

7 Acknowledgments

I want to devote my greatest gratitude to my supervisors; **Predrag Bakic,** your support and guidance has been invaluable. Thank you for always providing great discussions and pushing this project to a new level. **Magnus Dustler** and **Anders Tingberg,** thank you for your commitment into initiating this project and for encouraging me further. **Anna Bejnö** and **Gustav Hellgren**, thank you for your support and inspiration. I am immensely grateful for receiving such a warm welcome to the group by all of you.

Kristin Johnson, many thanks for showing enthusiasm and taking part in this work. Thanks to **Bruno Barufaldi** for your guidance regarding OpenVCT and **Daniel Förnvik** for providing great insight into your work.

I would like to direct my sincere thanks to the **LUCI** family for warmly welcoming me to be a part of your successful and outstanding group. I would also like to extend a thanks to the staff at the **Department of Radiation Physics, SUS, Malmö** for many wonderful chats during coffee breaks.

Lastly, thanks to my wonderful family. Mum, dad, my little brother and my grandmother, I would not be here without your love and encouragement. Thanks to my boyfriend, **Roberto Mastio**, for unceasingly having my back and for being there for me throughout my whole education. Your support means the world to me.

8 References

1. Swedish National Board of Health and Welfare and, Cancerfonden. Cancer i siffror 2018. 2018.

2. Swedish National Board of Health and Welfare. Statistikdatabas för dödsorsaker [Internet]. Available from: https://sdb.socialstyrelsen.se/if_dor/

3. Swedish National Board of Health and Welfare. Screening för bröstcancer - Rekommendation och bedömningsunderlag. 2014.

4. Zackrisson S, Lång K, Rosso A, Johnson K, Dustler M, Förnvik D, et al. One-view breast tomosynthesis versus two-view mammography in the Malmö Breast Tomosynthesis Screening Trial (MBTST): a prospective, population-based, diagnostic accuracy study. Lancet Oncol. 2018;

5. Zeng R, Park S, Bakic P, Myers KJ. Evaluating the sensitivity of the optimization of acquisition geometry to the choice of reconstruction algorithm in digital breast tomosynthesis through a simulation study. Phys Med Biol. 2015;

6. Das M, Kavuri A, Nisbett WH. Investigating the contributions of anatomical variations and quantum noise to image texture in digital breast tomosynthesis. In 2018.

7. Diaz O, Dance DR, Young KC, Elangovan P, Bakic PR, Wells K. Estimation of scattered radiation in digital breast tomosynthesis. Phys Med Biol. 2014;

8. Förnvik D, Timberg P, Zackrisson S, Tingberg A, Petersson H, Dustler M. Towards determination of individual glandular dose. In 2018.

9. Young S, Bakic PR, Myers KJ, Jennings RJ, Park S. A virtual trial framework for quantifying the detectability of masses in breast tomosynthesis projection data. Med Phys. 2013;

10. Bakic PR, Barufaldi B, Higginbotham D, Weinstein SP, Avanaki A, Espig K, et al. Virtual clinical trial of lesion detection in digital mammography and digital breast tomosynthesis. In 2018.

11. Bakic PR, Barufaldi B, Pokrajac D, Weinstein S, Maidment A. Optimized simulation of breast anatomy for virtual clinical trials. In 2018.

12. Barufaldi B, Bakic PR, Higginbotham D, Maidment ADA. OpenVCT: a GPU-accelerated virtual clinical trial pipeline for mammography and digital breast tomosynthesis. In 2018.

13. Pokrajac DD, Maidment ADA, Bakic PR. Optimized generation of high resolution breast anthropomorphic software phantoms. Med Phys. 2012;

14. Yaffe MJ. Mammographic density. Measurement of mammographic density. Breast Cancer Research. 2008.

15. Breast Density and Your Mammogram Report [Internet]. American Cancer Society. 2019 [cited 2020 Apr 8]. Available from: https://www.cancer.org/cancer/breast-cancer/screening-tests-and-early-detection/mammograms/breast-density-and-your-mammogram-report.html#written_by

16. Kopans DB. Breast Imaging. Third Ed. Lippincott Williams and Wilkins; 2006. 7–8, 35–42, 281–340, 422–426 p.

17. Azzopardi JG, Chepick OF, Hartmann WH. The world health organization. Histological typing of breast tumors. Neoplasma. 1983;806–16.

18. Bliss D. Breast Anatomy [Internet]. National Cancer Institute. [cited 2020 Apr 17]. Available from: https://commons.wikimedia.org/wiki/File:Breast_anatomy_-_spanish.jpg

19. Helvie M. Imaging Analysis: Mammography. In: Diseases of the Breast. Fifth Ed. 2014.

20. Deng, Francis, Gaillard et.al. F. Breast Neoplasms [Internet]. Radiopedia.org. [cited 2020 Apr 9]. Available from: https://radiopaedia.org/articles/breast-neoplasms

21. Suckling J, Boggis CRM, Hutt I, Astley S, Betal D, Cerneaz N, et al. The mini-MIAS database of mammograms. The Mammographic Image Analysis Society Digital Mammogram Database Exerpta Medica. International Congress Series 1069. 1994.

22. Woods RW, Sisney GS, Salkowski LR, Shinki K, Lin Y, Burnside ES. The mammographic density of a mass is a significant predictor of breast cancer. Radiology. 2011;

23. Deng F. Breast mass [Internet]. Radiopedia.org. [cited 2020 Apr 9]. Available from: https://radiopaedia.org/articles/breast-mass?lang=us

24. Haagensen CD. Diseases of the Breast. Third Ed. London: W B Saunders Co; 1971. 380–381 p.

25. Dilmen N. Wikicommons - Human female breast quadrants [Internet]. 2011 [cited 2020 May 5]. Available from: https://commons.wikimedia.org/wiki/File:Breast_quadrants.svg

26. Valenica E, Lee C. Two-View Asymmetry. In: Breast Imaging. Oxford University Press; 2018.

27. Eliyatkin N, Yalcin E, Zengel B, Aktaş S, Vardar E. Molecular Classification of Breast Carcinoma: From Traditional, Old-Fashioned Way to A New Age, and A New Way. J Breast Heal. 2015;

28. Förnvik D, Lång K, Andersson I, Dustler M, Borgquist S, Timberg P. Estimates of breastcancer growth rate from mammograms and its relation to tumour characteristics. Radiat Prot Dosimetry. 2016;

29. Ryu EB, Chang JM, Seo M, Kim SA, Lim JH, Moon WK. Tumour volume doubling time of molecular breast cancer subtypes assessed by serial breast ultrasound. Eur Radiol [Internet]. 2014 Sep 4;24(9):2227–35. Available from: http://link.springer.com/10.1007/s00330-014-3256-0

30. MacInnes E, Duffy S, Simpson J, Wallis M, Turnbull A, Wilkinson L, et al. Tumour growth rates in interval cancers. The Breast. 2020;(51):114–9.

31. Watanabe Y, Dahlman EL, Leder KZ, Hui SK. A mathematical model of tumor growth and its response to single irradiation. In: Theoretical Biology and Medical Modelling. 2016.

32. Oter Perrote E. Some simple mathematical models of tumor growth. University of Barcelona; 2015.

33. Perthame B. Some mathematical models of tumor growth. Lect Notes. 2015;

34. Fournier D V., Weber E, Hoeffken W, Bauer M, Kubli F, Barth V. Growth rate of 147 mammary carcinomas. Cancer. 1980;

35. Bartoszyński R, Edler L, Hanin L, Kopp-Schneider A, Pavlova L, Tsodikov A, et al. Modeling cancer detection: Tumor size as a source of information on unobservable stages of carcinogenesis. Math Biosci. 2001;

36. Isheden G. Statistical models of breast cancer tumour growth and spread. Karolinska Institutet, Stockholm; 2019.

37. Ikeda DM, Andersson I, Wattsgard C, Janzon L, Linell F. Interval carcinomas in the Malmo Mammographic Screening Trial: Radiographic appearance and prognostic considerations. Am J Roentgenol. 1992;

38. Duffy SW, Day NE, Tabár L, Chen HH, Smith TC. Markov models of breast tumor progression: some age-specific results. J Natl Cancer Inst Monogr. 1997;

39. Weedon-Fekjær H, Lindqvist BH, Vatten LJ, Aalen OO, Tretli S. Breast cancer tumor growth estimated through mammography screening data. Breast Cancer Res. 2008;

40. Socialstyrelsen. Screening för bröstcancer. 2014.

41. Lind H, Svane G, Kemetli L, Törnberg S. Breast cancer screening program in Stockholm county, Sweden-aspects of organization and quality assurance. Breast Care. 2010;

42. Perry N, Broeders M, de Wolf C, Tornberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. Eur Guidel. 2006;63–5, 172, 124.

43. Alnafea MA. Detection and Diagnosis of Breast Diseases. In: Breast Imaging. 2018.

44. Andersson I, Hildell J, Muhlow A, Pettersson H. Number of projections in mammography: influence on detection of breast disease. Am J Roentgenol. 1978;

45. Egan RL. The Technical Aspects of Mammography. CA Cancer J Clin. 1967;22–4.

46. International Agency for Research on Cancer. IARC handbooks of cancer prevention, volume 15: Breast cancer screening. 2016;15:113–37.

47. D. Dance, S. Christofides, M. Maidment IM. Diagnostic Radiology Physics: A handbook for teachers and students. Iaea. 2014;189.

48. Söderberg M, Gunnarsson M. Automatic exposure control in computed tomography an evaluation of systems from different manufacturers. Acta radiol. 2010;

49. Zackrisson S, Houssami N. Evolution of Mammography Screening: From Film Screen to Digital Breast Tomosynthesis. In: Breast Cancer Screening: An Examination of Scientific Evidence. 2016.

50. Sechopoulos I. A review of breast tomosynthesis. Part I. The image acquisition process.

Medical Physics. 2013.

51. Li J, Zhang H, Jiang H, Guo X, Zhang Y, Qi D, et al. Diagnostic Performance of Digital Breast Tomosynthesis for Breast Suspicious Calcifications From Various Populations: A Comparison With Full-field Digital Mammography. Comput Struct Biotechnol J. 2019;

52. Niklason LT, Christian BT, Niklason LE, Kopans DB, Castleberry DE, Opsahl-Ong BH, et al. Digital tomosynthesis in breast imaging. Radiology. 1997;

53. Skaane P, Bandos AI, Gullien R, Eben EB, Ekseth U, Haakenaasen U, et al. Comparison of digital mammography alone and digital mammography plus tomosynthesis in a populationbased screening program. Radiology. 2013;

54. Friedewald SM, Rafferty EA, Rose SL, Durand MA, Plecha DM, Greenberg JS, et al. Breast cancer screening using tomosynthesis in combination with digital mammography. JAMA - J Am Med Assoc. 2014;

55. Ciatto S, Houssami N, Bernardi D, Caumo F, Pellegrini M, Brunelli S, et al. Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): A prospective comparison study. Lancet Oncol. 2013;

56. Conant EF, Beaber EF, Sprague BL, Herschorn SD, Weaver DL, Onega T, et al. Breast cancer screening using tomosynthesis in combination with digital mammography compared to digital mammography alone: a cohort study within the PROSPR consortium. Breast Cancer Res Treat. 2016;

57. Bernardi D, Macaskill P, Pellegrini M, Valentini M, Fantò C, Ostillio L, et al. Breast cancer screening with tomosynthesis (3D mammography) with acquired or synthetic 2D mammography compared with 2D mammography alone (STORM-2): a population-based prospective study. Lancet Oncol. 2016;

58. Duijm L, Guit GL, Zaat JOM, Koomen AR, Willebrand D. Sensitivity, specificity and predictive values of breast imaging in the detection of cancer. Br J Cancer. 1997;

59. Lago MA, Maidment ADA, Bakic PR. Modelling of mammographic compression of anthropomorphic software breast phantom using FEBio. Int'l Symp Comput Methods Biomech Biomed Eng. 2013;

60. Sturgeon GM, Kiarashi N, Lo JY, Samei E, Segars WP. Finite-element modeling of compression and gravity on a population of breast phantoms for multimodality imaging simulation. Med Phys. 2016;

61. Lago MA, Abbey CK, Bakic PR, Maidment ADA, Weinstein SP, Eckstein MP, et al. Interactions of lesion detectability and size across single-slice DBT and 3D DBT. In 2018.

62. Siddon RL. Fast calculation of the exact radiological path for a three dimensional CT array. Medical Physics. 1985.

63. De Greef M, Crezee J, Van Eijk JC, Pool R, Bel A. Accelerated ray tracing for radiotherapy dose calculations on a GPU. Med Phys. 2009;

64. Feng SSJ, Sechopoulos I. Clinical digital breast tomosynthesis system: Dosimetric

characterization. Radiology. 2012;

65. SPSS TUTORIALS. Pearson Correlations – Quick Introduction [Internet]. [cited 2020 Apr 29]. Available from: https://www.spss-tutorials.com/pearson-correlation-coefficient/

66. The Editors of Encyclopaedia Britannica. Student's t-test. In: Encyclopedia Britannica [Internet]. 2005. Available from: https://www.britannica.com/science/Students-t-test

67. SPSS TUTORIALS. SPSS Shapiro-Wilk Test – Quick Tutorial with Example [Internet]. [cited 2020 Apr 29]. Available from: https://www.spss-tutorials.com/spss-shapiro-wilk-test-for-normality/

68. C. Thode H. Introduction. In: Testing For Normality. New York: Marcel Dekker Inc; 2002. p. 2.

69. NIST/SEMATECH e-Handbook of Statistical Methods. Kolmogorov-Smirnov Goodness-of-Fit Test [Internet]. 2012 [cited 2020 Apr 29]. Available from: https://www.itl.nist.gov/div898/handbook/eda/section3/eda35g.htm

70. Evren A, Tuna E. On Some Properties of Goodness of Fit Measures Based. IJRRAS. 2012;194.

71. McDonald JH. Wilcoxon signed-rank test. In: Handbook of Biological Statistics. 3d Edition. Baltimore, Maryland: Sparky House Publishing; 2014. p. 186–9.

72. Berg WA, Hendrick ER, Kopans DB, Smith RA. Frequently Asked Questions about Mammography and the USPSTF Recommendations : A Guide for Practitioners. 2009;1.

73. Michaelson J, Satija S, Moore R, Weber G, Halpern E, Garland A, et al. Estimates of the sizes at which breast cancers become detectable on mammographic and clinical grounds. J Women's Imaging. 2003;

74. Torlegård B, Tingberg A, Zackrisson S, Barufaldi B, Maidment A, Dustler M, et al. Identifying and Modelling Clinical Subpopulations from the Malmö Breast Tomosynthesis Screening Trial. Int'l Work Breast Imaging. 2020;15th.