Clinical Investigation Plan

Comparison of the Philips MicroDose Tomosynthesis System to 2D Digital Mammography

Security Classification:	**For internal use**
Author:	Britta Stenson/Users/Mamea
Document ID and Revision:	DOC-MTIN-9RWH8K-8.0
Approved by:	

SPONSOR

Sponsor's Legal Entity:	Philips Digital Mammography Sweden AB Box 1111 SE-171 22 Solna Sweden
Business Unit:	Mammography Solutions
Department:	Product Marketing
NCT number:	NCT02615509

APPROVAL PRINCIPAL INVESTIGATOR

I, Etta Pisano, M.D., Principal Investigator of this clinical investigation, approve the clinical investigational plan described in Document ID DOC-MTIN-9RWH8K dated 2015 August 12.

Name	Function	Date	Signature
Etta Pisano, M.D.	Principal investigator		

REVISION HISTORY

Date	Rev.	Author	Changes/Comments
2015 Jan 15	1.0	Miriam von Tiedemann	First PMA supplement version (based on DOC-VHAT-9JMGAX)
20015 July 3	2.0	Elin Moa/Sonia Souza	Updated protocol according to feedback from clinical investigators; Updated schematic flow chart of enrollment (fig 2)
2015 July 15	3.0	Sonia Souza/Elin Moa	Updated objectives and statistical section
2015 Aug 12	4.0	Jonas Rehn	Updated Study Design to ensure balanced enrollment from the different sources
2015 Aug 12	5.0	Jonas Rehn	Minor editorial change
2016	6.0	Jonas Rehn	No changes to protocol content Version signed by Dr Pisano added. Protocol content still equivalent to version 5.0.
2015 December 12	7.0	Britta Stenson	No changes to protocol content Document re-approved by clinical evaluator, project manager and regulatory affairs officer. Protocol content still equivalent to version 5.0.
2016 April 13	8.0	Britta Stenson	No changes to protocol content Clarified revision history Protocol content still equivalent to version 5.0.

TABLE OF CONTENTS

SUMMARY

Study Title (ID)	Comparison of the Philips MicroDose Tomosynthesis to 2D Digital Mammography
Subjects	Approximately 800 women who fulfill all inclusion and exclusion criteria will be enrolled in chronological order from up to 3 different US sites with one study investigator for each site. Two cohorts of patients will be enrolled (i.e., Biopsy and Screening Cohorts). Paired sets of a standard 4-view exam will be collected both from Philips Tomosynthesis system (Tomosynthesis and synthetic 2D images) and a FDA-cleared FFDM system. All enrolled cases will be quality controlled for completeness of image and patient information. Inclusion criteria: • Female and at least 40 years of age • No contraindication for routine bilateral mammography Screening Cohort • Patient presents for a routine screening mammogram Biopsy Cohort • Patient is scheduled for a biopsy • Patient has a screening detected abnormality The reading portion of the study will include approximately 300 normal cases (including some cases determined to be normal after call back after screening mammography), 65 cancer cases, and 40 biopsy proven benign cases. The primary objective of the study is to compare the safety and clinical performance of the Philips MicroDose Tomosynthesis system (Tomosynthesis and synthetic 2D) images to conventional 2D mammography images (FFDM).
Study Design	The study will be performed in two phases: Image Accrual and Image Reading. Image accrual will be performed in a multicenter setting. Image reading will be designed as a multi-case, multi-reader (MCMR) study with paired image acquisition and reader evaluation. Approximately 18 American Board of Radiology (ABR) certified, and Mammography Quality Safety Act (MQSA) qualified radiologists will review and score the cases. Outcome of the reader study for the two imaging groups will be compared using ROC methodology.
Treatment Groups	• Investigational Device: Philips MicroDose Tomosynthesis system (Tomosynthesis prototype Philips MicroDose SI S0) • Comparator: Digital mammography system (2D FFDM)
Primary Objectives	The primary objective of the study is to compare the safety and clinical performance of the Philips MicroDose Tomosynthesis system (tomosynthesis and synthetic 2D) images and 2D digital mammography (from any FDA cleared FFDM system) images.

	Clinical performance will be based on the Area under the Receiver Operating Characteristics curve (ROC AUC). The ROC curves will be generated using a confidence scores (CS) using a numeric scale. The ROC curves will be calculated for each reader.
Secondary Objectives	The secondary objectives are: • To compare the average glandular dose for exams performed with Philips MicroDose Tomosynthesis system and the 2D mammography system (2D FFDM. • To compare the non-cancer recall rate for Philips MicroDose Tomosynthesis system and 2D mammography exams. • To compare the sensitivity and specificity of the Philips MicroDose Tomosynthesis system (tomosynthesis and synthetic 2D) to that obtained with a 2D mammography system (2D FFDM). The average glandular dose (AGD) will be recorded per subject for each breast (right and left) and view (CC and MLO). Sensitivity will be estimated for each reader as the proportion of cancer cases with per-case BI-RADS category 4 or higher. Specificity will also be estimated by reader as the proportion of non-cancer cases with per-case BI-RADS category less than 4. Cancer determination will be based on biopsy results or the outcome of the one-year follow-up exam.
Supplementary Objectives	The following supplementary objectives will be evaluated: • The effect of breast density on AUC for images obtained using the Philips MicroDose Tomosynthesis system and 2D mammography system (2D FFDM). Breast density will be stratified as fatty (i.e., BI-RADS breast density type A and B) and dense (BI-RADS breast density type C and D). • To compare the average image interpretation time for exams performed with Philips MicroDose Tomosynthesis and 2D mammography system (2D FFDM).

1. STATEMENT OF COMPLIANCE

This clinical research shall be conducted in accordance with the Clinical Investigation Plan, and with the ethical principles that have their origin in the Declaration of Helsinki and all applicable regional and/or national regulations. Furthermore, all investigators will complete financial disclosures, as outlined in the 21 CFR part 54. Investigators located in the US shall follow the 21 CFR parts 50, 56 and 812.

This clinical research shall not be started prior to obtaining a favorable opinion from an Ethics Committee (EC)/Institutional Review Board (IRB) and Regulatory authority/Food and Drug Administration (FDA), if required. Any additional requirements imposed by the EC/IRB and/or regulatory authority/FDA shall be followed. Insurance shall be provided for the subjects participating in this clinical trial according to local law.

2. ABBREVIATIONS USED

Abbreviations	Explanation of abbreviation
ABR	American Board of Radiology
ACR	American College of Radiology
AGD	Average Glandular Dose measured in [mGy]
AUC	Area Under the Curve
BI-RADS	Breast Imaging- Reporting and Data System
CC	Craniocaudal
CI	Confidence Interval
CRF	Case Report Form
CRA	Clinical Research Associate
CS	Confidence Score
DBT	Digital Breast Tomosynthesis
DF	Degree of Freedom
EC	Ethics Committee
EUREF	European Reference Organisation for Quality Assured Breast Screening and Diagnostic Services
FDA	Food and Drug Administration
FFDM	Full Field Digital Mammography (all U.S. marketed FFDM units incl. MicroDose SI L50)
FOM	Figure-of-Merit
FROC	Free Receiver Operating Characteristics
IFU	Instructions for Use
IRB	Institutional Review Board
MLO	Medio-lateral oblique
MQSA	Mammography Quality Standards Act
MRMC	Multi-Reader, Multi-Case
NI	Non-inferiority
Synthetic 2D	A projection image calculated from tomosynthesis data. Does not require additional radiation exposure.
2D, FFDM	Conventional digital mammography
DBT plus synthetic 2D	Only tomosynthesis images are acquired, and a synthetic 2D image is calculated. Both images are read together.
ROC	Receiver Operating Characteristics

3. INTRODUCTION AND RATIONALE

3.1. Rationale

Breast cancer is one of the most common cancer types among women around the world. Approximately one in eight American women will be diagnosed with breast cancer in her life time [1]. Early detection of cancer leads to a better prognosis with a bigger chance to survive the disease. As such, screening programs have been developed to identify non-palpable breast cancer in asymptomatic women with x-ray imaging. While there have been numerous technologies used for breast cancer screening, the gold standard since the late 1970s has been x-ray based mammography [2].

Population based mammography screening has in some studies been shown to reduce breast cancer mortality by between 40% and 45% [3]. Digital mammography has supplanted film-screen mammography during the last ten years and diagnostic performance of the newer technology has now been proven in several large studies [4]-[6]. Digital mammography enables efficiency improvements and opportunities of future advanced applications such as spectral imaging, tomosynthesis and phase contrast imaging.

Tomosynthesis provides a three-dimensional reconstruction of the breast which is of interest both for screening and diagnostic procedures. The image reconstruction enables cross-sectional visualization of the breast tissue reducing difficulties of interpretation of projection images of overlapping tissue or superposition.

3.2. Literature review

Tomosynthesis is intensively marketed by the major mammography vendors, notably Hologic, GE and Siemens. A growing body of clinical data supports increased cancer detection rates, and also in some locations, reduced non-cancer recall rates [7]. The technology has passed the early adopter phase and is now moving into the mass market. Recently it was reported that 30% of the radiologists in the US used tomosynthesis in their practice and that 62% of the remaining radiologists report that they planned to obtain it soon [8].

In spite of the trend towards fast adoption of tomosynthesis there are still many major markets where the transition from 2D mammography to tomosynthesis is still expected to take many years. Guidelines for quality control as well as standards for safety and essential performance are still works in progress for tomosynthesis. Furthermore screening guidelines and reimbursement systems do not yet support tomosynthesis in screening.

3.3. Device description

The investigational device, Philips MicroDose SI version S0, Tomosynthesis prototype system (referred to as Philips MicroDose Tomosynthesis here) is developed by Philips Digital Mammography Sweden AB, a Philips Healthcare company. The proposed Philips MicroDose Tomosynthesis system is a based on the same technology as the FDA cleared FFDM device, Philips MicroDose SI.

The investigational device includes MicroDose SI S0 Tomosynthesis system with software version 11.0.

The manufacturer of the investigational device is:
Philips Digital Mammography Sweden AB
Box 1111
SE-171 22
Sweden

3.3.1. Intended use of the investigational device during the investigation / Indications for Use statement.

The intended medical purpose of the Philips MicroDose Tomosynthesis is to produce radiographic images of the human breast for the purpose of diagnostic and screening mammography. The Philips MicroDose Tomosynthesis system is intended to be used in the same clinical applications as 2D mammography systems.

The intended medical purpose of the investigational device is to collect clinical data to support the PMA submission for Digital Breast Tomosynthesis (DBT) device. The examination where the investigational device will be used is a standard 4-view mammogram, i.e., medio-lateral oblique (MLO) and craniocaudal (CC) views of each breast, performed on asymptomatic women to detect the presence of breast cancer. Repeat standard views might be done for technical reasons, such as inadequate compression, inadequate amount of included breast tissue, or over- or underexposure.

3.3.2. Populations and indications for which the device is intended, including contraindications

The digital tomosynthesis mammography system is intended to support examinations of the human breast of both women and men.

Women invited for screening are usually in the age of 40 to 70, depending on the local regulations. Patients coming for diagnostic examinations are of any adult age.

3.3.3. Device description

The design of the MicroDose Tomosynthesis system is based on the currently marketed MicroDose SI L50, which was cleared by FDA via K123995 – February 1, 2013. Due to the close similarity of the two devices, most design features that are described in the 510(k) are applicable to the MicroDose Tomosynthesis system as well.

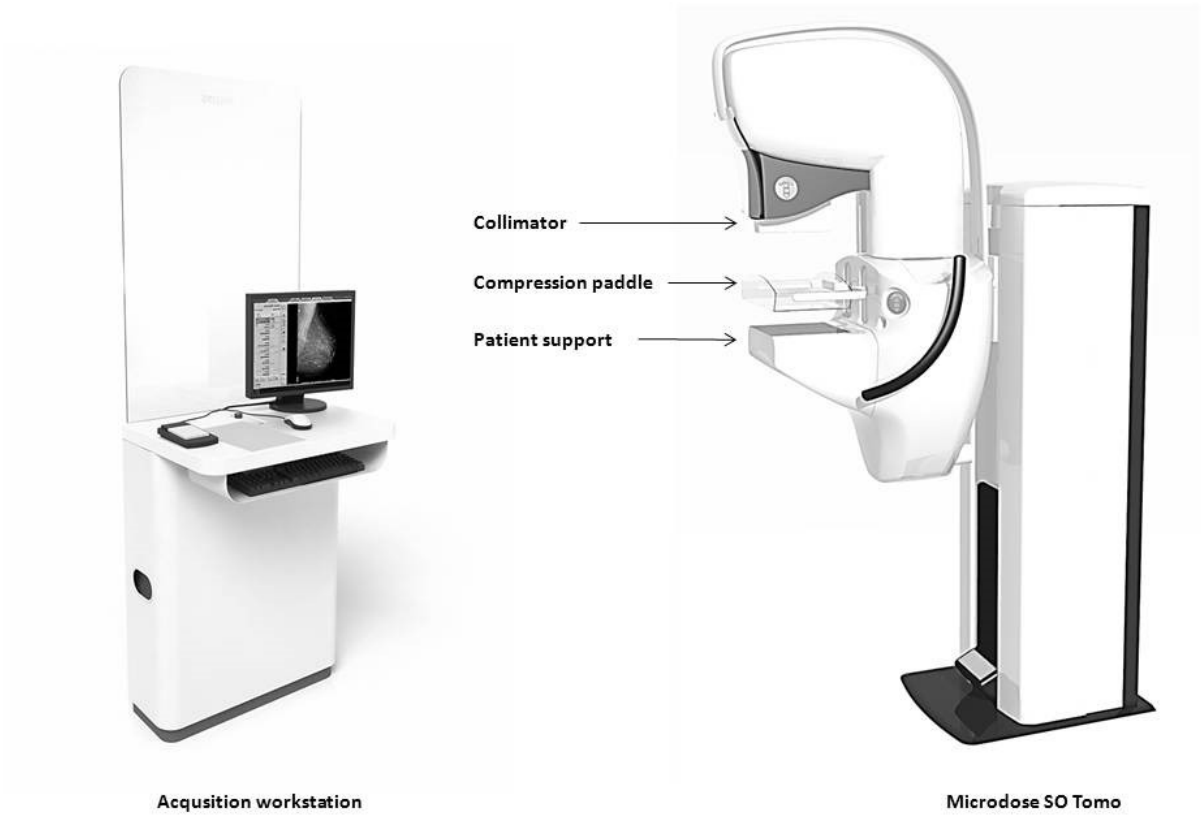

Figure 1. Overview of the MicroDose Tomosynthesis system.

Like the currently marketed MicroDose SI L50, the MicroDose tomosynthesis system consists of an acquisition workstation, a main cabinet and a side cabinet. With the exception of the side cabinet, these components are shown in Figure 1. In fact, these parts are the same in both systems. The main difference between the MicroDose Tomosynthesis system and the currently marketed MicroDose SI Model L50 is the design of the C-arm.

3.3.4. **Pre-clinical testing**

The product development project of the MicroDose Tomosynthesis system will be managed according to applicable parts of the Philips Digital Mammography's Product Development Process. The design will be verified according to the design verification processes on the first MicroDose Tomosynthesis system. The subsequent systems will be produced by limited production processes. To ensure that the subsequent systems will comply with the verified design, a predefined amount of test, selected via impact analysis, will be performed on each system. All risk related requirements will be verified on all systems.

The pre-clinical testing will put specific focus on image and radiation quality. The image quality tests include (but are not limited to) tests of exposure time, image areas, detector linearity, scatter rejection, spatial resolution, contrast-to-noise and signal-to-noise ratios, tests of the automatic exposure control, image homogeneity, contrast-detail resolution, and the Mammographic Accreditation Phantom. The radiation quality tests include (but are not limited to) tests of the tube output and generator accuracy, dose, half value layer, x-ray field, leakage radiation and protective shielding .

Regular quality control (QC) will be performed on the system according to a protocol aligned with the most recent version of the EUREF quality control procedures.

Change control will be performed according to applicable development processes. Issues found during verification and validation will be registered in a project issue database handled by a cross functional team (change control board). In addition to the internal testing described above, compliance verification will be conducted both in-house and by third party test house Intertek.

3.3.5. **Device risk analysis and risk assessment**

The risk management procedure is followed for the project. After performing risk analysis of the Philips MicroDose Tomosynthesis system, there are certain identified residual risks with justification. These risks are mainly related to software, x-ray exposure, moving parts during image acquisition, electronics of the photon-counting detector and high voltage used for X-ray.

A significant part of the identified justified risks are related to software. The technology of producing high quality digital images from a photon-counting imaging system can only be solved with software, for example image processing algorithms, image acquisition, data storage and data transfer, which are critical functions to the system performance. We still consider the risks of the system to be as-low-as–reasonably-practicable.

The scanning collimator is a necessary function when using multi-slit beam geometry to receive a tomosynthetic motion. This geometry reduces the dose to the examined patients. In summary, the benefit of Philips MicroDose overweighs the related risks.

3.3.6. **Conclusion**

The benefit of the study is to enable new technology that potentially could increase early detection and reduce breast cancer mortality. Furthermore the technology is also expected to reduce radiation dose level compared to existing commercially available flat panel systems.

As the subjects enrolled in the study will be imaged both with tomosynthesis and conventional 2D mammography, additional lesions could be detected in the screening patients that would not have been found only with conventional 2D mammography. However a direct benefit of the enrolled subjects is not expected. Possibly a missed cancer on the 2D mammogram could be detected on the tomosynthesis image. Another scenario is that a benign lesion is detected which could require additional imaging or even biopsy. The tomosynthesis images alone will not be used to make a decision to biopsy, but further work-up of such lesions with conventional imaging might indicate the need for biopsy. If a lesion is found on the tomosynthesis images but not detected on 2D, the radiologist at site might want to do additional imaging to assess the lesion. For the majority of the enrolled patients in the study, there will be no direct benefit as of participating in the study, though it is possible that a cancer that is missed on the 2D might be detected.

A risk that the enrolled subjects in the study are taking when participating in the study, is that they will be exposed to additional radiation. The radiation of the investigations device is comparable to Philips MicroDose SI, which is a low dose digital mammography system (approx. 0.8 mGy per view for a standard examination).

4. OBJECTIVES AND HYPOTHESES

4.1. Primary Objective

The primary objective of the study is to compare the safety and clinical performance of the Philips MicroDose Tomosynthesis system (Tomosynthesis and synthetic 2D) images to conventional 2D mammography images (FFDM).

Clinical performance will be based on the Area under the Receiver Operating Characteristics (ROC) Curve. The ROC curves will be generated using the confidence scores (CS) based on a numeric scale. ROC curves will be calculated for each reader.

The following hypothesis will be evaluated:

Null Hypothesis: Ho: $\Theta_2 - \Theta_1 \geq \Delta$
Alternative Hypothesis: $\Theta_2 - \Theta_1 < \Delta$

Where Θ_1 (Philips) and Θ_2 (FFDM) are the area under the Receiver Operating Characteristic (ROC) curve (AUC) for radiologists interpreting mammograms obtained on the Philips MicroDose Tomosynthesis system or FFDM, respectively. Delta (Δ) is the chosen non-inferior margin of 0.05.

4.2. Secondary Objectives

The secondary objectives are:

- To compare the average glandular dose for exams acquired on Philips MicroDose Tomosynthesis system and the 2D mammography system.
- To compare the non-cancer recall rate for Philips MicroDose Tomosynthesis system and 2D mammography exams.
- To compare the sensitivity and specificity of the Philips MicroDose tomosynthesis system (i.e., Tomosynthesis and synthetic 2D) to that obtained with 2D mammography.

The average glandular dose (AGD) will be recorded per subject for each breast (right and left) and view (CC and MLO). The following hypothesis will be evaluated:

- Null Hypothesis: Ho: $\mu_2 - \mu_1 > \delta$
- Alternative Hypothesis: $\mu_2 - \mu_1 \leq \delta$

where μ_1 and μ_2 are the mean average glandular dose (i.e., radiation dose) for the Philips MicroDose Tomosynthesis system and FFDM system, respectively, and δ is 10% of the FFDM radiation dose.

The non-cancer recall rate will be recorded per subject. The following hypothesis will be evaluated:

- Null Hypothesis: Ho: $\pi_2 - \pi_1 = 0$
- Alternative Hypothesis: $\pi_2 - \pi_1 \neq 0$

where π_1 and π_2 are the non-cancer recall rate for the Philips MicroDose Tomosynthesis system and FFDM system, respectively.

Sensitivity will be estimated for each reader as the proportion of cancer cases with per-case BI-RADS category 4 or higher. Specificity will also be estimated by reader as the proportion of non-cancer cases with per-case BI-RADS category less than 4. Ground truth will be based on biopsy results or the outcome for the woman after one year follow-up.

4.3. Supplementary Objectives

The following supplementary objectives will be evaluated:

- To evaluate the effect of breast density on AUC for images obtained using the Philips MicroDose Tomosynthesis system and 2D mammography system. Breast density will be stratified as fatty (i.e., BI-RADS breast density type A and B) and dense (BI-RADS breast density type C and D).
- To compare the average image interpretation time for exams performed with Philips MicroDose Tomosynthesis system and 2D mammography.
- To assess the diagnostic performance of the Philips MicroDose Tomosynthesis compared to the FFDM

5. STUDY DESIGN

5.1. General

The study is a multicenter, multi-case, multi-reader (MCMR) study with paired image acquisition and reader evaluation. The study will be performed in two phases: Imaging Accrual phase and Imaging Reading phase.

Approximately 800 women who fulfill all inclusion and exclusion criteria will be enrolled in chronological order from approximately 3 different US sites with one study investigator for each site. Two cohorts of patients will be enrolled (i.e., Biopsy and Screening Cohorts). Paired sets of a standard 4-view synthetic 2D and tomosynthesis slices exposed on Philips MicroDose Tomosynthesis system and 4-view mammograms on a FDA cleared FFDM system will be collected. All enrolled cases will be quality controlled for completeness of image and patient information.

A subset of the images collected as part of this study will also be used to validate the spectral breast density measurement. This validation is specified in a separate study protocol [9].

Depending on the experience at the sites, each site could also enroll up to 10 additional women to train the technologists on using the system and whose images will be used for training purposes. These cases will not be used in the reader study. In addition to these cases, images for reader training will be available that have been previously collected under another IRB approved protocol.

Approval for the study will be obtained from the ethics committees for all participating centers. All eligible subjects will provide written informed consent prior to entry into the study, which is defined as being imaged on the Philips MicroDose Tomosynthesis system and an FFDM system. Subject enrollment will be monitored by the sponsor at each site. All subjects will be consecutively enrolled.

All subjects interested in participating in the study will be assigned a unique study identification number. At the time of providing informed consent, subjects will be assigned the next number in sequence from the list of study identification numbers. After study enrollment, subjects will be identified solely by their study identification number on the Case Report Forms (CRF). In addition, all subject specific information will be removed from the images prior to them being electronically submitted to the study sponsor and only the subject's study identification number will be provided. Each participating site will maintain a database linking the study number to subject identification.

To protect against any bias during the imaging acquisition, the order in which the subject is imaged on either the investigational device or the FFDM will be randomized, except for patients that have undergone a screening exam including 2D. In this case, the patient will be imaged only with the Philips MicroDose Tomosynthesis system (since the 2D mammogram has been acquired previously). Reimaging will only be performed if more than 45 days have passed since the original acquisition. A randomization list will be generated for each cohort (i.e., the biopsy and screening cohort).

The mode of the eligibility screening exam (source) will be recorded. To protect from case selection bias, not more than 60% of the cases in the reader study can originally be screened with 2D or tomosynthesis. In this aspect combo screening (2D + tomosynthesis) is regarded as tomosynthesis.

The enrollment and randomization flow for each cohort is presented in Figure 2.

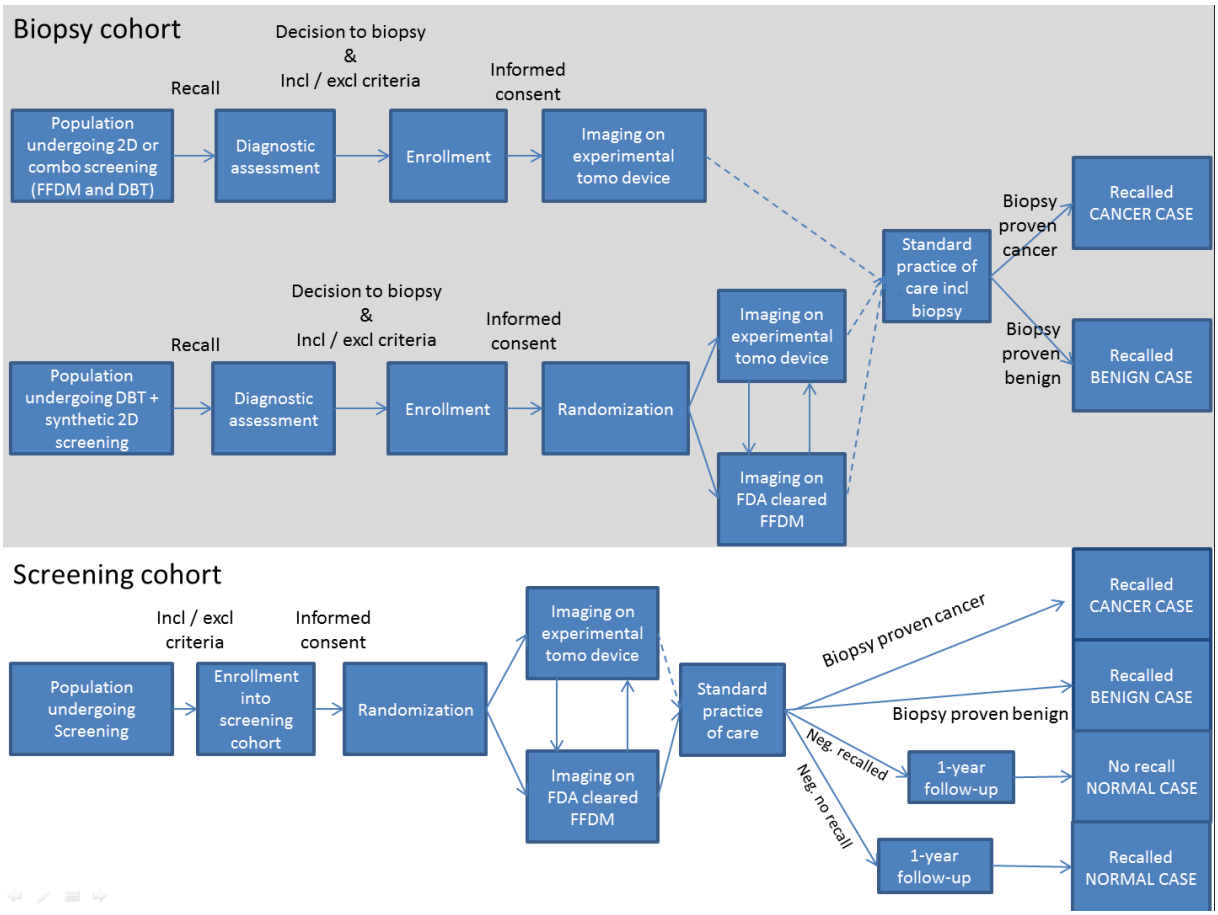

Figure 2: Patient Enrollment and Randomization by Cohort

5.2. Comparators

As of today mammography screening with 2D mammography is the global standard of care. Tomosynthesis is only recently approved for screening and is not reimbursed in many countries, and has no harmonized quality control procedure.

Also there are no established guidelines for Tomosynthesis in terms of angles, dose, number of slices required or how it should be used. Therefore this study uses 2D mammography as the comparator.

5.2.1. Maintenance and calibration

The system will be maintained and calibrated according to the systems manual.

5.3. Subjects

Subjects participating in the study will be selected based on the inclusion and exclusion criteria:

5.3.1. Inclusion criteria

- Patient is female and at least 40 years of age
- No contraindication for routine bilateral mammography

Screening Cohort

- Patient presents for a routine screening mammogram

Biopsy Cohort

- Patient is scheduled for a biopsy
- Patient has a screening detected abnormality

5.3.2. Exclusion criteria

- Patient is pregnant or believes she may be pregnant; OR,
- Patient is breast feeding; OR
- Patient is unable or unwilling to give informed consent, including consent to reuse data for future research; OR,
- Patient has breast implants; OR
- Patient has previous surgical biopsy; OR
- Patient has previous breast cancer; OR
- Patient presented to screening with clinical symptoms (e.g. palpable lump, nipple discharge, nipple retraction, skin irritation or dimpling)

5.3.3. Enrollment and duration

Subjects are considered to be enrolled in the study after they have signed the informed consent form. No study procedures will be performed before this moment.

The study period, per subject, is approximately up to 90 minutes and is limited to a single mammography session, except for the normal cases which will require a one-year follow-up exam (normal screening exam, not using the investigational device).

5.3.4. Number of Subjects and Readers

Approximately 800 patients who fulfill all inclusion and exclusion criteria and pass the quality control for completeness of image and patient information will be enrolled in the image accrual phase of the study. Among these patients approximately 300 normal, 65 cancer, and 40 biopsy-proven benign cases will be included in the image reading phase of the study.

The numbers of readers, and the overall sample size justification for the study is detailed in the Statistical considerations, section 6 of this protocol.

5.3.5. Procedure for the replacement of subjects

Subjects will not be replaced.

5.3.6. Subject withdrawal or discontinuation

Subjects can withdraw informed consent at any time during the trial.

5.4. Procedures

5.4.1. Image Accrual

The subjects will be recruited to either the biopsy cohort or the screening cohort as discussed in section 5.1. If the inclusion/exclusion criteria are met for either cohort, the subject will be asked to

participate in the study. The enrollment procedure will include study brief, inclusion/exclusion criteria, informed consent, randomization and data base entry (using a CRF). If the subject has no prior 2D screening exam, she will be randomized to either be imaged first on the investigational device and thereafter on the FFDM system, or the other way around. Imaging will be acquired by a trained technologist at site. Data management will be performed by a study assistant from each site; this includes anonymization, export, updated CRF etc.

5.4.2. **Image Reading**

The readers will independently review Philips MicroDose Tomosynthesis images and FFDM screening mammograms in two separate sessions using dedicated softcopy review workstations. The reading sessions will be separated by a wash-out period of at least one month. There will be equal numbers of cancer cases and equal mixture of FFDM and Tomosynthesis exams in each group.

Each reader will work at a designated workstation and will complete the case report forms independently. Readers will read in private reading rooms with very low ambient lighting. A study assistant will be available in the lab but not in the reading room to monitor each reader's progress and to resolve any issues that are unrelated to the actual image interpretation.

The review of Tomosynthesis will include reading of the Tomosynthesis (CC and MLO) plus synthetized 2D images (CC and MLO) generated from the tomosynthesis images. The review of FFDM exams will include reading the standard bilateral 2-view mammograms (CC and MLO). No patient history or other clinical information will be provided to the readers. Each reader will have a case lists displaying images in the prescribed order for each reader session.

The reader will review the images for a given modality then will follow the reporting workflow depending upon whether they would recall the patient or not based on the imaging data presented.

The Confidence Scores for recalled cases will have to fall between (51 and 100) and the confidence scores for the no recall cases will have to fall somewhere between (1 and 50). For the per-case ROC analysis, the confidence score for the case will be the overall score documented by the reader.

5.4.2.1. No Suspicious Findings (No Recall) Reporting

If the reader believes that a case contains no suspicious lesion(s), the reader will be asked to assign an overall forced BI-RADS® rating and a confidence score of the case being normal (1 to 50). The score 1 represents the least suspicious No-Recall case, whereas the score 50 represents the least suspicious No-Recall case. The allowed forced BI-RADS ratings are 1 and 2 for all cases that the radiologist interprets as containing NO suspicious, including benign lesion(s) that he or she is not recommending for additional imaging after screening.

5.4.2.2. Suspicious Findings (Recall) Reporting

For each of the image sets, the reader will examine the images and look for any area of concern. For each suspicious lesion(s) triggering the need to recall a patient, the reader will localize the lesion(s) by providing the laterality and an o'clock location, assign a forced BI-RADS® rating, and the confidence score of the lesion being malignant (CS) (51 to 100) will be reported. The score 51 represents the least suspicious recall case, whereas the score 100 represents the most suspicious recall case. The reader will also provide the lesion type and size. After marking and rating all lesions, the reader will be asked to assign an overall BI-RADS® rating and CS score for the whole case. The allowed forced BI-RADS® ratings for suspicious lesion(s) are 3, 4, or 5. BI-RADS® 0 rating is not allowed in this study.

The annotations of suspicious lesion(s) will be made on the softcopy review workstation using the software features of case review application.

5.4.2.3. Matching Reader Indicated Lesion Locations to Truth

For all cancer cases, the lesion location(s) provided by the readers will be assessed for correspondence to the known cancer locations. For purposes of analysis, the cases which the reader assigns a score of BI-RADS 3 or higher and which correspond to the known cancer location (as described below) will be considered as correctly localized.

5.4.2.4. Correct Localization:

A reader will get credit for correct localization of a cancerous lesion if the laterality matches the true laterality of the cancer and any of the three rules below apply:

- Reader recorded lesion at a specific o'clock location anywhere from 1 to 12 is the same as the o'clock location documented by the truther within 2 o'clock ticks.
- Lesions indicated by the truther as being located in the Subareolar region that is reported as Subareolar by the reader.
- Lesions indicated by the truther as in the Axillary Tail region must be reported in either the Axillary Tail or 10, 11, or 12 o'clock (right breast) or 12, 1, 2 (left breast) by the reader.

5.4.3. **Case selection - Image Reading**

Cases for the reader study will be selected from all of the cases acquired through the case acquisition phase. A total of approximately 65 cancers, 40 cases with benign biopsies and 300 normal cases will be selected in sequential order to fulfill the requirements for case stratification. The 300 normal cases will include some cases that were recalled from screening but ultimately determined to be normal after additional work-up, as well as cases that were not recalled from screening. Approximately 10% additional cases will be included to protect from loss to follow-up prior to 1 year status verification.

Case stratification is described in detail in 5.4.6. In summary, cancer cases will be stratified by lesion type and size, and breast density. Benign cases will be stratified by lesion type and breast density. Normal cases will be stratified by breast density and by the recommendation for additional work-up after screening. Once the target case number is reached data collection will stop.

Approximately 18 American Board of Radiology (ABR) certified and Mammography Quality Safety Act (MQSA) qualified radiologists will review and score the cases according to Image Review methods described Image Reading, Section 5.4.2. All readers will undergo training on interpretation using the Philips MicroDose tomosynthesis system images to become familiar with the images and the review workstation to be used in the study.

To avoid recall bias, the reading will be divided in two sessions structured using a crossover design with a washout period of one month between sessions as presented in Figure 3.

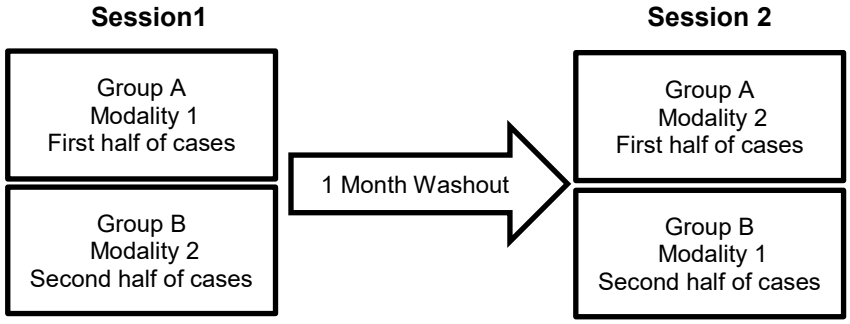

Figure 3: Reading Scheme

Each reader will independently review 2D mammograms and MicroDose images (i.e., Tomosynthesis and synthetic 2D) of the same patient in separate sessions to minimize memory effects. The cases will be randomly divided into two groups, Group A and Group B. In the first reading session, the readers will read Group A with modality 1 and Group B with modality 2. In the second reading session, the readers will read Group A with modality 2 and Group B with modality 1. Within each reading session, the order in which the cases will be read will be randomized to avoid confounding the device effect with reading order effects.

Each reading session may be divided into several visit times to adapt to the readers' schedule. Furthermore, to avoid bias, no patient history or other clinical information will be provided to the reader.

5.4.4. **Reader Selection**

The readers will be a mixture of general radiologists and breast imaging subspecialists to reflect the pool of potential users in clinical practice. All readers will be MQSA-qualified, ABR Board-certified radiologists. The readers' experience level with reading tomosynthesis images will be documented.

The following inclusion criteria must be met for readers:

- MQSA qualified radiologist
- ABR Board Certified
- At least 2 years in clinical practice including interpretation of digital mammograms
- All potential readers must complete the Philips tomosynthesis training and test session in order to be a valid reader in the actual study.

The radiologists will provide the following documentation in support of determining their eligibility reader study coordinator and maintained in the reader's study folder:

- Curriculum Vitae
- Medical License
- Qualification Survey which will record the name and address of their radiology practice, the years of experience in interpreting digital mammograms.

Documentation of completion of Philips tomosynthesis training by each reader will be included in the reader's study folder kept on site with a copy in the Philips study master file.

5.4.5. **Reader Training**

The goal of the reader training sessions is to assure that the readers are sufficiently trained in the interpretation of Philips tomosynthesis images. Training will be conducted by an expert radiologist with an application specialist experienced in the softcopy review workstation operation will be on hand to assist readers with use of the review workstation in order for the readers become familiar with its functionality.

The training will consist of two parts, Part 1: an initial overview of tomosynthesis and synthetic 2D presented by radiologist training with subsequent discussion of a series of 20 training cases and Part 2: Independent review by the reader of 30 additional cases.

5.4.5.1. Part 1 Training

The training data set will consist of 10 biopsy-proven cancer cases (5 calcifications, 5 mass) and 10 biopsy-proven benign lesions (5 calcifications, 5 mass) cases and 10 normal cases. During the initial review, the cases will be presented side by side for comparison of tomosynthesis and synthetic 2D and FFDM images. The trainer will point out the range of typical anatomy and abnormalities identified with tomosynthesis. Training will also include protocol procedure training on reading images from

tomosynthesis and FFDM systems, review station software tools, instructions for use of the Confidence Scales, instruction regarding completion of the Case Report Forms (CRFs) and a hands-on session at the review workstation to provide the Readers with an overview of their functionality and standard image review tools for tomosynthesis and FFDM viewing. Reader trainees will also be given written instructions on image interpretation, which they may refer to during the training sessions.

5.4.5.2. Part 2 Training

After this initial training, all readers will required to independently review and score thirty tomosynthesis and synthetic 2D case sets. These sets will consist of 10 malignant cases, 10 benign cases, and 10 normal cases and complete the CRFs. Following the independent read of this case set, a study assistant will review the reader's completion of all CRFs in training to make sure they understand the process.

No cases used in training will be used in the reader study.

5.4.6. Case Selection

5.4.6.1. Selection of Normal Cases

A normal – no recall case is defined as a case classified as BIRADS 1 or 2 by initial screening clinical reporting from the sites with confirmation that no cancer has been diagnosed after 1 year of follow-up past enrollment to the study. A normal – recall case is defined as a case classified as BIRADS 1 or 2 after completion of the diagnostic workup following an abnormal screen triggering recall.

A total of approximately 300 normal cases, 270 normal-no recall and 30 normal-recall will be sequentially selected by the statistician from all cleared cases which approximately represents the distribution of breast densities in the general population, including almost entirely fatty, scattered fibroglandular elements, heterogeneously dense and extremely dense breast as presented in Table 2.

Stratification criteria	*Target*	*Min (-30%)*
Breast density		
A: Almost entirely fat	*11% / 32*	*22*
B: Scattered fibroglandular densities	*43% / 129*	*90*
C: Heterogeneously dense	*39% / 117*	*82*
D: Extremely dense	*7% / 22*	*16*

Table 2: Stratification Criteria for Normal Cases (n=300) in accordance with [6]

5.4.6.2. Selection of Cases that are benign at biopsy

In addition, cases that were benign at biopsy will be selected among those cases that had benign proven biopsy. A biopsy benign case is defined as a case for which the subject undergoes biopsy of one or more breast lesions that were all determined pathologically to be benign with confirmation of non-cancer status after 1 year of follow-up past enrollment of the study.

27 benign cases (by imaging) will be sequentially selected by the statistician from all cleared cases which approximately represents the distribution of breast densities in the general population, including almost entirely fatty, scattered fibroglandular elements, heterogeneously dense and extremely dense breast. Both mass and calcifications findings type will be included.

13 biopsy benign cases will be sequentially selected by the statistician from all cleared cases which approximately represents the distribution of breast densities in the general population, including almost

entirely fatty, scattered fibroglandular elements, heterogeneously dense and extremely dense breast. Both mass and calcifications findings type will be included. One year follow-up will be documented prior to the final statistical analysis. The stratification for benign cases is presented in Table 3.

Stratification criteria	*Target*	*Min (-30%)*
Lesion type		
Calcification	*43% / 17*	*12*
Mass	*57% / 23*	*16*
Breast density		
A: Almost entirely fat	*11% / 4*	*3*
B: Scattered fibroglandular densities	*43% / 17*	*12*
C: Heterogeneously dense	*39% / 16*	*11*
D: Extremely dense	*7% / 3*	*2*

Table 3: Stratification criteria Benign Cases (n=40)

5.4.6.3. Selection of Cancer Cases

The final classification of a cancer case for potential inclusion in the reader study will be performed by an expert radiologist. The expert radiologist will be provided with the Philips MicroDose tomosynthesis system images, digital mammograms, and relevant clinical data including confirmation of a positive biopsy for cancer. The expert radiologist will record the visibility and location of the lesion on the CRF.

65 positive cases will be selected to represent an enriched sample similar to the distribution of cases in a screening population which might be seen in clinical practice in the U.S. regarding lesion type, breast density distribution, and lesion size as presented in Table 4.

Stratification criteria	*Target*	*Min (-30%)*
Lesion type		
Calcification	*37% / 24*	*17*
Mass	*48% / 31*	*22*
Architectural distortion or asymmetry	*15% / 10*	*7*
Breast density		
A: Almost entirely fat	*11% / 7*	*5*
B: Scattered fibroglandular densities	*43% / 28*	*20*
C: Heterogeneously dense	*39% / 25*	*18*
D: Extremely dense	*7% / 5*	*3*
Lesion size		
Minimal cancer (≤ 10 mm or DCIS)	*52.7% / 34*	*24*

Table 4: Stratification criteria CANCER CASES (n=65)

5.5. Endpoints

5.5.1. Clinical Endpoint: Image Accrual Phase

For all participating subjects, the following information, as applicable will be collected during the imaging acquisition and recorded on the case report form (CRF):

- Demographics: Date of informed consent, date of birth, race, ethnicity, and menopausal status
- Medical History
- Breast Imaging-Reporting and Data System (BI-RADS) score will be performed by the interpreting radiologists for both the Philips MicroDose tomosynthesis mammograms and the FFDM. The following ratings will be used:
  - 0 – Incomplete,
  - 1 – Negative,
  - 2 – Benign finding(s),
  - 3 – Probably Benign,
  - 4 – Suspicious abnormality,
    - 4 a - Low suspicion for malignancy
    - 4 b – Moderate suspicion for malignancy
    - 4 c – High suspicion for malignancy
  - 5 – Highly suggestive of Malignancy
- Screening Population cases will be classified as No Recall if both Tomo and FFDM exams have final assessment of BIRADS 1 or 2. Screening population cases will be classified as Recall if either Tomo or FFDM exam is classified as BI-RADS 0, requiring additional diagnostic imaging workup.
- Screening population cases that were recalled, 0, for additional diagnostic work-up must have a final BIRADS assessment of 1, 2, 3, 4a, 4b, 4c, or 5.
- The study imaging standard exam (Tomo or FFDM) that is performed prior to biopsy in the biopsy cohort arm must be interpreted and assigned a final BIRADS assessment of 1, 2, 3, 4a, 4b, 4c, or 5. Breast density score: The breast density score will be performed at the acquisition site by at least one radiologists using the standard of care FFDM images. Cases will be categorized using BI-RADS breast density scale as either non-dense (BI-RADS breast density type A or B) or dense (BI-RADS breast density type C or D).
- Whether the case is scheduled for a screening mammogram or a biopsy
- Lesion type: Lesion type will be defined at the acquisition site using BI-RADS scale and classified as mass, calcifications, asymmetry, or architectural distortion.
- Biopsy results
- Images from the screening examination for the screening cohort

5.5.2. Clinical Endpoint: Image Reading Phase

The Readers will independently review Philips MicroDose Tomosynthesis images and FFDM screening mammogram for each case in two separate reading sessions separated by a wash-out period of at least one month. The cases to be included in the reader study will be divided into two groups with an equal number of cancer cases and an equal number of FFDM and tomosynthesis exams in each group.

The review of FFDM exams will include reading the standard bilateral 2-view mammograms (CC and MLO). The review of Tomosynthesis will include reading of the Tomosynthesis (CC and MLO) plus synthetized 2D images (CC and MLO) generated from the tomosynthesis images. No patient history

or other clinical information will be provided to the Readers. The following data points will be collected using case report forms which will be completed directly by the readers:

- Forced BI-RADS score of 1, 2, 3, 4a, 4b, 4c, or 5
- Confidence Score (CS): Readers will be asked to assign a confidence score for the BI-RADS ratings as follows: For BI-RADS 3 to 5, the confidence score represents the reader's interpretation of the likelihood of the lesion being malignant on a scale of 51 to 100. The score 51 represents the least suspicious Recall case, whereas the score 100 represents the most suspicious Recall case. For BI-RADS 1 or 2, the confidence score represent the reader´s confidence that the case is normal and will range from 1 to 50. The score 1 represents the most normal No-Recall case, whereas the score 50 represents the least normal No-Recall case. For the per-case ROC analysis, the confidence score for the case will be the overall score documented by the reader.
- Readers will mark all lesions by location (breast and clock location).
- For the FROC analysis, each lesion or foci identified by the reader will rated using a confidence score that the abnormality in question representing a cancer "if it was actually present". For this analysis confidence score will range from 1 to 100. A research associate (RA) located at the reader study facility will check the progress of the readers during each reading session. The RA will also check for missing data elements on the form and check all cases for completeness for each group of cases during the reading session. The RA will generate a query list to resolve any missing data points or cases prior to the reader ending a session. The RA will also assist readers with workstation problems and CRF issues should they arise.

5.5.3. **Procedure Endpoint**

- Radiation dose: Average Glandular radiation dose will be read out from the DICOM header for each mammogram.
- Image interpretation time: Interpretation time will be measured during the reading study for each reader. The interpretation time includes the time the reader takes reviewing the case and completing the case report form.

5.5.4. **Safety Endpoint**

- All expected or unexpected adverse events will be reported.

5.6. Monitoring Plan

Monitoring will be performed by a trained person appointed by Philips to ensure compliance with the Clinical Investigation plan, applicable national regulations and international standards, patient safety and data validity. The Sponsor may designate one or more individuals to monitor the progress of a clinical study. The Sponsor may also delegate the monitoring responsibilities to a third party. However, the Sponsor remains ultimately responsible for the conduct of the study. The Institution is responsible for the appropriate de-identification of subject data. The investigational site should provide access to the source data of the subjects.

The monitor selected by the Sponsor will visit each site prior to study start and periodically during data collection phase. During each visit the monitor will review subject records including informed consent forms and case report form data. After each visit an on-site report will be written. A close-out visit for sites that have enrolled subjects will be conducted once the site has completed collecting data for the study.

The monitor will ensure that all adverse device effects (anticipated and unanticipated) and complaints are reported to the sponsor in a timely manner and that unanticipated adverse device effects are reported to the IRB.

A detailed plan for monitoring arrangement will be available separate from this protocol.

Names of the monitor(s) can be found in *Appendix II: List of monitor(s)* of this protocol. An update of this list can be provided to the site under separate cover.

5.7. Deviations from the clinical plan

The Investigator is not allowed to deviate from the Clinical Investigation Plan or to enroll subjects that do not comply with all inclusion and exclusion criteria. Under emergency circumstances, deviations from the Clinical Investigation Plan to protect the rights, safety and well-being of human subjects may proceed without prior approval of the sponsor and the EC. Such deviations shall be documented and reported to the sponsor and the EC as soon as possible.

All deviations from the Clinical Investigation Plan will be documented with date, subject, reason, actions taken and if the deviation affects subject's rights, safety and wellbeing or the scientific integrity of the clinical study. The deviation shall be notified to the Sponsor as soon as possible via the CRF. Deviations will be reviewed by the sponsor and in case of serious or repetitive deviations a corrective action plan may represent a need to initiate a corrective action plan with the principal investigator. In some cases, necessitate suspension of enrollment at the site or ultimately the principal investigator will be disqualified.

6. STATISTICAL CONSIDERATIONS

6.1. Sample Size and Number of Readers Justification

6.1.1. Sample Size Justification for Primary Efficacy Analysis

The primary endpoint is the Receiver Operator Characteristics curve analysis for both tests. In this trial, the 2D is the control test, and the tomosynthesis and synthetic 2D is the experiment test. The plan is to enroll approximate 800 patients, to include 65 patients with cancer, 40 biopsy-proven benign cases and 300 normal cases. The evaluations will be based on paired-cases and paired-readers that is all readers will interpret the test results of all patients in both tests (i.e., Philips MicroDose tomosynthesis system and FFDM).

The sample size (number of readers) for the Reader Study is estimated based mixed-effects model for the diagnostic accuracy of a reader *j* using a particular diagnostic test *i* on reading occasion *q* as proposed by Obuchowski [11]. The mixed model is specified as follows:

$$\hat{\theta}_{ijq} = \mu + \mu_i + r_j + (ur)_{ij} + \epsilon_{ijq}$$

Here, $v_i = \mu + \mu_i$ is the mean accuracy of test *i* (fixed effect), r_j is a random effect for reader *j*, $(ur)_{ij}$ is an interaction term for readers and tests (random effect), and ϵ_{ijq} is an error term.

An F-statistic with 1 and (*J*-1) degree of freedom (df) will be used for testing the null hypothesis that the mean diagnostic accuracies of the tests are equal, where *J* is the number of readers. Under alternative hypothesis, the test statistic follows a non-central F distribution with 1 and (*J*-1) df, and the non-centrality parameter is estimated as following:

$$\hat{\lambda} = \frac{J(\mu_1 - \mu_2)^2}{2\{\hat{\sigma}_b{}^2(1-\hat{\rho}_b) + \frac{\hat{\sigma}_w{}^2}{Q} + \hat{\sigma}_c{}^2[(1-\hat{\rho}_1) + (J-1)(\hat{\rho}_2 - \hat{\rho}_3)\}}$$

Thus, to calculate the statistical power, multiple parameters in the equation above need to be estimated. For this study, these estimates are based on the following assumptions:

- The type I error: α = 0.05
- AUC for control test: A = 0.85
- Number of positive patients: n_1 = 65
- Number of negative patients: n_2 = 300
- Non-inferiority margin: Δ = 0.05
- Variability between readers: σ_b^2 = 0.0009
- Variability within the reader: σ_w^2 = 0.000225
- Correlations: ρ_1 = 0.46, $\rho_2 - \rho_3$= 0
- Correlation: ρ_b = 0.5
- Power for the non-inferiority test: 1 - β ≥ 0.8

The variance between different patients σ_c^2 will be calculated based on the method of Obuchowski (1994). It is suggested that for the ROC analysis, σ_c^2 can be obtained as follows.

$$\sigma_c^2 = \frac{0.0099 \times e^{-\frac{a^2}{2}} \times [5a^2 + 8 + (a^2 + 8)/R]}{n_1}$$

Where $a = \Phi^{-1}(\mathrm{A}) \times 1.414$ and Φ^{-1} is the inverse of the cumulative normal distribution function; n_1 is the number of positive patients; and R is the ratio of the number of negative patients (n_2) to positive patients (n_1).

The assumed AUC for control test (A) is 0.85. In a previous study on the Philips L30 system, the area under the ROC curve was 0.947, with a 95% C.I. (0.920, 0.974)[19]. In another similar study for the Hologic Tomosynthesis Mammography System, the AUC based on Probability of Malignancy for a 2D system is 0.867 [20]. Therefore for this study it is expected that the AUC for the control test will be approximately 0.85.

The variance between readers interpreting the same patients σ_b^2 is based on a study performed by Beam, Layde and Sullivan [16]; and it was widely used in previous studies. In a previous study on Philips L30 system [19], the estimated σ_b^2 with 95% C.I. is 0.0005 (0.0003, 0.0013), thus a value of 0.0009 is a reasonable estimation for σ_b^2 in the proposed study.

The variance within each reader interpreting the results of the same patients σ_w^2 is estimated as one-fourth of σ_b^2, thus 0.00025.

The four correlations are usually estimated from a pilot study or previous studies. Correlation for the same patients evaluated by the same reader using different tests (ρ_1), for the same patients evaluated by different readers with the same tests (ρ_2) and for the same patients evaluated by different readers and different tests (ρ_3) are based on suggestions from Obuchowski [11]. The correlation for the same reader to evaluate patients using different tests ρ_b ranges from 0.44 to 0.86, based on Rockette et al. (1999). It is assumed that a correlation equal to 0.5 is a conservative estimate.

Based on these values, 18 readers will be required to achieve statistical power of at least 0.80.

6.1.2. **Sample Size Justification for Average Glandular Dose**

Another secondary objective of this study is to compare the average radiation dose between the Philips MicroDose tomosynthesis system and FFDM systems. Radiation dose will be evaluated based on the subject overall dose received during the examination. For the null hypothesis, no difference in radiation dose delivered to the patient during a mammographic exam between two systems, will be evaluated.

Using historical data from the study site, the average radiation dose per breast for 20 subjects, 2 sample (MLO + CC) was of 3.6 (standard deviation of 1.1), ranging from 2.4 to 6.7 for the GE FFDM, and of 2.6 (standard deviation of 0.7) ranging from 1.6 to 4.0 for Hologic mammography system. Also, in a previous study, comparing the three systems reported radiation dose per breast of 1.86, 2.91, and 3.03, for Philips MicroDose, Hologic and GE, respectively representing a reduction of radiation of approximately 40% for MicroDose compared to Hologic and or the GE mammography system.

Assuming that the standard deviation for the difference in radiation dose will be approximately 1.5mGy, a sample size of approximately 100 subjects will allow for greater than 80% power to detect a difference in radiation between the two systems of approximately 0.5mGy (i.e., 10% of the overall patient radiation) using a paired t-test with a 0.025 one-sided significance level.

6.1.3. **Sample Size Justification for Recall Rate**

Based on the outcome of the most recent tomosynthesis PMA (i.e., MAMMOMAT Inspiration with Tomosynthesis, P140011), the recall rate for A FFDM alone was 0.438. For this study, a similar recall rate is expected. Assuming that a 5% reduction in recall rate for tomosynthesis is clinically significant and adjusting for the multiple endpoints, a total of 2300 cases will allow for approximately 80% to detect an improvement in recall rate of approximately 5%. Furthermore, assuming that the correlation among readers reading the same case in approximately 0.5, and that all subject will be read by all 18 readers, approximately 250 subjects will be required.

6.1.4. **Sample Size Justification for Sensitivity Non-inferiority Analysis**

For the sensitivity analysis, the sample size estimation is based on the following assumptions:

- The type I error: $\alpha = 0.05$
- Sensitivity for control test: Se = 0.85
- Number of positive patients: $n_1 = 65$
- Non-inferiority margin: $\Delta = 0.10$
- Variability between readers: $\sigma_b^2 = 0.009$
- Variability within the reader: $\sigma_w^2 = 0.00225$
- Correlations: $\rho_1 = 0.46$, $\rho_2 - \rho_3 = 0$
- Correlation: $\rho_b = 0.8$
- Power for the non-inferiority test: $1 - \beta \geq 0.8$

The variance between different patients σ_c^2 will be calculated based on a binomial approximation as follows:

$$\sigma_c^2 = \frac{S_e(1 - S_e)}{n_1}$$

Where S_e is the expected sensitivity and n_1 is the number of positive patients (i.e., patients with cancer).

In a previous study on the Philips MicroDose L30 system, the estimated sensitivity was 0.936, with a 95% C.I. (0.897, 0.976) (see Statistics Collaborative, Inc., 2011). In another similar study for the Hologic Tomosynthesis Mammography System, the sensitivity for a 2D system was 0.862 Hologic, Inc. 2012). Based on these previous results, it is expected that the sensitivity for both tests will be approximately 0.85.

In a previous study on Philips MicroDose L30 [19] the estimated σ_b^2 with 95% C.I. was 0.0035 (0.0019, 0.0084). In another similar study for the Hologic Tomosynthesis Mammography System [20] the estimated σ_b^2 with 95% C.I. was 0.0046 (0.0024, 0.0114). Thus a value of 0.009 is a reasonable estimation for σ_b^2 in the proposed study.

The variance within each reader interpreting the results of the same patients σ_w^2 it is assumed to be one-fourth of σ_b^2, thus 0.0025.

Correlation for the same patients evaluated by the same reader using different tests (ρ_1), for the same patients evaluated by different readers with the same tests (ρ_2) and for the same patients evaluated by different readers and different tests (ρ_3) are based on suggestions from Obuchowski (2000). The correlation for the same reader to evaluate patients using different tests ρ_b ranges from 0.44 to 0.86 in AUC analysis, according to Rockette et al. (1999 such correlation are usually higher for sensitivity analysis. Thus, it is expected that the value ρ_b will be approximately 0.80 for this study.

Based on these assumptions, for the sensitivity analysis, 18 readers will be required to achieve statistical power of at least 0.80, after adjusting for multiple endpoints.

6.1.5. **Sample Size Justification for Specificity Non-inferiority Analysis**

For specificity analysis, similar to sensitivity analysis, the following assumptions were made:

- The type I error: α = 0.05
- Specificity for control test: S_p = 0.70
- Number of positive patients: n_2 = 300
- Non-inferiority margin: Δ = 0.10
- Variability between readers: σ_b^2 = 0.012
- Variability within the reader: σ_w^2 = 0.003
- Correlations: ρ_1 = 0.46, $\rho_2 - \rho_3$= 0
- Correlation: ρ_b = 0.8
- Power for the non-inferiority test: 1 - β ≥ 0.8

Similar to sensitivity analysis, the variance between different patients σ_c^2 will be calculated based a binomial assumption, as described above.

In a previous study on the Philips MicroDose L30, the estimated specificity was 0.764, with a 95% C.I. (0.688, 0.841) [19]. In another similar study for the Hologic Tomosynthesis Mammography System, the specificity for a 2D system was 0.651 [20]. Based on these previous results, it is assumed that the specificity for the control test is approximately 0.70.

Also, in a previous study on Philips MicroDose L30 [19], the estimated σ_b^2 with 95% C.I. was 0.0087 (0.0047, 0.0208). In another similar study for the Hologic Tomosynthesis Mammography System [20], the estimated σ_b^2 with 95% C.I. was 0.0144 (0.0077, 0.0357). Thus a value of 0.012 is a reasonable estimation for σ_b^2 for the proposed study.

The variance within each reader interpreting the results of the same patients σ_w^2 is estimated as one-fourth of σ_b^2, thus 0.003.

Correlation for the same patients evaluated by the same reader using different tests (ρ_1), for the same patients evaluated by different readers with the same tests (ρ_2) and for the same patients evaluated by different readers and different tests (ρ_3) are based on suggestions from Obuchowski [11]. The correlation for the same reader to evaluate patients using different tests ρ_b ranges from 0.44 to 0.86 in AUC analysis, according to Rockette [18] such correlation are usually higher for sensitivity/specificity analysis. Thus, it is safe to assume the expected value ρ_b to be 0.80 in the proposed study.

Based on these assumptions, for the specificity analysis, 18 readers will be required to achieve statistical power of at least 0.80, after adjusting for multiple endpoints.

6.2. General Statistical Considerations

The primary analysis will be performed including all subjects selected for the reading portion of the study with confidence score for both Philips MicroDose tomosynthesis system and FFDM evaluation (modified intent to treat, MITT). For this analysis patient will be the basis for the analysis. Non-cancer cases, with missing one year follow-up will be assumed as cancer free case. Missing values will be assumed to be missed at random.

The analysis of safety includes all enrolled subjects in the acquisition portion of the study.

Continuous variables will be summarized using the number of non-missing observations, mean, standard deviation, 95% Confidence Interval (CI) for the mean, median, minimum, and maximum; categorical variables will be summarized using the frequency count and the percentage of subjects in each category. Analyses will be conducted using SAS® software and OR-DBM MRMC software [12].

6.3. Subject Disposition

Subject disposition, including total screened, enrolled in the acquisition and reading portion of the study, early terminations and withdrawals, will be presented. In addition, a listing will be provided with the reasons for discontinuation.

6.4. Demographics and Baseline Characteristics

Standard subject baseline characteristics (e.g., age, race and origin) will be summarized for all subjects enrolled in the imaging acquisition and imaging reading portion of the study. For continuous subject characteristics, means, standard deviation, median and range will be provided. The number of events and percentage of events will be presented for all categorical variables.

6.5. Readers Compliance

Reader compliance will be listed by image type. If applicable, comparability between the two image types (i.e., Philips MicroDose tomosynthesis system and FFDM) groups will be evaluated using a Chi-square or Fisher's exact test as appropriate. A listing of all deviations will be presented.

6.6. Primary Efficacy Analysis

The primary objective of the study is to compare the clinical performance of the Philips MicroDose tomosynthesis system (tomosynthesis with a synthetic 2D) images and 2D FFDM system. The primary endpoint is the Area under the ROC curve (AUC) using patient as the unit of analysis. The ROC curves will be generated using the confidence scores (1 to 100) and the reference standard (i.e., normal/benign or cancer). All subjects with both Philips MicroDose tomosynthesis system and FFDM evaluation (modified intent to treat, MITT). For subjects with negative outcome for whom a 1 year follow-up is required, missing values will be removed from the study database.

Smooth ROC curves and area under the curve will be estimated from the CS for each study reader interpreting images acquired with the Philips MicroDose tomosynthesis system and the FFDM

systems. If smooth ROC curves cannot be fitted for each reader in each condition (e.g., if there is perfect performance), nonparametric methods may be used. Results will be summarized across readers using the method of [10] which includes random effects for study readers and uses the jackknife to handle correlation within cases, as obtained in the corresponding OR-DBM MRMC software. Robustness of the statistical method may be explored through use of the ANOVA model of Obuchowski and Rockette [11]; this model includes random effects for study readers, and adjusts variances and hypothesis tests for correlations due to having each reader interpret the same case sample. The degrees of freedom may be adjusted as suggested by Hillis [14]. The average AUC for Philips MicroDose tomosynthesis system and the FFDM, the average difference between the two systems, and its standard error for use in computing confidence interval and to performing hypothesis tests for the difference will be estimated.

The following hypothesis will be evaluated:

- Null Hypothesis: Ho: $\Theta_2 - \Theta_1 \geq 0.05$
- Alternative Hypothesis: $\Theta_2 - \Theta_1 < 0.05$

Where Θ_1 is the average AUC for readers using the Philips MicroDose System (tomosynthesis + synthetic 2D) and Θ_2 is the average AUC for readers using the FFDM System.The Philips MicroDose tomosynthesis system will be considered non-inferior to the FFDM if the upper limit one-sided 95% CI for the difference in AUCs is less than 0.05. AUC for Philips MicroDose tomosynthesis system and the FFDM will be also be presented by readers.

6.7. Effect of Missing Values on Primary Efficacy Analysis

To evaluate the effect of possible missing assessment for the reference standard (e.g., subjects does not return for one 1 year follow-up exam) sensitivity analysis will be performed to with missing values assigned as negative outcome, n – 1 assumed to be negative and 1 positive in random order, or excluded from analysis.

6.8. Secondary Efficacy Analysis

The secondary objectives are:

- To compare the average glandular dose for exams acquired on Philips MicroDose Tomosynthesis system and the 2D mammography system.
- To compare the non-cancer recall rate for Philips MicroDose Tomosynthesis system and 2D mammography exams.
- To compare the sensitivity and specificity of the Philips MicroDose tomosynthesis system (i.e., Tomosynthesis and synthetic 2D) to that obtained with 2D mammography.

Secondary objectives will only be tested if the null hypothesis for the primary endpoint is rejected. Four secondary endpoints will be evaluated. To control the Type I error, the following pre-specified order using a gatekeeper procedure as described by Dmitrienko et al. (2007) will be used:

Family	Endpoint	Procedure
F1	AUC ROC	Test at the 0.05 level, no adjustment to the p-value
F2	Radiation Dose	Test only if Ho for AUC ROC is rejected, p-value adjusted using HOLM procedure
	Recall Rate	Test only if Ho for AUC ROC is rejected, p-value adjusted using HOLM procedure
F2	Sensitivity	Test only if any Ho for F2 is rejected, p-value adjusted using HOLM procedure
	Specificity	Test only if any Ho F2 is rejected, p-value adjusted using HOLM procedure

The average glandular dose (i.e., radiation dose) refers to the total radiation dose during the exam and is expressed in mGy. The goal is to compare total radiation dose between the Philips MicroDose tomosynthesis system and the FFDM systems. The following hypothesis will be tested:

The average glandular dose (AGD) will be recorded per subject for each breast (right and left) and view (CC and MLO). The following hypothesis will be evaluated:

- Null Hypothesis: Ho: $\mu_2 - \mu_1 > \delta$
- Alternative Hypothesis: $\mu_2 - \mu_1 \leq \delta$

where μ_1 and μ_2 are the mean average glandular dose (i.e., radiation dose) for the Philips MicroDose Tomosynthesis system and FFDM system, respectively, and δ is 10% of the FFDM radiation dose.

Analysis will be performed using an appropriate analysis of variance. If the required assumptions for the ANOVA are violated, transformation of the data or non-parametric procedures may be used. The 95% confidence interval for the mean difference in radiation dose between the two systems will be constructed. Under this Under this proposal Philips tomosynthesis will be declared as emitting less radiation then the FFDM if the upper bound of the 95% confidence interval for the difference in radiation dose is less than 10% of the FFDM.

The non-cancer recall rate will be recorded per subject. All negative screening, recalled and benign biopsy cases will be included. The following hypothesis will be evaluated:

- Null Hypothesis: Ho: $\pi 2 - \pi 1 = 0$
- Alternative Hypothesis: $\pi 2 - \pi 1 \neq 0$

where π_1 and π_2 are the non-cancer recall rate for the Philips MicroDose Tomosynthesis system and FFDM system, respectively.

Recall rate will be analyzed using the concept of clustered binary data as described by Zhou et al (2011). In this case, each Reader will be considered as multiple observations within the same subject (cluster). The 95% confidence interval for the difference in recall rate will be constructed. If the upper limit to the one-sided 95% Confidence Interval for the difference (Philips minus FFDM) in recall rates among non-cancers is less than 0.05 (delta of 5%), Philips tomosynthesis will be considered as resulting in a similar recall rate to that of 2D FFDM. In addition, a bootstrapping method with replacement will be used to compare average recall rates among all Readers.

The sensitivity or specificity within a reader between the two systems will be estimated, and tested for statistical significance using binomial test for paired data. Results will be summarized across readers using the ANOVA model of Obuchowski and Rockette [11], which applies to any metric of diagnostic accuracy under general conditions. The following hypothesis will be evaluated:

- Null Hypothesis: Ho: $S_2 - S_1 \geq 0.10$

- Alternative Hypothesis: $S_2 - S_1 < 0.10$

Where S_1 is average sensitivity/specificity for readers using the Philips MicroDose tomosynthesis system and S_2 is the average sensitivity/specificity for readers using the FFDM. The Philips MicroDose tomosynthesis system will be considered non-inferior to the FFDM if the upper limit one-sided 95% CI for the difference in sensitivity/specificity is less than 0.10.

To evaluate the effect of breast density on ROC AUC for images obtained using Philips MicroDose tomosynthesis system and FFDM, breast density will be stratified as fatty (i.e., BI-RADS breast density type A and B) and dense (BI-RADS breast density type C and D). Analysis will be similar to the described for the primary efficacy analysis.

6.9. Supplementary Efficacy Analysis

The following supplementary objectives will be evaluated:

- To evaluate the effect of breast density on ROC AUC for images obtained using Philips MicroDose tomosynthesis system and FFDM system. Breast density will be stratified as fatty (i.e., BI-RADS breast density type 1 and 2) and dense (BI-RADS breast density type 3 and 4).
- To compare the average image interpretation time for exams performed with Philips MicroDose SI Tomosynthesis and the FFDM system.
- To assess the diagnostic performance of the Philips MicroDose Tomosynthesis compared to FFDM

To evaluate the effect of breast density on ROC AUC for images obtained using Philips MicroDose tomosynthesis system and FFDM, breast density will be stratified as fatty (i.e., BI-RADS breast density type A and B) and dense (BI-RADS breast density type C and D). Analysis will be similar to the described for the primary efficacy analysis.

The diagnostic performance of the Philips MicroDose Tomosynthesis compared to the FFDM will be assessed using the free-response receiver operating characteristic (FROC) paradigm. The analysis will use Confidence score for each suspicious focus, along with its location relative to any biopsy-proven malignancy; and will allow multiple suspicious foci per-case. Analysis will be performed using JAFROC (jackknife AFROC) as described by Chakraborty and Berbaum (2004). Truth will be established by a MQSA qualified radiologist.

7. DATA MANAGEMENT

7.1. Access to Source Documents

Study records, including each subject's signed informed consent, and other study-related documents pertaining to the conduct of the study will be kept in a secure area. Subject information will remain confidential. However, consent forms, medical history and image records that identify subjects may be inspected by the sponsor, its authorized designees and regulatory agencies including but not limited to, the Department of Health and Human Services (DHHS), the United States Food and Drug Administration (FDA), other foreign regulatory bodies and the Institutional Review Board/Ethics Committee for this study. Because of the need to release information to these parties, absolute confidentiality cannot be guaranteed. The results of this research project may be presented at meetings or in publications; however, subject identity will not be disclosed in such publications.

7.2. Retention period

The sponsor and investigator shall maintain the records related to this study during the investigation and for a period of 5 years after the latter of the following two dates: The date on which the Research is terminated or completed, or the date that the records are no longer required for purposes of supporting a premarket approval application in the US or a notice of completion of a product development protocol. Or longer if national regulation requires this.

The sponsor and principal investigator shall take measures to prevent accidental or premature destruction of these documents.

7.3. Data Management

Data will be collected in a case report form (CRF) and transcribed in a central database system or in an electronic CRF (eCRF). After the data are entered in the database, edit checks, programmed to ensure the collection of consistent and complete data, may be raised.

In addition, Philips (or contracted) Monitors, Project Statistician, and/or the Philips data management group will review data listings generated at different time points during the study. Data queries will be generated to resolve any discrepancies or concerns.

Additionally, Data transcribed to the CRFs will be source data verified by the Sponsor/designee(s) on a percentage of the subject population (e.g., 10% or more as appropriate per a risk-based monitoring approach). Data queries will be generated to resolve any discrepancies or concerns.

It is the responsibility of the Site designated personnel to respond to all edits checks and queries. Submitted data as well as all data modifications to submitted data will be documented by the system and available in audit trails.

Upon conclusion of the study, after all data are marked as complete and all discrepancies are resolved, the Principal Investigator will be notified to review and sign the case books. Subsequently, a member of the Philips Health Management Group or designee will lock the database. The final data set will be transferred to SAS for analysis at Philips.

8. DEVICE ACCOUNTABILITY

Access to the investigational device shall be controlled and the investigational devices shall be used only in accordance with the clinical investigational plan and the contract with the sites.

The sponsor shall keep records to document the physical location of all investigational devices from shipment of investigational device to the Research sites until de-installation.

9. INFORMED CONSENT PROCESS

Informed consent will be obtained from every subject in writing by the Investigator or his authorized designee before the clinical trial is started. The subject will be informed both orally and in writing about all aspects that are relevant to the subject's decision to participate in the trial, including the trial procedures and risks and benefits of participation in the clinical trial. Ample time should be provided for the subject to read and understand the informed consent form and to consider participation. The informed consent will include personally dated signatures of the subject and the principal investigator or an authorized designee responsible for conducting the informed consent process. A copy of the signed and dated informed consent form and any other written information will be provided to the subject.

Subjects who are unable or unwilling to provide informed consent will not be included in the trial.

If new information becomes available that might significantly affect the subject's future health and medical care, it shall be provided to the subjects in written form. If relevant, subject shall be asked to reconfirm their continuing informed consent in writing.

The patient information and consent will follow the guidelines set forth in the Code of Federal Regulations (21 CFR 50) and ISO 14155 for patient information and informed consent. To be eligible for enrollment in the study, all subjects or their authorized representatives must first understand and sign a patient information and consent form approved by the study center's Institutional Review Board / Ethics Committee or provide informed consent verbally when allowed by IRB/EC.

10. ADVERSE EVENT REPORTING

An Adverse Event (AE) is defined as any untoward medical occurrence, unintended disease or injury or any untoward clinical signs (including an abnormal laboratory finding) in subjects, users or other persons whether or not related to the investigational device.

NOTE 1: This definition includes events related to the investigational medical device or the comparator.
NOTE 2: This definition includes events related to the procedures involved.
NOTE 3: For users or other persons, this definition is restricted to events related to investigational medical devices.

An Adverse Device Effect (ADE) is defined as an adverse event related to the use of a medical device. This includes any adverse event resulting from insufficiencies or inadequacies in the instructions for use, the deployment, the implantation, the installation, the operation, or any malfunction of the medical device, and any event that is a result of a use error or intentional misuse.

An AE or ADE is classified as serious when:
- it results in death
- it is life threatening
- it requires inpatient hospitalization or prolongation of existing hospitalization
- it results in persistent or significant disability/incapacity

NOTE: Planned hospitalization for a pre-existing condition, or a procedure required by the Clinical Investigational Plan, without serious deterioration in health, is not considered a serious adverse event.

Unanticipated adverse device effect (UADE) means any serious adverse effect on health or safety or any life-threatening problem or death caused by, or associated with, a device, if that effect, problem, or death was not previously identified in nature, severity, or degree of incidence in the investigational plan or application (including a supplementary plan or application), or any other unanticipated serious problem associated with a device that relates to the rights, safety, or welfare of subjects.

10.1. Complaint

Any written, electronic, or oral communication that alleges deficiencies related to the identity, quality, durability, reliability, safety, effectiveness, or performance of a device after it is released for distribution.

A security event is defined as an observable occurrence leading to either or both of the following:
- Software or any data that are managed by a PH product are suspected of being maliciously altered, misused or lost (includes virus, worm, hackers, etc.); and/or
- A system or component has a customer-reported security vulnerability that could result in alteration, misuse or loss.

The investigator shall report any related or unrelated serious adverse (device) event within 48 hours to the sponsor and IRB/EC and shall be confirmed by using the adverse event form within one week.

Non-serious adverse (device) events shall be recorded on the adverse event forms. The Sponsor and Monitor can request access to this information at any time.

10.2. Reporting

The investigator shall report via the adverse event form in the CRF any anticipated and unanticipated adverse device effects to the sponsor in a timely manner.

The investigator shall report any complaint via the device deficiency form.

An investigator shall submit to the sponsor and to the reviewing IRB a report of any unanticipated adverse device effect occurring during the study as soon as possible, but in no event later than 10 working days after the investigator first learns of the effect.

The Sponsor and Monitor can request access to this adverse event information at any time.

Please see Appendix II: List of monitor(s) for emergency contact for reporting UADE.

The CRF will include the following information for adverse events: date of the adverse event, description, actions taken, resolution, assessment of both the seriousness and the relationship to the Philips MicroDose Tomosynthesis and procedure. Information collected for device deficiencies are: date of device deficiency, whether this could have led to a Serious Adverse Device Effect (SADE) if a) suitable action had not been taken or b) intervention had not been made or c) if circumstances had been less fortunate.

10.3. Anticipated and unavoidable adverse device effects

Anticipated adverse device effects are described below in Table 5.

Table 5: Anticipated adverse device effects

Anticipated adverse device effects	**Likely incidence**	**Mitigation or Treatment**
Pinch or cut injury	Not likely	Treatment of injury
Loss of images	Not likely	Retake of image or patient excluded from study
Collision with rotating C-arm	Not likely	Treatment of injury
Image artefacts	Not likely	Retake of image or patient excluded from study

The following unavoidable adverse events are very common according during the procedures performed in this clinical trial. Unavoidable adverse events do not need to be reported.

Table 6: Unavoidable adverse events

Unavoidable adverse event	**Time frame**
Discomfort due to compression	During examination

11. EARLY TERMINATION OR SUSPENSION OF THE RESEARCH

There are no provisions or interim analyses planned that can result in an early termination of the trial.

Any signs of unknown or increased risks for the subjects will be discussed by the sponsor and investigator to assess the impact on the subjects and clinical research. Serious or repetitive

occurrence of deviations from study protocol or non-compliance with regulations may also be reason for early termination or suspension of a study site.

12. PUBLICATION POLICY

This study will be registered on clinicaltrials.gov before first enrollment. Results from the study may be published in peer reviewed journals and presented at scientific conferences. Study sites may use data from the study for other research as long as they have prior agreement with the study sponsor and have IRB approval as appropriate.

13. REFERENCES

[1] Surveillance Research Program. Lifetime risk of breast cancer. in National Cancer Institute [database online]. Washington, DC, 2009Available from http://seer.cancer.gov/studies/surveillance/study7.html (accessed May 16, 2009).

[2] Feig, S. A., G. S. Shaber, G. F. Schwartz, A. Patchefsky, H. I. Libshitz, J. Edeiken, R. Nerlinger, R. F. Curley, and J. D. Wallace. 1977. Thermography, mammography, and clinical examination in breast cancer screening. review of 16,000 studies. *Radiology* 122, (1) (Jan): 123-7.

[3] The Swedish Organized Service Screening Evaluation Group. Duffy SW, Tabar L, Tony H.H. Chen et al. Reduction in Breast Cancer Mortality from Organized Service Screening with Mammography: 1. Further Confirmation with Extended Data. Cancer Epidemiol Biomarkers Prev 2006;15(1):45-51

[4] Skaane P, Young K, Skjennald A.: “Population-based mammography screening: comparison of screen-film and full-field digital mammography with soft-copy reading--Oslo I study.” Radiology. 2003 Dec; 229(3):877-84. Epub 2003 Oct 23

[5] Skaane P. and Skjennald A.: “Screen-Film Mammography versus Full-Field Digital Mammography with Soft-Copy Reading: Randomized Trial in a Population-based Screening Program--The Oslo II Study” Radiology, July 1, 2004; 232(1): 197 - 204.

[6] Etta D. Pisano, M.D. et al.: Diagnostic Performance of Digital versus Film Mammography for Breast-Cancer Screening, New England Journal of Medicine 2005; 353

[7] Houssami N, Skaane P.: Overview of the evidence on digital breast tomosynthesis in breast cancer detection. Breast. 2013 Apr; 22(2):101-8. doi: 10.1016/j.breast.2013.01.017

[8] Hardesty et al, Digital Breast Tomosynthesis Utilization in the United States: A Survey of Physician Members of the Society of Breast Imaging., J Am Coll Radiol. 2014 Apr 5. pii: S1546-1440(13)00807-7. doi: 10.1016/j.jacr.2013.11.025. [Epub ahead of print]

[9] Spectral tomosynthesis breast density validation - reader study protocol, DOC-JREN-9LYJYC

[10] Dorfman DD, Berbaum KS, Metz CE. Receiver Operating Characteristic Rating Analysis: Generalization to the Population of Readers and Patients with the Jackknife Method. *Investigative Radiology* 1992; 27:723-731.

[11] Obuchowski NA, Rockette HE. Hypothesis testing of diagnostic accuracy for multiple observers and multiple tests: an ANOVA approach with dependent observations. *Comm Stat* 1995; 24:285-308

[12] OR-DBM MRMC 2.4 *http://perception.radiology.uiowa.edu/Software/ReceiverOperatingCharacteristicROC/MRMCAnalysis/ORDBMMRMCDownloads/tabid/160/Default.aspx*

[13] Hillis SL, Obuchowski NA, Schartz KM, Berbaum KS. A comparison of the Dorfman-Berbaum-Metz and Obuchowski-Rockette Methods for receiver operating characteristic (ROC) data. Statistics in Medicine 2005; 24:1579-1607.

[14] Hillis SL. A comparison of denominator degrees of freedom methods for multiple observer ROC analysis. *Statistics in Medicine* 2007; 26:596-619

[15] Hillis SL, Berbaum KS, Metz CE. Recent developments in the Dorfman-BerbaumMetz procedure for multireader ROC study analysis. Academic Radiology2008; 15:647-661

[16] Beam, Layde and Sullivan. Variability in the Interpretation of Screening Mammograms by US Radiologists, Findings from a National Sample. *Arch Intern Med* 1996; 156(2):209-213

[17] Obuchowski NA. Sample size tables for receiver operating characteristic studies. *AJR Am J Roentgenol* 2000; 175(3):603-8

[18] Rockette, Campbell, Britton Holbert, King and Gur. Empiric assessment of parameters that affect the design of multireader receiver operating characteristic studies. *Academic Radiology* 1999; 6:723-729

[19] Statistics Collaborative, Inc. Protocols DOC-SRYL-7VN99P and DOC-SRYL-7W5FTC Sectra MicroDose Mammography L30: Accuracy and Feature Analysis for the Multi-Reader, Multi-Case Comparison of the Sectra MicroDose Mammography L30 to the GE Digital Mammography System. June 6, 2011

[20] Hologic, Inc. P080003/S001 Hologic Selenia Dimensions 3D System, FDA Executive Summary. Oct 24, 2012.

[21] Gallas BD, Chan HP, D'Orsi CJ, Dodd LE, Giger ML, Gur D, Krupinski EA, Metz CE, Myers KJ, Obuchowski NA, Sahiner B, Toledano AY, Zuley ML. Evaluating imaging and computer-aided detection and diagnosis devices at the FDA. Acad Radiol. 2012 Apr; 19(4):463-77. doi: 10.1016/j.acra.2011.12.016. Epub 2012 Feb 3. PubMed PMID: 22306064.

[22] Dmitrienko, A.; Tamhane, A. C.; Bretz, F. Multiple Testing Problems in Pharmaceutical Statistics. Chapman & Hall/CRC Biostatistics Series, 2007, 320pages.

[23] P. Chakraborty, P., and Berbaum , K. S. Observer studies involving detection and localization: Modeling, analysis, and Validation. Med. Phys. 37, 5054 (2004): 2313 – 2330.

APPENDIX I: LIST OF INVESTIGATORS AND SITES

Update of this list can be provided to the Research site under separate cover.

Table 1: List of principal Investigators

Name Principal Investigator(s)	**Name and address investigation site(s)**
Etta Pisano, M.D.	N/A

APPENDIX II: LIST OF MONITOR(S)/CLINICAL SCIENTIST

Update of this list can be provided to the Investigational sites under separate cover.

Table 2: List of monitor/clinical scientist

Name Monitor(s)/Clinical Scientist	**Contact Information of Monitors**
Ingrid Schulze-Wenck	Hamburg, Germany, Telephone: +49 40 5078-1750 Mobile: +49 172 299 88 73 ingrid.schulze-wenck@philips.com
Elin Moa	Solna, Sweden Telephone: +46 8 623 52 00 Elin.moa@philips.com