SUMMARY OF SAFETY AND EFFECTIVENESS DATA (SSED)

I. GENERAL INFORMATION

Device Generic Name:	Digital Breast Tomosynthesis
Device Trade Name:	MAMMOMAT Inspiration with Tomosynthesis Option
Device Procode:	OTE
Applicant's Name and Address:	Siemens Medical Solutions USA, Inc. 51 Valley Stream Parkway Malvern, PA 19355-1406
Date(s) of Panel Recommendation:	None

Premarket Approval Application (PMA) Number: P140011

Date of FDA Notice of Approval:	April 21, 2015
Priority Review:	Not Applicable

II. INDICATIONS FOR USE

The MAMMOMAT Inspiration with Tomosynthesis Option is indicated for acquisition of 2D as well as 3D digital mammography images to be used in screening and diagnosis of breast cancer.

Each screening examination may consist of CC and MLO views in:

- a 2D image set or

- a 2D and 3D image set.

Note:
The screening examination may consist of 2D FFDM images set with or without the 3D image set.

III. CONTRAINDICATIONS

None.

IV. WARNINGS AND PRECAUTIONS

The warnings and precautions can be found in the MAMMOMAT Inspiration system labeling.

V. DEVICE DESCRIPTION

The MAMMOMAT Inspiration with Tomosynthesis option consists of the MAMMOMAT Inspiration FFDM System (originally approved under P030010 and cleared in its most recent version under k122286) with an additional software upgrade. The additional software upgrade enables the acquisition of Digital Breast Tomosynthesis (DBT) images that can be used for screening or diagnostic mammography.

The MAMMOMAT Inspiration system consists of an examination stand with an integrated, microprocessor-controlled, high-frequency generator. The swivel arm contains the X-ray tube on the top end and the object table with the detector on the bottom end. It can be rotated isocentrically between -180º (clockwise) and + 180º (counter-clockwise) to provide angulated views.

In Tomosynthesis mode the MAMMOMAT Inspiration system produces 25 low-dose exposures of the compressed breast by moving the X-ray source over a 50° arc above the breast. The acquisition of the 25 projections takes up to 25 seconds. The set of projections is used to reconstruct the compressed breast volume, called Digital Breast Tomosynthesis (DBT) image, displayed in a series of 1 mm slices parallel to the detector. The DBT acquisition can also be combined with a FFDM acquisition under a single breast compression.

The MAMMOMAT Inspiration Acquisition Workstation (AWS) is provided with the Mammography stand and consists of a PC including with syngo graphical user interface. It receives and processes the digital image data, and reconstructs the DBT images from the low-dose projections. Images acquired for screening or diagnostic imaging are displayed at the acquisition workstation for quality control purposes. The processed images are then sent via DICOM protocol to either an archive and/or a diagnostic workstation intended for diagnostic mammography.

VI. ALTERNATIVE PRACTICES AND PROCEDURES

There are several other alternatives for breast cancer screening and diagnosis. These include clinical breast examination, film-screen mammography, digital mammography, contrast enhanced spectral mammography, ultrasound, dedicated breast CT and magnetic resonance imaging. The Hologic Selenia Dimensions 3D System and the GE SenoClaire, respectively approved by FDA via PMA P080003 and PMA P130020, can also produce DBT images.

After detection of an abnormality, a biopsy and pathologic examination may be performed to diagnose the cancer. Each alternative has its own advantages and

drawbacks. Patients should fully discuss these alternatives with their physician to select the method that best meets expectations and lifestyle.

VII. MARKETING HISTORY

The MAMMOMAT Inspiration system with Tomosynthesis Option is commercialized internationally outside of the United States. It received CE mark on November 20, 2009 and licenses for commercialization in Brazil, China and Canada on June 15, 2014, July 15, 2014 and March 26, 2015 respectively.

MAMMOMAT Inspiration with Tomosynthesis Option has not been withdrawn from any market for any reason related to its safety or effectiveness.

VIII. POTENTIAL ADVERSE EFFECTS OF THE DEVICE ON HEALTH

Below is a list of the potential adverse effects (e.g., complications) associated with the use of MAMMOMAT Inspiration with Tomosynthesis Option. These potential adverse effects are common to all mammography systems:

- Excessive breast compression
- Excessive x-ray exposure
- Electrical shock
- Skin irritation, abrasion, or puncture wound
- Infection

One adverse event and one serious adverse event were reported for the patients enrolled in the clinical study, but none of these adverse events were related to the use of the MAMMOMAT Inspiration system with Tomosynthesis Option. For the specific adverse events that occurred in the clinical study, please see Section X below.

Failure of the device to perform as expected or failure to correctly interpret the images produced by the device may lead to improper patient management decisions. False positives could lead to additional exams that could result in a small risk of additional discomfort and complications such as infection or bleeding if a biopsy were performed. The risk of a serious complication is extremely low. False negatives would not be recalled which may result in delay in diagnosis and progression of disease up until the next screening exam or interval diagnosis.

IX. SUMMARY OF PRECLINICAL STUDIES

A. Physical Laboratory Testing

Where applicable to a digital breast tomosynthesis system, Siemens Medical Solutions USA, Inc. (also referred to as “the sponsor”) followed the physical laboratory testing methods mentioned in the FDA guidance, Class II Special Controls Guidance Document: Full-Field Digital Mammography System.

MAMMOMAT Inspiration Units are equipped with two versions of the Large Mammography (LMAM) detector: LMAM v1 and LMAM v2. The main difference between v1 and v2 is in power supply and data communication. The sponsor intends to offer the tomosynthesis option for systems equipped with either detector version. To support approval of the system with LMAM v1 and LMAM v2, bench testing was conducted on both detectors for system characterization and performance comparison.

Table 1: Physical Laboratory Testing

Test	Purpose	Acceptance Criteria	Results
Sensitometric response in tomosynthesis mode	Assess detector signal response vs. radiation exposure	Linearity of detector response at different angles	Detector output response was linear (0.999990) between 2 µGy and 90 µGy.
Modulation transfer function (MTF) in projection space for moving x-ray source	Quantitative measure of the spatial resolution properties of the image acquisition system; Assessment of angular exposure to MTF	System Characterization – No pass/fail criteria	See Table 2: Modulation transfer function in projection space for moving X-ray source.
Noise Power Spectrum in projection space for moving x-ray source	Quantitative measure of the noise properties of the image acquisition system	System Characterization – No pass/fail criteria	The pixel noise variance responds linearly to the incoming air kerma. Neither electronic noise nor fixed pattern (structural) noise plays a significant role.
Detective quantum efficiency (DQE) in projection space for moving x-ray source	Quantitative measure of the efficiency of signal-to-noise ratio (SNR) transfer of the image acquisition system	System Characterization – No pass/fail criteria	DQE show almost no dependence on the input air kerma in the range between 8 µGy and 90 µGy. For oblique irradiation (25° for tomo) the DQE values are about 10% lower than compared to the orthogonal irradiation used for 2D imaging.
Detector Lag	Assessment of effects of detector lag on successive projections (up to 12 steps) in an imaging sequence and in consecutive imaging sequences.	The measured residual signal should be negligible compared to the next acquisition.	The residual image intensity after the first step was about 2% for step 2 to 12.

Spatial Resolution	Quantitative spatial resolution of the tomosynthesis volume, perpendicular to the slices	Stable slice sensitivity profile for different object sizes (wax insert of the ACR phantom)	The visibility of structures in the ACR (American College of Radiology) phantom was evaluated in 3 different height positions with no noticeable deviation between variable object height positions.
ACR phantom image quality	Detectability of small structures in the breast	MQSA minimum requirement – Perfect phantom scores values are: 5, 5, 6. Passing phantom scores values are: 3, 3, 4.	Tomosynthesis scans using AEC (automatic exposure control) settings at W/Rh anode/filter at 28 kV were evaluated by a group of experienced human observers. The following are the average score values: 4.02, 3.50, 5.38 for masses, speck groups, fibers respectively.
Uniformity in reconstructed slices	Grayscale uniformity inside individual slices is assessed using spatially homogeneous PMMA blocks	System characterization – No pass/fail criteria	Slightly higher non-uniformity for LMAM v1.
Geometric accuracy	Assess fidelity in the mapping of geometrical lengths inside individual slices	Visual inspection of the selected slice	No significant variation of the detected object sizes with varying height position inside the reconstructed volume.
Visual limiting resolution in tomosynthesis plane	Assess limiting spatial resolution in the reconstructed image of a bar pattern phantom in variable height position is assessed.	Object recognizability	Both detectors have identical visual limiting resolution for all height positions, phantom orientations, and binning modes.
Average Glandular dose	Quantitative estimate of the patient radiation dose is provided	Dose estimation and comparison of two detectors.	Negligible deviations in the applied tube load and generated average glandular dose values from the two detector versions. See Table 3: Average Glandular Dose

Table 2: Modulation transfer function in projection space for moving X-ray source

Spatial Frequency Line Pairs	Horizontal	Vertical
1/mm	0.95	0.93
2/mm	0.88	0.84
3/mm	0.79	0.69
4/mm	0.68	0.56
5/mm	0.60	0.43
5.88/mm	0.53	0.32

Table 3: Average Glandular Dose

PMMA Thickness (mm)	Tube Voltage (kV)	Tube Load (mAs)	Average Glandular Dose (mGy)
19	26	69.75	1.108
37	26	190.00	2.095
56	26	523.24	4.207
19	28	55.00	1.065
37	28	139.00	1.913
56	28	363.29	3.705
19	30	55.01	1.275
37	30	106.25	1.750
56	30	268.75	3.333

Physical laboratory testing demonstrated that the LMAM v1 and LMAM v2 have similar performance characteristics.

Sample clinical images acquired with LMAM v1 and v2 were reviewed by board certified and MQSA qualified radiologists. The radiologists found the images to be of diagnostic quality and therefore suitable for clinical use.

B. Additional Studies

1. Conformance to Standards

Siemens Medical Solutions USA, Inc. provided certificates of conformance to the following standards:

IEC 60336: 2005: Medical electrical equipment – X-ray tube assemblies for medical diagnosis – Characteristics of focal spots

IEC 60601-1 Medical Electrical Equipment

- Part 1: 1988: General Requirements for Safety; 1991: Amendment 1; 1995: Amendment 2.
- Part 1-1: General Requirements for Safety – Collateral Standard: Safety requirements for medical electrical systems.
- Part 1-2: General Requirements for Safety – Collateral Standard: Electromagnetic Compatibility Requirements and Tests.
- Part 1-3: General requirements for basic safety and essential performance – Collateral Standard: Radiation protection in diagnostic X-ray equipment.
- Part 1-4: General requirements for safety – Collateral standard: Programmable electrical medical systems, edition 1.1.
- Part 2-7: 1998: Particular requirements for the safety of high-voltage generators of diagnostic X-ray generators.
- Part 2-28: 1993: Particular requirements for the safety of X-ray source assemblies and X-ray tube assemblies for medical diagnosis, edition 1.0.
- Part 2-32: 1994: Particular requirements for the safety of associated equipment of X-ray equipment, edition 1.0.
- Part 2-45: 2001: Particular requirements for the safety of mammographic X-ray equipment and mammographic stereotactic devices, edition 2.0.

IEC 61223-3-2: 2007 Evaluation and routine testing in medical imaging departments

- Part 3-2: Acceptance tests – Imaging performance of mammographic X-ray equipment.

IEC 61674: 1997: 2002: Amendment 1, Medical electrical equipment – Dosimeters with ionization chambers and/or semi-conductor detectors as used in X-ray diagnostic imaging

IEC 62304: Medical device software – Software life cycle processes, edition 1.0

IEC 62366: Medical devices – Application of usability engineering to medical devices

ISO 14971:2007 Medical devices – Application of risk management to medical devices

AAMI ANSI ISO 10993-1:2009 Biological evaluation of medical devices

- Part 1: Evaluation and testing within a risk management process

NEMA PS 3.1 – 3.18: 2009 Digital Imaging and Communications in Medicine (DICOM) Set

2. Software Design and Testing Documentation

Siemens Medical Solutions USA, Inc. provided design and software testing documentation consistent with FDA's Guidance for the Content of Premarket Submissions for Software Contained in Medical Devices. The sponsor conducted software unit testing and integration testing to verify that all the sub-systems satisfy the software requirements and integrated successfully. System testing was also conducted to validate that the software specifications

conform to its intended use and user requirements. The sponsor conducted regression testing to ensure that new software features introduced by the tomosynthesis option do not create problems with previous version of the software. Impact analysis was also provided as a justification for test case selection in the regression testing. All the test activities were completed successfully. The impact of the unresolved anomalies on device safety and effectiveness were properly assessed. The mitigations for the unresolved anomalies were provided and acceptable.

Conclusion of Non-Clinical Testing

Physical laboratory testing and examination of sample clinical images demonstrates that the MAMMOMAT Inspiration system with Tomosynthesis Option can be used to produce diagnostic quality DBT images.

X. SUMMARY OF PRIMARY CLINICAL STUDY

The applicant performed a clinical study to establish a reasonable assurance of safety and effectiveness of MAMMOMAT Inspiration with Tomosynthesis Option for breast cancer screening and diagnosis in the US under IDE #G100247. Data from this clinical study were the basis for the PMA approval decision. A summary of the clinical study is presented below.

The study consisted of a prospective case collection study in which patients were imaged using the standard of care and the tested device; and of a retrospective Multi-Case Multi-Reader (MRMC) study.

Study	Study Design	Study Objective	Number of Sites/Readers	Number of Subjects
Accrual	Prospective subject accrual	Subject accrual for blinded reader study Evaluate the safety of the device	7 enrollment Sites	800 patients
MRMC study	Retrospective reader study	Evaluate the safety and effectiveness of the device Establish that FFDM + 2-view DBT (MLO and CC) has superior diagnostic accuracy to FFDM alone.	22 readers	300 cases (50 cancer cases, 85 recall cases and 165 normal cases)

A. Prospective Case Accrual

Siemens Medical Solutions USA, Inc. designed and conducted a prospective case accrual study to collect Full Field Digital Mammography (FFDM) images and Digital Breast

Tomosynthesis (DBT) images of the patients undergoing regular breast cancer screening imaging or undergoing diagnostic workup after a potential anomaly was detected at screening. Patients were enrolled and imaged between May 2011 and February 2014.

The pivotal reader study uses a subset of the images accrued under this case collection protocol by September 30, 2013. At that time, a total of 698 patients were enrolled from 7 United States clinical sites under IRB approved clinical case acquisition protocol:

- SMS-SP09-01: Retrospective, Multi-Reader, Multi-Case Study (MRMC) to demonstrate the superior accuracy of Siemens Digital Breast Tomosynthesis system and Full Field Digital Mammography (FFDM) Systems to FFDM alone for Screening and Diagnostic Mammography. ClinicalTrials.gov protocol number: NCT01373671.

All enrolled subjects had routine 2-view Mediolateral Oblique (MLO) and Craniocaudal (CC) screening mammograms on a commercially available FFDM system before or upon enrollment. In addition subjects were imaged with a prototype of the MAMMOMAT Inspiration system with Tomosynthesis Option under MLO and CC positioning.

The 2D FFDM images from screening mammography and/or diagnostic mammography and the DBT images were provided in digital format to the sponsor.

1. Clinical Inclusion and Exclusion Criteria

Inclusion Criteria

Enrollment in the Case Acquisition study was limited to patients who met the following inclusion criteria:

- Provided signed informed consent after receiving a verbal and written explanation of the purpose and nature of the study;

- Women, 40 years of age or older at the screening mammographic evaluation or age 30 or older presenting for a biopsy and have one of the following mammograms:
    - Normal cases at screening (BI-RADS 1,2, and 3)
        - Have a screening mammogram that includes the 4 standard screening views (RCC, RMLO, LCC, LMLO), as well as have both MLO and CC DBT scans of each breast;
    - Actionable cases at screening (BI-RADS 0, 4 or 5) with final BI-RADS 1,2,3,4, or 5:
        - Have a screening mammogram with 4 screening views and any clinically necessary diagnostic mammographic views, plus 4 screening views repeated at the diagnostic or biopsy visit if the screening images are unavailable or were acquired more than 45 days prior to DBT acquisition.

- Have supporting ground-truth documentation for the final BI-RADS assessment:
    - A 1 year FFDM follow-up without evidence of cancer for normal cases not undergoing biopsy;
    - A 6-month or 12-month follow up confirming benign status for biopsy proven benign cases, except for surgical biopsies; or
    - Pathology report for either benign or malignant biopsy finding.

Exclusion Criteria

Patients were <u>not</u> permitted to enroll in the prospective image accrual study if they met any of the following exclusion criteria:

- Pregnant women or women who believe they may be pregnant or are trying to become pregnant;
- Mastectomy patients;
- Patients who have had previous lumpectomy (cancer) within 5 years of the study entry;
- Inmates (45 CFR 46.306) or mentally disabled individuals;
- BI-RADS category 6 (e.g. for which mammogram was performed for the purpose of planning cancer therapy);
- BI-RADS 4 or 5 without confirming pathology reports;
- Subjects with mammograms that lack the required views or with views judged to be technically inadequate;
- Subjects accrued from the screening population who know they will not be in the United States or available for follow-up mammograms in one year.

2. <u>Follow-up Schedule</u>

All cancer mammograms were confirmed by a biopsy proven pathology report. Subjects with findings deemed benign through a non-open biopsy procedure had a 6 or 12 month FFDM follow up to confirm their benign status. Subjects that underwent surgical biopsy were not required to have the FFDM follow up done in 6 or 12 months. All subjects with normal mammograms were asked to return for confirmatory mammograms on an FFDM system at 1 year to confirm non-cancer status. All applicable radiology reports were collected. Additionally FFDM images obtained during the follow up of normal and biopsy proven benign cases were collected if possible.

B. <u>Accountability of PMA Cohort</u>

Patient Accountability: By September 30, 2013, 698 patients were enrolled in the case collection study. Since the collection study was still ongoing at the time of image selection for the MRMC study, a number of cases were not eligible for inclusion because of lack of 1-year follow-up or because the DBT images had not been retrieved from the collection sites at the time of case selection. Table 4: Patient Accountability summarizes the patient disposition as of September 30, 2013. Three hundred and thirty-eight (338)

cases were available for the selection into the MRMC study including: 67 malignant cases, 85 biopsy proven benign cases and 186 negative cases.

Table 4: Patient Accountability

Case accountability	**As of September 30, 2013**
Enrolled in Case Collection Study	**698**
Screen Failure and malfunctions	17
Negative Mammograms but no follow-up yet	197
Subject whose follow-up mammogram was not negative (other than BI-RADS 1 or 2	13
- Cancer	4
- Biopsy Benign	3
- Lost to follow up	1
- BI-RADS 3 at follow up	5
Subjects whose images were not yet retrieved from clinical sites	45
Subjects whose images were not reprocessed	69
Subjects enrolled recently (Jul to Sep 2013)	19
Subjects Available for MRMC study (cancer, benign, normal)	**338 (67, 85, 186)**
Subjects Enrolled in the MRMC study (cancer, benign, normal)	**300 (50, 85, 165)**

The case set used for the reader study included 300 cases and was composed of 165 negative cases, 85 biopsy-proven benign cases and 50 cancer cases. The sponsor enrolled all available recall patients with initial BI-RADS score of 0 and follow up BI-RADS 1 or 2 (22 cases) and a random sample of cases with BI-RADS 1 or 2 and negative follow up (143 cases). All available biopsy proven benign cases were enrolled into the study regardless of whether follow up had been completed or not. All available cancer cases presenting with mammography findings of microcalcifications, architectural distortions or asymmetric densities were enrolled into the reader study and cancer cases with masses were randomly selected.

Breast Accountability: The study protocol defined a negative case/breast as a negative case/breast at baseline confirmed with a negative 1-year-follow up mammographic exam. As of September 30, 2013, 110 out of 600 breasts included in the MRMC study did not have 1-year follow-up information and were excluded from the primary analysis. The excluded breasts consisted of 47 contralateral breasts of patients with cancer in one breast, 1 breast found to be abnormal at 1-year follow-up, and 62 breasts of biopsy proven benign cases whose follow-up was incomplete. The primary analysis included 490 breasts (53 with biopsy-proven cancer breasts, 90 biopsy-proven benign breasts, and 347 normal breasts).

C. Study Population Demographics and Baseline Parameters

The demographics of the study population in terms of breast density and cancerous lesion characteristics are typical of a US screening population (cf. Table 5 and Table 6).

Table 5: BI-RADS score at enrollment and Breast Density Distribution

	Categories	Cancer cases (N=50)	Non-Cancer Cases (N=250)
BI-RADS score at enrollment	0	26	78
	1	1	92
	2	1	51
	3	0	2
	4	8	27
	5	14	0
Breast Density	Almost entirely fat	10.0%	4.4%
	Scattered fibroglandular	38.0%	41.2%
	Heterogeneously dense	46.0%	49.2%
	Extremely dense	4.0%	4.8%
	Missing Density Data	2.0%	0.4%

Table 6: Cancerous Lesion Characteristics

	Categories	# of lesions	Percentage of # of lesions
Lesion Type	Mass	43	65.2%
	Calcifications	16	24.2%
	Architectural Distortion	6	9.1%
	Asymmetric Density	1	1.5%
Cancer Type	Invasive Cancer	52	78.8%
	Ductal Carcinoma In Situ	14	21.2%
Lesion Size	< 10 mm	16	24.2%
	10 to 19 mm	18	27.3%
	20 to 29 mm	18	27.3%
	More than 30 mm	11	16.7%
	Missing size information	3	4.5%
Total number of cancerous lesions		66	100.0%

D. Reader Study Design and Methods

Siemens Medical Solutions USA, Inc. conducted a Multi-Reader Multi-Case (MRMC) study using an enriched subset of cases accrued in the aforementioned prospective case collection study. The objective was to demonstrate the superiority of FFDM and 2-view DBT to FFDM only for the screening and diagnosis of breast cancer.

1. Reference standards

All cancer mammograms were confirmed by a biopsy proven pathology report. Subjects with findings deemed benign through a non-open surgical biopsy procedure had a 6 or 12 month FFDM follow up to confirm their benign status. Subjects that underwent open surgical biopsy were not required to have the FFDM follow up done in 6 or 12 months. All subjects with normal mammograms were asked to return for confirmatory mammograms on an FFDM system at 1 year to confirm non-cancer status. All applicable radiology reports were collected. Additionally FFDM images obtained during the follow up of normal and biopsy proven benign cases were collected if possible.

Malignant Lesions: Ground Truth (GT) for the type and location of malignant lesions was based on the mammography findings described by the radiologist at the clinical site and supported by the radiology and the pathology report from biopsy procedures. GT FFDM and DBT images with lesion locations marked electronically were created by an independent radiologist based on review of the electronic Case Report Forms (eCRFs), radiology and pathology reports.

2. Readers

22 MQSA qualified radiologists were involved in the MRMC study. Each reader had at least 5-year experience in reading film screen mammograms and at least 1-year experience in reading digital mammograms. The readers had a variety of experience ranging from breast imagers to general radiologists and represented academic and nonacademic institutions.

Prior to the blinded reading experiment, all readers took part in a training, which provided an overview of DBT physics, the specific features of DBT and the differences between FFDM and DBT. After being trained on the use and features of MammoReport a multimodality reading and reporting software, the readers reviewed approximately 65 FFDM images with corresponding DBT images with the trainer. Additionally, the flow of the sequential read (FFDM, followed by FFDM plus DBT MLO, and then FFDM plus 2-view DBT) was explained as well as how interpretation data should be recorded in the eCRFs.

The following day, the training was followed by a test to determine the readers' accuracy in detecting cancers on FFDM and DBT. Each reader individually reviewed and scored 70 image sets (FFDM and 2-view DBT) as part of a test to determine the readers' cancer detection rate with FFDM and with FFDM+2-view DBT. The passing criteria were a cancer detection rate of at least 75% for FFDM + DBT and a detection rate of at least 75% for FFDM. All 22 radiologists met the FFDM + DBT criterion but only 20 out of 22 radiologists met the FFDM detection rate criterion. Although 2 of the 22 readers did not pass the FFDM detection rate criterion, all readers provided ratings on the 300 cases enrolled in the MRMC study. None of the readers were excluded from the primary analysis since readers are not subjected to any cancer detection test in clinical practice.

3. Blinded Image Reading

Blinded Image Interpretation*:* The study design is fully crossed, i.e. all readers read all cases under each modality. The cases were randomized, presented as in normal reading practice on the MammoReport and without any history data or marks on the mammograms. No prior films, subject histories or any other demographic information accompanied the interpretation. Readers were informed that the study was enriched with cases with cancer. In the same session, the readers reviewed the cases sequentially in the following three modality configurations:

- Arm 1: FFDM image sets with bilateral MLO and CC views;
- Arm 2: FFDM image sets plus 1-view DBT (MLO) image sets. The dataset included reconstructed DBT images of planes over the full breast volume;
- Arm 3: FFDM image sets plus 2-view DBT (MLO and CC) image sets.

The readers first assigned a BI-RADS score (0, 1 or 2) at the breast level. If an actionable finding was identified, the readers recorded the type of finding, the slice number if applicable, a forced BI-RADS score (1, 2, 3, 4, 5) and a Probability of Malignancy (POM) confidence score between 0 and 100 for each suspicious finding. If a breast was read as negative (BI-RADS score 1 or 2), the readers only recorded the BI-RADS score and were not allowed to describe a finding in the eCRF. All findings identified during the first read with FFDM were transferred to the FFDM plus 1-view DBT and FFDM plus 2-view DBT eCRFs to allow readers to make adjustments in the BI-RADS and POM scores based on the additional information presented in the following reading arms.

Ground Truth and Adjudication: Two (2) board certified and MQSA qualified radiologists (who were not part of the reader study) served as Ground Truth (GT) Judges to determine whether the marks noted by readers in the study coincide with the biopsy-proven lesions identified at the clinical sites. A third board certified and MQSA qualified radiologist (who was not part of the reader study either) served as an Adjudicator GT Judge to resolve differences between the two GT Judges. Ground Truth was defined as a suspicious area identified at the clinical collection site, which was biopsied and confirmed to be positive by pathology. The results of the Adjudication process served as the final judgment. In 6 instances the Adjudicator did not agree with either Judge's assessment to identify the correct lesion.

4. Conversion of Lesion Scores into Breast and Case Scores

During the blinded reading experiment, radiologists were asked to provide a BI-RADS score at the breast level and a POM score for each suspicious finding in the breast. Primary and secondary ROC based analyses were conducted at the breast or case level. The following rules were established to assign a POM score to each breast based on the lesion POM and breast BI-RADS assessments provided by the reader.

If the reader assigned a breast BI-RADS score of 1 or 2 to the breast, the breast was assigned a POM score of 0. For a breast to be considered a true positive the reader had to

correctly locate at least one malignant lesion within the breast. If the reader had not correctly located at least one malignant lesion, then the breast was assigned a POM of 0 in the statistical analysis. For breasts with more than one biopsy-proven malignant lesion, the highest POM score assigned to the correctly located malignant lesions was used in the analysis. If a breast had no biopsy proven malignancies, the highest POM score assigned to any false findings was used in the analysis.

Similar rules were used to assign a POM score to each case.

5. Statistical Method for Primary Analysis

The primary null hypothesis was that the readers' average breast level AUC ROC with FFDM plus 2-view DBT is equal to the readers' average breast level AUC ROC with FFDM only. The alternative hypothesis was that readers' average breast level AUC ROC with FFDM plus 2-view DBT is different from the readers' average breast level AUC ROC with FFDM alone.

Nonparametric methods [1] were used to estimate reader ROC curves and their areas for each reader using the inferred breast-level POM values. The sponsor used the method proposed by Obuchowski and Rockette [2], adjusted for the clustered data [3], to test the null hypothesis, with the recommended adjustment to the degrees of freedom by Hillis et al [4]. This statistical method was developed for the analysis of MRMC studies to account for the inter-reader variability and the correlations between and within readers, allowing the results to be generalized to both the subject and reader populations. The sponsor used a two-tailed test, with 0.05 significance level, and reported p-value for the test of the null hypothesis and a 95% CI for the difference in the mean breast level AUC ROC with FFDM alone vs. FFDM plus 2-view DBT.

E. Safety and Effectiveness Results

1. Safety Results

Adverse Events: The analysis of safety was based on the 698 enrolled patients as of September 30, 2013. The sponsor reported one adverse event and one serious adverse event during the case acquisition study. None of these events were related to the use of the device.

Adverse effects that occurred in the PMA clinical study:

Serious Adverse Event: One subject was enrolled in the study in May 2011 and underwent the DBT exam on the day of study enrollment. The subject passed away from pneumonia in July 2012. The Principal Investigator at the site determined that the Serious Adverse Event was not device related and the subject was classified as "lost to follow up".

Adverse Event: One subject was enrolled in February 2012 and underwent DBT exam on the day of study enrollment. The following day the subject reported lightheadedness, which lasted for 3 days. The adverse event was mild in intensity and the subject fully recovered. The Principal Investigator at the site determined that the Adverse Event was not related to the device.

Average Glandular Dose*:* With the addition of DBT views to the mammography screening examination, the radiation dose delivered to the patient increases significantly. DBT radiation dose for one view equals 1.5 - 2 mGy (0.15 - 0.2 rad) for a standard breast [50% fibroglandular tissue, 50% fat tissue], whereas a 2D mammogram with appropriate x-ray spectrum applies 1 mGy (0.1 rad). Figure 1 presents the distribution of the estimated Average Glandular Dose delivered to the 300 patients enrolled in the MRMC study.

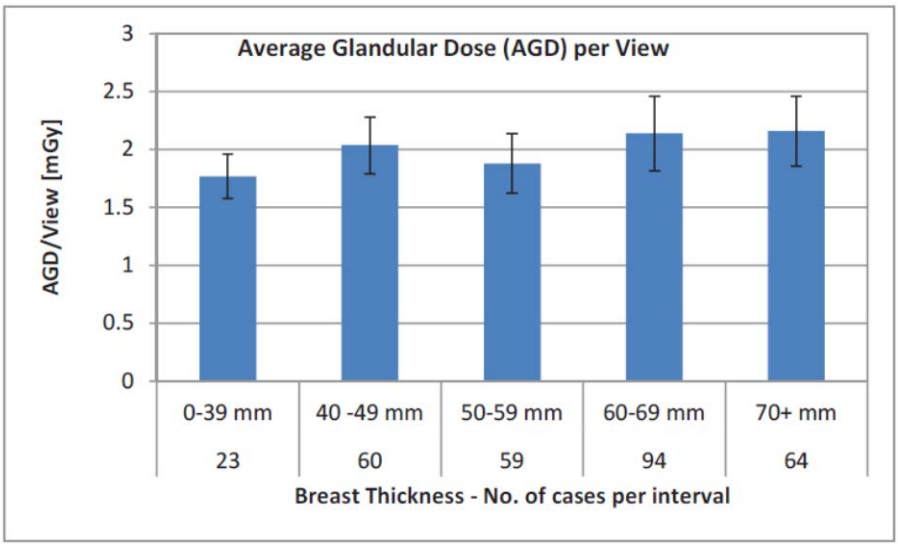

Figure 1: Average Glandular Dose (for a standard breast with 50% fibroglandular tissue and 50% fat tissue) per view stratified by breast thickness in the 300 patients enrolled in the MRMC study.

2. Effectiveness Results: Primary Endpoints

The analysis of the Probability of Malignancy scores provided by 22 readers on 490 breasts (53 cancer breasts, 90 biopsy benign breasts, 437 normal breasts) resulted in an AUC ROC of 0.7516 for the control arm (FFDM) and an AUC ROC value of 0.8527 for the tested arm (FFDM + 2-view DBT). The mean difference in breast level AUC ROC is 0.1011 with 95% confidence interval [0.063, 0.139]. The null-hypothesis is rejected with p-value < 0.0001. The primary endpoint was met and demonstrates superiority in terms of diagnostic accuracy of FFDM + 2-view DBT over FFDM alone.

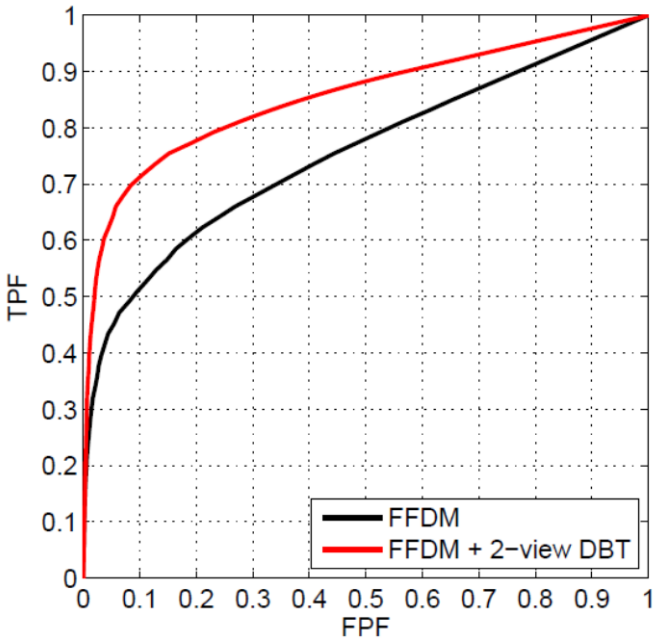

Figure 2: Average reader breast level ROC Curves. Difference in AUC ROC between FFDM plus 2-view DBT (MLO and CC) and FFDM is 0.1011 with confidence interval [0.063, 0.139]. Calculation based on ratings of 22 readers on 490 breasts (110 breasts excluded for lack of follow-up information).

Figure 2: Average reader breast level ROC Curves provides a graphical representation of the average breast level ROC curves for FFDM and FFDM + 2-view DBT using Chen's [5] nonparametric method for calculating the average ROC curve.
Several complimentary analyses were conducted to assess the robustness of the study results. Please refer to Section X.E.4 for more information on these additional analyses.

3. Effectiveness Results: Secondary Endpoints

Seven secondary analyses were conducted in a pre-specified hierarchical order.

The readers' parametric and nonparametric estimates of the patient-level AUC ROC comparing FFDM alone with FFDM plus 2-view DBT also showed statistically significant improvement with the addition of 2-view DBT.

Non-cancer recall rate was lower for FFDM plus 2-view DBT than for FFDM alone. The readers' mean recall rate with FFDM alone was 0.438 (SE=0.030) and with FFDM plus 2-view DBT was 0.355 (0.022), which is a statistically significant reduction.

Reader's mean sensitivities and specificities observed in the MRMC are summarized in Table 7. For all subgroups and both definitions of a positive test result, readers' sensitivity increased with the addition of 2-view DBT. No increase in specificity was observed.

Because sensitivity and specificity were tested jointly for superiority, this secondary endpoint was not met. The statistical significance of the results of the remaining pre-specified secondary analyses could not be assessed due to the failure one of the previous secondary endpoint in the hierarchy of pre-specified analysis and due to the lack of multiple testing correction. The values reported below were observed in the MRMC study but cannot be used to determine the significance of the differences between the controlled and study arms.

Table 7: Sensitivity and Specificity observed in the study (22 readers)

	FFDM	**FFDM + 2-view DBT**
Sensitivity with BI-RADS 4,5 as a positive test		
Case level (n=50)	0.5945 (0.0408)	0.7855 (0.0357)
Breast level (n=53)	0.5858 (0.0486)	0.7710 (0.0444)
Dense Breast (n=27)	0.5572 (0.0585)	0.7677 (0.0529)
Fatty Breast (n=25)	0.6382 (0.0711)	0.7982 (0.0692)
Breast with calcs (n=15)	0.5909 (0.0901)	0.6758 (0.0867)
Breast with masses (n=38)	0.6304 (0.0539)	0.8289 (0.0498)
Specificity with BI-RADS 4,5 as a positive test		
Case level (n=250)	0.7324 (0.0364)	0.7298 (0.0289)
Breast level (n=437)	0.8293 (0.0280)	0.8302 (0.0219)
Dense Breasts (n=239)	0.8389 (0. 277)	0.8391 (0.0231)
Fatty Breasts (n=197)	0.8168 (0.0326)	0.8196 (0.0256)
Sensitivity with BI-RADS 3, 4, 5 as a positive test		
Case level (n=50)	0.6718 (0.0367)	0.8218 (0.0345)
Breast level (n=53)	0.6638 (0.0439)	0.8087 (0.0418)
Specificity with BI-RADS 3, 4, 5 as a positive test		
Case level (n=250)	0.5616 (0.0295)	0.5931 (0.0242)
Breast level (n=437)	0.7181 (0.0251)	0.7411 (0.0204)

Table 8 summarizes the stratified breast level ROC analysis.

The following was observed in the MRMC study conducted by Siemens Medical Solutions USA, Inc. The readers' breast level AUC ROC increased with the addition of 2-view DBT for dense and fatty breasts, for microcalcifications and masses. The statistical significance of these observations could not be determined.

The patient-level AUC ROC increased with the addition of 1-view DBT (MLO only) to FFDM. The AUC ROC is 0.7091 (SE 0.0329) and 0.8036 (0.033) for FFDM alone and FFDM + 1-view DBT respectively. The statistical significance of this observation could not be determined.

Table 8: Breast level ROC AUC stratified by breast density or lesion type.

Stratification	Number of breasts in analysis	FFDM only AUC ROC (SE)	FFDM + 2-view DBT AUC ROC (SE)
Dense Breast	22 readers; 27 cancer breasts; 239 non-cancer breasts	0.7336 (0.0332)	0.8505 (0.0315)
Fatty Breast	22 readers; 25 cancer breasts; 197 non-cancer breasts	0.7865 (0.0428)	0.8704 (0.0423)
Microcalcification	22 readers; 16 lesions in 15 cancer breasts; 437 non-cancer breasts	0.7439 (0.0539)	0.7895 (0.0520)
Masses	22 readers; 41 lesions in 38 cancer breasts; 437 non-cancer breasts	0.7820 (0.0309)	0.8907 (0.0300)

4. Additional Analyses

The sponsor conducted several analyses to assess the robustness of the results of the primary analysis.

Max ROC analysis: During the blinded image evaluation experiment readers were asked to provide a POM score for suspicious findings in each breast and in each arm of the study. These scores were combined following the rules described in Section X.d.4 to produce a single breast score. Similarly POM score and BI-RADS scores were combined to produce a single case score. There are several ways to combine lesion-level POM scores and little consensus on what is the best method to combine individual lesion scores into a single breast score or a single case score. A complementary ROC analysis was conducted, in which each breast is assigned the highest POM score provided by the reader, regardless of mark localization or BI-RADS assessment. The primary endpoint, i.e. the difference in AUC ROC between the control and tested arms remains significant and confirms the superiority of FFDM plus 2-view DBT over FFDM alone.

Free Response Operating Characteristic (FROC) Curves: Since both the primary analysis and the max ROC analysis use breast level summary measures, FROC curves were produced to visually assess the trade-off between readers' Sensitivity and False Positive

Marks with FFDM alone and with FFDM + 2-view DBT. For most readers the FROC curve with FFDM + 2-view DBT is superior to the FROC of FFDM alone which suggests an improvement in reader performance.

Exclusion of Cancers found at 1-year follow-up: A subset of cases were selected from the accrued collection of cases on September 30, 2013. Twelve cases were excluded from the pool of eligible cases due to an abnormal 1-year follow-up. These 12 cases included 4 cancer cases found at follow-up, 3 benign cases, and 5 BI-RADS 3 cases. In order to demonstrate that the result of the study is not impacted by the exclusion of the cancer cases detected during follow-up, the sponsor conducted a sensitivity analysis using a non-informative and worst-case scenario to impute missing reader scores on the cancer cases excluded from the MRMC study. In both scenarios the effect size of using DBT as an adjunct to FFDM is reduced, but does not reach a tipping point, demonstrating the robustness of the superiority conclusion to the exclusion of the 4 cancer cases detected during follow up.

Missing Follow-up Information: The study protocol defined a negative case/breast as a negative case/breast at baseline confirmed with a negative 1-year-follow up mammographic exam. As of September 30, 2013, 110 out of 600 breasts included in the MRMC study did not have 1-year follow-up information and were excluded from the primary analysis. The excluded breasts consisted of 47 contralateral breasts of patients with cancer in one breast, 1 breast found to be abnormal at 1-year follow-up, and 62 breasts of biopsy proven benign cases whose follow-up was incomplete. The primary analysis included 490 breasts (53 with biopsy-proven cancer breasts, 90 biopsy-proven benign breasts, and 347 normal breasts). Upon reexamination of the follow-up data collected by the end of 2014, the sponsor verified the disease status of 22 additional breasts. Eighty-eight breasts remained without follow-up information. To demonstrate the robustness of the study results to missing follow-up information, the sponsor reported the results of the primary endpoint using all 600 breasts assuming no additional cancer would be detected during follow-up. The analysis of the Probability of Malignancy scores provided by 22 readers on 600 breasts (54 cancer breasts, 90 biopsy benign breasts, 456 normal breasts) resulted in an AUC ROC of 0.7547 for the control arm (FFDM) and an AUC ROC value of 0.8548 for the tested arm (FFDM + 2-view DBT). The mean difference in breast level AUC ROC is 0.1001 with 95% confidence interval [0.061, 0.139]. The null-hypothesis was rejected with p-value < 0.0001. In addition, the sponsor conducted a sensitivity analysis by varying the rate of cancer detection among the 88 breasts without follow-up information. The sponsor showed that, for plausible rates of cancer detection during follow-up, the study results would not change.

5. Other Important Considerations regarding the MRMC study

ROC analysis of inferred breast POM scores: The primary endpoint of the MRMC study assessed the difference in AUC ROC between the control arm (FFDM) and the tested arm (FFDM + 2-view DBT). In the primary analysis, the sponsor converted lesion based POM scores into breast level POM scores while enforcing correct lesion localization (cf. Section X.D.4). The task at the breast level was in fact more complex

than deciding to recall or not recall a patient since the reader was required to correctly locate the lesion. ROC curves are mathematical tools to characterize a binary classification task (recall vs. not recall) frequently summarized by the area under the ROC curve. Localization ROC (LROC) curves is however mathematically defined to account for lesion localization. Instead of using an LROC method, the sponsor chose to set the POM scores assigned to each breast to 0 with incorrect lesion localization, artificially forcing the localization-specific sensitivity to intersect the (1, 1) point on the sensitivity versus (1-specificity) axes, so the resulting curve looks like an ROC curve, the inferred ROC curve. LROC curves do not have to intersect (1, 1) point, since the probability of correctly localizing a malignant lesion will generally be less than one for all values of the decision threshold. Because the inferred ROC curve does not have all the properties of a rigorously defined ROC curve, parametric curve estimation is likely to lead to an improper fit of the data.

The examination of the empirical individual reader curves showed that the extrapolation to the (1,1) point was generally large due to the concentration of data in the first half of the ROC domain. This is a frequent issue in MRMC studies which is typically mitigated with parametric estimation of the ROC curves. For reasons noted above, this was not possible with inferred ROC curves. It is therefore possible that the lack of well distributed data on the ROC domain and the use of inferred ROC curves slightly skewed the differences in AUC ROC leading to greater uncertainties on the differences in AUC ROC between the control and studied arms. The examination of individual reader inferred ROC curves showed that in the left hand side of the ROC plot (i.e. 1-Specificity < 0.5; where most data is usually collected) the inferred ROC curve of FFDM + 2-view DBT is superior to that of FFDM alone for all readers. While there is some uncertainty on the 2-view DBT effect size due to the large extrapolation area to the (1,1) corner, the study provides good evidence of the superiority of FFDM + 2-view DBT over FFDM.

Missing FFDM + 2-view DBT lesion scores: In some instances, readers provided a POM score for the FFDM + 1-view DBT (MLO) but indicated that the lesion was not visible in the DBT CC view and omitted to provide a POM score based on all views. In a sequential study design, i.e. consecutive readings of the same case with different imaging modalities, the reader is asked to report its rating on a lesion or breast based on all examined images, and not only based on the last image that was presented. In the primary analysis, the sponsor chose to use the POM score of the intermediate arm (FFDM + 1-view DBT) if the reader did not provide a score for that lesion in the third viewing mode (FFDM + 2-view DBT) and checked the electronic Case Report Forms (eCRF) box indicating that the lesion was not visible in DBT CC view. Since a radiologist is (1) not likely to dismiss a lesion seen in DBT MLO but not visible on the DBT CC view and (2) the FFDM + 1-view DBT favorably compares to FFDM only in the study, then the correction that consists of using the intermediate arm score as if it was the complete case review score in some instances is unlikely to have significantly changed the study results.

Combined Acquisition of FFDM and DBT images: While the MAMMOMAT Inspiration system with Tomosynthesis Option enables the combined acquisition under a single breast compression of a FFDM image and a DBT image, only 4 of the 300 cases included in the MRMC analysis had been acquired using the combined acquisition mode. All other cases were acquired with patient repositioning between the FFDM image acquisition and the DBT acquisition. While the combined acquisition of the DBT and FFDM images is not likely to impact reader performance, the data collected in the study could not be used to demonstrate that the acquisition mode (combined FFDM and DBT acquisition vs. independent FFDM and DBT acquisitions with patient repositioning) had no impact on reader performance.

F. Financial Disclosure

The Financial Disclosure by Clinical Investigators regulation (21 CFR 54) requires applicants who submit a marketing application to include certain information concerning the compensation to, and financial interests and arrangement of, any clinical investigator conducting clinical studies covered by the regulation. The pivotal clinical study included 9 principal investigators and 25 board certified and MQSA qualified radiologists. None of the clinical investigators or radiologists had disclosable financial interests/arrangements as defined in sections 54.2(a), (b), (c), and (f). The information provided does not raise any questions about the reliability of the data

XI. PANEL MEETING RECOMMENDATION AND FDA'S POST-PANEL ACTION

In accordance with the provisions of section 515(c)(2) of the act as amended by the Safe Medical Devices Act of 1990, this PMA was not referred to the Radiological Advisory Panel, an FDA advisory committee, for review and recommendation because the information in the PMA substantially duplicates information previously reviewed by this panel.

XII. CONCLUSIONS DRAWN FROM PRECLINICAL AND CLINICAL STUDIES

A. Effectiveness Conclusions

The Multi-Reader Multi-Case study showed that when 2-view DBT (MLO and CC), acquired with the MAMMOMAT Inspiration system with Tomosynthesis Option, is used as an adjunct to FFDM images reader performance on average increases 0.1011 AUC ROC units with 95% CI [0.063, 0.139].

Combined with physical laboratory test results and sample image evaluation, the pivotal study results demonstrate that the MAMMOMAT Inspiration system with Tomosynthesis Option used as an adjunct is superior to FFDM alone.

B. Safety Conclusions

The risks of the device are based on physical laboratory testing as well as data collected in a clinical study conducted to support PMA approval as described above.

The risk of direct harm to the patient is minimal. There were two adverse events during the collection study, but none of the events were related to the proposed device.

The risk posed by the proposed device is similar to that of other screening and diagnostic mammography devices.

C. Benefit-Risk Conclusions

The probable benefits of the device are based on data collected in a clinical study conducted to support PMA approval as described above.

The MAMMOMAT Inspiration system with Tomosynthesis Option is used to reconstruct the breast volume from limited angle projections while eliminating the tissue overlapping effect observed in 2D projections. It is likely to benefit a substantial number of screening patients whose cancers could have otherwise been missed due to tissue superimposition (false negatives), or who may otherwise have been unnecessarily referred for additional workup (false positives).

The proposed device has no significant risk of direct harm to the patient.

The primary risk of the device comes from the possibility of false positive and false negative clinical decisions when using the images produced by the MAMMOMAT Inspiration with Tomosynthesis system. The sponsor conducted an MRMC study to compare the performance of readers with FFDM alone and with FFDM plus 2-view DBT. The study design is consistent with other mammography studies. Because MRMC studies are conducted outside of the clinical setting, with an enriched case set, and without patient history, the generalizability of some figures of merit such as recall rate, sensitivity and specificity is limited. The design is considered acceptable in order to reduce the size of the trial and avoid confounders.

Adding 2-view DBT as an adjunct to FFDM requires additional exposure to ionizing radiation. The overall FFDM plus 2-view DBT mammography exam remains a low dose examination. The risk associated with exposure to low dose radiation is theoretical and long-term, while an undetected breast cancer, particularly of the invasive type, is an immediate risk to a patient.

Additional factors to be considered in determining probable risks and benefits for the MAMMOMAT Inspiration system with Tomosynthesis Option included a few study design choices that added some uncertainty to the estimated difference of reader performance between FFDM and FFDM plus 2-view DBT. The most notable issues were the method used to assign a POM score to a breast by combining the lesion level scores,

the limited sampling of the ROC domain, the exclusion of a subset of cancers not detected at baseline but diagnosed at the time of follow-up, and the exclusion of breasts from the MRMC analyses due to the lack of ground truth verification. Due to the large performance improvement that results from adding 2-view DBT to FFDM, the impact of these design choices is unlikely to be large enough to alter the study conclusions.

In conclusion, given the available information described above, the data support that the probable benefits of using the MAMMOMAT Inspiration system with Tomosynthesis Option as an adjunct to FFDM (in accordance with the indications for use) outweigh the probable risks.

D. Overall Conclusions

The data in this application support the reasonable assurance of safety and effectiveness of the MAMMOMAT Inspiration device with Tomosynthesis Option when used in accordance with the indications for use.

XIII. CDRH DECISION

CDRH issued an approval order on April 21, 2015. There were no conditions of approval.

The applicant's manufacturing facility has been inspected and found to be in compliance with the device Quality System (QS) regulation (21 CFR 820).

XIV. APPROVAL SPECIFICATIONS

Directions for use: See device labeling.

Hazards to Health from Use of the Device: See Indications, Contraindications, Warnings, Precautions, and Adverse Events in the device labeling.

Post-approval Requirements and Restrictions: See approval order.

XV. REFERENCES

[1] DeLong E, DeLong D, Clarke-Pearson D. "Comparing areas under two or more correlated receiver operating characteristic curves: A nonparametric approach." Biometrics. 1988; 44: 837-845.

[2] Obuchowski NA, Rockette HE. Hypothesis testing of the diagnostic accuracy for multiple diagnostic tests: An ANOVA approach with dependent observations. Communications in Statistics: Simulation and Computation. 1995; 24: 285-308.

[3] Obuchowski NA. "New methodological tools for multiple-reader ROC studies." Radiology. 2007; 243:10-12.

[4] 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; 24:1579-607.

[5] Chen W and Samuelson FW, “The average receiver operating characteristic in multireader multicase imaging studies”, British Journal of Radiology. 2014; 87 (1040); Aug. 2014.