

SUMMARY OF SAFETY AND EFFECTIVENESS DATA (SSED)

I. GENERAL INFORMATION

Device Generic Name: PowerLook® Tomo Detection Software

Device Trade Name: PowerLook® Tomo Detection Software

Device Procode: MYN

Applicant's Name and Address:

John DeLucia, VP Regulatory Affairs, Clinical Affairs and Quality Assurance
iCAD, Inc.
98 Spitbrook Road, Suite 100
Nashua, NH 03062

Date(s) of Panel Recommendation: None

Premarket Approval Application (PMA) Number: P160009

Date of FDA Notice of Approval: March 24, 2017

II. INDICATIONS FOR USE

The iCAD PowerLook® Tomo Detection Software is a computer-assisted detection (CAD) software device intended to be used concurrently by radiologists while reading GE Senoclaire breast tomosynthesis exams. The system detects up to five soft tissue densities (masses, architectural distortions and asymmetries) in the 3D tomosynthesis images. The detections are blended with the standard 2D synthetic image and the CAD-enhanced 2D synthetic image is viewed on a mammography review workstation.

The CAD-enhanced 2D synthetic image assists radiologists in identifying densities (masses, architectural distortions and asymmetries) that may be confirmed or dismissed by the radiologist in the digital breast tomosynthesis (DBT) images.

III. CONTRAINDICATIONS

There are no known contraindications.

IV. WARNINGS AND PRECAUTIONS

- Blended areas on the Enhanced V-Preview image may obscure other areas of interest that may have been visible on the unenhanced V-Preview image.

- Incorrect implementation of the CAD-enhanced synthetic image can lead to an inability to view the CAD-enhanced image or the inability to navigate into the 3D dataset using the product.
- The CAD does not detect any microcalcifications and also may miss some soft tissue densities (masses, architectural distortions and asymmetries) so a complete review of the whole set of images indicated by the GE Senoclaire DBT system is necessary to identify calcifications and other breast lesions that CAD may have missed.
- The safety and effectiveness in patients with breast implants has not been established for views that include the implant. When non-displaced implant views are analyzed by the system, the CAD enhanced V-Preview images should not be used.
- The safety and effectiveness of CAD has not been established for non-standard mammographic views (e.g., magnification/compression views). When these views are analyzed by the system, the CAD enhanced V-Preview images should not be used.
- The safety and effectiveness of CAD has not been established for the following standard mammographic views (CC, ML, LM, SIO, XCCL, XCCM).

V. **DEVICE DESCRIPTION**

The iCAD PowerLook® Tomo Detection Software in conjunction with the GE Senoclaire digital breast tomosynthesis (DBT) and a mammography review workstation create a CAD-enhanced 2D synthetic image that can allow radiologists to review tomosynthesis images more quickly without a statistically significant decrease in radiologist's performance. The product detects soft tissue densities (masses, architectural distortions and asymmetries) in the 3D tomosynthesis images. A blending algorithm then processes the CAD detections from the 3D planes and merges them onto GE's existing 2D volume preview ("V-Preview") synthetic image and the mammography review workstation displays the CAD-enhanced 2D synthetic image. The PowerLook® Tomo Detection Software does not introduce any CAD marks onto the V-Preview image.

PowerLook® Tomo Detection Software uses pattern recognition technology which is applied to the 3D dataset to identify potential regions of interest which are soft tissue density findings (masses, architectural distortions and asymmetries). Once these regions are identified, there is an enhancement step that aims to project any identified 3D structure onto the V-Preview image through a blending process. This enhancement step has the effect of removing overlapping structures above and below the structure of interest without showing any CAD marks on the CAD-enhanced 2D synthetic view. The blending is on top of the existing V-Preview image, in order to be consistent with the existing V-Preview image.

The pattern recognition technology used is similar to the technology used in computer-aided detection devices and involves training classifiers using sample soft tissue densities. Five detections per 3D volume are enhanced, where a DBT volume is the set of 3D planes or slabs from a DBT acquisition, e.g. a right breast MLO DBT view. Five detections per DBT volume was chosen based on customer feedback with prototype

versions of the PowerLook® Tomo Detection Software. Using demonstration cases radiologists felt that more than five detections per DBT volume could be distracting, while five would be helpful.

VI. ALTERNATIVE PRACTICES AND PROCEDURES

As an alternative to the use of this device, a radiologists may elect to review GE Senoclaire breast tomosynthesis exams without the assistance of CAD. This should have no effect on the patient but may increase the time spent by the radiologist to review the study.

No alternate computer-assisted detection concurrent read software either for 2D mammography or for digital breast tomosynthesis currently exists.

Each alternative has its own advantages and disadvantages. A patient should fully discuss these alternatives with his/her physician to select the method that best meets expectations and lifestyle.

VII. MARKETING HISTORY

The PowerLook® Tomo Detection Software has not been marketed in the United States or any foreign country.

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 the device.

- The PowerLook® Tomo Detection Software may increase false-positive or false-negative rates for both screening and diagnostic mammography. Increased false-positives may lead to unnecessary additional imaging radiation exposure, biopsy, patient anxiety, etc. Increased false-negative rates may lead to missed that would otherwise have been found.

For the specific adverse events that occurred in the clinical study, please see Section X below.

IX. SUMMARY OF NONCLINICAL STUDIES

In compliance with the requirements of 21 CFR 820.30 Design Controls, iCAD, Inc. performed verification and validation testing on the PowerLook® Tomo Detection Software. Verification testing consisted of software unit testing, software integration testing, and software system testing. The purpose of the verification test was to assure that the software application satisfied the software requirements. Verification testing was successfully completed. Several minor anomalies remain in the software application but

these anomalies have no impact on the safety and effectiveness of the device and no impact on the operator usage.

Device validation testing consisted of measuring CAD standalone performance, including sensitivity and false positives rate. The results of CAD standalone performance testing, using an internal clinical image evaluation database consisting of soft tissue densities (52 malignant lesions with 43 soft tissue density only lesions and 9 mixed lesions in 48 cases with 39 soft tissue density only cases and 9 mixed cases) and soft tissue densities and calcifications (68 malignant lesions with 43 soft tissue density only lesions, 16 calcifications only lesions and 9 mixed lesions in 61 cases with 39 soft tissue density only cases, 13 calcifications only cases and 9 mixed cases) are provided below:

CAD Standalone Performance		
	Lesion-Level with 95% CIs**	Case-Level with 95% CIs***
CAD Sensitivity for Soft Tissue Densities	0.923 (48/52) 0.849~0.997	0.917 (44/48) 0.804~0.967
CAD Sensitivity for Soft Tissue Densities and Calcifications*	0.706 (48/68) 0.588~0.824	0.721 (44/61) 0.598~0.818
CAD Detection Rate Per Image: 5****	-	-
*Note: calcifications are not detected by iCAD's PowerLook® Tomo Detection Software Computation of confidence intervals: ** Rao JN, Scott AJ. A simple method for the analysis of clustered binary data. <i>Biometrics</i> 1992;48(2):577-85. *** Wilson EB. Probable inference, the law of succession, and statistical inference. <i>J Am Stat Assoc</i> 1927;22:209-212. **** The system provides up to 5 marks, so the false positive rate for cases without disease is 100%.		

Truthing Radiologist's Visibility Sensitivity*		
	Lesion-Level with 95% CIs**	Case-Level with 95% CIs***
Standard 2D V-Preview Sensitivity for Soft Tissue Densities	0.519 (27/52) 0.372~0.667	0.521 (25/48) 0.383~0.655
Standard 2D V-Preview Sensitivity for Soft Tissue Densities and Calcifications	0.632 (43/68) 0.506~0.758	0.623 (38/61) 0.497~0.734
CAD-enhanced 2D V-Preview Sensitivity for Soft Tissue Densities	0.904 (47/52) 0.824~0.984	0.917 (44/48) 0.804~0.967
CAD-enhanced 2D V-Preview Sensitivity for Soft Tissue Densities and Calcifications	0.912 (62/68) 0.844~0.979	0.918 (56/61) 0.822~0.964

*Note: If the lesion was judged by the truthing radiologist to be visible enough on the standard 2D V-preview image to be detected in clinical practice by a radiologist without knowledge that the lesion was present in the case, then the lesion was considered visible (i.e., a TP in the sensitivity computation).
 Computation of confidence intervals:
 ** Rao JN, Scott AJ. A simple method for the analysis of clustered binary data. *Biometrics* 1992;48(2):577-85.
 *** Wilson EB. Probable inference, the law of succession, and statistical inference. *J Am Stat Assoc* 1927;22:209-212.

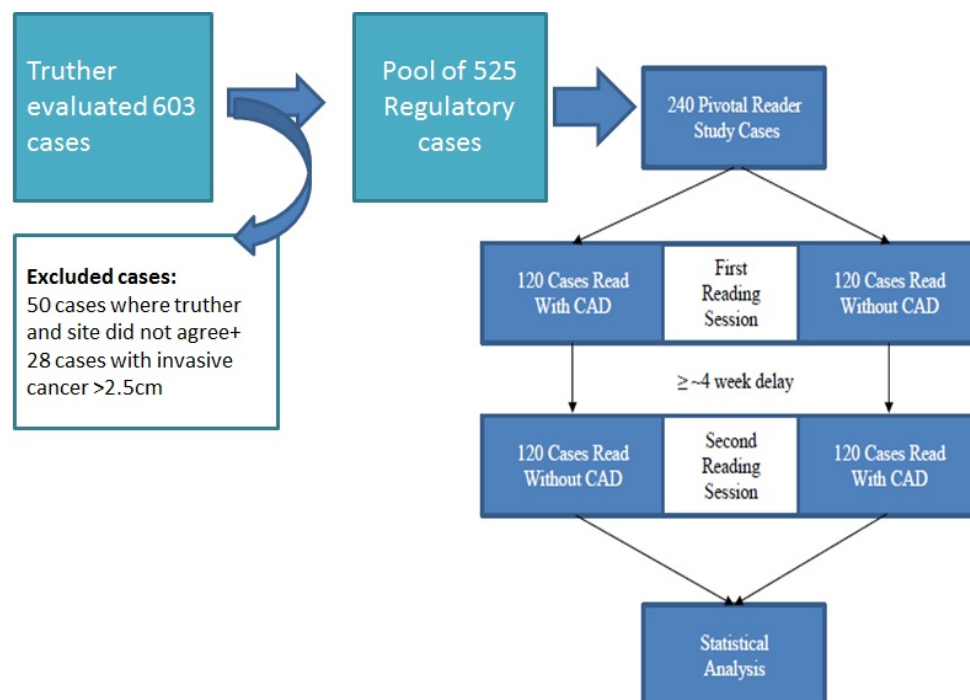
X. SUMMARY OF PIVOTAL CLINICAL STUDY

The applicant performed a clinical study to establish a reasonable assurance of safety and effectiveness with the PowerLook® Tomo Detection Software for review of GE Senoclaire breast tomosynthesis exams in the US and France. This was a non-significant risk investigation that did not require FDA approval of an IDE. Data from this clinical study were the basis for the PMA approval decision. A summary of the clinical study is presented below.

A. Study Design

The Pivotal Reader Study was a retrospective, multi-reader, multi-case study. Patient data was acquired between 2/7/2014 and 7/30/2015 from five investigational sites (2 in the US, 3 OUS). The dataset collected for this PMA included 603 cases.

The truther evaluated all 603 acquired cases; 78 cases were excluded from the regulatory case pool either because the truthing radiologist's assessment of the case differed from the site's assessment (50 cases) or the cases had invasive cancer >2.5cm (28 cases). 240 of the remaining 525 cases in the regulatory case pool were selected for inclusion in the pivotal reader study.



The study had a cross-over design (reading with/without CAD) so that each subject served as their own control.

1. Clinical Inclusion and Exclusion Criteria

Enrollment in the Pivotal Reader study was limited to patients who met the following inclusion criteria:

- Female subjects;
- Subjects 18 or more years old;
- Subjects undergoing a bilateral screening or diagnostic tomosynthesis exam with GE Senoclaire, which must include at least bilateral 2D Craniocaudal (CC) FFDM images and bilateral Mediolateral Oblique (MLO) DBT images (2D V-Preview, 3D slabs, 3D planes and 2D projections);
- Imaging characteristics consistent with Mammography Quality Standards Act (MQSA) in terms of positioning, compression, exposure level, dose, contrast, sharpness, noise and artifacts;
- Subjects with breast implants were permitted if implant-displaced views are available for the required images: 2D CC FFDM images and MLO DBT images (2D V-Preview, 3D slabs, 3D planes and 2D projections);
- “Regulatory” cases, which are clinical and image data that is controlled by Quality & Regulatory Affairs (QA/RA) and were blinded to R&D (i.e., unavailable to R&D); and
- Recalled, Benign and Cancer cases were required to have additional mammographic views showing the location of the lesion (e.g., magnification and/or spot compression views, screenshots, needle-localization or biopsy images, etc.) and/or annotations by the contributing site radiologists outlining the lesion with the GE MammoWorkstation and/or radiology reports describing the location of the lesion for the truthing process

Patients were not permitted to enroll in the Pivotal Reader Study study if they met any of the following exclusion criteria:

- Subjects with a personal history of breast cancer;
- Subjects with imaging evidence of previous surgery (e.g., surgical clips visible on images); these women were excluded because readers were not provided patient history or prior exams during the pivotal reader study which would have been important in the assessment of these women’s images; subjects with breast reduction surgery were not excluded and were permitted; or
- Were included in the “Development” cases, which are clinical and image data that were not blinded to Research & Development (R&D) (i.e., available to R&D)

2. Follow-up Schedule

As this was a retrospective clinical study, no additional patient follow up was indicated. All patients whose images were included in the case pool were scheduled to return for additional imaging per current clinical screening guidelines.

B. Study Readers

The Pivotal Reader Study included 20 readers.

All 20 readers had American Board of Radiology certification, qualified to interpret mammograms under MQSA and had completed eight hours of initial training in breast tomosynthesis as required by the FDA.

Readers had a range of experience in the interpretation of breast images with 55% (11/20) of readers devoting less than 75% of their professional time to breast imaging for the last 3 years and 45% (9/20) devoting 75% or more of their professional time to breast imaging for the last 3 years.

A summary of reader characteristics is provided below:

Summary of Reader Experience : N(%) unless otherwise noted	
<i>Years in Practice</i>	
Mean (IQR)	14.5 (8.0 - 20.5)
Range	1.0 - 43.0
Mean (SD)	15.45 (10.86)
<i>Specialized Mammography Training</i>	
Yes	5 (25)
No	15 (75)
<i>Average Hours Spent in a Clinical Day</i>	
7	1 (5)
8	2 (10)
9	1 (5)
9.5	16 (80)
<i>75% or more Professional Time Devoted to Breast Imaging for the Last 3 years</i>	
Yes	9 (45)
No	11 (55)
IQR = interquartile range, 25th percentile through 75th percentile. SD = standard deviation.	

Readers were trained in study reading procedures with 30 tomosynthesis cases. Since the readers were fully certified to interpret tomosynthesis exams and were currently reading tomosynthesis exams in their clinical practices, no additional tomosynthesis interpretation training was provided.

C. Study Endpoints

The Pivotal Reader Study had two co-primary study endpoints:

- Whether radiologist performance when using CAD with GE DBT images is non-inferior to radiologist performance when using GE DBT images without CAD [measured as the lower limit of the two-sided 95% confidence interval for

the difference in average AUC with CAD – without CAD lies above the negative of the non-inferiority margin, -0.05]

and

- Whether radiologist reading time when using CAD with GE DBT images is superior to (shorter than) radiologist reading time when using GE DBT images without CAD [measured as the upper limit of the two-sided 95% confidence interval for the difference in average reading time with CAD – without CAD lies below zero, i.e., if reading time decreases]

Both co-primary hypotheses must be supported to declare that the primary objectives of the pivotal study were met.

D. Study Cases

The dataset collected to support this PMA included 603 cases retrospectively collected from 5 institutions (2 US, 3 OUS) over the course of a 17 month period (2/17/2014 – 7/30/2015).

A single truther evaluated all 603 acquired cases. The same individual truthing radiologist truthed every case. The reference standard for Cancer and Benign cases was biopsy proof. The reference standard for Recalled cases was BI-RADS 0 without biopsy and normal imaging at least one year (320 or more days inclusive) later. The reference standard for 262 of the 264 Negative cases was normal one-year follow-up imaging (320 or more days inclusive). Two Negative cases did not have one-year follow-up for the following reasons:

- One case was initially misclassified as Benign during data collection. Since Benign cases were not required to have one-year follow-up, this case does not have one-year follow-up.
- One Negative case in which the one-year follow-up images from 2015 were mistakenly used for the reader study instead of the images from 2014. As of November 7, 2016, the patient had not returned to the site for any additional mammography or DBT so this case does not have one-year follow-up.

78 cases were excluded from the regulatory case pool either because the truthing radiologist's assessment of the case differed from the site's assessment (50 cases) or the cases had invasive cancer >2.5cm (28 cases). This left 525 cases in the regulatory case pool. Characteristics of the 525 regulatory cases were as follows:

- **264 Negative Cases:** No BI-RADS 0, 3, 4 or 5 lesions detected by radiologist on tomosynthesis exam interpreted as BI-RADS 1 or 2 and no cancers were found. Negative cases had at least one confirming normal tomosynthesis exam or 2D mammogram (BI-RADS 1 or 2) at least one year (320 or more days) after the study tomosynthesis exam.
- **71 Recalled Cases:** Lesion(s) detected by radiologist on screening tomosynthesis exam interpreted as BIRADS 0 where no biopsy was warranted due to subsequent negative imaging work-up. Data collection sites may not have

used BI-RADS 0 if they performed an imaging workup, when needed, immediately after the screening the exam while the patient was still in the facility. Since the patient had an immediate diagnostic work-up, a final BI-RADS assessment category was used (which does not include BI-RADS 0). Recalled cases had at least one confirming normal tomosynthesis exam or 2D mammogram (BI-RADS 1 or 2) at least one year (320 or more days) after the study tomosynthesis exam.

- **86 Benign Cases:** Lesion(s) detected by radiologist on tomosynthesis exam interpreted as BI-RADS 3, 4 or 5 in a patient who had a breast biopsy within one year (365 or less days) of the tomosynthesis exam that was benign. Benign lesions included high-risk lesions lobular carcinoma in-situ, atypical ductal hyperplasia, atypical lobular hyperplasia, and papillomas.
- **104 Cancer Cases:** Lesion(s) detected by radiologist on tomosynthesis exam interpreted as BI-RADS 3, 4, or 5 in a patient who had a breast biopsy within one year (365 or less days) of the tomosynthesis exam that was malignant. Malignant lesions did not include high-risk lesions including lobular carcinoma in-situ, atypical ductal hyperplasia, atypical lobular hyperplasia and papillomas.

240 cases were randomly selected for inclusion in the Pivotal Reader Study with a stratification tool to meet the stratification targets based on the case and lesion characteristics established by the truthing process. All required targets were met. Case and lesion characteristics for the 240 selected cases are shown below. Selected cases were from women aged 29 – 86 with a median age of 57 (interquartile range 49 – 64) and a mean age of 56.5 (standard deviation 10.8).

Case Characteristic Criteria (240 selected cases)	
<i>Patient Criteria – Breast Density</i> • Almost entirely fatty or scattered areas of fibroglandular density • Heterogeneously dense or Extremely dense breasts	• 121/240 (50.4%) • 119/240 (49.6%)
<i>Negative Cases Criteria</i> • BI-RADS 1 or 2: not suspicious, no recall, no biopsy, not cancer - o BI-RADS 1 - o BI-RADS 2	• 102/240 (42.5%) - o 75/102 (73.5%) - o 27/102 (26.5%)
<i>Recalled Cases Criteria</i> • BI-RADS 0: suspicious, recall, but no biopsy warranted • Soft tissue densities - o Soft tissue densities and calcifications • Calcifications only	• 20/240 (8.3%) • 16/20 (80.0%) - o 1/20 (5.0%) • 4/20 (20.0%)
<i>Benign Cases Criteria</i> • BI-RADS 3, 4 or 5: suspicious, recall, biopsy-proven benign, not cancer • Soft tissue densities	• 57/240 (23.8%) • 40/57 (70.2%)

- o Soft tissue densities and calcifications • Calcifications only	- o 5/57 (8.8%) • 17/57 (29.8%)
<i>Cancer Cases Criteria</i> • BI-RADS 3, 4 or 5: suspicious, recall, biopsy-proven cancer • Soft tissue densities - o Soft tissue densities and calcifications • Calcifications only • Invasive Cancer (all ≤ 2.5cm in size) - o Proportion ≤ 1.4cm in size - o Proportion invasive lobular cancer • DCIS (no size restriction)	• 61/240 (25.4%) • 48/61 (78.7%) - o 9/61 (14.8%) • 13/61 (21.3%) • 50/61 (82.0%) - o 27/50 (54.0%) - o 7/50 (14.0%) • 11/61 (18.0%)

The cases included in the Pivotal Reader Study are typical for digital breast tomosynthesis studies performed in the US. It is noted that the Pivotal Reader Study dataset was enriched to include a higher percentage of cancer cases than would be observed in a screening population.

E. **Study Execution**

Study readings were conducted by Intrinsic Imaging, LLC (Bolton, MA), an imaging Clinical Research Organization (CRO), at their reading facility in San Antonio, TX. The twenty study readers assessed 2D FFDM and DBT exams from women presenting for bilateral screening or diagnostic tomosynthesis exams. The FFDM and DBT exams were assessed with and without CAD with a cross-over design such that each case was read by each reader both with and without CAD.

Reading was separated into two reading sessions. Half of the 240 cases were read by each reader (either with or without CAD) during each reading session. There was a 4-week or more washout period in between readings of the same case with and without CAD.

The fully-crossed study design meant that all 240 cases were read by each reader both with and without CAD. The case reading order was randomized separately for each reader.

The following images from both breasts of each case were available to each study radiologist in the control arm (cases read without CAD) of the study:

- 2D CC FFDM + standard 2D MLO V-Preview images
- 3D MLO planes
- 3D MLO slabs

The following images from both breasts of each case were available to each study radiologist in the intervention arm (cases read with CAD) of the study:

- 2D CC FFDM + standard 2D MLO V-Preview images
- 2D MLO CAD-enhanced V-Preview with navigation to 3D MLO DBT planes at locations of potential lesions

- 3D MLO slabs
- 3D MLO planes

Readers were not provided with prior images or patient history. The GE MammoWorkstation was used by the readers to review and assess the cases.

Readers completed a case report form for each of the two reads (with/without CAD). Readers were asked to document whether or not any lesions suspicious enough to warrant recall, short-term follow up, or tissue diagnosis were noted. For suspicious lesions, readers were asked to further provide:

- Location. The reader used a tool in the GE MammoWorkstation to annotate the specific location of the finding on the right CC, left CC for FFDM; right MLO, left MLO for standard V-Preview and 3D Planes.
- Mammographic appearance as soft tissue density, calcifications or mixed soft tissue density and calcifications. If soft tissue density or mixed, also whether non-spiculated mass, spiculated mass, architectural distortion or asymmetry.
- “Forced” BI-RADS assessment category 1, 2, 3, 4a, 4b, 4c, or 5. The case-level forced BIRADS was computed as the highest lesion-level forced BI-RADS in the case.
- Level of Suspicion (LOS) score 0 through 100%. The case-level LOS was computed as the highest lesion-level LOS in the case.

For cases not involving suspicious lesions, readers were asked to provide a case-level forced BI-RADS assessment category of 1, 2, 3, 4a, 4b, 4c, or 5 and a case-level LOS score (0 through 100%).

Readers were also asked to specify breast density (fatty, scattered fibroglandular, heterogeneously dense, extremely dense) for each case. For measuring reading time, the reader clicked “start” when he/she began viewing the images in the case, and “end” when he/she had finished reviewing the images. Reading time in seconds was calculated as the time between these entries. The reading time did not include the time needed to document the lesion location with the GE MammoWorkstation tool or to document the lesion and case characteristics in the CRF. Readers were informed that reading time was being measured and that the PowerLook® Tomo Detection Software was intended to reduce reading time, but readers were blinded to the reading time measurements for each case.

Consistency of BI-RADS scores and LOS scores was not forced – i.e., readers were permitted to use the full range of LOS scores for a finding no matter what BI-RADS score they assigned to it.

Lesion matching was performed by a validated electronic matching software tool. A reader’s finding was considered to match a breast cancer if the center point of the lesion location identified by the reader as a finding was within or coincident with the final outline of the malignant lesion, extended to the top and bottom delimiting planes

of the lesion, as determined by the truthing process. Visual confirmation of this matching was used to confirm that the electronic results were accurate.

F. Statistical Analysis Methods

The primary unit of analysis on this study is the subject (case). Malignant lesion is the secondary unit of analysis for per-lesion sensitivity. All cases were evaluated by all study readers; there were no treatment assignments or groups.

Co-Primary Endpoint 1: Non-inferior AUC

Radiologist performance for the detection of malignant lesions was assessed by measuring patient-level area under the receiver operating characteristic (ROC) curve (AUC). Malignant lesion localization was required for a reader to correctly detect cancer in a case. The general principle is that even at the subject level, credit is only given for identifying a subject with cancer if the reader marked findings in at least one location with cancer.

ROC curves and AUC values for each reader and their averages over readers were plotted/computed with/without CAD based on the per-subject LOS scores accounting for localization. The non-parametric (trapezoidal) method was used to compute AUC and the variance-covariance matrix of individual reader AUCs was obtained using the method of DeLong, et al. **Error! Reference source not found.** as extended by Zhou and Gatsonis [2] to allow for the missing reading data and accounting for partially paired data when estimating covariances. Statistical comparison of the reader-averaged AUCs between modalities accounting for both the reader variability and case variability was performed per the model of Obuchowski and Rockette [3] with adjusted F-statistic [4]. Two-sided 95% confidence intervals (CIs) were used to quantify uncertainty in the within-modality estimates and the between-modalities difference.

Multiple comparisons: The pre-defined plan to control the study's type I error rate at $\alpha = 0.05$ was conducted by performing hypothesis testing hierarchically in the pre-specified fixed sequence. The sequence was continued because non-inferiority of AUC was proved.

Co-Primary Endpoint 2: Superior Radiologist Reading Time

Obuchowski and Rockette's ANOVA model [3] was extended to analyze radiologist reading times. Like the MRMC analysis of AUC, this analysis also treated both readers and cases as random effects in assessing the uncertainty of the estimated reading time difference. The percent difference data were normalized before they were analyzed by the OR model using this transformation: $\text{natural log of } (\% \text{ difference} + 100) - \text{natural log of } 100$.

Treatment of Missing Data in Co-Primary Endpoints:

A number of reading times for the Pivotal Reader Study were reported as missing. Missing data resulted from readers not entering Start Times (75 cases) or End Times (94 cases) or the Start Time was reported to occur later than the End Time (2 cases). The CRO did not report these start and end times due to human reader error. These

account for about 4% of the differences in reading times (with CAD – without CAD). Robustness to treatment of missing times was evaluated through a worst-case analysis. Such analysis indicated that the conclusions reported were robust to treatment of missing data.

Secondary Endpoints associated with potential marketing claims:

Secondary objectives associated with potential marketing claims were to determine non-inferiority and possible superiority of radiologist per-lesion sensitivity, and possible superiority of AUC (subject-level), with CAD compared to without CAD.

1. Non-inferiority of per-lesion sensitivity (5% margin): iCAD used BI-RADS scores assigned to each finding and TP (BI-RADS 3 or higher), FN in the subgroup of cases with cancer to analyze lesion-level sensitivity. iCAD used methods for clustered data from MRMC studies that take into account the correlation between lesions in the same case. Rao and Scott (1992) method for estimating proportions from clustered data was used to obtain estimates for each reader in each reading condition [5], and Obuchowski (1998) extension of this to a pair of correlated proportions was used to estimate the variance-covariance matrix of all possible pairs of proportions [6]. Analysis accounted for the lack of data on Reader 2's interpretations of a single Cancer case with 1 lesion. The usual Obuchowski and Rockette method [3] was then applied to perform inferences that generalize to the population of readers and the population of cases while also taking into account within-case correlations between lesions. Non-inferiority was demonstrated.
2. Superiority of lesion-level sensitivity. This was not proved.
3. Superiority of AUC (subject-level). iCAD did not test for superiority of AUC because the preceding step in the pre-specified, fixed sequence of hypothesis testing, superior lesion-level sensitivity, was not proved. This was not proved.

Multiple comparisons: To protect the study's type I error rate from inflation, secondary study endpoints continued to be evaluated hierarchically in a pre-specified, fixed sequence. This hypothesis testing occurred because non-inferiority of AUC **and** superiority of radiologist reading time were both proved.

Secondary objectives to be reported but not associated with marketing claims:

Estimation of differences (with corresponding two-sided 95% confidence intervals) in sensitivity, specificity, recall rate, between WITH and WITHOUT CAD based on BI-RADS reader's assessment using BI-RADS 3 as a positivity threshold. Analysis used methods for MRMC studies similar to analysis of AUC. Secondary endpoints for diagnosis performance include per-subject and per-lesion sensitivity and specificity based on BI-RADS categories. Analysis of case-level sensitivity was location-specific.

In per-subject analysis:

- A true positive (TP) occurs when a case contains one or more cancerous lesions, and the reader records one or more findings that match the location of one or more cancerous lesions, and the reader assigns a BI-RADS score of 3 or higher to one of those findings, such that the per-case BI-RADS score as derived above is 3 or higher.
- A true negative (TN) occurs when a case does not contain any cancerous lesions, and the reader records no findings in the case and assigns a per-case BI-RADS of 1 or 2; or the reader records findings in the case and the maximum BI-RADS score assigned to these findings is 1 or 2; such that the per-case BI-RADS score as derived above is 1 or 2.
- A false negative (FN) occurs when a case contains one or more cancerous lesions, and the reader records no findings in the case and assigns a per-case BI-RADS of 1 or 2; or the reader records findings in the case and the maximum BI-RADS score assigned to these findings is 1 or 2; or the reader records findings in the case but no findings match the location(s) of any proven malignancies in this case; such that the per-case BI-RADS score as derived above is 1 or 2.
- A false positive (FP) occurs when a case does not contain any cancerous lesions, and the reader records no findings in the case but assigns a per-case BI-RADS of 3 or higher; or the reader records findings in the case and the maximum BI-RADS score assigned to these findings is 3 or higher; such that the per-case BI-RADS score as derived above is 3 or higher.

G. Safety and Effectiveness Results

The co-primary endpoints of the pivotal reader study were met. The primary pivotal study results are summarized as follows:

- **Area Under the ROC Curve (AUC): Non-inferiority:** Radiologist performance when using CAD with GE DBT images is non-inferior to radiologist performance when using GE DBT images without CAD, for non-inferiority margin $\Delta = 0.05$. Average AUC increased by 0.009 (95% CI: (-0.012, 0.030); non-inferiority $p < 0.01$), from 0.841 without CAD to 0.850 with CAD.
- **Radiologist reading time: Superiority:** Radiologist reading time when using CAD with GE DBT images is superior to (shorter than) radiologist reading time when using GE DBT images without CAD. Reading time improved 29.2% with CAD (95% CI: (21.1%, 36.5%); $p < 0.01$). These values were obtained through back-transformation from statistical inferences performed using the normalizing transformation $\ln(\% \text{ difference} + 100) - \ln(100)$, because % difference was not bell-shaped. The average decrease in reading time was 19.6 seconds (95% CI: (12.6 seconds, 26.6 seconds); $p < 0.01$).

Secondary endpoints associated with potential marketing claims:

- **Lesion-level sensitivity: Non-inferiority:** The study demonstrated non-inferior lesion-level sensitivity using CAD-enhanced 2D V-Preview images with GE DBT

compared to using GE DBT without CAD-enhanced 2D V-Preview images, for non-inferiority margin $\delta = 0.05$. Average lesion-level sensitivity increased by 0.021 (95% CI: -0.007, 0.050; non-inferiority $p < 0.01$), from 0.823 without CAD to 0.845 with CAD.

- **Lesion-level sensitivity: Superiority:** Non-inferior lesion-level sensitivity was proved, so superiority was tested. Superior lesion-level sensitivity was not proved ($p = 0.14$).
- **AUC: Superiority:** Superiority of AUC was not tested because the preceding step in the pre-specified, fixed sequence of hypothesis testing, superior lesion-level sensitivity, was not proved.

Adverse effects that occurred in the PMA clinical study:

- No adverse events were reported by the Radiologists (Readers) or Principal Investigator as part of the Pivotal Clinical Retrospective Reader Study.

Subgroup / Additional Analyses:

The following secondary endpoints were explored. The following were observed in the clinical study:

- **Sensitivity (per-subject):** Average sensitivity was 0.847 without CAD, and 0.871 with CAD.
- **Sensitivity for Calcifications:** With respect to sensitivity in the subgroup of 13 cancer cases with only calcifications, 0.881 was observed without CAD, and 0.873 with CAD.
- **Sensitivity for Soft Tissue Densities:** Performance in the subgroup of 48 cancer cases with at least one soft tissue density or mixed lesion was observed to be 0.837 without CAD and 0.871 with CAD.
- **Specificity:** Average specificity was observed to be 0.527 without CAD and 0.509 with CAD.
- **Recall rate for non-cancer cases:** Average recall rate was observed to be 0.474 without CAD, and 0.492 with CAD.
- **Patients with breast reduction surgery:** At FDA's request, iCAD performed a sub-analysis of device performance in patients with breast reduction surgery. FDA was concerned that breast reduction surgery can create imaging findings similar to some of the soft tissue densities the device is intended to detect. Six of 240 cases included in the Pivotal Reader study had prior breast reduction surgery. These 6 subjects represented one cancer case and 5 negative cases. No evidence was presented by iCAD to indicate that more false positives were generated for

patients with prior breast reduction surgery: for the 5 negative cases, there were 23 readings with raised BI-RADS ratings, 23 readings with lowered ratings, and 54 readings with no change.

Pediatric Extrapolation:

In this premarket application, existing clinical data was not leveraged to support approval of a pediatric patient population.

H. Study Limitations

The study carried out by iCAD had several significant limitations, including lack of one year follow up for a number of study cases and exclusion of some study cases from the regulatory case pool. Left unaddressed, these study limitations would have precluded approval of the PMA. Additional analyses were required to address these issues, which delayed approval of the device.

50 cases from the ‘intend to diagnose’ and intended use population were excluded from the eligible regulatory case pool because the truthing radiologist’s assessment of the case differed from the site’s assessment. To investigate the concern of whether exclusion of these 50 cases biased the study conclusions, iCAD performed analyses [7-9] to incorporate these 50 cases into the regulatory case pool using two Multiple Imputation methodologies. These analyses supported the robustness of the study conclusions with respect to the co-primary study endpoints.

In the analysis of the Primary Study endpoints, reader performance and reading times for individual readers were not linked. That is, non-inferiority was determined via a pooled analysis of reader performance and reading times – meaning, there was no penalty for a faster reads that were incorrectly read.

One year follow up data was not available for any of the 57 biopsy-proven benign cases included in the Pivotal Reader Study. FDA determined that biopsy-confirmed benign cases may be included in the study in the absence of one year imaging follow up provided the biopsy was surgical/excisional, or a large-core needle biopsy with a specific benign finding (for example, fibroadenoma, fibrocystic change, etc.,) that was concordant with the suspicious pre-biopsy imaging finding. Biopsies with fine needle aspiration cytology or with pathology results returned as ‘benign’ or ‘breast tissue’ continue to require one year imaging follow up. In iCAD’s study case pool, 39/57 cases met the above criteria and were acceptable for inclusion in the study without one year imaging follow up; 18/57 cases did not. iCAD performed sensitivity-tipping point analyses to evaluate the robustness of the study results to missing one-year follow-up verification in these subjects. The results of these analyses supported the robustness of the study conclusions and hence it was deemed acceptable to include these cases in the study in the absence of one year follow up data.

I. 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 23 investigators (20 of whom were readers). None of the clinical investigators 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. SUMMARY OF SUPPLEMENTAL CLINICAL INFORMATION

None

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

In accordance with the provisions of section 515(c)(3) of the act as amended by the Safe Medical Devices Act of 1990, this PMA was not referred to the Radiological Devices Panel, an FDA advisory committee, for review and recommendation because the information in the PMA substantially duplicates information previously reviewed by this panel.

XIII. CONCLUSIONS DRAWN FROM PRECLINICAL AND CLINICAL STUDIES

A. Effectiveness Conclusions

In the Pivotal Reader study average radiologist performance in ROC AUC when using CAD-enhanced 2D V-Preview images with GE DBT images was non-inferior to average radiologist performance in AUC when using GE DBT images without CAD-enhanced 2D V-Preview images (non-inferiority margin $\delta = 0.05$). Average AUC increased by 0.009 (95% CI: -0.012, 0.030; non-inferiority $p < 0.01$), from 0.841 without CAD to 0.850 with CAD. Average radiologist reading time was improved with CAD.

B. Safety Conclusions

The risks of the device are based on data collected in the clinical studies conducted to support PMA approval as described above. The PowerLook® Tomo Detection Software is used as an aid in the interpretation of existing DBT images. It does not modify the existing DBT image acquisition process nor alter the native DBT images. Additionally, the use of the software does not directly involve the patient. Therefore, there are no known direct safety or health risks caused by, or related to, the use of the device.

The PowerLook® Tomo Detection Software does not detect calcifications, so these could potentially be missed if the radiologist is relying solely on the CAD results.

The risk that the device presents is increased false positives or false negatives;

The risks from false-positive or false-negative results for diagnostics are:

- i. When reading DBT images with the aid of PowerLook® Tomo Detection Software, there is a chance that users will be influenced in their final decision by the CAD results, potentially leading to additional diagnostic work-up on CAD findings to determine true positive or negative. The potential additional workup may include diagnostic imaging and breast biopsy.
- ii. When reading DBT images with the aid of CAD, there is a chance that users may rely too heavily on the absence of CAD findings without sufficiently assessing the native DBT images. This may result in missing cancers that may have otherwise been found.

However, this risk is mitigated by the fact that the device is an aid and the radiologist is always making the final decision.

C. Benefit-Risk Determination

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

This device does not introduce unacceptable risk to the patient. The device allows radiologists to review tomosynthesis images more quickly without a statistically significant decrease in radiologist performance. The shorter reading time provides an aggregate benefit in that the radiologist is able to review more patient images.

This submission did not include specific information on patient perspectives for this device.

In conclusion, given the available information above, the data support that for

The iCAD PowerLook® Tomo Detection Software is a computer-assisted detection (CAD) software device intended to be used concurrently by radiologists while reading GE Senoclaire breast tomosynthesis exams. The system detects up to five soft tissue densities (masses, architectural distortions and asymmetries) in the 3D tomosynthesis images. The detections are blended with the standard 2D synthetic image and the CAD-enhanced 2D synthetic image is viewed on a mammography review workstation.

The CAD-enhanced 2D synthetic image assists radiologists in identifying densities (masses, architectural distortions and asymmetries) that may be confirmed or dismissed by the radiologist in the digital breast tomosynthesis (DBT) images.

the probable benefits outweigh the probable risks.

D. Overall Conclusions

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

The Pivotal Reader study demonstrated that the iCAD PowerLook® Tomo Detection software helped radiologists achieve a statistically-significant reduction in average interpretation time.

XIV. CDRH DECISION

CDRH issued an approval order on March 24, 2017. The final conditions of approval cited in the approval order are described below.

The applicant's manufacturing facilities were not inspected as a Quality System (QS) regulation (21 CFR 820) had recently been conducted and the applicant's manufacturing facilities were previously found to be in compliance with Quality System regulations.

XV. 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.

XVI. REFERENCES

- [1] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing areas under two or more correlated receiver-operating characteristic curves: a non-parametric approach. *Biometrics* 1988 44:837-845.
- [2] Zhou XH, Gatsonis CA. A simple method for comparing correlated ROC curves using incomplete data. *Statistics in Medicine* 1996; 15(15):1687-1693.
- [3] 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.
- [4] Hillis SL. A comparison of denominator degrees of freedom methods for multiple observer ROC analysis. *Stat Med* 2007; 26:596-619.
- [5] Rao JNK, Scott AJ. A simple method for the analysis of clustered binary data, *Biometrics* 1992; 48:577-585.
- [6] Obuchowski NA. On the comparison of correlated proportions for clustered data. *Statistics in Medicine* 1998; 17(13):1495-507.]
- [7] Little RJA, Rubin DB. *Statistical analysis with missing data*, 2nd ed. New York: John Wiley, 2002, Section 10.2, pp 209-211.
- [8] Rubin DB. *Multiple imputation for nonresponse in surveys*. New York: John Wiley, 1987.
- [9] Schafer JL. Multiple imputation: A primer, *Stat Methods Med Res* 1999; 8: 3-15.