EVALUATION OF AUTOMATIC CLASS III DESIGNATION FOR
QuantX

DECISION SUMMARY

A. DEN Number

DEN170022

B. Purpose for Submission

De novo request for evaluation of automatic class III designation for QuantX.

C. Applicant

Quantitative Insights, Inc.

D. Proprietary and Established Names

QuantX

E. Regulatory Information

1. Regulation section

   21 CFR 892.2060

2. Classification

   Class II (Special Controls)

3. Product code

   POK

4. Panel

   90 (Radiology)

F. Indications for Use

1. Indications for Use

   QuantX is a computer-aided diagnosis (CADx) software device used to assist radiologists in the assessment and characterization of breast abnormalities using MR image data. The software automatically registers images, and segments and analyzes user-selected regions of interest (ROI). QuantX extracts image data from the ROI to provide volumetric analysis and computer analytics based on morphological and enhancement characteristics. These imaging (or radiomic) features are then synthesized by an artificial intelligence algorithm into a single

value, the QI score, which is analyzed relative to a database of reference abnormalities with known ground truth.

QuantX is indicated for evaluation of patients presenting for high-risk screening, diagnostic imaging workup, or evaluation of extent of known disease. Extent of known disease refers to both the assessment of the boundary of a particular abnormality as well as the assessment of the total disease burden in a particular patient. In cases where multiple abnormalities are present, QuantX can be used to assess each abnormality independently.

This device provides information that may be useful in the characterization of breast abnormalities during image interpretation. For the QI score and component radiomic features, the QuantX device provides comparative analysis to lesions with known outcomes using an image atlas and histogram display format.

QuantX may also be used as an image viewer of multi-modality digital images, including ultrasound and mammography. The software also includes tools that allow users to measure and document images, and output in a structured report.

Limitations: QuantX is not intended for primary interpretation of digital mammography images.

2. Special conditions for use statement(s)

   For prescription use only

3. Warnings, precautions, and limitations

   QuantX is not intended for primary interpretation of digital mammography images.

   Please refer to the labeling for a more complete list of warnings, precautions and contraindications.

G. Device Description

The device is a software-only post-processing system for patient breast images that includes analysis of MR images, and viewing ultrasound and mammographic images.

MR images are acquired from a third-party acquisition device. The images can be loaded into the QuantX device manually or automatically if connected to a DICOM-compatible device. Users select and load the patient case to use the QuantX software tools in the examination of the images. Different types of MR sequences (T1, DCE, T2, DWI, etc.) can be viewed at the same time as mammography or ultrasound images from the same patient.

QuantX includes image registration, and automated segmentation and analysis functions, based on a seed point indicated by the user. Users can select a ROI manually from the MR image, or use the automatic segmentation tool to obtain and accept a ROI, for input to the QuantX analytics. The QuantX analytics display the QI Most Enhancing Curve, the Average Enhancing Curve, and volume of the specified region.

QuantX provides users the QI Score, based on the morphological and enhancement characteristics of the region of interest. The QuantX package provides comparative analysis for the QI score and its component element features to lesions with known ground truth (either biopsy- proven diagnosis or minimum one year follow-up negative scan for non-biopsied lesions) using an image atlas and histogram display format.

A user experienced with the significance of such data will be able to view and interpret this additional information during the diagnosis of breast lesions.

Users may select from a variety of information sources to make the diagnosis. The key features of the device are related categorization of lesions include the display of similar cases and the histogram of known lesions for various analytic features (included the QI score). The QI Score is not a "probability of malignancy," but is intended for the organization of an online atlas (reference database) provided to the user as the Similar Case Database. The QI score is based on a machine learning algorithm, trained on a subset of features calculated on a segmented lesions.

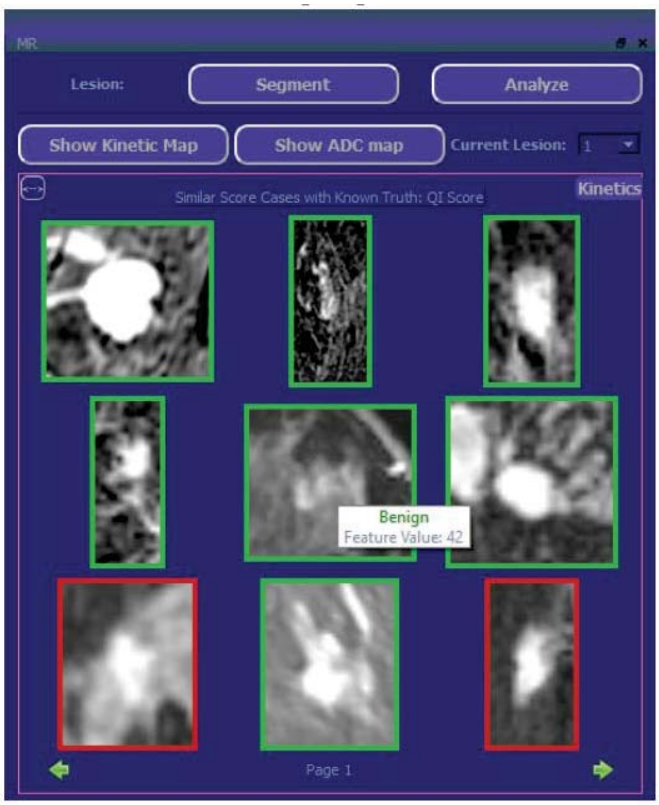

Figure 1: The User Manual that demonstrates how the device provides the 45 most similar cases for the currently selected feature. The cases chosen from the Similar Case Database are those with the smallest absolute difference for the selected lesion feature. The 45 most similar cases are displayed for the user without any restrictions on minimum or maximum similarity value.

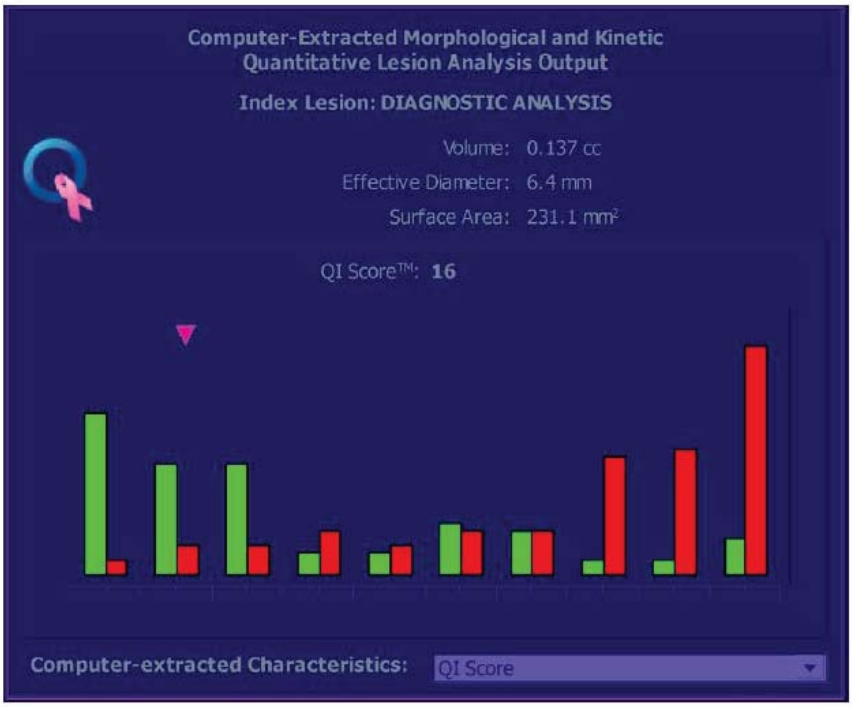

Figure 2: The histogram shows the distribution of lesions in the similar case database and the score of the current lesion relative to the database (pink arrow).

In both of the above screenshots (Figures 1 and 2) red represents malignant and green represents benign.

1. Image-based analytic features

   Each of the features available in the software has been well-described in mathematical language in the software documentation and labeling. A summary of each is provided below:

   QI Score™ Combines multiple lesion features to obtain a single measure for the lesion and is used to find similar cases.
   Max Uptake Is the maximum normalized uptake value of the lesion.
   Time to Peak Is the post-contrast time at which enhancement is maximum.
   Uptake Rate The uptake rate of contrast enhancement.
   Washout Rate Is the rate at which contrast enhancements washes out of the lesion.
   Curve Shape Index Is the difference of late and early enhancement relative to the initial enhancement of the lesion.
   Enhancement$_{Mid}$ (E2) Is the normalized uptake at the mid timepoint (1 1/2–2 minutes post-contrast injection).
   SER (Signal Enhancement Ratio) The ratio of initial enhancement to overall enhancement.
   Sphericity A measure of the conformity of the lesion to a spherical shape.
   Irregularity The deviation of the lesion's surface area from a spherical surface.
   Effective radius Is the radius of a sphere having the same volume as the lesion.
   Contrast A measure of the lesion's local image variations.
   Correlation A measure of the lesion's image linearity.

Diff Entropy A measure of the randomness of the difference of neighboring voxel values.
Diff Variance A measure of variation of the difference of voxel values between voxel pairs in the lesion.
Uniformity A measure of image homogeneity of the lesion.
Entropy A measure of the randomness of the voxel values in the lesion.
Homogeneity A measure of the local homogeneity of the lesion.
IMC1 A measure of non-linear voxel value dependence of the lesion.
IMC2 A measure of non-linear voxel value dependence of the lesion.
Max CC A measure of non-linear voxel value dependence to neighboring lesion voxels.
Sum Average A measure of the overall brightness of the lesion.
Sum Entropy A measure of the randomness of the sum of the voxel values of neighboring voxels in the lesion.
Sum Variance A measure of the spread of the distribution of the sum of the voxel values of voxel pairs in the lesion.
Variance Is a measure of how spread out the distribution of voxel values in the lesion are.
Margin Feature 1 Is the average gradient of the lesion's margin.
Margin Feature 2 Is the standard deviation of the gradient of the lesion's margin.
Variance of the Radial Gradient Histogram (vRGH) Is a measure of the variation in the margin sharpness of the lesion.
Max Uptake Var Is the maximum of the variance in enhancement of the lesion.
Peak Timepoint Var Is the post-contrast time at which *Max Uptake Var* occurs.
Uptake Rate Var Is the rate at which the variance in enhancement increases to its maximum.
Washout Rate Var Is the rate at which the variance in enhancement decreases from its maximum.
Volume Is the volume of the lesion.
Surface Area Is the surface area of the lesion.

2. Similar Case Database

The similar case database included in the device was collected with a range of acquisition parameters as detailed in the table below. A patient's case was included if the scan contained a lesion for which pathology had been obtained. There were no explicit age or ethnicity inclusion or exclusion criteria. A summary description of the cases in the database is detailed in the table below. Cases were included if 1) the scan contained a lesion for which biopsy-proven truth had been obtained or for non-biopsied benign lesions, clinical and radiology reports and a negative follow-up MRI study at a minimum of 12 months and 2) lesion type had been determined by multidisciplinary review.

Table 1: Description of the Similar Case Database included in the device labeling. The summary of MR acquisition parameters and other descriptors of the cases are included.

		Philips	GE	Siemens
Acquisition Parameters	**Manufacturer/Model**	Philips Achieva	GE Signa	Siemens Avanto
	Field Strength	1.5T (359 lesions) 3.0T (70 lesions)	1.5T (14 lesions) 3.0T (34 lesions)	1.5T (66 lesions)
	# Coil Channels	16	7, 16	16
	Acq. Plane	Axial	Coronal	Axial
	Pulse Sequence	3D gradient echo (THRIVE)	Enhanced Fast Gradient Echo 3D (efgre3D)	FLASH 3D (fl3d1)
	TR (ms)	5.4 (4.8 - 7.5)	8.2 (5.0 – 11.2)	4.5 (4.5 – 5.0)
	TE (ms)	2.7 (2.4 – 4.6)	2.7 (2.1 - 4.2)	1.4 (1.4 – 1.5)
	Flip Angle (degrees)	10 (71 lesions) 12 (353 lesions) 15 (3 lesions) 20 (2 lesions)	10 (26 lesions) 12 (3 lesions) 15 (19 lesions)	10 (66 lesions)
	Slice spacing (mm)	0.8 (70 lesions) 1.0 (356 lesions) 2.5 (3 lesions)	3.0 (48 lesions)	1.0 (66 lesions)
	In-plane resolution (mm)	0.72 (0.54 – 0.97)	0.55 (0.39 – 1.05)	0.73 (0.71 – 0.83)
	# Postcontrast	5	3	3
	Fat Suppression	Yes	Yes	Yes
	Parallel Imaging	Yes	Yes	Yes
Cases Description	**Date Range**	11/2008 - 12/2013	01/2009 - 12/2013	08/2013 - 12/2014
	Total # Lesions	429	48	66
	# Benign	190	27	41
	Nonbiopsied	78	6	1
	# Malignant	239	21	25
	DCIS	43	8	7
	IDC	158	11	15
	ILC	21	2	3
	Mixed (IMC)	4	0	0
	Unknown/Other	13	0	0
	Age (± std dev)	54 ± 13 (25 – 85)	50 ± 10 (37 – 75)	55 ± 13 (22 – 86)

3. Image analysis algorithm (Special Control 1.i.)

   The QI Score is calculcated using a combined feature score algorithm based on literature described in detail within the submission. Individual features, feature selection process, algorithm training, algorithm inputs, major algorithm components, algorithm outputs, and algorithm limitations are included in the device description, software description, standalone performance and testing documentation, and literature references.

   Study limitation: Multiple candidate classifiers were evaluated on a dataset intended to serve as an independent validation of standalone classifier performance. Although

the resulting differences in performance between the candidate classifiers based on AUC were small, we note that in this diagnostic device space, the effect size for meaningful difference is also small. Small differences in AUC can translate into significant clinical differences.

H. Standard/Guidance Document Referenced

NEMA PS 3.1 - 3.20 (2016), Digital Imaging and Communications in Medicine (DICOM) set. FDA recognition number 12-300.

Guidance for Industry and FDA Staff: Guidance for the Content of Premarket Submissions for Software Contained in Medical Devices (issued May 11, 2005)

Guidance for Industry and Food and Drug Administration Staff: Computer-Assisted Detection Devices Applied to Radiology Images and Radiology Device Data – Premarket Notification [510(k)] Submissions (issued July 3, 2012)

I. Performance Characteristics

The device is a software-only device. Some common performance characteristics for other device types are included below with a note that these characteristics are not applicable to this type of software-only device.

1. Biocompatibility/Materials

Not applicable

2. Shelf Life/Sterility

Not applicable

3. Electromagnetic Compatibility and Electrical Safety

Not applicable

4. Magnetic Resonance (MR) Compatibility

Not applicable

Nonclinical performance data were provided to address the following areas:

5. Software (Special Control 1.v.)

The device is a software-only device.

The sponsor provided software documentation at a Moderate Level of Concern according to the "Guidance for the Content of Premarket Submissions for Software Contained in Medical Devices" (May 11, 2005).

Version: 1.0.2022
Level of Concern: Moderate

Software description:

The sponsor provided a general description of the features in the software documentation and in the device description. The software runs on Windows or OS X and hardware requirements are included in the User Manual. The programming languages were described.

Device Hazard Analysis:

The device hazard analysis includes

- identification of the hazardous event
- severity of the hazard
- probability of the hazard
- cause(s) of the hazard
- method of control or mitigation
- corrective measures taken, including an explanation of the aspects of the device design/requirements, that eliminate, reduce, or warn of a hazardous event
- verification of the control implementation is traceable through the enumerated traceability matrix

Software Requirements Specifications (SRS):

The SRS includes hardware requirements, programming language requirements, interface requirements, functional requirements, performance requirements, and safety requirements. Performance and functional requirements are described within the Quantitative Insights QuantX Software Device – Software Requirements Specification document. High-level requirements are also included in the SRS.

Functional requirements are described in greater detail in individual SRS references for specific features (such as kinetic colormap, ADC calculation,

Architecture Design Chart:

The architecture design chart provides the software overview and includes flow diagrams representative of process flow for various features of the QuantX software.

Software Design Specifications (SDS):

Detailed non-functional requirements are included in the SRS that pertain to the software design such as data structure and behavior requirements.

The SDS include:

- an introduction
- system overview
- design map, architecture design
- database schema
- high level design
- low level design
- user interface design

Some of the SDS elements such as implementation were also included within the SRS. Low level design documents were included for the graphical user interface (GUI), data, database, and MR analysis.

Traceability Analysis/Matrix: A series of traceability tables link enumerated requirements, hazards, and test results.
Software Development Environment: The software development environment includes a summary of the software development life cycle plan and the processes that are in place to manage the various life cycle activities.
Verification & Validation Testing: The validation and system level verifications procedures are based upon the requirements with clearly defined test procedures and pass/fail criteria. All tests passed. Use case validation results were provided. Unit level test procedures, actual, and expected results are included for specific feature requirements. Enumerated test results were included for each test. V&V testing included an assessment of software controls included to confirm the appropriate image type and the image characteristics were in the case prior to processing the images.
Revision level history: Development from alpha version through the version used in the reader study is included with a description of the changes between versions and dates.
Unresolved anomalies: Unresolved anomalies are described with the problem, impact on device performance and plans for correcting the problem.

Cybersecurity

The cybersecurity documentation is consistent with the recommendations for information that should be included in premarket submissions outlined in the FDA guidance document Content of Premarket Submissions for Management of Cybersecurity in Medical Devices: Guidance for Industry and Food and Drug Administration Staff (issued October 2, 2014). Information related to cybersecurity reviewed included:

a. Hazard analysis related to cybersecurity risks,
b. Traceability documentation linking cybersecurity controls to risks considered,
c. Summary plan for validating software updates and patches throughout the lifecycle of the medical device,
d. Summary describing controls in place to ensure that the medical device will maintain its integrity, and
e. Device instructions for use and product specifications related to recommended cybersecurity controls appropriate for the intended use of the device.

The software documentation is acceptable.

6. Standalone performance testing protocols and results (Special Control 1.iv.)

The sponsor provided standalone performance testing and results of the QI Score in distinguishing between the benign and malignant cases. The QI score is a combination of multiple lesion features, each calculated after lesion segmentation.

a. Segmentation in the standalone testing

For the included lesions, an initial seed point was selected by referring to the radiology report from the exam to obtain the approximate location of the lesion. In practice, the radiologist selects the seed point location; however, the segmentation algorithm includes dependency on the seed point location. The automatic lesion segmentation was performed using the QuantX software without any manual correction in the standalone performance assessment.

b. Case distribution in the standalone testing

For the similar case database

Cases were collected were retrospectively collected over a 7-year span (2008-2014) of dynamic contrast enhanced (DCE) breast MRI studies from the University of Chicago Medical Center, Memorial Sloan Kettering Cancer Center, and the X-Ray Associates of New Mexico. The dataset included a total of 652 lesions, 314 benign and 338 malignant.

Case inclusion criteria:

1. The lesion reported in a radiology report
2. The lesion not having a previous biopsy/excision that removed a large portion of the lesion
3. The scan contained a lesion for which biopsy-proven truth had been obtained and pathological truth for the lesion, or for non-biopsied benign lesions, clinical and radiology reports and a negative follow-up MRI study at a minimum of 12 months and lesion type had been determined by multidisciplinary review.

Cases were included from multiple MR system manufacturers and field strengths. The cases used in the standalone testing are the same as those included in the similar case database (see Table 1 for more details).

For the reader study test database

Please refer to the case information under Section J. Summary of Clinical Information for details.

c. Testing and results

Similar Case Database

A 0.632 bootstrap method was used to evaluate the performance of the QI Score on the data included Similar Case Database. The area under the curve (AUC) was calculated using the trapezoidal method, with the confidence interval estimated empirically using bootstrap results.

Overall AUC performance based on the bootstrap method was 0.86 ± 0.02 (mean ± standard error)

Study limitation: The standalone study that used the Similar Case Database (i.e., training database) cannot be considered an independent validation study, and consequently results from such study cannot be considered confirmatory evidence. However, this standalone assessment does contain cases from important cohorts for the intended use population.

Reader study testing database

Overall AUC performance based on the reader study testing database using the automated segmentation without manual input was 0.75 ± 0.05 (mean ± standard error).

Overall AUC performance based on the reader study testing database using the segmentation from the clinical reader study (i.e., including variability of different seed point locations) was 0.71 ± 0.05 (mean ± standard error).

Study limitation: The standalone studies that used the reader study testing dataset included 'challenge cases' as defined in the reader study protocol, so it is expected that the test database AUCs may be less than the bootstrap AUCs and not precisely reflect expected clinical performance. So, while these results can be considered independent validation, the performance may not adequately characterize the performance of the device for the intended use population.

7. Animal and/or Cadaver testing

   None provided.

J. Summary of Clinical Information

A multiple reader, multiple case (MRMC) clinical study including a sequential reading design was used to determine the impact on Reader Performance in diagnosing breast cancer, as characterized by the area under the receiver operating characteristic (ROC) curve (AUC), when QuantX is used during breast MRI interpretation (SECOND READ), compared with conventional MRI interpretation without the use of QuantX (FIRST READ).

This study tested the following hypothesis regarding use of QuantX in improving reader performance for the diagnosis of breast cancer. The hypothesis is: diagnostic performance, as characterized by the area under the receiver operating characteristic (ROC) curve (AUC), of qualified interpreting radiologists (readers) improves (i.e., ΔAUC = $[AUC_{SECOND\ READ} - AUC_{FIRST\ READ}] > 0$) when using QuantX as an aid for interpretation (SECOND READ), compared with when not using QuantX for interpretation (FIRST READ).

The FIRST READ modality of image interpretation consisted of the interpretation of diagnostic breast MR images using a set of features that are equivalent to those of commercially available computer-aided evaluation software products. These systems

display color maps and kinetic curves of contrast enhancement, noting whether within the first minute or so the lesion has enhanced pass some threshold (e.g., 50%, or 100%, of the initial value) set by the user, and whether the subsequent patterns of enhancement represent increasing, plateau, or decreasing (i.e., washout).

The SECOND READ modality of image interpretation consisted of the interpretation of diagnostic breast MR images, displayed on the QuantX interface with all of the QuantX functionality. The available functionality of QuantX (i.e., the CADx output) during the SECOND READ interpretation included: the conventional CAE kinetic information, a QI Score, a volumetric analysis, a similar case database, and additional values quantifying various morphological and kinetic features. With QuantX, the user can choose to view several feature values/outputs (examples listed below) in addition to color maps and kinetic curves for the case being read, as well as images presented within the similar case database.

The QuantX system was used in both the FIRST READ and SECOND READ workflow by controlling the functionality available to the reader during specific points in the study.

1. Primary endpoint (Special Control 1.ii. and 1.iii.)

   The primary endpoint was the expected difference in the AUC between the FIRST READ image interpretation and the SECOND READ image interpretation. These ROC curves were estimated from the readers' likelihood-of-malignancy responses. Statistical estimation of the primary endpoint was made by using the Dorfman-Berbaum-Metz method of MRMC analysis.

   $$\Delta AUC = [AUC_{SECOND\ READ} - AUC_{FIRST\ READ}]$$

   $$H_0: \Delta AUC = 0$$

   $$H_1: \Delta AUC > 0\ [AUC_{SECOND\ READ} - AUC_{FIRST\ READ}]$$

   Rejection of the null hypothesis was with respect to the critical value of $\alpha = 0.05$.

   Note: The AUC can be interpreted as the average sensitivity over all possible specificities. If the area under the ROC curve is greater for diagnostic A than for diagnostic B, and the two curves do not cross, then diagnostic A is clearly superior diagnostic B. If the curves cross, then relative performance is not clear over the entire range of operating points.

2. Secondary analyses

   Secondary analyses consist of estimation of the expected differences in sensitivity and the expected differences in specificity, both between the FIRST READ image interpretation and the SECOND READ image interpretation. Sensitivity and specificity were calculated from the readers' responses for the 7-point BI-RADS assessment categories. Two different cut points were used as the definition for a positive call for cancer diagnosis:

   1) a BI-RADS assessment of 4a or higher (i.e., 4a, 4b, 4c, and 5) defines a positive call for cancer diagnosis and, conversely, a BI-RADS assessment

of 3 or lower (i.e., 3, 2, and 1) defines a negative call for cancer diagnosis; and

2) a BI-RADS assessment of 3 or higher (i.e., 3, 4a, 4b, 4c, and 5) defines a positive call for cancer diagnosis and, conversely, a BI-RADS assessment of 2 or lower (i.e., 2 and 1) defines a negative call for cancer diagnosis.

This sensitivity and specificity analyses are included to ensure there is not an unintended reduction in either sensitivity or specificity. Statistical estimation of the uncertainties in the expected differences in sensitivity and specificity was made by using the method of bootstrapping on the reader data (with appropriate sampling of both readers and cases), together with estimated 95% CIs. These endpoints were not adjusted for multiplicity, but may be considered descriptive results.

3. Case inclusion for the reader study testing dataset (Special Control 1.ii. and 1.iv.)

The Reference Standard (Ground Truth) for assessment of malignancy was used as an objective standard against which the rating data obtained from the readers was analyzed. Ground truth for the biopsied cancers and biopsied non-cancers was directly from the associated final pathology reports. Ground truth for the non-biopsied non-cancers was from clinical and radiology reports and a negative follow-up MRI study at a minimum of 12 months.

Study limitation: Cases determined to be biopsy benign lacked 1-year follow-up. Instead, the sponsor provided documentation that the benign result is concordant with the suspicious imaging appearance that prompted the biopsy. We defined concordance as a determination that a tissue biopsy result is compatible with (i.e., is a plausible explanation for) the abnormal pre-biopsy imaging appearance which prompted the performance of the biopsy.

Anonymized breast imaging cases were retrospectively collected by the sponsor from three different institutions, an academic breast imaging center, a dedicated cancer imaging center and a community based imaging center. Breast MR images were collected representing cancers (including invasive and DCIS) and non-cancers (including benign lesions and non-biopsied suspect regions), and included a varying distribution of lesion descriptors according to the BIRADS lexicon for MRI. All cases in the reader study testing data were independent from the QuantX similar case database.

All the cases included in the study satisfied the indications for breast MRI according to the ACR practice guidelines for CE-MRI of the breast, revised 2013. Cases in the reader study/evaluation dataset included MR cases from clinical practice where an abnormality was detected, thus requiring workup. These cases were accrued from patients that presented for clinical indications such as high-risk screening (38%), diagnostic imaging workup (including follow-up diagnostic imaging workup) (35%), or evaluation of extent of known disease (27%). Several case selection criteria and quality control mechanisms were applied. Cases were included:

1) if evaluable T2 and a minimum of 2 post-contrast DCE MR image sets were available for the diagnostic interpretation,

2) for non-cancer cases, if negative biopsy or follow-up MRI study at a minimum of 12 months was negative,
3) for cancer cases, if positive biopsy, and case meets the cancer subtype requirements, and
4) if source records were available for clinical status verification purposes.

Cases were excluded if cases demonstrated administrative or technical errors, such as exam incomplete or cases not meeting the minimum acquisition image quality requirements per the QuantX User Manual.

A total of 111 breast MR images were included, with a total of 54 cancer and 57 non-cancerous breast lesions (i.e., an enriched set). Of the non-cancer lesions, 40 were biopsied non-cancers and 17 were non-biopsied non-cancers. Cases were not collected consecutively but rather collected to satisfy the distribution of cases according to scanner, vendor, cancer subtype, and benign cases. Sixty-three cases with multiple lesions were included in the reader study; however, the study used one lesion per case in order to enable ROC analysis.

Note: The dataset was enriched for the more challenging cases. The final study case distribution included a total of 64 cases that were rated by the source institution as BIRADS 3 or 4. Interval cancer cases (i.e., cases where current diagnosis is positive and where the previous diagnosis was negative) were neither explicitly included nor excluded.

Lesion classification

		Philips 1.5T	Philips 3T	Siemens 1.5T	GE 1.5T	GE 3T	Totals
Cases Description	**Date Range**	02/2009-01/2014	05/2010-12/2013	10/2013-12/2014	03/2011	01/2009-09/2013	
	Total # Cases	24	35	27	1	24	**111**
	# Benign	11	19	15	1	11	**57**
	Biopsied	2	12	15	1	10	**40**
	Non-biopsied	9	7	0	0	1	**17**
	# Malignant	13	16	12	0	13	**54**
	DCIS	3	3	4	0	3	**13**
	IDC	10	11	6	0	8	**35**
	ILC	0	2	2	0	1	**5**
	Other	0	0	0	0	1	**1**
	Age (mean ± 1 SD)	50 ± 13	52 ± 15	58 ± 13	52	48 ± 11	52 ± 13

Data acquisition parameters for the reader study testing database

	P1.5	P3	S	G1.5	G3
Manufacturer/Model	Philips Achieva	Philips Achieva	Siemens Avanto	GE Signa	GE Signa HDx
Field Strength	1.5T	3T	1.5T	1.5T	3T
# Coil Channels	16	16	16	16	7, 16
Acq. Plane	Axial	Axial	Axial	Sagittal	Sagittal
Pulse Sequence	3D gradient echo (THRIVE)	3D gradient echo (THRIVE)	FLASH 3D (fl3d1)	Enhanced Fast Gradient Echo 3D (efgre3D)	Enhanced Fast Gradient Echo 3D (efgre3D)
TR (ms)	5.5 (5.4-5.6)	4.8 (4.6-5.0)	4.5 (4.5-5.0)	6.6	9.6 (4.5-11)
TE (ms)	2.7 (2.7-2.7)	2.4 (2.3-2.5)	1.4 (1.4-1.5)	4.2	2.1 (2.1-2.3)
Flip Angle (degrees) [# of cases] **10**	0	35	27	1	4
12	23	0	0	0	1
15	1	0	0	0	19
Slice spacing (mm) [: # of cases]	0.8: 3 1.0: 21	0.8	1.0: 24 1.2: 3	3	3
In-plane resolution (mm) (range)	0.72 (0.56-0.79)	0.65 (0.56-0.79)	0.75 (0.71-0.83)	0.43	0.43 (0.39-0.51)
# Postcontrast	5, 6	4, 5	3	3	3
Fat Suppression	Yes	Yes	Yes	Yes	Yes
Parallel Imaging	Yes	Yes	Yes	Yes	Yes

Study limitation: 13 cases were re-used from a previous reader study of a similar device. This data re-use might have biased the study results.

4. Reader selection criteria (Special Control 1.ii.)

   Readers were recruited from practices in academia and private practice. The readers satisfied minimum qualifications and experience requirements in interpreting breast MRI.

   The inclusion criteria for all readers in the study were as follows:

   - Signed financial disclosures
   - Signed reader study agreement, including non-disclosure
   - Informed consent
   - Current medical license
   - American Board of Radiology or equivalent certification
   - Interpreting radiologist with at least 1 year of breast MRI interpretation experience (including a breast imaging fellowship, if applicable)
   - Fellowship-trained in breast imaging or 2 years' experience in breast imaging
   - Currently qualified as a Mammography interpreting physician under MQSA
   - Successful training on the use of study software

5. Pre-specified analysis plan (Special Control 1.ii.)

   Assumptions used to estimate the number of cases and readers included an AUC of 0.80-0.84, an expected difference in AUC of 0.03 to 0.05, and a 1:1 prevalence (cancers to noncancers).

For each reader, the likelihood of malignancy (LOMs) obtained from all study cases were used to estimate two ROC curves: one for the FIRST READ, and one for the SECOND READ, conditions. The respective estimated AUC values and standard errors are reported. The average AUC values across all readers and the associated standard errors were estimated by using the Dorfman-Berbaum-Metz ANOVA after jackknife (DBM) method. Proper-binormal AUC estimates and non-parametric (trapezoidal) AUC estimates were obtained. From the output of the DBM analysis, two-tailed p-values and 95% CIs for estimated AUC- value differences between the SECOND READ and FIRST READ conditions were obtained.

For the secondary analyses, statistical estimation of the differences in sensitivity and in specificity was made by using the method of bootstrapping on the reader data, together with estimated 95% CIs.

Study limitation: The Reader Study Protocol indicated that both non-parametric (trapezoidal) and proper-binormal AUC estimates would be obtained, but the study protocol did not specify which of the two methods would be considered definitive in case of disagreement for the primary analysis.

6. Reader training (Special Control 1.ii.)

Prior to initiation of the MRMC study, all Readers met the Reader Qualification criteria, gave informed consent, and were trained, in (a) the use of QuantX in breast MRI interpretation (functionality and "knobs"), (b) the 7-point forced BI-RADS, and (c) the role and use of the reader rating data (LOM) in ROC analysis.

Following training, a short proficiency test was given to the readers. The proficiency test was only used as a mechanism to ensure the readers paid attention during the training and were sufficiently proficient with the study software to proceed with the study. This proficiency assessment was only with respect to the software functions and an understanding of the reader study protocol. It in no way sought to assess their proficiency in diagnosis. All of the study readers successfully completed the proficiency test. Per study protocol, no readers were excluded from the study based on the results of the proficiency test.

Study limitation: Training materials included an optimistic assessment of the device performance that may have biased the readers in favor of the device. In other words, readers in the study were provided information about QuantX standalone performance results that may not represent true standalone performance (estimates may be biased). Estimates of performance included in training materials may influence reader performance; however, in this instance, it is not possible to definitively predict the effect of this bias.

7. Data collection (Special Control 1.ii.)

Readers interpreted a single lesion in each case that was identified to them as lesion location information. For each lesion, readers clicked on the lesion that invoked the software automated lesion segmentation algorithm and were provided information for that lesion consistent with the FIRST READ or SECOND READ condition (described above).

The readers were directed to report their estimate of the likelihood of malignancy and their opinion on BI-RADS final assessment for only the study-designated lesion. Such an arrangement allowed the readers to read cases that they would encounter in clinical practice, while still allowing for ROC analysis. The "QI score" depends to some extent on the lesion seed point that a reader selects because the lesion segmentation results can be affected by the choice of this seed point, which in turn could affect the "QI score."

For this sequential reader study, the reader first interprets and scores the case under the FIRST READ interpretation condition, after which the reader's FIRST READ responses are immediately locked. Then the CADx output becomes available to the reader, and then the reader interprets and scores the case under the SECOND READ interpretation condition, after which the reader's SECOND READ responses are immediately locked.

The order in which the cases were interpreted was randomized uniquely for each reader in order to minimize potential bias due to case-presentation order. Each reader read all cases in both the FIRST READ and SECOND READ interpretation conditions.

Note: In this study, there were no prior images available for comparison and no clinical information of patient history was provided.

8. MRMC study results (Special Control 1.iii.)

A total of 19 readers completed the study of 111 breast MR cases. Readers included a mix from private practice and academia. Eight completed a breast imaging fellowship.

Table 1. Reader experience

Years experience	Breast imaging	Breast MRI
0 – 5	6	8
5 – 10	4	7
Greater than 10	9	4

a. Primary endpoint analysis

Table 2. MRMC proper-binormal estimate of summary modality-specific area under the ROC curve (AUC)

Modality-specific AUC	AUC	Standard Error	95% CI	P-value
$AUC_{1^{st}\,read}$	0.7055	0.0308	[0.6450, 0.7660]	
$AUC_{2^{nd}\,read}$	0.7575	0.0350	[0.6889, 0.8261]	
$AUC_{2^{nd}\,read} - AUC_{1^{st}\,read}$	0.0520	0.0254	[0.0022, 0.1018]	0.0408

Table 3. MRMC estimate of summary modality-specific trapezoidal area under the ROC curve (AUC)

Modality-specific AUC	AUC	Standard Error	95% CI	P-value
$AUC_{1^{st}\,read}$	0.7090	0.0305	[0.6491, 0.7689]	
$AUC_{2^{nd}\,read}$	0.7577	0.0352	[0.6887, 0.8268]	
$AUC_{2^{nd}\,read} - AUC_{1^{st}\,read}$	0.0487	0.0254	[-0.0011, 0.0985]	0.0550

The results of the reader study indicate that the study marginally met the primary endpoint.

b. Secondary analyses (Special Control 1.iii.)

Two sets of sensitivity and specificity calculations were conducted with two different criteria for a positive call (by the readers). One of these criteria was with BI-RADS 3 as the cut-point, i.e., ≥ 3 indicates a positive call. The other criterion was with BI-RADS 4a as the cut-point, i.e., ≥ 4a indicates a positive call.

Table 4. Sensitivity and specificity for BI-RADS cut-point of 3 (≥3 indicates a positive call)

	1st READ	2nd READ	2nd READ – 1st READ Difference	95% CI
Sensivity	90.4	94.2	3.8	[0.8, 7.4]
Specificity	28.6	27.6	-1.0	[-6.5, 4.3]

Table 5. Sensitivity and specificity for BI-RADS cut-point of 4a (≥4a indicates a positive call)

	1st READ	2nd READ	2nd READ – 1st READ Difference	95% CI
Sensitivity	79.7	84.8	5.1	[-0.9, 10.9]
Specificity	52.2	51.7	-0.5	[-7.3, 6.0]

Note that because a detailed, formal hypothesis testing plan for sensitivity and specificity was not pre-specified, these results are considered descriptive.

c. Additional analyses (Special Control 1.iii.)

Individual reader ROC plots were evaluated for consistency between the fit of the proper-binormal ROC plot and the nonparametric ROC plots. The sponsor provided several supplemental analyses not pre-specified in the study protocol including analysis by site/MR system manufacturer, analysis by magnetic field strength, kappa analysis of reader agreement, and reader QI score variability. These analyses were intended to provide a qualitative assessment of potential trends in the data, evaluate the robustness of the study, and inform labeling. *Note that caution should be exercised in interpreting these types of results.*

9. Other notes

   One additional reader participated in the study. However, this reader interpreted the LOM as "confidence in his/her diagnostic decision" rather than "likelihood of malignancy," which consequently produced estimated ROC curves that resemble the "chance" line with AUC values of approximately 0.5. It is obvious that those LOM data are not appropriate for ROC analysis. In consultation with the FDA review team, the study principle investigator, and the sponsor, the FDA review team agreed that it is appropriate to exclude the data of this reader from the analysis of the study results.

10. Pediatric Extrapolation

    In this de novo request, existing clinical data were not leveraged to support the use of the device in a pediatric patient population.

K. Labeling:

The labeling is sufficient and it satisfies the requirements of 21 CFR Part 801, including 21 CFR Part 801.109 for prescription devices, and the special controls for this device type. The QuantX MRI User Manual provides the detailed instructions for use (Special Control 2.vii.). Other elements of labeling for QuantX related to the special controls for labeling of this device type are noted below.

1. Indicated patient population (Special Control 2.i.)

QuantX is indicated for evaluation of the assessment and characterization of breast abnormalities from MRI data in patients presenting for high-risk screening, diagnostic imaging workup, or evaluation of extent of known disease.

2. Intended reading protocol (Special Control 2.ii.)

The User Manual includes instructions for opening a new case, lesion segmentation (manual and automatic), viewing information such as features for the selected lesion and comparison to lesions in the similar case database (including histograms), information about the data provided, and reporting.

3. Intended user and recommended training (Special Control 2.iii.)

The indications for use note that QuantX is a computer-aided diagnosis (CADx) software device used to assist radiologists.

Quantitative Insights recommends that new users of the QuantX software successfully complete appropriate training courses in both the use of the software for the particular types of cases they will be reading as well as in the broader related areas of multi-modality breast image diagnosis using advanced visualization and analysis tools.

Quantitative Insights also recommends that each site using the QuantX platform maintain a certification program for all new users of the software and all Radiologists who are new to multimodality breast imaging using advanced visualization and analysis tools.

4. Device inputs and outputs(Special Control 2.iv.)

The User Manual describes appropriate image datasets for input and analysis. The User Manual includes instructions for the Report Generation tools and other tools that may be used to characterize lesions.

5. Compatible imaging hardware and imaging protocols (Special Control 2.v.)

Recommended image MR acquisition parameters are included in the labeling with along with ranges of acceptable criteria for each specified acquisition parameter.

6. Warnings, precautions, and limitations incluing situations in which the device may fail or may not operate at the expected performance level (Special Control 2.vi.)

Warnings and cautions related to the potential device failure or low performance are included in the User Manual including comments on image quality, patient population, and segmentation accuracy. Warnings and cautions related to the device are included in the User Manual. Information about the Similar Case Database and cases used in the MRMC study provided in labeling also provide information to users about potential limitations of the device related to specific subpopulations.

7. Summary of the performance testing (Special Control 2.viii.)

The User Manual contains summary information about the standalone performance and MRMC study test methods, dataset characteristics, results, and sub-analyses (including MR system manufacturer).

L. Identified Risks to Health and Mitigation Measures

Identified Risks to Health	Mitigation Measures
Incorrect lesion(s) characterization leading to false positive results may result in incorrect patient management with possible adverse effects such as unnecessary treatment, unnecessary additional medical imaging and/or unnecessary additional diagnostic workup such as biopsy.	Certain design verification and validation activities identified in special control (1) Certain labeling information identified in special control (2)
Incorrect lesion(s) characterization leading to false negative results may lead to complications, including incorrect diagnosis and delay in disease management.	Certain design verification and validation activities identified in special control (1) Certain labeling information identified in special control (2)
The device could be misused to analyze images from an unintended patient population or on images acquired with incompatible imaging hardware or incompatible image acquisition parameters, leading to inappropriate diagnostic information being displayed to the user.	Certain design verification and validation activities identified in special control (1) Certain labeling information identified in special control (2)

Device failure could lead to the absence of results, delay of results or incorrect results, which could likewise lead to inaccurate patient assessment.	Certain design verification and validation activities identified in special control (1) Certain labeling information identified in special control (2)

M. Benefit/Risk Analysis

Summary	
Summary of the Benefit(s)	This device provides a systematic automated analysis of breast MRI to assist users in characterizing breast lesions as a concurrent read of breast MRI. The clinical MRMC study demonstrated a marginally statistically significant improvement in reader performance diagnosing breast cancer when QuantX is used during breast MRI interpretation compared with conventional MRI interpretation without the use of QuantX. The systematic automated analysis may also provide an increased degree of consistency to the reader analysis. The improved diagnostic accuracy and consistency in analysis may improve the performance of some users. Therefore, an improved, more uniform interpretation may be expected across some radiology practices using the device.
Summary of the Risk(s)	There are minimal potential risks associated with the use of the device. The risks include: • Incorrect lesion(s) characterization leading to false positive results may result in incorrect patient management with possible adverse effects such as, unnecessary additional medical imaging and/or unnecessary additional diagnostic workup such as biopsy. Less likely, but possible would be unnecessary treatment. • Incorrect lesion(s) characterization leading to false negative results may lead to complications, including incorrect diagnosis and delay in disease management. • The device could be misused to analyze images from an unintended patient population or on images acquired with incompatible imaging hardware or incompatible image acquisition parameters, leading to inappropriate diagnostic information being displayed to the user. • Device failure could lead to the absence of results, delay of results or incorrect results, which could likewise lead to inaccurate patient assessment. However, based on the performance data and the application of mitigating measures (general controls and special controls established for this device type), use of the device is unlikely to decrease diagnostic performance of the user and possible misuse of the device does not present additional risks compared with misuse of other types of radiological image processing devices.

Summary of Other Factors	The study was enriched and individual readers in practice may not experience a significant improvement in diagnosing breast cancer lesions.
Conclusions Do the probable benefits outweigh the probable risks?	Yes, the probable benefits outweigh the probable risks, given the combination of required general controls and the special controls established for this device. The Special Controls will sufficiently assist in managing risks associated with incorrect lesion(s) characterization, application of the device results to the wrong patient population, analysis of incompatible images, and/or device failure by insuring proper performance and use of the device. By providing a systematic automated analysis of a multiple factors in reviewing breast MRI lesions, the device may marginally improve diagnostic performance and consistency in evaluating a multitude of clinical factors. The QuantX analytics calculate the morphological and enhancement characteristics of a breast lesion the QuantX package uses pattern recognition techniques and displays similar lesions from a fixed database of biopsy-proven non-cancer and cancer. A user will be able to view and interpret the lesion characteristics and similar cases during their diagnosis of a patient's breast lesions. By demonstrating an ability to assist users in the classification of breast MRI lesions as BIRADS 2, or 3 or 4a, the device may provide improved diagnostic accuracy as assessed by an improvement in AUC. In particular, secondary analyses suggest improved sensitivity based on BI-RADS 3 as the cut-point without decreased specificity for some readers. The device provides information that may be useful in the characterization of breast abnormalities, but does not replace the physician. The physician determines the diagnosis and the information provided by the device is unlikely to decrease diagnostic performance of the user. Not all users may experience a substantial improvement in diagnostic performance based on the use of the device. However, the device is not likely to lead to a substantial decline in reader performance either. Therefore, given the available information concerning the benefits, risks, and supporting data; the probable benefits outweigh the probable risks, given the combination of required general controls and special controls established for this device.

Patient Perspectives

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

N. Conclusion

The information provided in this *de novo* submission is sufficient to classify this device into class II under regulation 21 CFR 892.2060.

FDA believes that stated special controls, in combination with the applicable general controls, provide a reasonable assurance of the safety and effectiveness of the device type. The device is classified under the following:

Product Code: POK

Device Type: Radiological computer-assisted diagnostic (CADx) software for lesions suspicious for cancer

Class: II (special controls)

Regulation: 21 CFR 892.2060

(a) *Identification*: A radiological computer-assisted diagnostic (CADx) software for lesions suspicious for cancer is an image processing prescription device intended to aid in the characterization of lesions as suspicious for cancer identified on acquired medical images such as magnetic resonance, mammography, radiography, or computed tomography. The device characterizes lesions based on features or information extracted from the images and provides information about the lesion(s) to the user. Diagnostic and patient management decisions are made by the clinical user.

(b) *Classification*: Class II (special controls). A radiological computer-assisted diagnostic (CADx) software for lesions suspicious for cancer must comply with the following special controls:

1. Design verification and validation must include:
    i. A detailed description of the image analysis algorithms including, but not limited to, a detailed description of the algorithm inputs and outputs, each major component or block, and algorithm limitations.
    ii. A detailed description of pre-specified performance testing protocols and dataset(s) used to assess whether the device will improve reader performance as intended.
    iii. Results from performance testing protocols that demonstrate that the device improves reader performance in the intended use population when used in accordance with the instructions for use. The performance assessment must be based on appropriate diagnostic accuracy measures (e.g., receiver operator characteristic plot, sensitivity, specificity, predictive value, and diagnostic likelihood ratio). The test dataset must contain sufficient numbers of cases from important cohorts (e.g., subsets defined by clinically relevant confounders, effect modifiers, concomitant diseases, and subsets defined by image acquisition characteristics) such that the performance estimates and confidence intervals of the device for these individual subsets can be characterized for the intended use population and imaging equipment.
    iv. Standalone performance testing protocols and results of the device.

v. Appropriate software documentation (e.g., device hazard analysis; software requirements specification document; software design specification document; traceability analysis; description of verification and validation activities including system level test protocol, pass/fail criteria, results, and cybersecurity).

2. Labeling must include:
   i. A detailed description of the patient population for which the device is indicated for use.
   ii. A detailed description of the intended reading protocol.
   iii. A detailed description of the intended user and recommended user training.
   iv. A detailed description of the device inputs and outputs.
   v. A detailed description of compatible imaging hardware and imaging protocols.
   vi. Warnings, precautions, and limitations, including situations in which the device may fail or may not operate at its expected performance level (e.g., poor image quality or for certain subpopulations), as applicable.
   vii. Detailed instructions for use.
   viii. A detailed summary of the performance testing, including: test methods, dataset characteristics, results, and a summary of sub-analyses on case distributions stratified by relevant confounders (e.g., lesion and organ characteristics, disease stages, and imaging equipment).