

**510(k) SUBSTANTIAL EQUIVALENCE DETERMINATION
DECISION SUMMARY**

A. 510(k) Number:

k080564

B. Purpose for Submission:

Marketing product in the U.S.

C. Manufacturer and Instrument Name:

Aperio Technologies, Inc.

ScanScope® XT System, IHC HER2 Breast Tissue Tunable Image Analysis

D. Type of Test or Tests Performed:

Computer-assisted image analyzer for immunohistochemistry HER2 slides

E. System Descriptions:

1. Device Description:

The ScanScope® XT System is an automated digital slide creation, management, viewing and analysis system which consists of an automated digital microscope, slide scanner, computer, color monitor, keyboard, digital pathology information management software and image analysis software. In this particular application, IHC HER2 Breast Tissue Tunable Image Analysis, the image analysis software assists the pathologist in quantitative assessment of immunohistochemistry stained histological specimens for human epidermal growth factor Receptor 2 (HER2) using a tuneable algorithm. This algorithm determines the cell classification thresholds based on a set of 20 training slides. The system software makes no independent interpretations of the data.

2. Principles of Operation:

The ScanScope® XT System is intended to provide quantitative input to the pathologist to supplement the quantitative interpretation of HER2 immunohistochemistry stained breast cancer specimens. Formalin-fixed, paraffin embedded breast cancer specimens are stained with the Dako HercepTest™ according to the package insert. Slides are then scanned and digitized at high resolution using the ScanScope XT digital slide scanner. The pathologist then outlines tumor cell only regions and runs the image analysis algorithm. Between 15 and 20 regions and a minimum of 1000 tumor cells should be analyzed to maximize analysis results.

Intensity thresholds for the IHC HER2 Breast Tissue Tuneable Image Analysis algorithm are established by a previously scanned training set. Twenty training slides are annotated with a representative set of tumor regions and the “right” scores the algorithm should provide for the slides. Only the three cell classification intensity thresholds (0 to 1+, 1+ to 2+, and 2+ to 3+) are automatically determined by the training algorithm. The training algorithm is a deterministic optimization algorithm that uses a complete search of all possible threshold combinations. The algorithm generates the scores for all possible threshold combinations and compares the algorithm scores to the “right” scores for all slides. The threshold combination that yields the best overall agreement

between algorithm scores and “right” scores is the optimum parameter set.

The IHC HER2 Breast Tissue Tuneable Image Analysis algorithm then detects the membrane staining for the individual tumor cells in the selected regions and quantifies the intensity and completeness of the membrane staining. Tumor cells are individually classified as 0, 1+, 2+, and 3+ based on their membrane staining intensity and completeness. The HER2 score is then calculated based on the percentages of 0, 1+, 2+, and 3+ cells according to the HER2 scoring scheme. A markup image highlights the detected cell features (black = nuclei and membrane) and the membrane staining which is color-coded according to the cell classification (blue = 0, yellow = 1+, orange = 2+, red = 3+). The pathologist is then provided with HER2 score and the percentages of 0, 1+, 2+, and 3+ cells.

The pathologist makes a final call based on both the qualitative and quantitative information and should follow all appropriate instructions in the Dako HercepTest™ product insert.

3. Modes of Operation:
Computer-assisted interpretation
4. Specimen Identification:
Specimens are identified by slide label (a digital image is taken of the slide label and stored with the digital slide) or by barcode, if provided by the user’s laboratory information system.
5. Specimen Sampling and Handling:
Immunohistochemical stained microslides can be loaded in the ScanScope XT manually (one at a time) or automatically. The ScanScope XT can automatically scan 120 slides contained in slide racks.
6. Calibration:
Calibration of the ScanScope XT is an automated process which is re-verified as part of the scanning process for every scanned slide. If the calibration is not within predefined limits, then the user is prevented from scanning the slide and must take steps to assure that the scan is within acceptable limits.

When the user scans a slide, the controller software automatically performs a “prescan”. The prescan is a scan of a small region of the slide which contains clear glass or “white space”. The brightness and color characteristics of the image are used to correct the resulting scanned image. The main functions of the prescan process are to automatically verify that no significant tissue is present, flatten the illumination field, correct the white balance, and measure bulb brightness.

7. Quality Control:
The accuracy of the system depends on the laboratory Dako HercepTest™ kit.
8. Software:
FDA has reviewed applicant’s Hazard Analysis and Software Development processes for this line of product types:
Yes or No

F. Regulatory Information:

1. Regulation section:
21 CFR §864.1860 Immunohistochemistry reagents and kits
2. Classification:
Class II
3. Product code:
NOT (microscope, automated, image analysis, operator intervention)
4. Panel:
Pathology 88

G. Intended Use:

1. Indication(s) for Use:
The ScanScope® System is an automated digital slide creation, management, viewing and analysis system. It is intended for in vitro diagnostic use as an aid to the pathologist in the display, detection, counting and classification of tissues and cells of clinical interest based on particular color, intensity, size, pattern and shape.

The IHC HER2 Tunable Image Analysis application is intended for use as an aid to the pathologist in the detection and semi-quantitative measurement of HER2/neu (c-erbB-2) in formalin-fixed, paraffin-embedded normal and neoplastic tissue.

The IHC HER2 Tunable Image Analysis application is intended for use as an accessory to the DakoHercepTest™ to aid in the detection and semi-quantitative measurement of HER2/neu (c-erbB-2) in formalin-fixed, paraffin-embedded normal and neoplastic tissue. When used with the Dako HercepTest™, it is indicated for use as an aid in the assessment of breast cancer patients for whom HERCEPTIN® (Trastuzumab) treatment is being considered. Note: The IHC HER2 Image Analysis application is an adjunctive computer-assisted methodology to assist the reproducibility of a qualified pathologist in the acquisition and measurement of images from microscope slides of breast cancer specimens stained for the presence of HER-2 receptor protein. The accuracy of the test result depends upon the quality of the immunohistochemical staining. It is the responsibility of a qualified pathologist to employ appropriate morphological studies and controls as specified in the instructions for the Dako HercepTest™ to assure the validity of the IHC HER2 Image Analysis application assisted HER-2/neu score. The actual correlation of the Dako HercepTest™ to Herceptin® clinical outcome has not been established.

2. Special Conditions for Use Statement(s):
For prescription use only.

H. Substantial Equivalence Information:

1. Predicate Device Name(s) and 510(k) numbers:
ChromaVision Medical Systems, Automated Cellular Imaging System (ACIS), k032113
Applied Imaging Ariol™, k031715
2. Comparison with Predicate Device:

Similarities			
Item	Device	Predicate K032113	Predicate K031715
Device type	Examines formalin-fixed paraffin-embedded breast cancer specimens stained by DakoCytomation HercepTest™ for Her2/neu receptor protein.	Same	Same
Specimen Type	Formalin-fixed, paraffin-embedded stained by immunohistochemistry	Same	Same
Device Components	Automated digital slide scanner, computer, color monitor, keyboard, image analysis software and digital pathology information management software	Same	Same

Differences			
Item	Device	Predicate	Predicate
Image algorithm training	Intensity thresholds are established by a previously scanned manually scored training set. The three cell classification intensity thresholds (0 to 1+, 1+ to 2+, and 2+ to 3+) are automatically determined by the training algorithm.	Image algorithm is permanently set and cannot be modified. No training is involved.	User trains the classifiers with breast cancer slides previously manually scored by a pathologist as 1+ and 3+.

I. Special Control/Guidance Document Referenced (if applicable):

Guidance for Industry and FDA Staff: Format for Traditional and Abbreviated 510(k)s

Guidance for Industry and FDA Staff: Guidance for the Content of Premarket Submissions for Software Contained in Medical Devices

J. Performance Characteristics:

1. Analytical Performance:

a. *Accuracy:*

A multi-site study was conducted at two clinical sites to compare the performance of Aperio's IHC HER2 Breast Tissue Tunable Image Analysis system to manual microscopy. One hundred and eighty (180) formalin-fixed, paraffin-embedded breast tissue specimens from two clinical sites were used

for this study; 80 specimens with approximately equal HER2 score distribution from site 1 and 100 routine specimens from site 2. The specimens were immunohistochemically stained using Dako's HercepTest.

At each site, the IHC HER2 Breast Tissue Tunable Image Analysis system was set up using the automatic algorithm training procedure on a training data set of 20 HER2 slides with approximately equal HER2 score distribution scored independently by three pathologists.

At each site, three pathologists performed a blinded read of the glass slides using a microscope and reported the HER2 score for each of the slides. The glass slides were scanned at Aperio using a different ScanScope for each site, and after a wash-out period and randomization of the slides, the same three pathologists remotely viewed and outlined a representative set of tumor regions to be analyzed by the IHC HER2 Breast Tissue Tunable Image Analysis application. The algorithm itself was run in batch mode blinded from the pathologists to avoid any influence of the pathologists in their choice of the tumor regions. The algorithm reported the HER2 score for each of the three pathologists for each of the slides.

The statistical analyses are presented across all slides for each of the methods: manual microscopy and image analysis, and comparatively between methods for manual microscopy against image analysis.

The pair wise observations of the HER2 Score categories are summarized in 4x4 tables.

		Pathologist 1				Total
		0	1+	2+	3+	
Pathologist 2	0	24	1			25
	1+	9	15	3		27
	2+	1	1	10	2	14
	3+			0	14	14
	Total	34	17	13	16	80

		Pathologist 1				Total
		0	1+	2+	3+	
Pathologist 3	0	19	1	1		21
	1+	12	2			14
	2+	3	14	12		29
	3+				16	16
	Total	34	17	13	16	80

		Pathologist 2				Total
		0	1+	2+	3+	
Pathologist 3	0	17	4			21
	1+	6	8			14
	2+	2	15	12		29
	3+			2	14	16
	Total	25	27	14	14	80

Manual Microscopy – Clinical Site 1 – Inter-Pathologists

		Pathologist 1				Total
		0	1+	2+	3+	
Pathologist 2	0	15	3			18
	1+	5	30	5		40
	2+		3	19	6	28
	3+			2	12	14
	Total	20	36	26	18	100

		Pathologist 1				Total
		0	1+	2+	3+	
Pathologist 3	0	17	8	2		27
	1+	3	28	9		40
	2+			14	6	20
	3+			1	12	13
	Total	20	36	26	18	100

		Pathologist 2				Total
		0	1+	2+	3+	
Pathologist 3	0	16	10	1		27
	1+	2	30	8		40
	2+			19	1	20
	3+				13	13
	Total	18	40	28	14	100

Manual Microscopy – Clinical Site 2 – Inter-Pathologists

		Pathologist 1				
		0	1+	2+	3+	Total
Pathologist 2	0	18	0	0	0	18
	1+	7	23	2	0	32
	2+	0	0	10	0	10
	3+	0	0	5	15	20
	Total	25	23	17	15	80

		Pathologist 1				
		0	1+	2+	3+	Total
Pathologist 3	0	18	1	0	0	19
	1+	7	20	1	1	29
	2+	0	2	14	0	16
	3+	0	0	2	14	16
	Total	25	23	17	15	80

		Pathologist 2				
		0	1+	2+	3+	Total
Pathologist 3	0	16	3	0	0	19
	1+	2	26	0	1	29
	2+	0	3	9	4	16
	3+	0	0	1	15	16
	Total	18	32	10	20	80

Image Analysis – Clinical Site 1 – Inter-Pathologists

		Pathologist 1				
		0	1+	2+	3+	Total
Pathologist 2	0	15	2	0	0	17
	1+	3	35	8	0	46
	2+	0	5	16	2	23
	3+	0	0	0	14	14
	Total	18	42	24	16	100

		Pathologist 1				
		0	1+	2+	3+	Total
Pathologist 3	0	14	1	0	0	15
	1+	4	39	2	0	45
	2+	0	2	21	1	24
	3+	0	0	1	15	16
	Total	18	42	24	16	100

		Pathologist 2				
		0	1+	2+	3+	Total
Pathologist 3	0	13	2	0	0	15
	1+	4	37	4	0	45
	2+	0	7	17	0	24
	3+	0	0	2	14	16
	Total	17	46	23	14	100

Image Analysis – Clinical Site 2 – Inter-Pathologists

Pathologist 1		Image Analysis				Total
		0	1+	2+	3+	
Manual Microscopy	0	22	11	1	0	34
	1+	2	10	5	0	17
	2+	1	2	10	0	13
	3+	0	0	1	15	16
	Total	25	23	17	15	80

Pathologist 2		Image Analysis				Total
		0	1+	2+	3+	
Manual Microscopy	0	17	8	0	0	25
	1+	1	23	3	0	27
	2+	0	1	7	6	14
	3+	0	0	0	14	14
	Total	18	32	10	20	80

Pathologist 3		Image Analysis				Total
		0	1+	2+	3+	
Manual Microscopy	0	17	4	0	0	21
	1+	2	12	0	0	14
	2+	0	12	15	2	29
	3+	0	1	1	14	16
	Total	19	29	16	16	80

Manual Microscopy vs. Image Analysis – Clinical Site 1 – same Pathologist

Pathologist 1		Image Analysis				Total
		0	1+	2+	3+	
Manual Microscopy	0	15	5	0	0	20
	1+	3	32	1	0	36
	2+	0	5	20	1	26
	3+	0	0	3	15	18
	Total	18	42	24	16	100

Pathologist 2		Image Analysis				Total
		0	1+	2+	3+	
Manual Microscopy	0	14	4	0	0	18
	1+	3	32	5	0	40
	2+	0	10	15	3	28
	3+	0	0	3	11	14
	Total	17	46	23	14	100

Pathologist 3		Image Analysis				Total
		0	1+	2+	3+	
Manual Microscopy	0	13	12	2	0	27
	1+	2	33	5	0	40
	2+	0	0	16	4	20
	3+	0	0	1	12	13
	Total	15	45	24	16	100

Manual Microscopy vs. Image Analysis – Clinical Site 2 – same Pathologist

Statistical analyses are provided for a trichotomous categorization of the HER2 scores combining 0 and 1+ and leaving 2+ and 3+ uncombined. Percentage Agreement (PA) along with an exact 95% Confidence Interval (CI) are presented overall for all trichotomous HER2 score categories combined and for each of the trichotomous HER2 score categories separately using a dichotomous outcome of that category vs. the two other categories.

	Pathologist 1 v 2		Pathologist 1 v 3		Pathologist 2 v 3	
	PA	PA 95% CI	PA	PA 95% CI	PA	PA 95% CI
Clinical Site 1	91.3%	(82.8, 96.4)	77.5%	(66.8, 86.1)	76.3%	(65.4, 85.1)
Clinical Site 2	84.0%	(75.3, 90.6)	82.0%	(73.1, 89.0)	90.0%	(82.4, 95.1)

Manual Microscopy - Inter-Pathologists - Agreements.

	Pathologist 1 v 2		Pathologist 1 v 3		Pathologist 2 v 3	
	PA	PA 95% CI	PA	PA 95% CI	PA	PA 95% CI
Clinical Site 1	91.3%	(82.8, 96.4)	92.5%	(84.4, 97.2)	88.8%	(79.7, 94.7)
Clinical Site 2	85.0%	(76.5, 91.4)	94.0%	(87.4, 97.8)	87.0%	(78.8, 92.9)

Image Analysis - Inter-Pathologists - Agreements.

	Pathologist 1		Pathologist 2		Pathologist 3	
	PA	PA 95% CI	PA	PA 95% CI	PA	PA 95% CI
Clinical Site 1	87.5%	(78.2, 93.8)	87.5%	(78.2, 93.8)	80.0%	(69.6, 88.1)
Clinical Site 2	90.0%	(82.4, 95.1)	79.0%	(69.7, 86.5)	88.0%	(80.0, 93.6)

Manual Microscopy vs Image Analysis – same Pathologist - Agreements.

The percent agreements between the pathologists’ manual microscopy and Image Analysis ranged from 79.0% to 90.0% with confidence bounds from 69.7% to 95.1%; the inter-pathologists agreements for manual microscopy ranged from 76.3% to 91.3% with confidence bounds from 65.4% to 96.4% and the inter-pathologists agreements for Image Analysis ranged from 85.0% to 94.0% with confidence bounds from 76.5% to 97.8%.

Note that these image analysis results were obtained by having the Pathologists choose and outline a representative set of tumor regions anywhere on the entire slide, completely blinded from each other, and blinded from the image analysis results (there was no influence on the Pathologists in their choice of the tumor regions).

b. Precision:

Eight HER2 slides with two slides per HER2 score 0, 1+, 2+ and 3+ were sampled from one of the clinical sites that used Dako’s HercepTest to be used in a suite of precision studies. The slides were sampled in sequential order using the rounded average score of the manual microscopy scores provided by the three pathologists.

	Slide 1	Slide 2	Slide 3	Slide 4	Slide 5	Slide 6	Slide 7	Slide 8
Pathologist 1	0	0	0	1	2	2	3	3
Pathologist 2	0	0	1	1	2	2	3	3
Pathologist 3	1	0	1	0	2	2	3	3
Average	0	0	1	1	2	2	3	3

HER2 scores provided by 3 Pathologists for the sampled slides.

Separate studies were conducted to analyze the system variability separately from the variability introduced by the pathologists outlining the tumor regions for the analysis. System precision studies used the same tumor regions for analysis over all runs to eliminate the influence by the pathologists. Pathologist precision studies used the same digital slides to outline tumor regions and run the analysis to eliminate the influence of the system.

The precision studies analyzed the changes in the system response by extending the analysis of the coarse HER2 score to the underlying cumulative percentages of 3+, 2+ and 1+ cells on which the HER2 score calculations are based; allowing detecting and quantifying smaller changes of the system.

Intra-Day/Intra-System

The eight HER2 slides were scanned 10 times on the same ScanScope system. The image analysis results show perfect agreement (100%) for the calculated HER2 scores and an overall average standard deviation of 0.70% (maximum 2.46%) and average range (maximum – minimum) of 1.30% (maximum 7.14%) for the cumulative percentages of 3+, 2+ and 1+ cells (range from 0.0 to 100.0%) across all runs.

Inter-Day/Intra-System

The eight HER2 slides were scanned on the same ScanScope system over 20 times on different days. The image analysis results show perfect agreement (100%) for the calculated HER2 scores and an overall average standard deviation of 0.69% (maximum 2.43%) and average range of 1.75% (maximum 12.07%) for the cumulative percentages of 3+, 2+ and 1+ cells across all runs.

Inter-System

The same eight HER2 slides were scanned 10 times on three different ScanScope systems. The image analysis results show perfect agreement (100%) for the calculated HER2 scores across all systems and all runs. The image analysis results on each of the three ScanScope systems show an overall average standard deviation of 0.70%, 0.56% and 0.57% (maximum 2.46%, 1.65% and 1.34%) and average range of 1.30%, 1.07% and 1.17% (maximum 7.14%, 5.09% and 4.70%) for the cumulative percentages of 3+, 2+ and 1+ cells respectively over all runs. The image analysis results of the three ScanScope systems combined show an overall average standard deviation of 0.80% (maximum 2.41%) and average range of 1.94%

(maximum 8.95%) for the cumulative percentages of 3+, 2+ and 1+ cells respectively over all runs.

Intra-Pathologist

One pathologist outlined the tumor regions for analysis on the same eight HER2 slides 5 times. The image analysis results show 4 cases out of 40 (10%) where the HER2 scores differed from the median HER2 score and an overall average standard deviation of 2.69% (maximum 8.08%) and average range of 3.90% (maximum 18.61%) for the cumulative percentages of 3+, 2+ and 1+ cells.

Inter-Pathologists

Three pathologists outlined the tumor regions for analysis on the same eight HER2 slides as part of the clinical comparison to manual microscopy study. The image analysis results show 3 cases out of 24 (12.5%) where the HER2 scores differed from the median HER2 score and an overall average standard deviation of 10.03% (maximum 27.09%) and average range of 11.74% (maximum 48.26%) for the cumulative percentages of 3+, 2+ and 1+ cells (range from 0.0 to 100.0%).

A summary of the overall average and maximum Standard Deviation (SD) and range for the different precision studies is shown in the following table.

		Average SD	Maximum SD	Average Range	Maximum Range
Intra-Run/Intra-System		0.70%	2.46%	1.30%	7.14%
Inter-Run/Intra-System		0.69%	2.43%	1.75%	12.07%
Inter-System	ScanScope #1	0.70%	2.46%	1.30%	7.14%
	ScanScope #2	0.56%	1.65%	1.07%	5.09%
	ScanScope #3	0.57%	1.34%	1.17%	4.70%
	Combined	0.80%	2.41%	1.94%	8.95%
Intra-Pathologist		2.69%	8.08%	3.90%	18.61%
Inter-Pathologist		10.03%	27.09%	11.74%	48.26%

Algorithm Training Set

100 HER2 slides from clinical site 1 were stratified into 0, 1+, 2+ and 3+ classes based on the average HER2 score provided by three pathologists using manual microscopy.

Three different algorithm training and evaluation runs were conducted. Each time, the 100 slides were separated into a training data set and an evaluation data set. The training data set consisted of 5 slides for each 0, 1+, 2+ and 3+ HER2 class that were selected randomly from the available slides within the HER2 classes (stratified-random selection)—a total of 20 slides. The remaining 80 slides were used as the evaluation data set. The training data set was used to tune the IHC HER2 Breast Tissue Tunable Image Analysis

application according to the procedure outlined in the previous sections of this chapter. The tuned HER2 image analysis application was then run on the 80 slides of the test data set using the tumor region outlines provided by three pathologists during the digital read in the substantial equivalence study. The inter-pathologists variations for manual microscopy and image analysis as well as the inter-method variations are reported as previously in the substantial equivalence study.

The agreements between the pathologists' manual microscopy and Image Analysis ranged from 75.0% to 88.8% with confidence bounds from 66.5% to 95.7%; the inter-pathologists agreements for manual microscopy ranged from 75% to 90% with confidence bounds from 66.5% to 96.6% and the inter-pathologists agreements for Image Analysis ranged from 86.3% to 92.5% with confidence bounds from 78.7% to 97.7%.

- c. *Linearity*:
Not applicable.
- d. *Carryover*:
Not applicable.
- e. *Interfering Substances*:
Not applicable.

2. Other Supportive Instrument Performance Data Not Covered Above:

Precision study to assess variability due to different training sets:

The sponsor did not perform studies to assess how the use of different training sets affects the performance of the device. FDA requested a precision study to demonstrate that using different training sets does not affect performance. While these results did show different threshold estimations for each of the three training sets the difference was not reflected in the image analysis HER2 score determination. Disagreements in HER2 scoring more than 2 degrees, i.e. 0 to 2+, or 1+ to 3+, could be a cause for concern as the results determine patient treatment. There were no such differences for the image analysis system while there were 4 such disagreements for manual scoring (reference method). The different threshold estimations lack of disagreement for the image analysis slides could be an artifact of the sample set selected; still there is less disagreement compared to the reference method. The risk of a miscall due to variation in the training sets is mitigated by mandatory pathologist review of all results and labeling stressing the importance of test quality control and validation.

K. Proposed Labeling:

The labeling is sufficient and it satisfies the requirements of 21 CFR Part 809.10.

L. Conclusion:

The submitted information in this premarket notification is complete and supports a substantial equivalence decision.