White Paper

PowerLook®
Density Assessment

CLINICAL PERFORMANCE STUDY

3 PowerLook® Density Assessment Clinical Performance Study

Abstract

PowerLook® Density Assessment automated breast density software was tested for standalone performance in measuring breast density during routine mammography. The results demonstrate that the PowerLook Density Assessment Automated Breast Density Software accurately estimates percent breast density (PBD), area of dense tissue, total breast area, and BI-RADS breast density category – irrespective of the (FDA-approved) imaging sensor.

Introduction

PowerLook Density Assessment (formally branded as iReveal) automated breast density software automatically analyzes "for processing" digital mammograms and calculates the dense tissue area of each breast. The measured dense tissue area and a dispersion feature computed on the available mammographic views are then used to provide a Calibrated Density Category which maps the percentage of breast density and dispersion to a BI-RADS breast density category (a-d).

PowerLook Density Assessment is a stand-alone software application designed to interoperate with all digital radiography (DR) and computed radiography (CR) mammography systems. PowerLook Density Assessment is displayed in the form of a DICOM mammography structured report or secondary capture and reports the following for output:

- Breast Area (cm^2) for each breast
- Dense Area (cm^2) for each breast
- Percent Breast Density for each breast
- BI-RADS Breast Density Category for each case

The results of PowerLook Density Assessment are designed to display on a mammography workstation, on a high resolution monitor or in a printed case report. PowerLook Density Assessment is designed to process approximately 60-120 cases per hour. PowerLook Density Assessment is designed to operate on iCAD's PowerLook® platform.

The goal of this study was to determine the accuracy of the four PowerLook Density Assessment outputs: percent breast density (PBD), area of dense tissue, total breast area, and BI-RADS breast density category.

Methods and Materials

One production CAD Station was utilized to process all data. This Station was configured as follows:

- Dell Precision T1600, Service Tag number/HASP key:
  - Test Unit 1: 6J52LS1/2039563047
- Windows 7, service pack 1
- CAD Station software version 6.2.0.0
- PowerLook Density Assessment 2.1.0.0

Additional software tools developed internally were utilized:

- Density Truth Tool

- Breast Segmentation Tool
- PacsSim Tool, Version 6.5

This clinical standalone performance study included cases selected from an image database to create validation test datasets (Table 1). Set 1 contains two cancer cases, allowing it to mimic the prevalence of cancer in a typical screening population. Set 2 contains 9 pairs of cancer cases. Production PowerLook Density Assessment test systems were used to process data from the clinical test datasets to evaluate algorithm stand-alone performance. The clinical test dataset was processed through the PowerLook Density Assessment test units.

Table 1. Validation Test Datasets

Dataset	Sensor	Total Cases
Set 1		**502**
	Kodak (Carestream) DirectView CR Mammography System	73
	Konica Minolta Xpress Digital Mammography System	75
	Hologic Selenia (Lorad Selenia) Dimensions 2D	74†
	GE Senographe 2000D, GE Senographe DS, GE Senographe Essential	74†
	Sectra (Philips) MicroDose	51
	FUJIFILM Aspire HD, FUJIFILM Aspire HD Plus, FUJIFILM Aspire HD-s	55
	Siemens MAMMOMAT Inspiration	50
	Philips Mammodiagnost DR	50
Set 2		**146**
	Hologic Selenia (Lorad Selenia) Dimensions 2D	84††
	GE Senographe 2000D, GE Senographe DS, GE Senographe Essential	62†††
Set 3		**150**
	Kodak (Carestream) DirectView CR Mammography System	22
	Konica Minolta Xpress Digital Mammography System	19
	Hologic Selenia (Lorad Selenia) Dimensions 2D	22
	GE Senographe 2000D, GE Senographe DS, GE Senographe Essential	22
	Sectra (Philips) MicroDose	22
	FUJIFILM Aspire HD, FUJIFILM Aspire HD Plus, FUJIFILM Aspire HD-s	21
	Siemens MAMMOMAT Inspiration	22

†Contains one cancer case
†† Contains eight cancer cases
††† Contains ten cancer cases

The data set was processed as follows:

1. Data Acquisition: cases that met all inclusion and exclusion criteria were put through a series of quality control checks before being imported into the database.

2. Establishing and Releasing the Clinical Test Dataset: the Study Data Manager selected qualified cases from an image database. Cases were organized into groups, and the groups were assigned a validation test dataset number and released for use.
3. Image Processing:
    a. Qualified cases are transmitted to the PowerLook Density Assessment CAD Station.
    b. PowerLook Density Assessment processes cases.
    c. The PowerLook Density Assessment CAD Station stores results which are extracted for analysis.

The set was divided and run on one PowerLook Density Assessment test system. Data processing was performed as the datasets were released. All datasets were processed successfully. During processing, there were no issues noted and there were no excluded cases.

Results

1. Percent Breast Density

Ground Truth

The following procedure was used to establish ground truth for PBD. Ten (10) radiologists analyzed all images in Set 1 using PowerLook Density Assessment Density Truth Tool. This tool allowed the radiologists to segment the dense regions in the mammogram by adjusting an intensity threshold. PowerLook Density Assessment's Breast Segmentation Tool automatically determined the total area of each breast. PBD for a specified view was calculated as 100 times the ratio of the dense area to the total breast area. To arrive at a single PBD ground truth for each view, the PBD measurements (for each view) were averaged over all ten radiologists. By further averaging over views (CC and MLO) and then breasts (right and left), we arrived at ground truth estimates for breasts and cases, respectively.

Comparison of Radiologist Visual PBD to PowerLook Density Assessment PBD

Ten radiologists visually estimated the PBD for all cases in Set 1. PowerLook Density Assessment also estimated the PBD. Tables 2-4 compare the efficacy of PowerLook Density Assessment PBD estimates to the visual PBD estimates from the radiologists. Efficacy is measured by the absolute difference between the PBD estimates and the associated ground truth. (Note that PBD is a percentage: 0%-100%) Figure 1 plots PBD estimates for PowerLook Density Assessment and the ten radiologists versus ground truth.

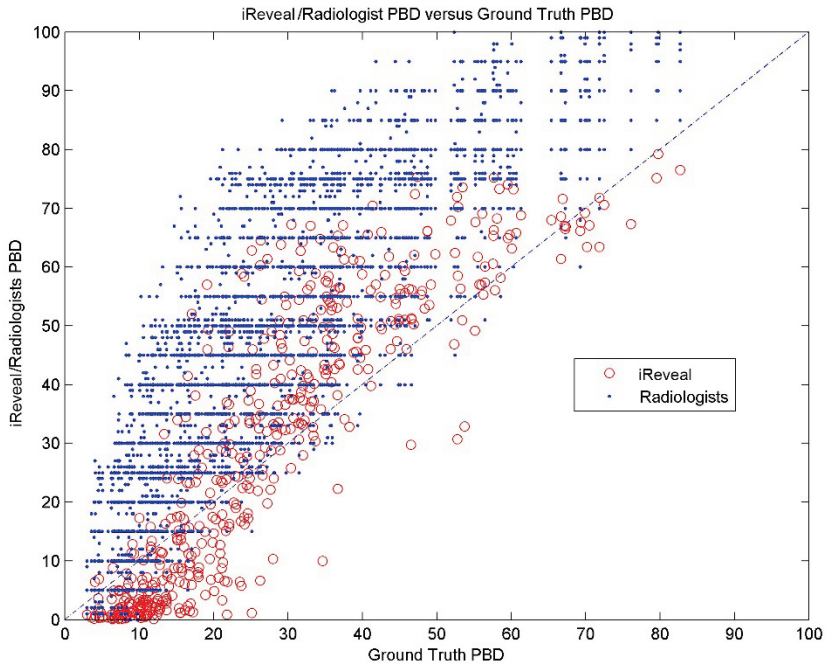

Figure 1. Plot of PowerLook Density Assessment PBD and Radiologist PBD versus ground-truth PBD

Table 2 compares PowerLook Density Assessment PBD performance to the performance of each radiologist. The second column of Table 2 reports the mean and standard deviation of the absolute error between the PowerLook Density Assessment/radiologist PBD estimates and the associated ground truth. The third column provides the mean and standard deviation of radiologist absolute error subtracted from the PowerLook Density Assessment absolute error. The final two columns indicate whether the mean differences in the third column are statistically significant under the null hypothesis that the radiologists have the same or less absolute error than PowerLook Density Assessment. Statistical significance was determined using a one-sided paired t-test with a significance level of 0.01.

Note that all mean differences in the third column in Table 2 are negative, indicating that the PowerLook Density Assessment mean absolute error is less than that of every radiologist. Further note that all these mean differences are statistically significant, except that comparing PowerLook Density Assessment to Radiologist 10.

The mean absolute error of PowerLook Density Assessment PBD is less than the mean absolute error of 9 out of 10 radiologists by a margin that is statistically significant (alpha of 0.01).

PowerLook® Density Assessment Clinical Performance Study

Table 2. Percent breast density analyses comparing PowerLook Density Assessment to radiologist visual estimation

Raters	Absolute Error from Truth (%)	Difference of Absolute Errors from PowerLook Density Assessment	P-Value[1]	Upper One-Sided 99% Confidence Interval[1]
	Mean (SD)	Mean (SD)		
PowerLook Density Assessment	9.566 (7.791)	--	--	--
Radiologist 1	26.62 (15.40)	-17.06 (16.51)	< 0.001	-15.34
Radiologist 2	10.97 (7.173)	-1.403 (9.707)	0.001	-0.392
Radiologist 3	25.89 (11.65)	-16.32 (14.13)	< 0.001	-14.85
Radiologist 4	29.94 (10.22)	-20.37 (12.70)	< 0.001	-19.05
Radiologist 5	21.72 (9.617)	-12.16 (10.79)	< 0.001	-11.03
Radiologist 6	13.90 (10.35)	-4.335 (12.21)	< 0.001	-3.063
Radiologist 7	17.71 (12.73)	-8.146 (14.25)	< 0.001	-6.661
Radiologist 8	25.83 (13.39)	-16.27 (14.74)	< 0.001	-14.73
Radiologist 9	12.07 (8.365)	-2.507 (12.10)	< 0.001	-1.247
Radiologist 10	10.20 (7.406)	-0.632 (10.40)	0.087	0.451
Radiologist Average	18.97 (8.027)	-9.407 (10.53)	< 0.001	-8.310

1Paired t-test

Table 3 recapitulates the results of Table 2; however, the results are now stratified by sensor type and aggregated across all radiologists. The second/third columns of Table 3 report – for each sensor – the mean and standard deviation of the absolute error between the PowerLook Density Assessment/radiologist PBD estimates and the associated ground truth. The fourth column provides the mean and standard deviation of radiologist absolute error subtracted from the PowerLook Density Assessment absolute error. The final two columns indicate whether the mean differences in the fourth column are statistically significant (under the null hypothesis that the mean radiologist has the same or less absolute error than PowerLook Density Assessment) using a one-sided paired t-test with a significance level of 0.01.

Note that all mean differences in the fourth column in Table 3 are negative, indicating that the PowerLook Density Assessment mean absolute error is less than that of the radiologists for all sensors. Also note that all these mean differences are statistically significant.

For each sensor, the mean absolute error of PowerLook Density Assessment PBD is less than the mean absolute error of the radiologists by a margin that is statistically significant (alpha of 0.01).

Table 3. Stratification of absolute error of PowerLook Density Assessment/radiologist PBD by sensor type

Sensor Name	Absolute Error from Ground Truth (%) Mean (SD)		Difference of Absolute Errors from PowerLook Density Assessment	P-Value[1]	Upper One-Sided 99% Confidence Interval[1]
	PowerLook Density Assessment	Radiologist	Mean (SD)		
Konica Minolta	6.607 (5.537)	21.34 (7.099)	-14.73 (8.254)	< 0.001	-12.46
GE	8.436 (5.364)	21.53 (7.487)	-13.09 (11.09)	< 0.001	-10.03
Hologic	11.03 (7.642)	15.09 (7.959)	-4.062 (8.270)	< 0.001	-1.775
Siemens	5.784 (4.525)	17.53 (9.886)	-11.74 (10.86)	< 0.001	-8.050
Philips	8.097 (5.571)	17.54 (7.806)	-9.447 (9.648)	< 0.001	-6.166
Carestream	9.817 (5.719)	17.26 (7.686)	-7.442 (9.576)	< 0.001	-4.776
Fuji-DR	13.95 (11.54)	18.18 (7.031)	-4.229 (10.55)	0.002	-0.819
Sectra	13.49 (11.07)	23.54 (5.656)	-10.05 (11.22)	< 0.001	-6.279

1Paired t-test

Table 4 also recapitulates the results of Table 2; however, the results are now stratified by BI-RADS breast density category and aggregated across all radiologists. The BI-RADS category of each case is defined as the mode of the distribution of BI-RADS categories (for that case) as determined by the 10 radiologists. The second/third columns of Table 4 report – for each BI-RADS category – the mean and standard deviation of the absolute error between the PowerLook Density Assessment/radiologist PBD estimates and the associated ground truth. The fourth column provides the mean and standard deviation of radiologist absolute error subtracted from the PowerLook Density Assessment absolute error. The final two columns indicate whether the mean differences in the fourth column are statistically significant (under the null hypothesis that the radiologists have the same or less absolute error than PowerLook Density Assessment) using a one-sided paired t-test with a significance level of 0.01.

Note that all mean differences in the fourth column in Table 4 are negative, indicating that the PowerLook Density Assessment mean absolute error is less than that of the radiologists for all BI-RADS categories. Further note that all these mean differences are statistically significant.

For each BI-RADS category, the mean absolute error of PowerLook Density Assessment PBD is less than the mean absolute error of the radiologists by a margin that is statistically significant (alpha of 0.01)

Table 4. Stratification of absolute error of PowerLook Density Assessment/radiologist PBD by BI-RADS breast density category

BI-RADS Density Category	Absolute Difference from Ground Truth (%) Mean (SD)		Difference of Absolute Differences from PowerLook Density Assessment	P-Value[1]	Upper One-Sided 99% Confidence Interval[1]
	PowerLook Density Assessment	Radiologist	Mean (SD)		
a	6.453 (2.878)	7.989 (3.908)	-1.537 (4.935)	0.001	-0.410
b	10.47 (8.591)	18.69 (4.749)	-8.229 (9.800)	< 0.001	-6.542
c	12.02 (8.871)	25.73 (4.800)	-13.71 (11.52)	< 0.001	-11.46
d	6.499 (5.392)	22.78 (6.526)	-16.28 (7.403)	< 0.001	-14.06

1Paired t-test

Relationship between PowerLook Density Assessment PBD and Patient Age

PBD is known to decrease with patient age [1, 2]. Figure 2 plots PowerLook Density Assessment PBD as a function of patient age for Set 1. Note that 265 of the 502 cases in Set 1 were omitted from this plot since age information was not available. The Spearman's rank correlation test confirms a negative correlation: -0.530. This value is statistically significant (alpha level of 0.01) with a p-value of < 0.001 (MATLAB 7.9.0). The Spearman's rank correlation coefficient is negative and statistically significant (alpha of 0.01).

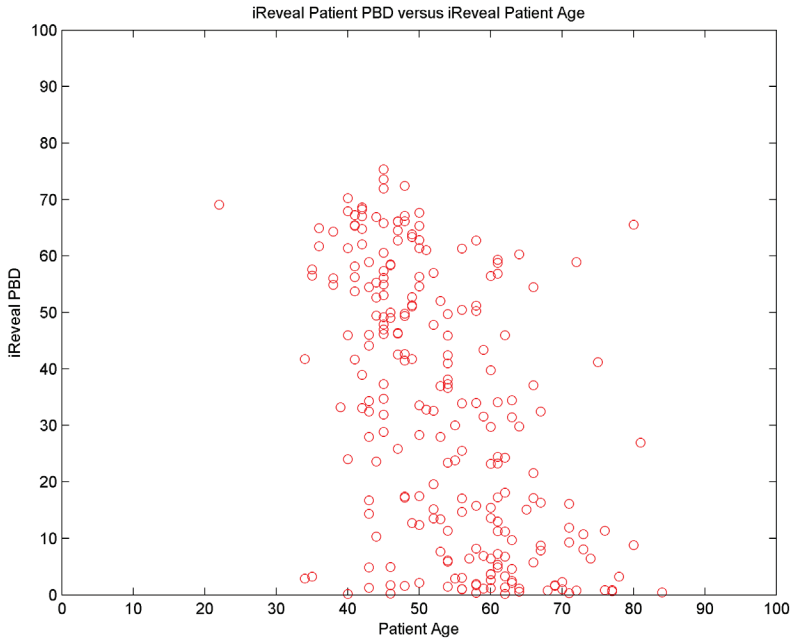

Figure 2. Plot of PowerLook Density Assessment PBD versus patient age

PBD Measurements for Currents and Priors

PBD measurements from the same patient should remain reasonably consistent over small changes in time. Figure 3 plots the PowerLook Density Assessment PBD measurements of all patients in Set 2 at two different time points. Measurements made at the earlier time point are called “PBD priors”; corresponding measurements made at the later time point are “PBD currents”. The maximum separation between the time points is two years. The Pearson’s correlation coefficient is 0.965 (p-value < 0.001), indicating high linear dependence. The Pearson’s correlation coefficient is statistically significant and indicates high linear correlation (0.8 or greater).

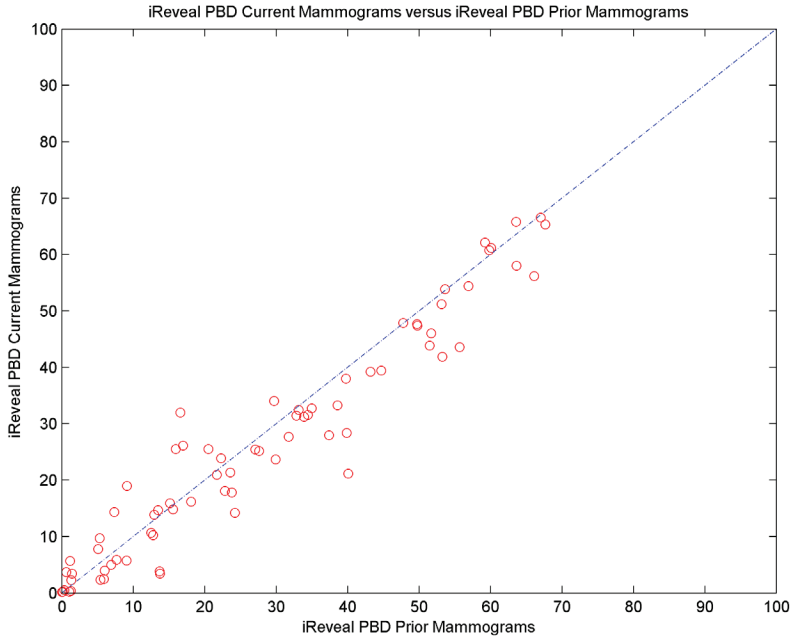

Figure 3. Plot of PowerLook Density Assessment PBD measurements of 73 patients made at two different time points. The separation in time between successive measurements is no more than two years.

PBD Measurements for Left and Right Breasts

PBD measurements from the left and right breasts of the same patient should be very similar. Figure 4 plots the PowerLook Density Assessment PBD of the right breast versus the PowerLook Density Assessment PBD of the left breast from Set 1. The Pearson's correlation coefficient of 0.970 (p-value of < 0.001) indicates strong linear dependence. The Pearson's correlation coefficient is statistically significant and indicates high linear correlation (0.8 or greater).

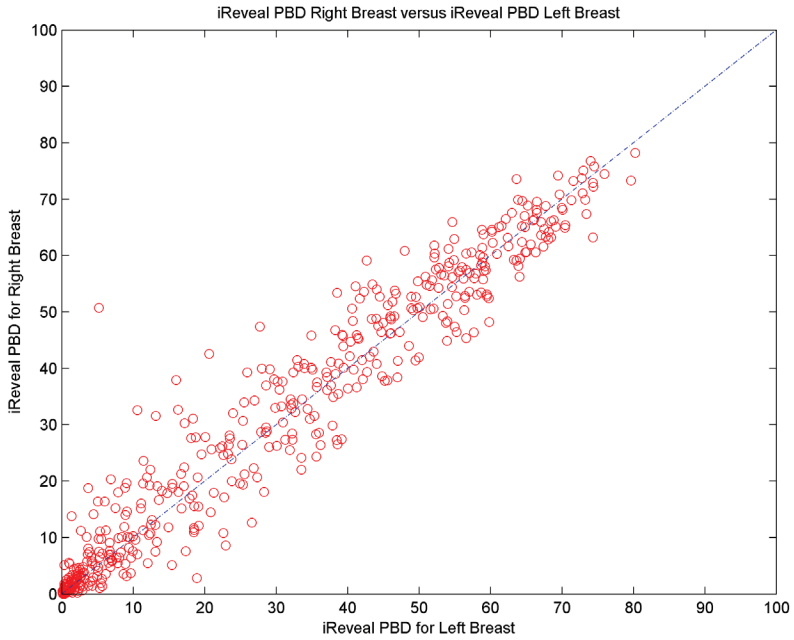

Figure 4. Plot of PowerLook Density Assessment PBD estimates of the right breast versus the PowerLook Density Assessment PBD estimates of the left breast.

PBD Measurements for MLO and CC Views

PBD measurements from the MLO and CC views of the same breast should be very similar. Figure 5 plots the PowerLook Density Assessment PBD of the MLO view versus the PowerLook Density Assessment PBD of the CC view from Set 1. The Pearson's correlation coefficient of 0.935 (p-value of < 0.001) indicates strong linear dependence. The Pearson's correlation coefficient is statistically significant and indicates high linear correlation (0.8 or greater).

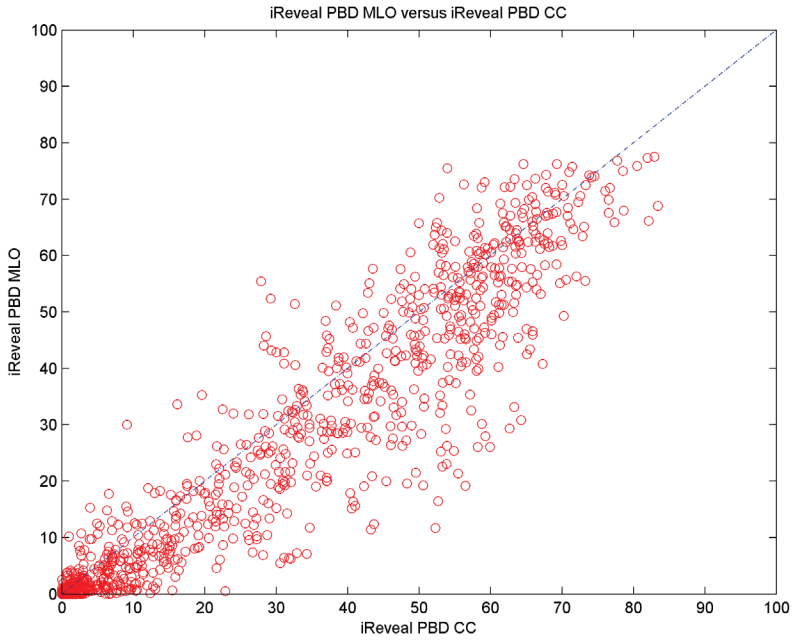

Figure 5. Plot of PowerLook Density Assessment PBD estimates of the MLO view versus PowerLook Density Assessment PBD estimates of the CC view

2. Breast Area

Ground Truth

To establish ground truth for breast area, a single radiologist examined 50 cases (200 images) randomly selected from Set 1. The truth procedure was as follows: 1) the radiologist viewed the Breast Segmentation Tool's automatically-generated breast boundary overlaid on the corresponding view in the mammogram and 2) if the radiologist did not agree with the breast boundary, she manually delineated an alternative boundary using the Truth Tool. The breast area (ground truth) was subsequently computed using the corrected (or uncorrected) boundaries.

Comparison of PowerLook Density Assessment Breast Area to Ground Truth

To assess that PowerLook Density Assessment correctly estimates breast area, we calculated (for each breast) the mean and standard deviation of the absolute relative difference in area between the PowerLook Density Assessment breast area and ground truth (Table 5). These statistics are determined over the 50 cases randomly selected from Set 1. The PowerLook Density Assessment breast area estimates differ from the ground truth by less than 0.6%. The mean absolute relative error is very small (less than 1%).

Table 5. Mean and standard deviations of absolute relative difference between PowerLook Density Assessment breast area estimates and ground truth

	Left Breast	Right Breast
Mean Abs. Relative Difference	0.004	0.007
Standard Deviation Abs. Relative Difference	0.015	0.024

3. Dense Area

PBD equals the ratio of dense area to overall breast area. Since we assessed PowerLook Density Assessment PBD and PowerLook Density Assessment breast area, the performance of PowerLook Density Assessment dense area follows by implication.

4. BI-RADS Breast Density Category

PowerLook Density Assessment chooses a BI-RADS density category by mapping the PowerLook Density Assessment PBD and a dispersion feature computed on the available mammographic views to the letters a, b, c or d. The specific mapping was calibrated to 13 expert radiologists. This calibration allows PowerLook Density Assessment to yield BI-RADS categories that mimic the consensus of these 13 expert radiologists – where the consensus (for each case) is defined as the mode of the distribution of BI-RADS categories as determined by the 13 radiologists.

The 13 expert radiologists also provided BI-RADS breast density categories for the 150 cases in Set 3. Additionally, PowerLook Density Assessment produced BI-RADS breast density categories for the set of 150 cases. The following experimental results over Set 3 confirm the success of calibration. Note that Set 3 is independent of the dataset used for calibration.

Inter-Rater Agreement (Kappa Statistic) of PowerLook Density Assessment Breast Density Category to Expert Radiologists

Using a weighted kappa statistic, we assess the inter-rater agreement among all pairs of radiologists and between each radiologist and PowerLook Density Assessment. The results are shown in Table 6. Note that the agreement between PowerLook Density Assessment and radiologists – which ranges from 0.537 to 0.744 – indicates "moderate" to "substantial" agreement [3], and is comparable to the kappa statistics among radiologists.

Table 6. Weighted kappa statistic measuring inter-rater agreement among 13 radiologists and PowerLook Density Assessment

Raters	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	PowerLook Density Assessment
R1	--	0.71	0.66	0.72	0.51	0.65	0.66	0.79	0.76	0.69	0.67	0.52	0.70	0.67
R2	0.71	--	0.48	0.67	0.42	0.61	0.63	0.76	0.68	0.79	0.53	0.44	0.61	0.68
R3	0.66	0.48	--	0.51	0.39	0.54	0.49	0.59	0.57	0.48	0.70	0.39	0.53	0.55
R4	0.72	0.67	0.51	--	0.69	0.72	0.80	0.78	0.83	0.78	0.64	0.71	0.83	0.69
R5	0.51	0.42	0.39	0.69	--	0.68	0.70	0.56	0.62	0.58	0.59	0.79	0.67	0.54
R6	0.65	0.61	0.54	0.72	0.68	--	0.77	0.69	0.65	0.75	0.69	0.71	0.67	0.68
R7	0.66	0.63	0.49	0.80	0.70	0.77	--	0.74	0.80	0.76	0.58	0.73	0.82	0.72
R8	0.79	0.76	0.59	0.78	0.56	0.69	0.74	--	0.80	0.74	0.71	0.57	0.73	0.73
R9	0.76	0.68	0.57	0.83	0.62	0.65	0.80	0.80	--	0.71	0.63	0.65	0.85	0.71
R10	0.69	0.79	0.48	0.78	0.58	0.75	0.76	0.74	0.71	--	0.59	0.63	0.69	0.74
R11	0.67	0.53	0.70	0.64	0.59	0.69	0.58	0.71	0.63	0.59	--	0.58	0.62	0.57
R12	0.52	0.44	0.39	0.71	0.79	0.71	0.73	0.57	0.65	0.63	0.58	--	0.66	0.57
R13	0.70	0.61	0.53	0.83	0.67	0.67	0.82	0.73	0.85	0.69	0.62	0.66	--	0.70
PowerLook Density Assessment	0.67	0.68	0.55	0.69	0.54	0.68	0.72	0.73	0.71	0.74	0.57	0.57	0.70	--

More important than the agreement between PowerLook Density Assessment and the individual radiologists is the agreement between PowerLook Density Assessment and the mode of the radiologists (to which PowerLook Density Assessment was calibrated). Figure 6 provides a box-whisker plot of PowerLook Density Assessment PBD versus the mode of the radiologists BI-RADS estimates. Each box in the plot corresponds to one BI-RADS category; the central mark indicates the median; the edges correspond to the 25th and 75th percentiles. The whiskers of each box extend from the edges of the box to 1.5 times the interquartile range. The outliers – i.e. points outside the whiskers – are plotted individually using cross marks. The Spearman's rank correlation, which measures the agreement between PowerLook Density Assessment PBD and the mode of the 13 experts, is 0.856 (p-value of < 0.001); this indicates a high degree of correlation. The weighted kappa statistic between the PowerLook Density Assessment and the mode of the radiologists BI-RADS is 0.746, indicating "substantial" agreement.

The weighted kappa statistic relating PowerLook Density Assessment to the mode of the radiologists BI-RADS categories indicates good agreement (0.6 or higher).

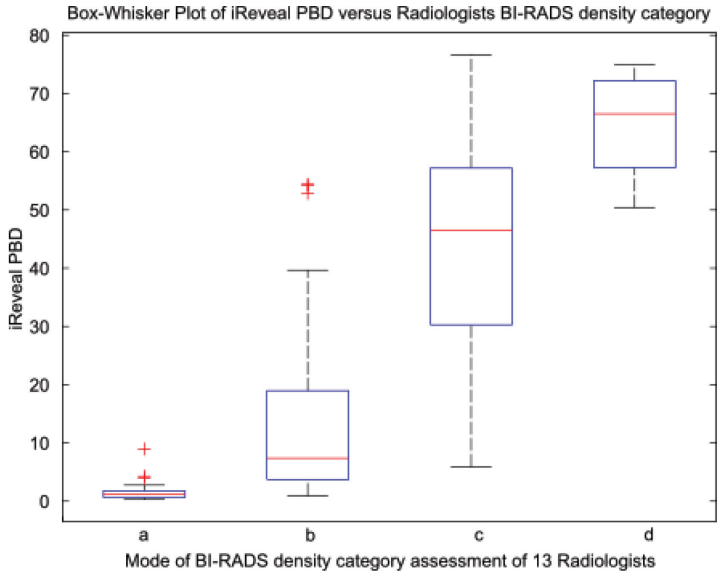

Figure 6. Box-Whisker plot of PowerLook Density Assessment BDE PBD versus mode of the BI-RADS breast density category from the 13 expert radiologists

Observation Matrices of Breast Density Categories

Table 7 shows the observation matrices between 1) the PowerLook Density Assessment density category and the associated mode of the radiologists BI-RADS categories and 2) each radiologist and the mode. Table 8 summarizes these matrices by reporting the proportion of cases in which PowerLook Density Assessment/radiologist yields a category within one level of the mode. Note that for 100% of the cases in Set 3, the PowerLook Density Assessment density category falls within one level of the mode. The one-sided lower 95% confidence interval for the above proportion is 1.000 (determined via bootstrapping).

Table 7. Observation matrices relating density categories from PowerLook Density Assessment/radiologist to radiologists' BI-RADS mode

		Radiologist BI-RADS (Mode)				Totals
		a	b	c	d	
PowerLook Density Assessment Density Category	a	20	4	0	0	24
	b	4	39	5	0	48
	c	0	10	41	2	53
	d	0	0	13	12	25
Totals		24	53	59	14	150

		Rad. BI-RADS (Mode without 1)				Totals
		a	b	c	d	
Rad 1	a	12	0	0	0	12
	b	20	34	0	0	54
	c	0	16	52	0	68
	d	0	0	3	13	16
Totals		32	50	55	13	150

		Rad. BI-RADS (Mode without 2)				Totals
		a	b	c	d	
Rad 2	a	15	0	0	0	15
	b	17	33	0	0	50
	c	0	17	25	0	42
	d	0	0	30	13	43
Totals		32	50	55	13	150

		Rad. BI-RADS (Mode without 3)				Totals
		a	b	c	d	
Rad 3	a	2	0	0	0	2
	b	30	36	3	0	69
	c	0	14	51	11	76
	d	0	0	0	3	3
Totals		32	50	54	14	150

		Rad. BI-RADS (Mode without 4)				Totals
		a	b	c	d	
Rad 4	a	24	10	0	0	34
	b	1	40	3	0	44
	c	0	4	48	1	53
	d	0	0	7	12	19
Totals		25	54	58	13	150

		Rad. BI-RADS (Mode without 5)				Totals
		a	b	c	d	
Rad 5	a	22	14	0	0	36
	b	0	39	27	0	66
	c	0	0	32	4	36
	d	2	0	0	10	12
Totals		24	53	59	14	150

		Rad. BI-RADS (Mode without 6)				Totals
		a	b	c	d	
Rad 6	a	20	6	0	0	26
	b	11	38	13	0	62
	c	0	2	39	2	43
	d	0	0	7	12	19
Totals		31	46	59	14	150

		Rad. BI-RADS (Mode without 7)				Totals
		a	b	c	d	
Rad 7	a	24	16	0	0	40
	b	0	31	4	0	35
	c	0	8	47	2	57
	d	0	0	6	12	18
Totals		24	55	57	14	150

Table 7. Observation matrices relating density categories from PowerLook Density Assessment/radiologist to radiologists' BI-RADS mode

		Rad. BI-RADS (Mode without 8)				Totals
		a	b	c	d	
Rad 8	a	17	0	0	0	17
	b	15	39	1	0	55
	c	0	11	46	0	57
	d	0	0	8	13	21
Totals		32	50	55	13	150

		Rad. BI-RADS (Mode without 9)				Totals
		a	b	c	d	
Rad 9	a	23	8	0	0	31
	b	3	34	0	0	37
	c	0	14	52	1	67
	d	0	0	3	12	15
Totals		26	56	55	13	150

		Rad. BI-RADS (Mode without 10)				Totals
		a	b	c	d	
Rad 10	a	23	8	0	0	31
	b	6	38	2	0	46
	c	0	5	34	0	39
	d	0	0	21	13	34
Totals		29	51	57	13	150

		Rad. BI-RADS (Mode without 11)				Totals
		a	b	c	d	
Rad 11	a	11	0	0	0	11
	b	21	44	12	0	77
	c	0	2	46	8	56
	d	0	0	0	6	6
Totals		32	46	58	14	150

		Rad. BI-RADS (Mode without 12)				Totals
		a	b	c	d	
Rad 12	a	24	26	0	0	50
	b	0	27	23	0	50
	c	0	0	36	2	38
	d	0	0	1	11	12
Totals		24	53	60	13	150

		Rad. BI-RADS (Mode without 13)				Totals
		a	b	c	d	
Rad 13	a	24	10	0	0	34
	b	1	32	4	0	37
	c	0	12	52	5	69
	d	0	0	1	9	10
Totals		25	54	57	14	150

At least 95% of the cases (using a one-sided lower 95% confidence interval) have PowerLook Density Assessment density categories that fall within one level of the mode of the radiologists' BI-RADS breast density categories as shown in Table 8.

Table 8. Proportion of cases (for each individual rater) wherein the density category lies within one level of the radiologists' BI-RADS mode

Rater	Percentage of cases within one category of the mode
PowerLook Density Assessment	100%
Radiologist 1	100%
Radiologist 2	100%
Radiologist 3	100%
Radiologist 4	100%
Radiologist 5	98.67%
Radiologist 6	100%
Radiologist 7	100%
Radiologist 8	100%
Radiologist 9	100%
Radiologist 10	100%
Radiologist 11	100%
Radiologist 12	100%
Radiologist 13	100%

Sensor-specific Performance

PowerLook Density Assessment's density category performance was evaluated for individual sensors. For each sensor in Set 3, Table 9 provides the observation matrix relating PowerLook Density Assessment to the mode of the 13 radiologists' breast density categories.

Table 9. Observation matrix between PowerLook Density Assessment density category and the radiologists' BI-RADS mode for each sensor

		Rad. BI-RADS Mode (Konica)				Totals
		a	b	c	d	
PowerLook Density Assessment density score (Konica)	a	4	1	0	0	5
	b	2	3	1	0	6
	c	0	4	2	0	6
	d	0	0	2	0	2
Totals		6	8	5	0	19

		Radiologist BI-RADS Mode (GE)				Totals
		a	b	c	d	
PowerLook Density Assessment density score (GE)	a	2	0	0	0	2
	b	1	5	2	0	8
	c	0	0	8	0	8
	d	0	0	0	4	4
Totals		3	5	10	4	22

		Radiologist BI-RADS Mode (Siemens)				Totals
		a	b	c	d	
PowerLook Density Assessment density score (Siemens)	a	6	0	0	0	6
	b	0	5	0	0	5
	c	0	2	4	0	6
	d	0	0	3	2	5
Totals		6	7	7	2	22

		Rad. BI-RADS Mode (Hologic)				Totals
		a	b	c	d	
PowerLook Density Assessment density score (Hologic)	a	4	2	0	0	6
	b	0	5	1	0	6
	c	0	0	8	0	8
	d	0	0	1	1	2
Totals		4	7	10	1	22

		Radiologist BI-RADS Mode (Sectra)				Totals
		a	b	c	d	
PowerLook Density Assessment density score (Sectra)	a	0	0	0	0	0
	b	0	8	1	0	9
	c	0	1	7	0	8
	d	0	0	4	1	5
Totals		0	9	12	1	22

		Rad. BI-RADS Mode (Carestream)				Totals
		a	b	c	d	
PowerLook Density Assessment density score (Carestream)	a	3	1	0	0	4
	b	0	7	0	0	7
	c	0	0	6	2	8
	d	0	0	1	2	3
Totals		3	8	7	4	2

		Radiologist BI-RADS Mode (Fuji-DR)				Totals
		a	b	c	d	
PowerLook Density Assessment density score (Fuji-DR)	a	1	0	0	0	1
	b	1	6	0	0	7
	c	0	3	6	0	9
	d	0	0	2	2	4
Totals		2	9	8	2	21

Table 10 summarizes these matrices by reporting the proportion of cases in which PowerLook Density Assessment yields a category within one level of the mode.

Table 10. Proportion of cases (for each individual sensor) wherein the density category lies within one level of the radiologists' BI-RADS mode.

Sensor	Percentage of cases within one category of the mode (PowerLook Density Assessment)
Carestream	100% (22/22)
Hologic	100% (22/22)
Konica	100% (19/19)
GE	100% (22/22)
Sectra	100% (22/22)
Fuji-DR	100% (21/21)
Siemens	100% (22/22)

Figure 7 illustrates – for each of the sensors – corresponding boxplots of PowerLook Density Assessment PBD versus the mode.

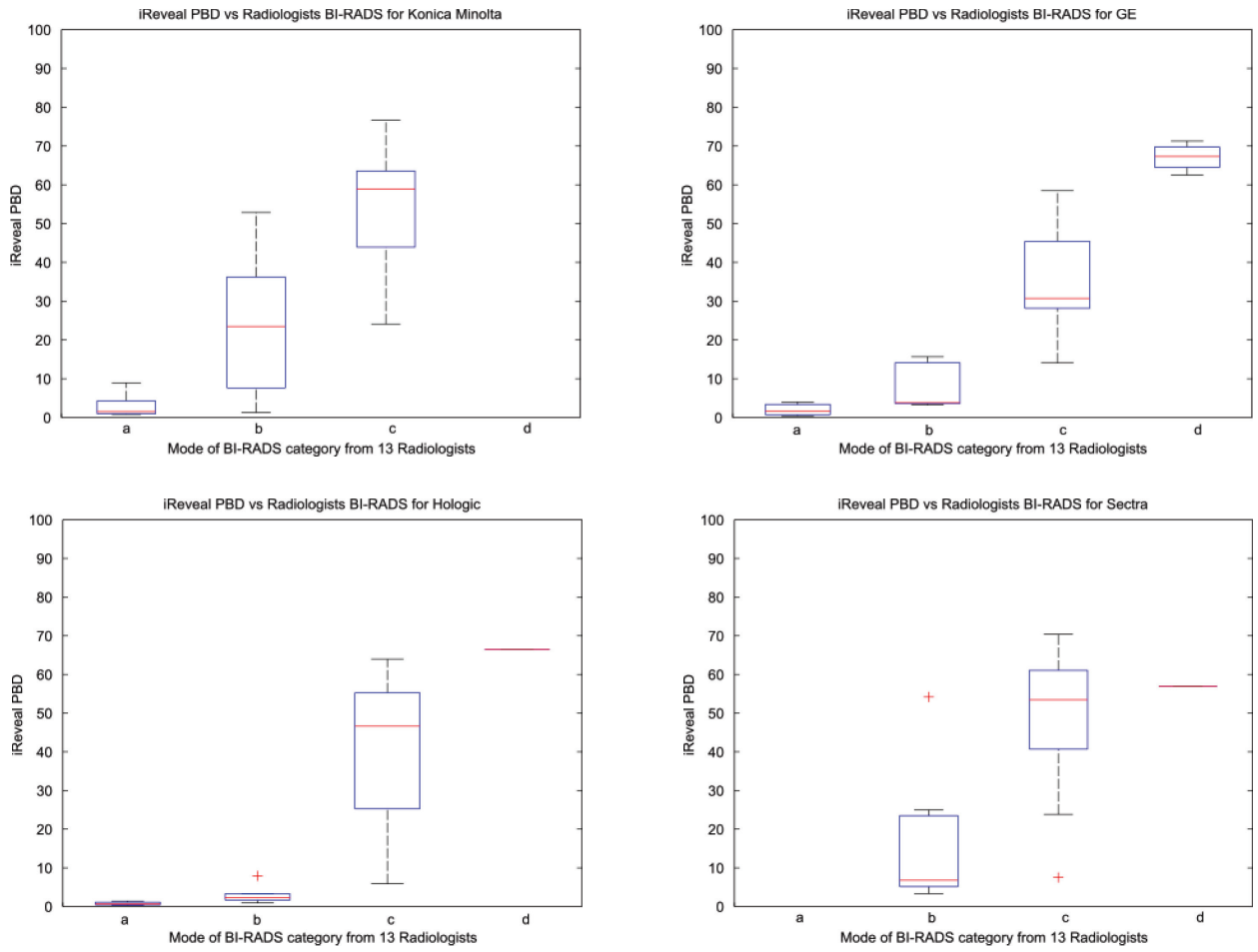

Figure 7. Box-Whisker plot of PowerLook Density Assessment PBD versus mode of the BI-RADS breast density category from the 13 expert radiologists for each sensor in Set 3.

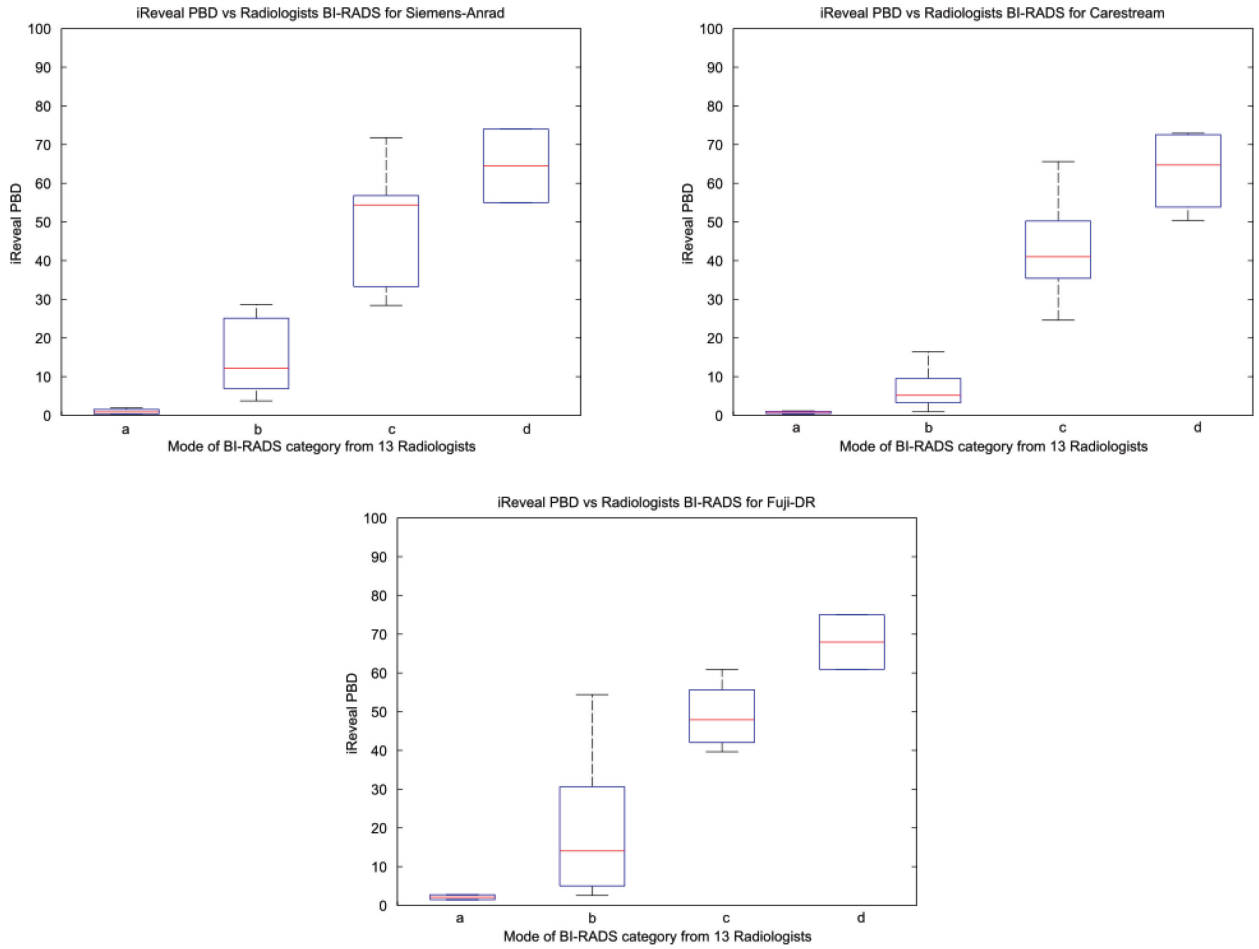

Figure 7. Box-Whisker plot of PowerLook Density Assessment PBD versus mode of the BI-RADS breast density category from the 13 expert radiologists for each sensor in Set 3, continued.

CONCLUSIONS

The clinical performance results indicate that the PowerLook Density Assessment has successfully achieved all study endpoints. Specifically, it demonstrates that PowerLook Density Assessment accurately estimates percent breast density (PBD), parenchymal area, total breast area, and BI-RADS breast density category irrespective of the imaging sensor.

REFERENCES

1. G. Maskarinec, I. Pagano, G. Lurie, L. N. Kolonel, “A longitudinal investigation of mammographic density: the multi-ethnic cohort,” Cancer Epidemiology, Biomarkers and Prevention, vol. 15, pp. 732-739, 2006.

2. C. Checka, M. Cristina, J. E. Chun, F. R. Schnabel, J. Lee, and H. Toth, “The relationship of mammographic density and age: implications for breast cancer screening,” American Journal of