Published in final edited form as:

Acad Radiol. 2014 August ; 21(8): 958–970. doi:10.1016/j.acra.2014.04.006.

Automated percentage of breast density measurements for full field digital mammography applications

E.E.E. Fowler¹, C.M Vachon³, C.G. Scott³, T.A. Sellers¹, and J.J. Heine^{1,2,*}

¹Department of Cancer Epidemiology, Division of Population Sciences, H. Lee Moffitt Cancer Center, 12901 Magnolia Drive, Tampa, FL 33612, USA

²Department Cancer Imaging and Metabolism, H. Lee Moffitt Cancer Center, 12901 Magnolia Drive, Tampa, FL 33612, USA

³Mayo Clinic College of Medicine, Department of Health Sciences Research, 200 First Street SW, Rochester, MN 55905, USA

Abstract

Purpose—Increased mammographic breast density is a significant risk factor for breast cancer. A reproducible, accurate, automated breast density measurement is required for full field digital mammography (FFDM) to support clinical applications. We evaluated a novel automated percentage of breast density measure (PD_a) and made comparisons with the standard operator-assisted measure (PD) using FFDM data.

Methods—We used a nested breast cancer case-control study matched on age, year of mammogram and diagnosis with images acquired from a specific direct x-ray conversion FFDM technology. PD_a was applied to the raw and clinical display (or processed) representation images. We evaluated the transformation (pixel mapping) of the raw image, giving a third representation (raw-transformed), to improve the PD_a performance using differential evolution optimization. We applied PD to the raw and clinical display images as a standard for measurement comparison. Conditional logistic regression was used to estimate the odd ratios (ORs) for breast cancer with 95% confidence intervals for all measurements; analyses were adjusted for body mass index (BMI). PD_a operates by evaluating signal dependent noise (SDN), captured as local signal variation. Therefore, we characterized the SDN relationship to understand the PD_a performance as a function of data representation and investigated a variation analysis of the transformation.

Results—The associations of the quartiles of operator-assisted PD with breast cancer were similar for the raw [OR: 1.00 (ref.); 1.59 (0.93, 2.70); 1.70 (0.95, 3.04); 2.04 (1.13, 3.67)] and clinical display [OR: 1.00 (ref.); 1.31 (0.79, 2.18); 1.14 (0.65, 1.98); 1.95 (1.09, 3.47)] images. PD_a could not be assessed on the raw images without preprocessing. However, PD_a had similar

© 2014 The Association of University Radiologists. Published by Elsevier Inc. All rights reserved.

*Corresponding Author: john.heine@moffitt.org, Phone: 813-745-6719, Fax: 813-745-6525.

Conflicts of Interest: The authors have a pending patent application for this work.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

associations with breast cancer when assessed on raw-transformed (a) [OR: 1.00 (ref.); 1.27 (0.74, 2.19); 1.86 (1.05, 3.28); 3.00 (1.67, 5.38)] and clinical display (b) [OR: 1.00 (ref.); 1.79 (1.04, 3.11); 1.61 (0.90, 2.88); 2.94 (1.66, 5.19)] images. The SDN analysis showed that a nonlinear relationship between the mammographic signal and its variation (i.e. the biomarker for the breast density) is required for PD_a processing. Although variability in the transform influenced the respective PD_a distribution, it did not affect the measurement's association with breast cancer.

Conclusions—PD_a assessed on either raw-transformed or clinical display images is a valid automated breast density measurement for a specific FFDM technology and compares well against PD (the standard). Further work is required for measurement generalization.

1. Introduction

Mammographic breast density is a significant breast cancer risk factor (1-3). Many of the breast density research studies to date have been based on an operator-assisted measure to estimate the percentage of breast density within a mammogram (herein referred to as PD). There are various methods under development to automate the estimation of breast density (4-21). Developing a fully automated and standardized breast density measurement has proven somewhat difficult, but at least two commercial standardized measures are available for raw FFDM images, Volpara and Quantra (19, 21-23). However, these have not been shown to be associated with breast cancer risk to date.

Although there are various FFDM manufacturers, the two predominant FFDM technologies used today consist of direct and indirect x-ray conversion systems (24-26) that produce images with different characteristics. The data representation produced by FFDM systems may vary due to the x-ray detection technology, x-ray generation, or post acquisition processing. FFDM systems produce both raw and clinical display (i.e. processed) representation mammograms. A given clinical display, or processed image, is derived from its respective raw image with methods developed by the unit's manufacturer. The raw images are normally not considered in the clinical evaluation. When applying automated methods, it is not clear if both representations result in similar breast density measurements, if there is a preferred representation, or what impact the technology plays.

Because of the long-standing merit of PD, we are developing an automated measure for FFDM applications referred to as PD_a that provides the same metric as PD. Our automated measurement evolved from earlier work in modeling the Fourier power spectra of digitized-film mammograms. In our prior work, we estimated the spectral form of a given mammogram and removed it with a deconvolution process, resulting in a *noise-field* (i.e. the filtered-image). The degree of local variation in the filtered-image (i.e. noise) corresponded to the degree of mammographic density (i.e. the signal) in the raw image at the same location (27), which is indicative of a signal dependent noise (SDN) relationship. We developed a statistical method for detecting these areas of increased variation in the filtered-image forming the basis of the PD_a technique (28). In subsequent work, the deconvolution process was replaced (approximated) with a high-pass wavelet filter, increasing the algorithm speed, and PD_a was validated using digitized film-mammograms with breast cancer status as the endpoint (29). As of yet, PD_a has not been evaluated in depth with FFDM images.

In this report, we are generalizing the PD_a algorithm for FFDM applications and developing metrics to evaluate the algorithm's performance relative to the data representation. Because our study focused on developing an automated density measure, we controlled for factors known to be related to breast density as well as breast cancer. We applied PD_a to a nested breast cancer case-control dataset for patients with images (raw and processed image representations) acquired from a specific direct x-ray conversion FFDM technology. We applied an empirically determined data transform to the raw images as a preprocessing step to improve the PD_a raw image processing (i.e. to improve the agreement with PD). This transform produces a third data format, defined as the raw-transformed representation. We used an evolutionary optimization strategy to determine the parameters of this transform. We applied PD_a to the three FFDM image representations and compared the respective associations with breast cancer. We compared these associations with those provided by PD (from the Cumulus program described below), considered as the standard for comparison. We also characterized the SDN as a function of the data representation using methods developed previously (30) to understand the impact of each representation on the automated PD_a processing.

2. Methods

2.1 Study Population and mammography

The patients for this study were derived from the Mayo Mammography Health Study (MMHS) cohort, Rochester MN, and described previously (31, 32). Briefly, the MMHS is a prospective cohort study of women living in Minnesota, Wisconsin, or Iowa, older than 35 years, who had a film screening mammography at the Mayo Clinic between 2003 and 2006, and no personal history of breast cancer at study entry. Participants completed a questionnaire and provided written informed consent to use their mammograms, medical records, and blood samples and to link their data to state cancer registries. The 19,924 subjects who participated (51% of the 38,883 subjects who were eligible) were followed up for incident cancer events through the tri-state cancer and Mayo Clinic tumor registries. Through December 31, 2010, a total of 492 incident and histologically confirmed primary breast cancers were identified. The analysis was restricted only to cases who had an FFDM exam at least 6 months prior to diagnosis, limiting our analysis to 228 breast cancer cases and 456 age and interval matched controls (2 per case), which formed our nested case-control study. All patient mammograms were acquired from Hologic Selenia FFDM units. This FFDM unit has 70 micron spatial resolution (pixel pitch), and a 24 cm $\times$ 29 cm field of view (FOV). Screening mammograms are most often acquired with two images sizes depending on the compression paddle choice, inducing a FOV change: 2560 $\times$ 3328 pixels (18 cm $\times$ 24 cm); and 3328 $\times$ 4098 pixels (24 cm $\times$ 29 cm). The raw and processed representation images (i.e. clinical display images) have 14 bit and 12 bit per pixel dynamic range, respectively. For cases, we used the non-cancerous breast, and for controls, the same side used for the matched case was analyzed. We used the cranial caudal (CC) views as the study images. This study was approved by the Mayo Clinic Institutional Review Board.

2.2 Statistical Analysis

Patient characteristics and breast measures were summarized with the distribution mean and standard deviation (SD), and differences between case/control groups were tested using conditional logistic regression. Quartiles and the SD of each breast density measure were defined based on the distribution of that density measure among the control subjects. Conditional logistic regression (33, 34) was used in the primary analysis to examine the association between quartiles or standard deviation of PD with breast cancer status. As the primary metric, the magnitudes of the associations were summarized by odds ratios (ORs) with 95% confidence intervals (CIs). Models were adjusted for body mass index (BMI) measured in kg/m². Missing BMI values for cases and controls were imputed by using the mean BMI of the respective distribution. Additionally as a secondary means to summarize the strength of association, the area under the receiver operating characteristic curve (Az) was computed as a summary of the ability of each model to discriminate between cases and controls. To match the study design, Az was calculated only within matched case-control pairs. A 95% confidence interval was calculated for each Az based on 1000 bootstrap samples and these samples were also used to compare Az. We also calculated Pearson correlation coefficients to measure association between continuous breast density measurements.

2.3 Operator Assisted Percentage of Density

We used the PD association with breast cancer as an established reference for the PD_a and breast cancer findings. PD was estimated from the raw and clinical display images in DICOM format with the Cumulus3 software (University of Toronto). We labeled these two representations because they are the images routinely available for the analysis with Cumulus. The dataset consisting of all cases and matched control images were de-identified and randomized. A reader (JH) with Cumulus experience (35) was blinded to the case-control status and original image identifiers. From our experience when using Cumulus, the operator sets window adjustments and thresholds for each image to separate the dense from non-dense tissue and remove the off breast area region from the analysis. PD was calculated as the total dense area normalized by the total breast area to give the percentage of dense breast tissue as the measure of PD. We have previously shown high intra-reader reliability as well as demonstrated associations with breast cancer using this measure on raw and processed FFDM data of differing technology (35).

2.4 Automated Percentage of Breast Density

The PD_a process is completely automated and is based on the wavelet expansion (28, 29), which takes this form for our application

$$r_{0i} = d_1 + f_1. \quad \text{Eq. (1)}$$

The subscript, *i*, defines the three data representations: *i* = *r* for raw image; *i* = *t* for the raw-transformed image; and *i* = *p* for the processed clinical-display image. The *d*₁ and *f*₁ images are complementary high and low half-band filtered versions (filter outputs) of *r*_{0i} (36-38).

When the raw image dimension is $n_x \times n_y$ pixels (in the x and y direction) the expansion images have the same dimension.

In the next step, the breast region is located automatically and a global-reference variance signal is estimated by using all the pixels in the d_1 image (corresponding to the breast region). We note the breast area detection is a relatively simple procedure because most raw images are intensity saturated, and the view annotation is the same spatial location in all images for a given view. A relatively small $n \times n$ pixel search window is maneuvered across the d_1 image constrained to the breast area with $n = 4$. At each window location, the local variance is calculated. The window is moved in box-width shifts (i.e. a grid) blanketing the breast region, giving the local variance image. We can manipulate this local variance image as a reduced spatial-resolution image with $n_x/n \times n_y/n$ pixels in each dimension. At each window location, the global reference (i.e. reference variance) is compared with the local variance. When the local variance deviates too far above the reference using a chi-square test, the respective $n \times n$ region is labeled as fibroglandular (i.e. dense). This process results in a binary image (on the breast area) with each pixel labeled as either dense or *other*. The global reference variance is then refined by restricting its second (repeated) estimation in the d_1 image to those locations that were initially labeled as *other* in the first search procedure. The search window process (described above) is repeated with the refined global reference variance resulting in the binary-labeled output image analogous to that provided by PD. In the final detection process, the *other* labeled pixels correspond to adipose regions in r_{0i} . The percentage of breast density is calculated in the same manner as PD (a ratio). Each search (detection) stage requires its own significance value of 0.1 and 0.0001, respectively, for these datasets. These parameters generally require modification when the data representation changes, as they affect the detection thresholds and automated density labeling.

All of the image processing was implemented on a Dell PowerEdge R415 (server-class) system. This unit has an AMD Opteron 4238 6 core 3.3 GHz 8MB L3 Cache processor, and 16GB DDR3 1600 MHz memory. For reference, we estimated the processing time for the PD_a supplication per large and small FOV image.

The PD_a algorithm is based on certain approximations and stochastic relationships. These details are discussed because they connect the algorithm's operation to breast tissue, breast density, and the imaging process. The signal dependent noise (SDN) analysis (discussed below in Section 2.6) is based on the validity of switching time (i.e. serial) averages (mean and variance for example) with spatial averages of limited spatial extent as described previously (30). A serial average in this context implies imaging the same object (patient) repeatedly and estimating distribution quantities from a fixed pixel location (or any fixed location). Because a given image is only acquired once and mammograms have long-range positive correlation, we make the assumption that summary measures from a small $n \times n$ region about a fixed location in one image approximates the distribution of taking n^2 serial acquisitions and examining the respective summary measures from a fixed location in the vicinity of the $n \times n$ region in the same image. These arguments with the underlying Poisson process connect the PD_a measure with underlying imaging physics.

2.5 Preprocessing

As an initial evaluation step, 10 case observations were selected randomly from the dataset with their 20 matched controls and processed with PD_a . The PD_a operation did not label properly when applied to the raw images, as assessed visually and compared with PD. The raw images are intensity reversed relative to clinical display or film mammograms (i.e. the format previously used for the PD_a development and validation). To boost the agreement with PD, we applied a linear transformation to reverse the intensities of the raw images, but the method was also ineffective. That is, PD_a did not produce a measurement from the raw data comparable to that of PD without the appropriate pre-processing (examples provided below). For these reasons, we preprocessed the raw images using an empirical based mapping given by

$$r_{0t} = \frac{a_0}{[(m_0 \times r_{0r}) + 1]^k}, \quad \text{Eq. (2)}$$

where a_0 and k are parameters determined with the optimization procedure. The pixel values of r_{0r} were first linearly mapped between (0, 1) before applying Eq. (2). The m_0 factor is an empirically determined scaling constant ≈ 101 , which constrains the pixel values to the allowable dynamic range, and was derived by generalizing the normalization method used in our calibration research work (13, 18, 39). The m_0 factor is the average current $\times$ time (mAs) system readout from a random sample of mammograms. The form of the denominator (addition of one), prevents pixel values in r_{0t} from reaching infinity. Equation (2) defines the raw-transformed image representation. We used this 30 image dataset to determine these two unknown parameters using an evolutionary optimization strategy (40).

Differential evolution (DE) optimization (40) was used to determine the two free-parameters in Eq. (2). These 30 (r_t representation) images were transformed with Eq. (2) and then processed with PD_a . We used the corresponding 30 PD quantities determined from Cumulus (using the raw images) for these images as the reference, or target values, and minimized the L_1 difference, $|PD_a(r_{0t}) - PD|$, between the respective PD_a and PD image pairs by summing over these differences (i.e. the error or fitness function). Our DE methods were described in detail previously (41). For reference, we used the same definitions for the DE parameters as its founders (40): the vector field population is $NP = 10$ random vectors, which is a rule of thumb for the population size for two unknown parameters, the crossover was $CR = 0.1$, and the evolutionary amplification factor was $F = 0.5$. The number of generations was fixed with $G = 100$. This process was initialized with 10 (i.e. NP) two-component random vectors with components corresponding to (k, A_0). A uniform random distribution defined over this range (0, 1) was used for the vector field initialization (i.e. 10 two-component parameter vectors with random values for the zero-generation population). A_0 was constrained to this range [1000, 25000] and k to this range [1, 10]. We matched the Cumulus method output above as the primary experiment because it minimized the amount of training data used in this report, limiting the possibility of over fitting. We determined the parameters of Eq. (2) with the optimization methods described above and used those parameters for the analysis after implementing the optimization process once (i.e. we used the first realization of these parameters) and evaluated the associations with breast cancer. To further evaluate the

variability in our approach while limiting over fitting, we implemented the optimization for 10 additional trials (30 samples in each trial). We estimated the variability in the parameters and evaluated its impact on the logistic regression analysis relative to the findings from the first realization. To induce variability, we selected 10 cases at random and selected the respective 20 controls without replacement. We then found the solution to Eq. (2) defining one additional trial. In the next trial these 10 cases and 20 controls were removed from the random selection process and so on. To assess the (k, A_0) parameter set variability impact on the breast cancer associations, we created three additional raw-transformed representations using the parameter set determined by the minimum and maximum values of k and the respective a_0 (that was determined with each k) and the set formed by the distribution averages of k and a_0 distributions. We transformed the raw data with each parameter set and processed them with PD_a .

2.6 Signal dependent noise analysis

Methods were developed to understand the PD_a performance (i.e. agreement with PD) relative to the data representation for (a) algorithm diagnostic purposes and (b) generalizing the automated processing to other imaging platforms in the future. This is a parallel analysis that does not affect the automated PD_a processing of the case-control dataset in this report. The connection with breast tissue and the SDN follows from considering the mammographic signal, which is described as a Poisson process. In this case, the signal dependency implies the mean and variance are equivalent (or at least bear a linear dependence for the raw data) when the x-ray exposure and pixel value response is linear (30). The PD_a density detection is influenced by the degree of measurable local variation in the d_1 image relative to the respective areas of increased density in r_{0i} . This relationship is captured by considering the signal dependent noise (SDN) characteristic of a given mammogram with a method developed previously (30), which is also uses the wavelet transform as expressed in Eq. (1). The local noise variance is estimated in the d_1 wavelet image, and the corresponding local signal average is estimated in the f_1 image. We modified the SDN functional relationship from the previous work (30), and fit the data from each of the r_{0i} representations (from a given patient) to this more empirically driven general model

$$y=c_0+c_1x+c_2x^2+c_3x^3+c_4\exp(-z)^2, \quad \text{Eq. (3)}$$

where y is the local noise variance, x is corresponding local signal, $z=\frac{x-c_5}{c_6}$, and the c_i are the fit coefficients. This expression is general and was tailored to each representation by initial observation using the 30 image dataset discussed above. For the raw data ($i = r$), we set $c_3 = c_4 = c_5 = c_6 = 0$, for the raw-transformed data ($i = t$) we set $c_4 = c_5 = c_6 = 0$, and for the processed data ($i = p$), we set $c_3 = 0$. Both the signal and noise values were mapped between (0, 1) before applying the curve-fitting analysis. Representative examples are provided to show the differences. We also summarized each representation's fit-coefficient distributions with the mean (taken over all images) and 95% confidence intervals. For our purpose, c_0 is a bias term used as degree of freedom, or flexibility, in the fitting processes, and is not discussed in detail.

In this SDN analysis, we used all case and control images (CC views only) excluding small image areas with calcifications in the applicable images. The presence of calcifications can skew the SDN relationship in this noise modeling [i.e. Eq. (3)] because they may appear as discontinuous abrupt localized spatial changes in the d_1 image. Calcification areas were marked manually, annotated in the automatically segmented mask images (qualified below), and then excluded from the SDN modeling analysis. This analysis was constrained to the interior breast area region corresponding to where the breast was in contact with the compression paddle. We estimate this region by eroding the breast area along a radial direction by 25% using an automated method described previously (18, 42) creating the mask defined above. Image examples of the segmentation and erosion processes are provided in the first figure of our previous work (43). We note, the manual exclusion of calcified areas was only used for the SDN modeling and not for the automated PD_a processing used in the case-control evaluation, where the presence of calcifications is irrelevant for its operation.

3. Results

The patient characteristics are summarized overall and by case-control status (Table 1). Some patients (cases and controls) were excluded from the study due to the presence of implants, bilateral cancer, or corrupt images (i.e. breast areas that were larger than that of the detector FOV) and their respective images were not processed with PD or PD_a . Thus, the final patient dataset used in the PD and automated PD_a processing and subsequent analysis was comprised of 192 cases and 358 controls giving a 1:2 case and control matching ratio for 166 cases and 1:1 ratio for 26 cases. Age and BMI were similar across the cases and controls. PD_a quantities from the raw data could not be calculated using this algorithm without preprocessing, as discussed below. Mean PD_a for cases from both the raw-transformed and clinical display representation images was larger than for controls. The respective means from PD follow the same trend, although the differences were not statistically significant. Although the respective inter-SD quantities vary, the SD findings across the case-control sets for a given measurement are similar.

As expected from previous work, PD assessed on the raw images was a significant risk factor for breast cancer (Table 2). This was seen when examining associations of breast cancer in the adjusted PD [OR for quartiles: 1.00 (ref.); 1.59 (0.93, 2.70); 1.70 (0.95, 3.04); 2.04 (1.13, 3.67); and $Az = 0.57$ (0.52, 0.62)], or continuous PD [OR per SD: 1.21 (0.97, 1.51); and $Az = 0.57$ (0.51, 0.62)]. PD from the clinical display images was also a risk factor in both the adjusted quartile [OR: 1.00 (ref.); 1.31 (0.79, 2.18); 1.14 (0.65, 1.98); 1.95 (1.09, 3.47); and $Az = 0.57$ (0.53, 0.62)] and continuous [OR: 1.22 (0.98, 1.51); and $Az = 0.55$ (0.50, 0.60)] models. The ROC curves for the continuous (unadjusted) PD models are shown in Figure 1.

For the PD_a measure, the raw images were pre-processed prior to estimation. The DE optimization process gave $k \approx 2.12$ and $a_0 \approx 24813$. The raw images were then processed with Eq. (2) to produce the r_{0t} images, which then take the position of the raw images. Figure 2 shows an example mammogram in the three representations: the raw image, r_{0r} , (left); raw-transformed image, r_{0t} , (middle); and the processed clinical display, r_{0p} , (right).

The outline in r_{0r} in Figure 2 defines a region of interest (ROI) used for illustration purposes below.

Examples of the breast density processing for PD_a are shown in Figure 3 and PD in Figure 4. Figure 3 (left) is representative of the PD_a output for the raw data without pre-processing indicating that percentage of breast density estimates were not accurate or comparable to those provided by PD, and therefore, not estimated. PD_a on the raw-transformed images was significantly associated with breast cancer (Table 2) in the adjusted [OR for PD_a quartiles: 1.00 (ref.); 1.27 (0.74, 2.19); 1.86 (1.05, 3.28); 3.00 (1.67, 5.38); and $Az = 0.61(0.55, 0.65)$] and continuous [OR: 1.40 (1.14, 1.71) per SD; and $Az = 0.60 (0.55, 0.65)$] models. PD_a on the clinical display images also resulted in significant risk factor (Table 2) and Az estimates in both the adjusted quartile [OR; 1.00 (ref.); 1.79 (1.04, 3.11); 1.61 (0.90, 2.88); 2.94 (1.66, 5.19); and $Az = 0.61 (0.56, 0.65)$] and continuous [OR: 1.43 (1.17, 1.74); and $Az = 0.60 (0.55, 0.65)$] models. The ROC curves for the continuous PD_a models (unadjusted) are also shown in Figure 1. Comparing the inter-measure Az for the continuous models for the PD_a -raw-transformed/PD-raw images gave $p = 0.18 (0.606 \text{ vs. } 0.567)$ and $p = 0.04 (0.603 \text{ vs. } 0.551)$ for the PD_a -processed/PD-processed images indicating PD_a provided marginally greater predictive capability than PD for the processed representation data. The PD_a processing takes approximately 15 seconds per small FOV image and 27 seconds per large FOV image with our server (average times taken over the dataset).

To assess variability in our raw-transform representation findings, we implemented the optimization process described in Section 2.5 multiple times without including the results from above. This gave (2.59, 22395) for the mean, (1.92, 20105) for the minimum, and (3.52, 23439) for the maximum (k , A_0) parameter sets. The PD_a findings from processing the raw images with these additional sets are provided in the Table 1A of the Appendix. The ORs and Az findings are similar to those shown in Table 2 for the raw-transformed images. Although the endpoint associations with breast cancer are similar, the quartile cutoffs vary inducing variation in the number of cases in each quartile.

Comparing the PD_a (applied to either r_{0t} or r_{0p}) automated labeling in Figure 3 with the corresponding PD results in Figure 4 shows the two methods operate differently. As noted within the broad dense regions, PD_a detects (labels) locally, whereas PD appears to use a contoured based threshold (i.e. a more global detection). This operational difference is noted by the salt and pepper appearance in the PD_a image in comparison with the respective broad uniformly dense regions in the PD image. The effect on the PD_a processing due to the different representations is further exemplified in the local-variance images shown in Figure 5, corresponding to the region of interest (ROI) marked in Figure 2 (left image). These ROIs were equalized by adjusting the window levels and widths similarly for comparison (see captions). The local variance in the raw representation (left) shows little contrast. The raw-transformed (middle) and clinical display (left) variance representations show distinct contrast differences, most notable in glandular-adipose transition regions. Intra and inter-measurement correlation analyses for these measures are provided in Table 3. The inter-measure correlation is less than 0.50 when considering the raw/raw-transformed and processed/processed comparisons between PD and PD_a . As expected, the intra-measure analysis provided better agreement when considering either the raw/processed comparison

for PD or the raw-transformed/processed comparison for PD_a. It is interesting to note, PD_a provided increased inter-measure agreement (0.87 compared with 0.73 from PD).

The differences, noted above, are further qualified by the corresponding SDN modeling analysis and representative plots shown in Figure 6. The raw data (left) shows a slight deviation from a linear model, whereas the raw-transformed data (middle) requires a third degree polynomial to capture the trend, and the processed data (right) exhibits a non-monotonic non-linear trend. Summary results (entire dataset) for the SDN models corresponding to each data representation are provided in Table 3. Both the raw and raw-transformed representations share a monotonic quality but differ in character. The linear trend of the raw data can be gauged by comparing the magnitudes of its c_1 (linear) and c_2 (quadratic) coefficient summaries (Table 3), whereas the raw-transformed data exhibits a stronger non-linear or quadratic tendency as noted by comparing its c_2 (quadratic), c_1 (linear) and c_3 (cubic) coefficients. The monotonic and non-linear form of the raw-transformed data is also a distinguishing characteristic. In contrast with the raw and raw-transformed representations, the clinical display representation data (right) exhibits non-linear but non-monotonic behavior as gauged by the relative magnitude of its c_4 coefficient in comparison with its linear and quadratic coefficients (i.e. c_1 and c_2). The non-linear and monotonic increase of the raw-transformed data relationship is distinguishing characteristics of its SDN characteristic. These findings suggest the nonlinear attributes are beneficial for the PD_a processing because they induce contrast in the variance representation images (see Figure 5), which gives the foundation for estimating the reference variance for a given mammogram with improved precision.

4. Discussion

This work provided an initial validation of the PD_a method for FFDM images acquired from Hologic technology. The PD_a findings from both the raw-transformed and processed images are similar to PD in breast cancer association with comparable ORs and greater Az. We, note the raw-transformed images and findings are used in place of the raw mammograms. The results in this report confirm our earlier findings from digitized film (29). However, the OR associations for this report provided by both PD and PD_a appear to be somewhat attenuated relative to other studies of digitized film, whereas the Az relationships are in agreement with reported values. This could be a function of the relatively small sample size or density assessed from digital mammograms.

There are several aspects of our work that merit further comment. The associations provided by PD_a from the raw data required a transformation. The form of this transformation was based on heuristic considerations and may not be optimal, although this transform produced a validated percentage of breast density measure in this report. The work involved training with a relatively small subset of images (30 from 550 images) and the PD_a algorithm is relatively straight forward. Therefore, we would expect the method to perform equally as well under similar circumstances with other datasets (i.e. from the same FFDM technology). There are two adjustable detection parameters, in addition to those parameters defined in Eq. (2), that will require modification when addressing other data representations (e.g. different detector technologies). There are also various methods that could be used to train the PD_a

algorithm each with its own limitations. Ideally, breast cancer status could be used as the endpoint when developing a measure for risk applications because it is the known without ambiguity, eliminating the requirement for matching to another density measure. We did not use breast cancer status as the endpoint for training because it would have required the entire dataset for training (i.e. increase the possibility of over fitting), and we had limited power to do so. Also, it would lead to protracted training times [i.e. using the 30 image dataset requires 18 days to solve Eq. (2)]. Our work did not consider either intra or inter operator variability influences. Considering operator influence may be important when designing a measure such as the BI-RADS breast composition descriptor, as this metric was designed to capture the radiologist's overall impression. As we have shown previously, the association with breast cancer and the overall radiologist's impression endpoints may not be equivalent (44), but are equally important. The (k, a_0) variability analysis shows there is latitude in the data representation indicating that the breast density assignments at the patient level can vary considerably across the dataset while keeping consistent associations with breast cancer. This may be attributable to a number of factors including an isolated reader, the ambiguous nature of breast density, or both. The correlation analysis between PD and our measures was somewhat less than shown previously with digitized-film data (29). These differences could be due to the way the two methods detect breast density, differences in the data representation (i.e. film compared to this form of FFDM), or to the limited number of training samples used in the report. In future work, we will train with the breast cancer endpoint status using the dataset in this report and then validate PD_a with an independent dataset of similar proportion that is under construction to develop a measure for risk applications. We have also developed a PD-type measure (defined as PD_c) using calibrated images from another FFDM technology (42) and are currently translating this work to the technology used in the current report (39). Alternatively, the applicable PD_c findings could be used to train the PD_a algorithm to reduce variation. We also used both clinical display and raw data in the analysis. Arguments against using the clinical display images for risk estimations with automated methods could be made due to their synthetic nature, although they are used clinically. Although we do not know the specific formula used to produce this representation, and it differs by manufacturer, the method of transforming the raw data (i.e. the acquired data) to the processed clinical data may be more involved than a non-linear memory-less mapping [i.e. a transform of the Eq. (2) form is memory-less]. From observation and informal discussions with radiologists at our center, these clinical images appear to have been frequency enhanced as well, which may raise an important concern when using these images for risk assessments. On the other hand, the enhancement is most likely based on the internal characteristics of the respective image. There is work showing consistency across the representations for density measures made semi-objectively when assisted by an operator (35, 45). However for automated techniques, comparisons of breast density measures from clinical display with the raw images will require further evaluation.

Additionally, we provided an analysis based on SDN to explain the PD_a performance. When the Poisson process approximation holds, we expect a linear relationship between the noise and signal as approximated with the raw data (Figure 6 on the left). The findings suggest that a near-linear relationship does not offer sufficient contrast between the glandular variation and that of the adipose variation (Figure 5 on the left) for the PD_a process to

operate effectively. These findings indicate that the non-monotonic characteristic of the clinical display images offers sufficient contrast but dense tissue may be lost on the right side of the crest because the variation from the adipose and glandular noise signals are mirrored across the crest midline. This non-monotonic behavior may be the cause for the inversion (or flat) of the respective second and third quartile ORs (Table 2) derived from PD_a using the clinical display images. It is interesting to note that similar OR inversion findings resulted from PD for the clinical display images, as well. Although there are marginal differences in the PD_a from the raw-transformed and processed images, the raw-transformed representation may be preferable because the quartile ORs monotonically increase as expected (Table 2) and the A_z is larger. This SDN modeling provided a coarse description that helps to understand a rather complicated process. Borrowing from Griffin (46) as justification, the best model may be a coarse approximation when it provides insight into a portion of the process that is both easily interpreted and manipulated.

Although there are many automated methods under development for breast density as discussed previously (4-6, 29), there are few studies showing that these measures show associations with breast cancer at least equivalent to that of PD. Standardized and calibrated approaches have given mixed results in producing a measure that provides equivalent or stronger breast cancer associations than given by PD (17, 20, 42-44, 47, 48). There are commercial products available that use standardized data for estimating breast density from raw FFDM images including studies that show their comparison with the clinical BI-RADS density measure (21, 22) and other risk factors (23). But, to date, there are no published data that demonstrates their association with breast cancer. Finally, we note as a limitation that this study focused only on the development and comparison of automated measures of percent density from FFDM for eventual use in risk models. Future work will focus on how much they add to other clinical risk factors, as well as how well they improve existing breast cancer risk models.

5. Conclusion

Breast density is an important risk factor for breast cancer and will likely play a role in future personalized mammographic screening recommendations (49). As such, there is a need for a reproducible automated density measure. This work we evaluated an automated approach for estimating the percentage of breast density for FFDM applications. We believe automated breast density measurements are likely required for large-scale clinical applications, as to increase clinical-throughput and to maintain measurement consistency. The PD_a measure is a valid candidate for further evaluation for possible inclusion in the clinical environment. This work is timely because film mammography has been essentially supplanted by FFDM. Over 82% of the accredited mammography facilities in the US operate FFDM units (50), and this particular (Hologic) FFDM technology accounts for about 70% of the market share in the US (51). Our work shows that either the raw (after transforming) or clinical display images can be processed with PD_a from this technology with minor loss. The work was performed with one dataset acquired with a specific FFDM technology. Future work includes exploring the possibility there is an optimal data representation for the PD_a measurement and evaluating our techniques with differing FFDM platforms, differing patient populations, and multiple readers.

Acknowledgments

This work was supported by the following National Institutes of Health, National Cancer Institute research grants: R21CA79947, R01CA140286, R01CA128931, R01CA97396, R01CA177150 and P50CA116201.

References

1. Boyd NF, Guo H, Martin LJ, et al. Mammographic density and the risk and detection of breast cancer. *N Engl J Med*. 2007; 356(3):227–36. [PubMed: 17229950]
2. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiol Biomarkers Prev*. 2006; 15(6):1159–69. [PubMed: 16775176]
3. Boyd NF, Martin LJ, Yaffe MJ, Minkin S. Mammographic density and breast cancer risk: current understanding and future prospects. *Breast Cancer Res*. 2011; 13(6):223. [PubMed: 22114898]
4. Yaffe MJ. Mammographic density. Measurement of mammographic density. *Breast Cancer Res*. 2008; 10(3):209. [PubMed: 18598375]
5. Keller BM, Nathan DL, Wang Y, et al. Estimation of breast percent density in raw and processed full field digital mammography images via adaptive fuzzy c-means clustering and support vector machine segmentation. *Medical Physics*. 2012; 39(8):4903–17. [PubMed: 22894417]
6. Manduca A, Carston MJ, Heine JJ, et al. Texture features from mammographic images and risk of breast cancer. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2009; 18(3):837–45.
7. Highnam, R.; Brady, M. *Mammographic Image Analysis*. Boston, MA: Kluwer Academic Publishers; 1999.
8. Pawluczyk O, Augustine BJ, Yaffe MJ, et al. A volumetric method for estimation of breast density on digitized screen-film mammograms. *Med Phys*. 2003; 30(3):352–64. [PubMed: 12674236]
9. Shepherd JA, Herve L, Landau J, Fan B, Kerlikowske K, Cummings SR. Novel use of single X-ray absorptiometry for measuring breast density. *Technol Cancer Res Treat*. 2005; 4(2):173–82. [PubMed: 15773786]
10. Tice JA, Cummings SR, Ziv E, Kerlikowske K. Mammographic breast density and the gail model for breast cancer risk prediction in a screening population. *Breast Cancer Res Treat*. 2005; 94(2): 115–22. [PubMed: 16261410]
11. Highnam R, Pan X, Warren R, Jeffreys M, Davey Smith G, Brady M. Breast composition measurements using retrospective standard mammogram form (SMF). *Phys Med Biol*. 2006; 51(11):2695–713. [PubMed: 16723760]
12. Jeffreys M, Warren R, Highnam R, Smith GD. Initial experiences of using an automated volumetric measure of breast density: the standard mammogram format. *Br J Radiol*. 2006; 79(941):378–82. [PubMed: 16632617]
13. Heine JJ, Thomas JA. Effective x-ray attenuation coefficient measurements from two full field digital mammography systems for data calibration applications. *Biomed Eng Online*. 2008; 7(1): 13. [PubMed: 18373863]
14. Jeffreys M, Warren R, Highnam R, Davey Smith G. Breast cancer risk factors and a novel measure of volumetric breast density: cross-sectional study. *Br J Cancer*. 2008; 98(1):210–6. [PubMed: 18087286]
15. Heine JJ, Cao K, Beam C. Cumulative Sum Quality Control for Calibrated Breast Density Measurements *Medical Physics*. 2009; 36(12):5380–90.
16. Malkov S, Wang J, Kerlikowske K, Cummings SR, Shepherd JA. Single x-ray absorptiometry method for the quantitative mammographic measure of fibroglandular tissue volume. *Med Phys*. 2009; 36(12):5525–36. [PubMed: 20095265]
17. Boyd N, Martin L, Gunasekara A, et al. Mammographic density and breast cancer risk: evaluation of a novel method of measuring breast tissue volumes. *Cancer Epidemiol Biomarkers Prev*. 2009; 18(6):1754–62. [PubMed: 19505909]

18. Heine JJ, Cao K, Thomas JA. Effective radiation attenuation calibration for breast density: compression thickness influences and correction. *BioMedical Engineering OnLine*. 2010; 9:73. [PubMed: 21080916]
19. Highnam, R.; Brady, M.; Yaffe, MJ.; Karsssemeijer, N.; Harvey, JA. Robust Breast Composition Measurement - Volpara™. In: Marti, J., editor. *IWDM 2010*. Springer-Verlag; 2010. p. 342-9.
20. Shepherd JA, Kerlikowske K, Ma L, et al. Volume of mammographic density and risk of breast cancer. *Cancer Epidemiol Biomarkers Prev*. 2011; 20(7):1473–82. [PubMed: 21610220]
21. Ciatto S, Bernardi D, Calabrese M, et al. A first evaluation of breast radiological density assessment by QUANTRA software as compared to visual classification. *Breast*. 2012; 21(4):503–6. [PubMed: 22285387]
22. Highnam, R.; Sauber, N.; Destounis, S.; Harvey, J.; McDonald, D. Breast Density into Clinical Practice. In: Maidment, ADA.; Bakic, PR.; Gavenonis, S., editors. *11th International Workshop of Digital Mammography*. Philadelphia, PA: Springer; 2012. p. 466-73.
23. Skippage P, Wilkinson L, Allen S, Roche N, Dowsett M, A'Hern R. Correlation of age and HRT use with breast density as assessed by Quantra. *The breast journal*. 2013; 19(1):79–86. [PubMed: 23230974]
24. Bick, U.; Diekmann, F. Medical Radiology Diagnostic Imaging and Radiation Oncology. In: Baert, AL.; Brady, LW.; Heilmann, HP.; Hricak, H.; Knauth, M.; Molls, M.; Nieder, C.; Reiser, MF., editors. *Continuation of Handbuch der medizinischen Radiologie Encyclopedia of Medical Radiology*. Berlin: Springer-Verlag; 2010.
25. Kim HK, Cunningham IA, Yin Z, Cho G. On the development of digital radiography detectors : A review. *Int J Precis Eng Man*. 2008; 9(4):86–100.
26. Mahesh M. AAPM/RSNA physics tutorial for residents: digital mammography: an overview. *Radiographics*. 2004; 24(6):1747–60. [PubMed: 15537982]
27. Heine JJ, Deans SR, Velthuisen RP, Clarke LP. On the statistical nature of mammograms. *Medical physics*. 1999; 26(11):2254–65. [PubMed: 10587206]
28. Heine JJ, Velthuisen RP. A statistical methodology for mammographic density detection. *Med Phys*. 2000; 27(12):2644–51. [PubMed: 11190946]
29. Heine JJ, Carston MJ, Scott CG, et al. An automated approach for estimation of breast density. *Cancer Epidemiol Biomarkers Prev*. 2008; 17(11):3090–7. [PubMed: 18990749]
30. Heine JJ, Behera M. Aspects of signal-dependent noise characterization. *J Opt Soc Am A*. 2006; 23(4):806–15.
31. Olson JE, Sellers TA, Scott CG, et al. The influence of mammogram acquisition on the mammographic density and breast cancer association in the mayo mammography health study cohort. *Breast Cancer Res*. 2012; 14(6):R147. [PubMed: 23152984]
32. Heine JJ, Scott CG, Sellers TA, et al. A novel automated mammographic density measure and breast cancer risk. *Journal of the National Cancer Institute*. 2012; 104(13):1028–37. [PubMed: 22761274]
33. Hosmer, DW.; Lemeshow, S. *Applied Logistic Regression Second Edition* ed. Hoboken: John Wiley & Sons Inc.; 2000.
34. Breslow N, Day N. The analysis of case-control studies. *IARC Sci Publ*. 1980; 32:147–78.
35. Vachon CM, Fowler EE, Tiffenberg G, et al. Comparison of percent density from raw and processed full-field digital mammography data. *Breast Cancer Res*. 2013; 15(1):R1. [PubMed: 23289950]
36. Heine JJ, Deans SR, Cullers DK, Stauduhar R, Clarke LP. Multiresolution statistical analysis of high-resolution digital mammograms. *IEEE Trans Med Imaging*. 1997; 16(5):503–15. [PubMed: 9368106]
37. Heine JJ, Deans SR, Cullers DK, Stauduhar R, Clarke LP. Multiresolution probability analysis of gray-scaled images. *Journal of the Optical Society of America A*. 1998; 15(5):1048–58.
38. Heine JJ, Deans SR, Gangadharan D, Clarke LP. Multiresolution analysis of two-dimensional 1/f processes: approximation methods for random variable transformations. *Optical Engineering*. 1999; 38:1505.
39. Fowler EE, Lu B, Heine JJ. A comparison of calibration data from full field digital mammography units for breast density measurements. *Biomed Eng Online*. 2013; 12:114. [PubMed: 24207013]

40. Price, KV.; Storn, RM.; Lampinen, JA. Differential evolution : a practical approach to global optimization. Berlin ; New York: Springer; 2005.
41. Behera M, Fowler EE, Owonikoko TK, et al. Statistical learning methods as a preprocessing step for survival analysis: evaluation of concept using lung cancer data. *BioMedical Engineering Online*. 2011; 10(1):97. [PubMed: 22067671]
42. Heine JJ, Cao K, Rollison DE, Tiffenberg G, Thomas JA. A Quantitative Description of the Percentage of Breast Density Measurement Using Full-field Digital Mammography. *Acad Radiol*. 2011; 18(5):556–64. [PubMed: 21474058]
43. Heine JJ, Cao K, Rollison DE. Calibrated measures for breast density estimation. *Acad Radiol*. 2011; 18(5):547–55. [PubMed: 21371912]
44. Fowler EE, Sellers TA, Lu B, Heine JJ. Breast Imaging Reporting and Data System (BI-RADS) breast composition descriptors: Automated measurement development for full field digital mammography. *Medical Physics*. 2013; 40(11):113502. [PubMed: 24320473]
45. Keller BM, Nathan DL, Gavenonis SC, Chen J, Conant EF, Kontos D. Reader variability in breast density estimation from full-field digital mammograms: the effect of image postprocessing on relative and absolute measures. *Acad Radiol*. 2013; 20(5):560–8. [PubMed: 23465381]
46. Griffin, WC. Transform Techniques for Probability Modeling. New York: Academic Press; 1975.
47. Ding J, Warren R, Warsi I, et al. Evaluating the effectiveness of using standard mammogram form to predict breast cancer risk: case-control study. *Cancer Epidemiol Biomarkers Prev*. 2008; 17(5): 1074–81. [PubMed: 18483328]
48. Heine JJ, Fowler EE, Flowers CI. Full field digital mammography and breast density: comparison of calibrated and noncalibrated measurements. *Academic radiology*. 2011; 18(11):1430–6. [PubMed: 21971260]
49. Schousboe JT, Kerlikowske K, Loh A, Cummings SR. Personalizing mammography by breast density and other risk factors for breast cancer: analysis of health benefits and cost-effectiveness. *Annals of internal medicine*. 2011; 155(1):10–20. [PubMed: 21727289]
50. FDA. MQSA National Statistics. Available at: <http://www.fda.gov/Radiation-EmittingProducts/MammographyQualityStandardsActandProgram/FacilityScorecard/ucm113858.htm>
51. Full-Field Digital Mammography Equipment - Global Pipeline Analysis, Competitive Landscape and Market Forecasts to 2017. MarketResearch.com. Available at: F:\Full-Field Digital Mammography Equipment - Global Pipeline Analysis, Competitive Landscape and Market Forecasts to 2017 by GlobalData in Mammography.mht.

Figure 1.

Breast density measurement receiver operating curve analysis. This shows sensitivity and 1 – specificity for unadjusted continuous density measures with breast cancer. For the raw-transformed PD_a $Az = 0.606$ (top-left), raw PD $Az = 0.567$ (top-right), processed PD_a $Az = 0.603$ (bottom-left), and processed PD $Az = 0.551$ (bottom-right). The ROC curves are bolder than the no-discrimination line.

Figure 2. Mammogram representation example. This shows one mammogram in three representations: raw (left), raw-transformed (middle), and processed clinical display (right). The raw image has inverted pixel values relative to mammograms used for clinical purposes (i.e. adipose tissue is bright and glandular tissue dark). The rectangular region of interest defined on the raw image is referenced in subsequent developments.

Figure 3.

Automated Percentage of breast density labeling examples. The automated percentage of breast density measure (PD_a) results for the images shown in Figure 2 are provided. From left to right, this shows the raw (left), the raw-transformed (middle), and clinical display images (right). The respective breast density measurements were: (i) not estimated; (ii) 25.5%; (iii) and 33.4%

Figure 4.

Operator-assisted percentage of breast density measure labeling examples. The operator-assisted breast density measure (PD) results for the raw (left) and clinical display (right) images shown in Figure 2 are provided. The respective breast density measurements were: (i) 19.8%; (ii) and 13.8%.

Figure 5.

Local variance region of interest. This shows the local variance region of interest image corresponding to the region marked in Figure 2 for the raw (left), raw-transformed (middle) and clinical display (right) representations. To make valid comparisons, the window level (WL) for each region is the respective median pixel value (skewed distributions) and the window-width was 256 gray values centered about the WL.

Figure 6.

Signal Dependent Noise Analysis Example. This illustrates the signal dependent noise relationships for the images shown in Figure 2 fitted with Equation 5 for the raw (left), raw-transformed (middle), and processed (right) representations. The fitted curves are shown with a solid line and dots represent the measured data. Ordered pairs have been suppressed to better show the trends. From left to right, the fitted coefficients are: $(c_0, c_1, c_2) \approx (0.209, 0.885, -0.133)$, $(c_0, c_1, c_2, c_3) \approx (0.016, 0.094, 0.607, 0.483)$, and $(c_0, c_1, c_2, c_5, c_6) \approx (-0.195, 3.92, -3.59, -0.050, 0.999, 0.002)$.

Table 1

Patient Characteristics: This table lists the distribution of relevant patient characteristics (variable) and breast density measures for the cases, controls, and overall. The breast density measures include: (1) the automated percentage of breast density measure (PD_a) applied to the raw-transformed images (trans); (2) PD_a applied to the clinical display processed images (proc); (3) operator assisted percentage of breast density measure (PD) applied to the raw images (raw); and (4) PD applied to the clinical display processed images. The mean and standard deviation (SD) are provided for each characteristic. The respective case-control quantities were compared using conditional logistic regression (Wald test).

variable	Case (n)	mean	SD	Control (n)	mean	SD	Total (n)	mean	std	P
Age (yrs.)	192	64.2	10.6	358	64.3	10.6	550	64.3	10.7	--
BMI (kg/m ²)	188	29.0	6.4	335	28.8	6.2	523	28.9	6.3	0.81
PD _a (trans)	192	21.0	7.3	358	19.1	7.3	550	19.8	7.4	0.002
PD _a (proc)	192	19.1	7.9	358	16.9	7.5	550	17.6	7.7	0.0005
PD (raw)	192	15.0	12.1	358	13.6	12.5	550	14.1	12.4	0.17
PD (proc)	192	18.1	10.3	358	16.9	10.0	550	17.4	10.1	0.15

Table 2

Percentage of breast density associations with breast cancer: This table provides the breast cancer quartile and continuous breast density associations with breast cancer for (1) the automated measure (PD_a) applied to the raw-transformed (top-left) and processed clinical display (bottom-left) representation images; and (2) the operator-assisted measure (PD) applied to the raw (top-right) and processed clinical display (bottom-right) representations images. Odds ratios (ORs) are cited with 95% confidence intervals (CIs) parenthetically and the area under the receiver operating characteristic curve (Az) is provided for each model with 95% CIs. CIs are cited below the respective quantities, parenthetically. Az was calculated within matched case-control pairs to utilize the design. SD is calculated from the control distribution. The quartile cutoff for each measure is also provided in the left hand column of each sub-table.

PD _a Raw Transformed	Control N	Case N	unadjusted	adjusted with BMI	PD Raw	Control N	Case N	unadjusted	adjusted with BMI
Quartile 1 [3.98, 13.41]	89	30	1.00	1.00	Quartile 1 [0.00, 4.81]	89	35	1.00	1.00
Quartile 2 [13.41, 18.14]	90	40	1.26 (0.73, 2.18)	1.27 (0.74, 2.19)	Quartile 2 [4.81, 10.09]	90	50	1.48 (0.88, 2.49)	1.59 (0.93, 2.70)
Quartile 3 [18.14, 23.82]	89	50	1.85 (1.05, 3.26)	1.86 (1.05, 3.28)	Quartile 3 [10.09, 18.69]	89	49	1.46 (0.85, 2.51)	1.70 (0.95, 3.04)
Quartile 4 [23.82, 38.47]	90	72	2.93 (1.64, 5.22)	3.00 (1.67, 5.38)	Quartile 4 [18.69, 76.84]	90	58	1.70 (1.00, 2.87)	2.04 (1.13, 3.67)
Az			0.588	0.595	Az			0.554	0.556
Per 1 SD increase			1.39 (1.13, 1.70)	1.40 (1.14, 1.71)	Per 1 SD increase			1.14 (0.94, 1.39)	1.21 (0.97, 1.51)
Az			0.580	0.579	Az			0.547	0.549
PD _a Processed	Control N	Case N	unadjusted	adjusted with BMI	PD Processed	Control N	Case N	unadjusted	adjusted with BMI
Quartile 1 [2.99, 10.72]	89	27	1.00	1.00	Quartile 1 [1.11, 9.63]	89	39	1.00	1.00
Quartile 2 [10.72, 15.80]	90	50	1.80 (1.04, 3.12)	1.79 (1.04, 3.11)	Quartile 2 [9.63, 15.07]	90	51	1.28 (0.77, 2.13)	1.31 (0.79, 2.18)
Quartile 3 [15.80, 22.14]	89	44	1.61 (0.90, 2.88)	1.61 (0.90, 2.88)	Quartile 3 [15.07, 21.39]	89	40	1.04 (0.61, 1.76)	1.14 (0.65, 1.98)
Quartile 4 [22.14, 35.37]	90	71	2.89 (1.64, 5.09)	2.94 (1.66, 5.19)	Quartile 4 [21.39, 67.10]	90	62	1.68 (1.00, 2.83)	1.95 (1.09, 3.47)
Az			0.588	0.591	Az			0.551	0.561
Per 1 SD increase			1.42 (1.16, 1.72)	1.43 (1.17, 1.74)	Per 1 SD increase			1.15 (0.95, 1.38)	1.22 (0.98, 1.51)
Az			0.581	0.581	Az			0.534	0.552

Table 3

Correlation Coefficients. The table lists the inter and intra measure Pearson correlation coefficients for PD and PD_a. PD was applied to the raw and clinical display processed (proc) images, whereas PD_a was applied to the raw-transformed (raw-trans) and processed (proc) images.

Measurement	PD raw	PD proc	PD _a raw-trans	PD _a proc
PD raw	1.00	0.73	0.37	0.46
PD proc	0.73	1.00	0.38	0.43
PD _a raw-trans	0.37	0.38	1.00	0.87
PD _a proc	0.46	0.43	0.87	1.00

Table 4

Signal Dependent Noise Model Coefficients: This table provides the summarized signal dependent noise modeling coefficients for the three data representations from modeled with Eq. (3): raw; raw-transformed (raw-trans); and processed clinical display (proc). For each coefficient, the mean over the entire dataset is cited. 95% confidence intervals are provided beneath each quantity parenthetically. Terms not included in the modeling are marked with dashes in the respective columns.

Coefficients	Image representation		
	raw	raw-trans	proc
c_0	0.140 (0.130, 0.150)	0.019 (0.017, 0.021)	-0.111 (-0.116, -0.106)
c_1	0.455 (0.410, 0.501)	0.274 (0.255, 0.293)	3.001 (2.926, 3.076)
c_2	0.131 (0.098, 0.164)	0.339 (0.265, 0.414)	-2.265 (-2.390, -2.141)
c_3	-	0.106 (0.047, 0.165)	-
c_4	-	-	-0.433 (-0.538, -0.328)
c_5	-	-	1.038 (1.016, 1.060)
c_6	-	-	0.083 (0.064, 0.102)