Should volumetric breast density be included in breast cancer prediction models? Proposal of an integrated quantitative and reproducible approach
[Studio Via San Giorgio; Veneto Institute of Oncology] Ciatti, S., Gennaro, G., Mungai, V. and Nannini, G.

SS 302b – Risk imaging and stratification; Wednesday March 4, 3.20 pm
Location: Room C

This study investigated the relationship between volumetric breast density (VBD) and breast cancer risk estimates using the Tyrer-Cuzick model. For 249 patients undergoing mammographic screening between January-July 2014, Tyrer –Cuzick lifetime risk estimates and Volpara Density Grades (VDG) based on VBD cut-off thresholds were obtained (i.e. VDG1 0-4.5%; VDG 2 4.5-7.5%; VDG3 7.5-15.5%; VDG4 >15.5%). Median lifetime risk estimates comparing VDG2 (11.0%) versus VDG3 (14.5%) and VDG4 (15.6%) were significantly different (p=0.0011 and p=0.0002, respectively). Lifetime risks were comparable between VDG3 and VDG4 categories (p=0.0931). Lifetime risk increased with increasing breast density, indicating that volumetric breast density measures could be used with existing risk prediction models to more accurately identify high-risk women.

Effect of volumetric mammographic density on performance of a breast cancer screening program using full-field digital mammography

SS 1802 – Population-based screening; Sunday March 8, 11.03 am
Location: Room C

This study looked at mammographic screening performance measures stratified according to volumetric breast density software (VolparaDensity). 69, 874 studies from women aged 50-75 undergoing biennial mammographic screening in the Dutch screening program were included in the analyses. Screening performance measures were obtained from the screening registration system and interval cancers identified through linkage with the Netherlands Cancer Registry. The distribution of women in each BI-RADS density categories was 19.7% (BI-RADS 1), 43.1% (BI-RADS 2), 29.4% (BI-RADS 3), and 7.7% (BI-RADS 4). For the 421 screen-detected and 150 interval cancers, cancer detection rates were 3.7%, 6.4%, 6.6%, and 6.3% in categories 1 through 4, respectively. The interval cancer rate increased with increasing breast density, with 0.7%, 1.9%, 3.0% and 4.5% in categories 1 through 4, respectively. Hence, sensitivity decreased as breast density increased, with 85%, 77.6%, 69.0% and 58.6% of cancer being detected by mammography across categories 1 through 4, respectively. The number of false-positives was also higher with increasing breast density (11.4%, 14.1%, 18.3% and 28.6% in categories 1-4, respectively). Breast cancer screening performance was significantly lower for women with denser breasts.
Volumetric breast density and BI-RADS 5th edition

Institute Jules Bordet
Grivegnée, A. MD, PhD

Link to EPOS poster: http://dx.doi.org/10.1594/ecr2015/C-1070

This educational exhibit looked at how the recent fifth edition updates to the American College of Radiology’s (ACR) Breast Imaging Reporting and Data System (BI-RADS) Atlas should impact the use of automated volumetric breast density software (VolparaDensity) at Institute Jules Bordet. The author assessed several changes and found: 1) the software can already be configured to output letters (a/b/c/d) rather than numbers (1/2/3/4) to denote the four BI-RADS density categories; 2) removal of the quantitative quartile ranges should not impact visual density assessment; and 3) breasts that are predominantly fatty but still contain regions in the breast that are sufficiently dense to obscure small masses (i.e. focal densities) tend to be scored as BI-RADS c by the automated software, in accordance with the ACR recommendations. Furthermore, use of the denser breast to assign the final density category (rather than the average) would align better with the ACR recommendations, and implementing this change into the software would result in <5% of women being re-classified from non-dense (BI-RADS 1 or 2) to dense (BI-RADS 3 or 4). Volumetric approaches are entirely consistent with the changes in the BI-RADS 5th edition provided that the software is configured to use the denser breast.

Impact of quantitative breast density on experienced radiologists’ assessment of mammographic breast density

Boca Raton Regional Hospital
Schilling, K., The, J., Griff, S., Oliver, L., Mahal, R., Saady, M., and Velasquez, M.V.

Link to EPOS poster: http://dx.doi.org/10.1594/ecr2015/C-1281

This study looked at whether the use of quantitative breast density software improved the consistency in breast density assessment between radiologists. In this study, eight readers assessed 88 digital mammographic studies and assigned each study into a BI-RADS density category. After a 2-week washout period, the radiologists re-read the 88 studies using the quantitative breast density software (VolparaDensity) as an interpretive aid. The use of VolparaDensity reduced the variability in the number of women allocated into each density category and significantly improved the interobserver agreement in radiologists’ assessment of BI-RADS (p=0.0374), with a mean kappa statistic of 0.5664 and 0.6266 without or with VolparaDensity, respectively. Most readers in this study accepted and used the automated scores to improve their readings.

Impact of objective volumetric breast density estimates on mean glandular dose calculations in digital mammography

Veneto Institute of Oncology; Emory University School of Medicine; Castelfranco Veneto Hospital; “Città della Salute” Hospital; Volpara Solutions
Gennaro, G., Sechopoulos, I., Gallo, L., Rossetti, V., and Highnam, R.

Link to EPOS poster: http://dx.doi.org/10.1594/ecr2015/C-1576

This study compared mean glandular dose (MGD) estimation using Volpara’s breast glandularity output (and David Dance’s model) with the MGD output by various X-ray system manufacturers. MGD estimates were obtained for 5076 patients who underwent digital mammography on GE DS, GE Essential, Hologic Selenia Dimension, IMS Giotto Image 3DL or Philips Microdose systems. The median MGD calculated by Volpara (1.28 mGy) was significantly higher (p<0.0001) compared to the median manufacturer MGD (1.18 mGy). Comparing individual systems, Volpara’s MGD was significantly higher compared to the MGD output by the GE DS, GE Essential, IMS Giotto, and one of the Hologic Selenia Dimension. Volpara’s MGD was significantly lower compared to the Philips Microdose and one of the Hologic Selenia Dimension systems. MGD values provided by manufacturers using different models and different assumptions of breast density are significantly different to MGD estimates obtained using the patient’s own breast density.
This study aimed to evaluate potential risk factors for the development of breast cancer, which may be more easily identified at the
time of the first (prevalent round) screen (i.e. breast density and history of high risk or atypical lesions) compared to other known risk
factors for breast cancer (e.g. parity, age at menarche, hormone replacement therapy use etc.). Women (n = 39,491) with raw
mammograms who presented for a screening mammogram between April 2013 and March 2014 were included in this study. Women
were categorized into either the prevalent or incident screening group. The number of biopsies and pathology results were obtained,
as well as volumetric breast density (VBD), as assessed by VolparaDensity software. A higher rate of B3 lesions (i.e. lesions of uncertain
malignant potential) with atypia was observed in the prevalent screening group compared to the incident screening group (0.19% and
0.09%, respectively; p=0.019). While women with cancer at the prevalent screen had higher mean VBD compared to other prevalent
screen women, women with cancer detected at incident screening had similar breast density compared to their peers. The prevalent
screen offers a particularly key opportunity to select women for more careful scrutiny over time and to educate women on their risk.

[Abstract Not Available]

Breast density across a screening population

This study looked at the variation in breast density across a screening population and the relationship of breast density with several
demographic factors. Screening studies obtained between March 2013 – September 2014, and which had raw mammograms
available, were processed using Volpara Density software. Breast volume (cm³), fibroglandular tissue volume (FGV; cm³), volumetric
breast density (VBD; %), and Volpara Density Grade (VDG) were analysed in respect to age, self-reported ethnicity, and socioeconomic
status (using a 5-point deprivation index 1-5, with 1 being the most deprived and 5 being the least deprived). The authors observed
little difference in breast volume with respect to age, particularly in the postmenopausal age group. As expected FGV and volumetric
breast density were negatively associated with age. The decline in FGV was most prominent prior to age 56, and may be attributed to
the perimenopausal/postmenopausal divide. Although the proportion of women with dense breasts is lower in older versus younger
women, there were women with very high breast density across all age groups. A relationship between breast density and ethnicity
was also observed, with Chinese women having the highest breast density and smallest breast volume, and Black women having
similar VBD to other ethnic groups and the widest variation in breast volume. Women in the most deprived quintile (deprivation index
= 1), tended to have larger and less dense breasts, aligning with previous studies that have shown a positive association between
socioeconomic status and breast density (due to a tendency for lower BMI in more affluent populations). Demographic factors may
impact breast density and are relevant for identification of future breast cancer risk.

[Abstract Not Available]
Robustness of automated volumetric breast density estimation for assessing temporal changes in breast density  
[Volpara Solutions] Wang, K., Chan, A., and Highnam, R.

Link to EPOS poster: http://dx.doi.org/10.1594/ecr2015/C-0737

This study evaluated the performance of automated volumetric breast density methods using serial mammograms from breast cancer patients. 112 women with raw digital mammograms from at least two screening mammography examinations were included in the study (i.e. 375 temporal pairs of screening examinations). Images were processed using VolparaResearch software to obtain BI-RADS scores and volumetric breast densities (VBD). As expected, an analysis of the changes in BI-RADS across each temporal series of mammograms indicated a trend in decreasing breast density. Aligning with previously published data, a comparison of the two most recent sequential screening examinations demonstrated that 3.2% of cases increased from non-dense to dense, 8.6% decreased from dense to non-dense, and 88% of cases did not change in density status. Further analysis found that including or excluding the breast side where the cancer was located did not significantly impact volumetric breast density assessments. Volpara demonstrated excellent temporal consistency and reproducibility in women undergoing screening and was robust to the presence of breast lesions.

Mammographic compression – a need for mechanical standardization  

Link to EPOS poster: http://dx.doi.org/10.1594/ecr2015/C-2052

This study compared current mammographic compression practices (i.e. compression force (daN) and compression pressure (kPa)) between one imaging facility in the United States (U.S.) with two sites in the Netherlands. The Dutch dataset comprised of 37,518 screening mammographic images (from 9,188 women) obtained from the Dutch national breast cancer screening program. The U.S. dataset comprised 7,171 screening and diagnostic mammographic images (from 1,851 women). All images were processed using VolparaAnalytics and VolparaDensity software, so that average force, pressure, breast thickness, breast volume, volumetric breast density and average glandular dose could be compared between populations. Large variation in compression was observed for the Dutch and U.S. datasets, with relative standard deviations of 19.6% and 41.9% for force and 43.1% and 50.6% for pressure, respectively. Significant differences (p<0.001) were observed for the mean forces and pressures used in both the Dutch and U.S. datasets (i.e. 13.8 daN versus 7.4 daN and 13.7 kPa versus 8.1 kPa, respectively). As a result of the lack of standardization, many extreme pressures were observed in both datasets, resulting in an inability to predict the compression pressure for a given individual. Standardization of compression has several benefits, including: 1) improving reproducibility; 2) minimizing unnecessary pain due to high pressure outliers; 3) reduce radiation dose and the risk of insufficient image quality due to very low pressure outliers; and 4) enhance quality control of mammographic compression procedures.

[Abstract Not Available]
Impact of compressed breast thickness on detectability of simulated lesions: a clinical trial


Link to EPOS poster: http://dx.doi.org/10.1594/ecr2015/C-1175

This study looked at the effects of compressed breast thickness on lesion detectability. Simulated lesions (microcalcification clusters and masses) were simulated into half of the randomly selected cranio-caudal lesion-free images (65/130), for each of the four thickness groups (T1<29 mm; T2=30-49 mm; T3=50-69 mm; T4>70 mm). Insertions were made into images with the same BIRADS score and at locations with the same Volpara density, so that the influence of breast thickness on lesion detectability could be isolated. Four radiologists performed a free search study using a five-point rating scale and ROC analysis was carried out. From two readers, the area-under-curve (AUC) was 0.85, 0.77, 0.64, and 0.67 for T1 to T4, respectively, indicating that lesion detectability decreased with increasing breast thickness.
Title: Should volumetric breast density be included in breast cancer prediction models? Proposal of an integrated quantitative and reproducible approach

Presenter: Ciatti, S.
Co-authors: Gennaro, G., Mungai, V. and Nannini, G.

PURPOSE
To analyze the relationship between volumetric breast density (VBD) and risk for breast cancer as estimated by prediction models.

MATERIALS AND METHODS
The study included 249 patients who underwent CR mammography in four views (RCC, LCC, RMLO, LMLO) between Jan 2014 and Jul 2014 within a screening program. For each patient the individual risk profile was determined using the Tyrer-Cuzick model, counting for familial and personal factors. A VBD value was computed from each mammogram (Volpara software), and averaged among the four views to obtain the mean VBD per patient. Differences in lifetime risk distributions for four groups of patients with increasing breast density (VG1: 0% - 4.5%; VG2: 4.5% - 7.5%; VG3: 7.5% – 15.5%; VG4 >15.5%) were compared.

RESULTS
The overall median VBD was 10.9%, ranging between 4.6% and 30%. There was no case in VG1, 58 cases in VG2 (median VBD: 6.0%), 135 cases in VG3 (median VBD: 10.35%), and 56 cases in VG4 (median VBD: 18.2%). The median lifetime risk was 11.0% for VG2, 14.5% for VG3, and 15.6% for VG4. Differences in lifetime risk between patients in VG2 and patients in VG3 and VG4 were significant (P-values equal to 0.0011 and 0.0002, respectively), while risk was comparable for patients in VG3 and VG4 (P = 0.0931).

CONCLUSION
Lifetime risk increases with breast density. Volumetric mammographic density measure might be used with existing risk prediction models to identify high-risk women more precisely.
Title: Effect of volumetric mammographic density on performance of a breast cancer screening program using full-field digital mammography

Presenter: Wanders, J.O.  

PURPOSE
We examined to what extent mammographic density affects screening performance when using full field digital mammography (FFDM).

METHODS AND MATERIALS
We collected a consecutive series of 69,874 FFDM examinations (2003-2009) from one screening unit of the Dutch biennial screening program (50-75 years). Volumetric mammographic density was automatically assessed with Volpara version 1.5.0 (Matakina, New Zealand). Recall and breast cancer detection information was obtained from the screening registration system. Interval cancers were identified through linkage with the Netherlands Cancer Registry. Within four density categories, comparable to ACR breast density categories, we determined screening performance measures and linear trends with a Chi Square linear trend test.

RESULTS
19.7% of the examinations was categorized as density category 1 (‘almost entirely fatty’), 43.1% as category 2, 29.4% as category 3 and 7.7% as category 4 (‘extremely dense’). In total 421 screen-detected and 150 interval tumors were identified. Cancer detection rates were 3.7‰, 6.4‰, 6.6‰ and 6.3‰ in categories 1 to 4 respectively (p=0.005). Interval cancer rates increased with increasing density categories: 0.7‰, 1.9‰, 3.0‰ and 4.5‰, respectively (p<0.001). As a result, the sensitivity (proportion of screen-detected tumors of screen-detected and interval tumors) was lower in higher density categories: 85.0%, 77.6%, 69.0% and 58.6% respectively (p<0.001). The number of false-positives was higher in women with dense breasts: 11.4‰, 14.1‰, 18.3‰ and 28.6‰ for categories 1 to 4, respectively (p<0.001).

CONCLUSION(S)
Also when FFDM is used in breast cancer screening higher interval cancer and false-positive rates are observed in women with mammographically dense breasts.
LEARNING OBJECTIVES
The 5th Edition of the BI-RADS Atlas was recently published. We consider how changes in breast composition assessment may affect automated, volumetric breast density (VBD) assessment.

BACKGROUND
The 5th Edition states that breast composition is a volumetric assessment of dense tissue, but also made several other changes, including: C1) using letters rather than numbers for the four breast composition categories; C2) removal of the quantitative area-based percentage ranges; C3) instructing that the denser breast be used for the final density assessment; and C4) suggesting increasing the density category for “breasts containing coalescent areas of fibroglandular tissue that are sufficiently dense to obscure small masses”.

Due to the improved reproducibility compared to visual readers, our clinic uses automated volumetric software, and it was important to understand how the changes should impact our BI-RADS density reading.

FINDINGS AND PROCEDURE DETAILS
C1 is simple to conform to, and the Atlas makes it clear that it does not expect C2 to have any impact. For C3, we found this would result in 4% of women being reclassified from “fatty” (BI-RADS a/b), to “dense” (BI-RADS c/d) based on our own data. For C4, we have noted that volumetric methods inherently tend to read high for “focal densities”. In this poster, we show six cases comprising largely fatty breasts with “focal densities” that the software already calls “dense”.

CONCLUSION
Volumetric approaches to breast density assessment are entirely consistent with the changes in the BI-RADS 5th edition, but the maximum (Left/Right) needs to be used for the final density score.
Title: Impact of quantitative breast density on experienced radiologists’ assessment of mammographic breast density

Co-authors: Schilling, K., The, J., Griff, S., Oliver, L., Mahal, R., Saady, M., and Velasquez, M.V.

AIMS AND OBJECTIVES
There is considerable interobserver variability in radiologists’ determination of mammographic breast density, an important breast cancer risk factor. We sought to study whether the use of quantitative volumetric breast density (VBD) software improves interobserver agreement.

METHODS AND MATERIALS
8 experienced breast imaging radiologists from one facility visually assigned 88 digital screening mammograms into one of four BI-RADS breast composition categories. Quantitative software (VolparaDensity™ v1.4, Matakina Technology, NZ) assigned an equivalent Volpara Density Grade (VDG) based on mapping of the VBD to BI-RADS. After a 2-week washout period, all radiologists re-assessed breast density with the assistance of VDG. Cohen’s kappa (k) coefficient was computed as a measure of interobserver agreement for each pair of radiologists separately, without and with VDG. Bootstrapping the mean difference in k without and with VDG was used to assess the superiority of interobserver agreement using quantitative VBD software.

RESULTS
The use of quantitative VBD significantly improved interobserver agreement in radiologists’ assessment of breast density (p=0.0426) with a mean k without VDG of 0.625 and with VDG of 0.688. Comparing radiologists’ assessments of low density (BI-RADS 1 or 2) versus high density (BI-RADS 3 or 4) without and with VDG showed a trend of improvement in interobserver agreement (p=0.180) with a mean k of 0.712 and 0.758, respectively.

CONCLUSION
Quantitative breast density software significantly improved interobserver agreement of experienced radiologists’ assessment of mammographic breast density. Consistency in the determination of breast density is important for clinical decision-making regarding breast cancer risk assessment and adjunctive imaging in women with dense breasts.
Title: Impact of objective volumetric breast density estimates on mean glandular dose calculations in digital mammography

Co-authors: Gennaro, G., Sechopoulos, I., Gallo, L., Rossetti, V., and Highnam, R.

PURPOSE
To evaluate the impact of volumetric breast density (VBD) estimates obtained from digital mammograms using automated software on the calculation of mean glandular dose (MGD).

METHODS AND MATERIALS
5076 patients who underwent digital mammography with one of five different systems (GE DS, GE Essential, Hologic Selenia Dimensions, IMS Giotto Image 3DL, Philips Microdose) were included in the study. VBD was calculated for 20247 images by Volpara 1.5.0 and those values were used to estimate mean glandular dose (Volpara MGD) according to the model proposed by Dance. MGDs were compared with those provided by each system (System MGD). A Wilcoxon test for paired samples was applied to each pair of medians; a P-value < 0.05 was considered statistically significant.

RESULTS
The mean compressed breast thickness was (53.4 ± 14.0) mm, while the overall volumetric breast density ranged between 1.6% and 52.5% with a median value of 8.7%. Median System MGD was 1.18 mGy (range 0.34- 6.70 mGy), while median Volpara MGD was 1.28 mGy (range 0.24-7.07 mGy), significantly higher (P< 0.0001). The absolute difference between the System MGD and Volpara MGD was mostly comprised between - 0.53 mGy and + 0.39 mGy, leading to a mean relative difference of -1.7% (range -41.3% - 44.6%). Median System MGD values were significantly lower than median Volpara MGD values for all five systems individually.

CONCLUSION
Dose values provided by manufacturers, using different models and assuming breast density a priori, are significantly different compared to those obtained after accounting for real differences among breast densities.

Volpara Solutions comment
The data presented in the final EPOS poster indicates that median system MGD values were significantly lower than median Volpara MDG values for all X-ray systems except for the Philips Microdose and one of the Hologic Selenia Dimensions.
Title: Robustness of automated volumetric breast density estimation for assessing temporal changes in breast density

Co-authors: Wang, K., Chan, A., and Highnam, R.

AIM AND OBJECTIVE
To evaluate the performance of an automated volumetric breast density (VDG) method, on serial screening mammograms from breast cancer patients.

METHODS AND MATERIALS
Raw mammograms were acquired from 112 women breast cancer cases, with at least two screening mammography examinations. Images were processed using VolparaResearch (Matakina Technology, New Zealand), a volumetric breast density method, which outputs a Volpara Density Grade (VDG; analogous to the BI-RADS density categories) and VDG percent.

The average VDG changes across each temporal series of mammograms was determined for each individual woman, as well as across the whole population. In addition, the temporal stability of Volpara was evaluated by comparing VDG scores of the most recent screen to the prior screen (4 categories or dichotomized into “dense” or “non-dense”). Reproducibility was assessed using weighted kappa statistics and Pearson Correlation Coefficient (PCC).

For screen-detected cases (57 cases), the consistency and robustness of Volpara was further investigated by comparing the average VDG per case, including or excluding the cancerous breast.

RESULTS
Volpara demonstrated stable VDG (kappa = 0.854; p = 0.44) and VDG (PCC = 0.936; p = 0.97) between two consecutive exams. Analyzed as dichotomous groups, 3.2% of cases increased from fatty to dense, 8.6% decreased from dense to fatty, and 88% of cases did not change density status. Moreover, VDG estimation was minimally affected by whether the cancerous breast was included or not.

CONCLUSION
Volpara demonstrated excellent temporal consistency and reproducibility in women undergoing screening over time, and the assessment of breast density was robust to the presence of breast lesions.
**Title:** Impact of compressed breast thickness on detectability of simulated lesions: a clinical trial

**Co-authors:** Salvagnini, E., Bosmans, H., Van Ongeval, C., Van Steen, A., Michilesen, K., Cockmartin, L., Struelens, L., and Marshall, N.W.

**PURPOSE**
To investigate lesion detectability as a function of compressed breast thickness in a clinical trial.

**METHOD AND MATERIALS**
Screening patient data, acquired on two identical digital mammography systems, were collected and divided into 4 thickness groups (T1<29mm; T2=30-49mm; T3=50-69mm; T4>70mm). For each group, 130 cranio-caudal lesion-free images were randomly selected. BI-RADS density scores and Volpara density maps were obtained for each selected image. Simulated microcalcification clusters and masses were inserted into half of the images with at least one lesion per image. Each lesion was included in each of the four thickness groups; insertions were made into images with the same BI-RADS score, at a location with the same Volpara density - crucial for isolating the influence of breast thickness on lesion detectability. Four radiologists performed a free search study using a five-point rating scale and inferred ROC analysis applied.

**RESULTS**
Results are currently available for two readers. In going from T1 to T4, the area-under-curve (AUC) was 0.85, 0.77, 0.64 and 0.67, respectively, indicating a decrease in detectability with increasing breast thickness. Comparison of AUC for T3-T4 showed no significant difference (p=0.49) while all the other differences were significant (p<0.05).

**CONCLUSION**
These clinical data confirm results found in simple QA results for many mammography systems - that detectability falls as breast thickness increases, especially for compressed breast thickness above 49mm. Investigation of alternative AEC set-ups in digital mammography is advised.