Comparison of breast density assessment between human eye and automated software on digital and synthetic mammography: Impact on breast cancer risk

M. Le Boulc'h, A. Bekhouche, E. Kermarrec, A. Milon, C. Abdel Wahab, S. Zilberman, N. Chabbert-Buffet, I. Thomassin-Naggara

Comparing visual and automated breast density assessment with on mammography. Impact on breast cancer risk.

Short title : Human and AI assessment of breast density

ABSTRACT

Purpose: To evaluate the agreement between automatic assessment software of breast density based on artificial intelligence (AI) and visual assessment by a senior and junior radiologist, as well as the impact on the assessment of breast cancer risk (BCR) risk at 5 years.

Materials and methods: We retrospectively included 311 consecutive women (40-74 years) without a personal history of breast cancer referred who underwent routine mammography between January 1, 2019 and February 28, 2019. Mammographic breast density (MBD) was independently evaluated by a junior and senior reader on digital mammography (DM) and synthetic mammography (SM) using BI-RADS (5th edition) and by an AI model. For each MBD, BCR at 5 years was estimated per woman by the AI model. Interobserver agreement for MBD between the two readers and the AI software were evaluated by quadratic κ coefficients. Reproducibility of BCR was assessed by intraclass correlation coefficient (ICC).

Results: Agreement for MBD assessment on DM and SM was almost perfect between senior and junior radiologists (κ=0.88 [95% CI: 0.84-0.92] and κ=0.86 [95% CI: 0.82-0.90], respectively) and substantial between the senior radiologist and AI (κ=0.79; 95% CI: 0.73-0.84). There was substantial agreement between DM and SM for the senior radiologist (κ=0.79; 95% CI: 0.74-0.84). BCR evaluation at 5 years was highly reproducible between the

two radiologists on DM and SM (ICC=0.98 [95% CI: 0.97-0.98] for both), between BCR evaluation based on DM and SM evaluated by the senior (ICC=0.96; 95% CI: 0.95-0.97) or junior radiologist (ICC=0.97; 95% CI: 0.96-0.98) and between the senior radiologist and AI (ICC=0.96; 95% CI: 0.95-0.97).

Conclusion: This preliminary study demonstrates a very good agreement for BCR evaluation based on the evaluation of MBD by a senior radiologist, junior radiologist and AI system.

Keys words: MammoRisk®, DenSeeMammo®, Breast density; Breast neoplasms; Artificial intelligence

Abbreviations: AI: artificial intelligence, BC: breast cancer; MBD: mammographic breast density; CNN: conventional neuronal network, DM: digital mammography, ICC: intra class correlation coefficient, SM: synthetic mammography

1. INTRODUCTION

Mammographic breast density (MBD) is an independent risk factor for breast cancer (BC) [1–8]. In addition, dense breast tissue decreases the sensitivity of mammography for BC detection [9,10]. The relative risk of developing BC is 4 to 6 times higher in women with extremely dense breast tissue as compared to women with fatty breast, and interval cancers are also more frequent [4,11,12]. Four categories of MBD are defined by the American College of Radiology's 5th edition lexicon based on visual estimation (BI-RADS A to D): A) The breasts are almost entirely fatty; B) There are scattered areas of fibroglandular density; C) The breasts are heterogeneously dense, which may obscure small masses; and D) The breasts are extremely dense, which lowers the sensitivity of mammography [13]. Patients with high MBD (BI-RADS C or D) may require additional screening by ultrasonography or magnetic

resonance imaging to improve sensitivity [14]. These exams generate extra costs but are deemed necessary: additional ultrasonography is mandatory for women with dense breasts in more than half of the states in the US and is strongly recommended in France.

Visual MBD assessments have shown significant intra- and interobserver variability that could lead to discordant assessments of dense and non-dense glandular tissue [15–17]. Moreover, the changes implemented by the fifth edition of the BI-RADS guidelines will likely result in an increase in the number of women classified as having dense breasts, because the new guidelines consider that even a small amount of heterogeneous fibroglandular tissue can hide an underlying mass [13], increasing interobserver variability [18,19].

In addition, the development of synthetic mammography (SM) from digital breast tomosynthesis adds another factor of variability with fewer cases of dense breast detected on SM than on digital mammography (DM) [20,21].

In order to improve reproducibility, standardized and automated breast density assessment software has been developed [17,22] with various degrees of correlation with human evaluation [23–25]. With the emergence of artificial intelligence (AI) models, new algorithms demonstrate better correlation with human assessment [23,24,26], and for predicting BC risk especially when using volumetric breast density [26–28].

The purpose of this study was to evaluate the agreement between automatic assessment software of breast density based on AI and visual assessment by a senior and junior radiologist, as well as the impact on the assessment of breast cancer risk (BCR) risk at 5 years.

2. MATERIALS AND METHODS

This single-center retrospective study was approved by the CERIM (Comité d'Ethique pour la Recherche en Imagerie Médicale) (IRB number: CRM-1903-004).

2.1. Study population

The database of our Institution was queried to retrieve all women between 40 to 74 year-old who underwent routine mammography from January 1, 2019 to February 28, 2019. A total of 650 consecutive women were initially identified. Of these, 181women were excluded because of unilateral or incomplete mammography and 469 women who had undergone bilateral mammography with two bilateral views and digital breast tomosynthesis were considered.

Exclusion criteria were: *i*), women with a high risk of breast cancer , including personal history of breast cancer (n = 100), personal history of high-risk breast lesions (*i.e.*, atypical ductal hyperplasia; typical columnar cell hyperplasia; atypical lobular hyperplasia and lobular carcinoma in situ; papillary lesions; flat epithelial atypia; radial scar, and complex sclerosing lesion) (n = 18), a genetic risk (BRCA mutations, serous ovarian cancer) or a personal history of radiotherapy for lymphoma (n = 16); *ii*) women with breast implants (n = 4); *iii*) women with incomplete risk factor data, mammographic artefacts or technical problems preventing application of the BC risk software (n = 20). Figure 1 shows the study flowchart.

The final study population included 311 women with a mean age of 55.6 ± 8,5 (SD) years (range: 40-74 years). Demographic and clinical characteristics of the study population are reported in Table 1.

2. 2. Acquisition technique

The imaging protocol consisted of full-field DM in both mediolateral oblique and craniocaudal views and breast tomosynthesis in craniocaudal views on two different systems: Selenia® Dimensions (Hologic) or Pristina® 3D (General Electric HealthcareGE). The mammography system, dose, number of views and breast thickness were collected for each patient. The Hologic model was used for approximately 85% of the patients and the GE for the remaining 15%.

2. 3. Mammographic breast density evaluation

Two radiologists, with 1 year (MLB) and 5 years (AB) of experience in breast imaging, independently rated the women's MBD according to the American College of Radiology BI-RADS 5th edition lexicon (category A, almost entirely fatty; category B, scattered areas of fibroglandular tissue; category C, heterogeneously dense; category D, extremely dense)[13]. Breast density assessment was made only on cranio-caudal views for each modality. Two different sessions at a minimum interval of 3 weeks were organized for both readers to independently evaluate MBD on DM and SM. MBD was also assessed by a fully automated Food and Drugs Administration approved software (DenSeeMammo®, Predilife) which had been trained on more than 10 000 DMs. DenSeeMammo® uses the convolutional neural network (CNN) and can only be used on DM.

For each MBD obtained from the senior reader, junior reader and automated software, the BC risk was estimated per woman using MammoRisk® software. The MammoRisk® score is a machine learning-based BC risk model developed on the United States Breast Cancer Screening Consortium cohort. The use of the MammoRisk® software in daily medical practice for BC risk consultation was validated in a French clinical study (RIVIERA, (NCT02997384) and is currently being used in a large European prospective study (My personalized Breast Screening [MyPeBS], Clinical trial NCT03672331) [29]. This algorithm

includes other risk factors including age, first-degree relatives with BC and their age at diagnosis, number of prior benign biopsies and their results. Results are expressed as a risk at 5 years in four categories: low risk, < 1%; moderate risk, 1%-1.66%; high risk, 1.67%-5.99%; very high risk, ≥ 6%.

2.4. Comparison of mammographic breast density evaluation

According to BI-RADS, the dense breast category corresponds to women with breast composition C or D while the non-dense breast category corresponds to women with breast composition A or B. Ultrasonography is required for women with breast composition C and D (dense breast). The senior radiologist's assessment of MBD was considered as the standard of reference.

2.5. Statistical analysis

Quadratic kappa (κ) coefficients were calculated to assess interobserver agreement between the two readers and the automatic software for breast density, using the following scale: 0.00–0.20 indicating slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; and 0.81–1.00, almost perfect agreement [32]. Reproducibility of BC risk at 5 years between the readers and automatic software was assessed using the intraclass correlation coefficient (ICC). A *P* value < 0.05 was considered to indicate a statistically significant difference. Statistical analyses were performed using MedCalc® software (MedCalc version 9.3.0.0; www.medcalc.be, Belgium).

3. RESULTS

3.1. Inter-reader agreement between radiologists and AI

Breast density agreement between the senior and junior radiologists on DM was almost perfect (weighted κ = 0.88; 95% CI: 0.84-0.92) (Table 2). To assess a potential impact on clinical routine (*i.e.*, indication of ultrasonography), we studied the difference of classification as dense / non-dense according to the reader and method. The junior radiologist misidentified 23 patients: 18 patients were upgraded from non-dense to dense breast and five were downgraded from dense to non-dense breast (Figure 2 and Table 2). Breast density agreement between the senior and junior radiologist was also almost perfect on SM (weighted κ = 0.86; 95% CI: 0.82-0.90) (Table 3). The junior radiologist misidentified 27 patients: 25 patients were upgraded from non-dense to dense breast and two were downgraded from dense to non-dense breast. We found substantial breast density agreement between DM and SM according to the senior radiologist's evaluation (weighted κ = 0.79; 95% CI: 0.74-0.84) (Figure 3 and Table 2A) and almost perfect according to the junior's evaluation (weighted κ = 0.88; 95% CI: 0.84-0.92) (Table 2B).

Breast density agreement between the senior radiologist and the AI model was substantial on DM (weighted κ = 0.79; 95% CI: 0.73-0.84). The AI model misidentified 25 patients as having dense breast and 14 as having non-dense breast (Figure 4, Table 2A). Breast density agreement between the junior radiologist and AI model was substantial on DM (weighted κ = 0.76; 95% CI: 0.71-0.82) (Table 2B). The AI model misidentified 23 patients as having dense breast and 25 as having non-dense breast. As the AI model was not trained on SM no evaluation of the agreement was performed between human and AI system assessment on synthetic views.

3.2. Impact on breast cancer risk evaluation

The BC risk as a function of MBD is shown in Table 4. On DM, the ICC of BC risk evaluation at 5 years between the senior (1.44 ± 0.6 [SD] %) and junior radiologists (1.45 ± 0.61 [SD] %) was very high (ICC = 0.98; 95% CI: 0.97-0.98) (Figure 5). Similarly, the ICC of BC risk evaluation at 5 years between the senior (1.32 ± 0.63 [SD] %) and junior radiologists (1.38% ± 0.60 [SD] %) was very high on SM (ICC = 0.98: 95% CI: 0.97-0.98).

On DM, the ICC of BC risk evaluation at 5 years between the senior radiologist (1.44% ± 0.60 [SD] %) and the AI model (1.47% ± 0.60 [SD] %) was very high (ICC = 0.96; 95% CI: 0.95-0.97) (Figure 6) as well as between the junior radiologist and AI model (ICC = 0.95; 95% CI: 0.94-0.96).

The intraclass correlation coefficient of BC risk evaluation at 5 years between DM (1.44% ± 0.60 [SD] %) and SM (1.32% ± 0.63 [SD] %) evaluated by the senior radiologist was very high (ICC = 0.96; 95% CI: 0.95-0.97) (Figures 7 and 8). Similar results were found for the junior radiologist with an ICC=0.96 (95% CI: 0.96-0.97).

4. DISCUSSION

Our study demonstrates the value of an AI model for predicting the risk of BC at 5 years, as well as the high concordance between senior and junior radiologists for breast density assessment. Moreover, our study reveals that differences regarding breast density evaluation exist between DM and SM even if this difference has a low impact on BC risk evaluation.

Breast density is currently a trending topic [33] : high breast density not only has a masking effect which decreases the sensitivity of mammography, but it also significantly impacts the evaluation of BC risk [34,35]. Sprague et al. reported an inter- and intraobserver variability of subjective assessment based on the BI-RADS density scale, with 17.2% discordant assessments of dense versus non-dense status for women with successive mammograms during the study period [16]. While our study confirms this variability, most of the discordance was observed on two contiguous categories of MBD, which explains why quadratic κ values were excellent independently of mammographic type (weighted κ for DM: 0.88; weighted κ for SM:0.86). -

Several automatic assessment techniques have been published mainly based initially on segmentation (volumetric MBD, for example) with a low rate of concordance with human evaluation [36,37]. Brandt et al. reported a moderate agreement between BI-RADS classification and two automatic assessement software Volpara® (weighted κ = 0.57; 95% CI: 0.55–0.59) and Quantra® (weighted κ = 0.46; 95% CI: 0.44–0.47), respectively [38]. Youk et al. found that the agreement of density category with visual assessment ranged from moderate to substantial in Quantra (weighted κ = 0.54–0.61) and fair to moderate in Volpara (weighted κ = 0.32–0.43) [39].

More recently, new techniques of breast density assessment have been published based on AI (mainly supervised model either with CNN or support vector model) and have shown a better concordance with BI-RADS evaluation [23–25]. The term ‘‘AI’’ covers many different training techniques including artificial neural networks (ANNs), machine learning (ML) and deep learning (DL). Supervised AI describes learning based on features from labeled images and consists of ANNs including CNNs, support vector machines, random forest, linear discriminant analysis, and decision trees [40]. Our study evaluated a new method with an AI model based on CNN (DenSeeMammo®) and confirms significant agreement between the

radiologists and the software on DM. Furthermore, the agreement we observed was higher than that found in previous studies (weighted κ: 0.79 with the senior radiologist and weighted κ: 0.76 with junior radiologist): Balleyguier et al. reported only a substantial agreement between their new software and the radiologists' consensus (unweighted κ = 0.68, 95% CI: 0.64 0.72) [23]. In our study, the AI model misidentified 14 patients (4.5%) as having non-dense breast on DM (breast density C misidentified as B). In comparison, the junior radiologists misidentified five patients (1.6%) as having non-dense breasts on DM, and two (0.6%) on SM (for both, breast density C misidentified as B). Even if this is lower than the standard error risk (5%), this misidentification could have an impact on routine practice and could exclude these patients from undergoing supplementary screening investigation such as ultrasonography [14,41].

The second question was related to the variability of BC risk evaluation at 5 years, as evaluated by the MammoRisk® software, depending on the way of MBD was rated. Our study confirms a high concordance between BC risk evaluation based on breast density evaluated on DM by the senior radiologist, junior radiologist and AI model (ICC = 0.96). Mammographers' acceptance of MammoRisk® for evaluating BC risk in the clinical setting is high [42] and the model is being increasingly used with training as part of the ambitious European clinical trial MyPeBS [29]. The primary objective of the MyPeBS study is the same as that of the WISDOM study conducted in the United States [43]: to evaluate the impact of personalized BC screening on the rate of diagnosis of BC stage 2B in the general population. The aim is to improve personalized screening, increase the interval between two mammograms for low-risk patients in order to reduce X-ray exposure and costs, and to maintain the benefits of screening [44]. In the MyPeBS study, BC risk estimation is based on the MammoRisk® software and on a single nucleotide polymorphism tested in a saliva sample.

Another issue was to assess the use of SM to evaluate breast density and the consequent impact on BC risk evaluation. The sensitivity and specificity of breast tomosynthesis has been shown to be increasingly good for BC detection with lower recall rates reported in several studies and a meta-analysis [45–47]. The main drawback of this technique is the potential extra radiation dose because of its use in addition to conventional DM. Several studies have demonstrated that a combination of SM and tomosynthesis has a better diagnostic performance than 2D mammography alone [48–51]. As several vendors are now able to obtain the same results, the question is "can we use breast density obtained with SM in our BC risk model with a significant impact?". In our study, we found a very good correlation between the BC risk assessment based on both DM and SM breast density evaluation (ICC=0.96; 95% CI: 0.95-0.97) even if the breast density agreement between DM and SM was only substantial (weighted $\kappa = 0.79$). 12.9% of the patients in our study who were rated as having dense breast by DM were rated as having non-dense breast on SM. This result is in line with other authors who reported that density differed from one imaging modality to another, with more breasts classified as non-dense on SM [20,52,53]. In this most recent study, 26% of the patients rated as having dense breast on DM were rated as having non-dense breast on SM [20]. In our study, this change in MBD category (dense vs non-dense breast) had a very low impact on the evaluation of BC risk at 5 years but may impact the use of additional screening techniques such as ultrasonography which is strongly recommended for all women with dense breasts in France and mandatory in more than half of the states in the United States.

Our study has several limitations. First, it was a single-center retrospective study which may limit its reproducibility. Furthermore, although we used two models from different manufacturers, as the Hologic model was used for 5 times as many patients than the GE model, we did not evaluate the impact of each system even if the technology used to generate

SM images can be quite different from one model to another [54]. Thus, our results need to be confirmed in a larger multicenter study, ideally during a prospective study. Second, we did not evaluate the actual occurrence of subsequent BC to validate the BC risk model. However, previous studies have been published in this setting [23]. Finally, DenSeeMammo® software cannot currently be used for SM.

In conclusion, our results demonstrate a good correlation between the evaluation of breast density by radiologists and AI systems and on BC risk evaluation. Further studies need to be performed to implement the evaluation of SM in a larger multicenter cohort.

DECLARATION OF INTEREST:

M. Le Boulc'h, None

I. Thomassin-Naggara: paid reading: GE, Hologic, Canon, Guerbet; expert meeting participation: SIEMENS

Asma Bekhouche, Edith Kermarrec, Audrey Milon, Cendos Abdel Wahab, Sonia Zilberman, Nathalie Chabbert-Buffet : None

REFERENCES

[1] Harvey JA, Bovbjerg VE. Quantitative assessment of mammographic breast density: relationship with breast cancer risk. Radiology 2004;230:29–41.

[2] Vacek PM, Geller BM. A prospective study of breast cancer risk using routine mammographic breast density measurements. Cancer Epidemiol Biomarkers Prev 2004;13:715–22.

[3] Boyd NF, Rommens JM, Vogt K, Lee V, Hopper JL, Yaffe MJ, et al. Mammographic breast density as an intermediate phenotype for breast cancer. Lancet Oncol 2005;6:798–808.

[4] Mccormack VA, Silva S. Breast density and parenchymal patterns as markers of breast cancer risk : a meta-analysis. Cancer Epidemiol Biomarkers Prev 2006;15:1159–70.

[5] Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, et al. Mammographic density and the risk and detection of breast cancer. N Engl J Med 2007;356:227–36.

[6] Moshina N, Sebuødegård S, Lee CI, Akslen LA, Tsuruda KM, Elmore JG, et al. Automated volumetric analysis of mammographic density in a screening setting: worse outcomes for women with dense breasts. Radiology 2018;288:343–52.

[7] van der Waal D, Ripping TM, Verbeek ALM, Broeders MJM. Breast cancer screening effect across breast density strata: a case–control study. Int J Cancer 2017;140:41–9.

[8] Chiu SYH, Duffy S, Yen AMF, Tabár L, Smith RA, Chen HH. Effect of baseline breast density on breast cancer incidence, stage, mortality, and screening parameters: 25-year follow-up of a Swedish mammographic screening. Cancer Epidemiol Biomarkers Prev 2010;19:1219–28.

[9] Tesic V, Kolaric B, Znaor A, Kuna SK, Brkljacic B. Mammographic density and estimation of breast cancer risk in intermediate risk population. Breast J 2013;19:71–8.

[10] Saarenmaa I, Salminen T, Geiger U, Heikkinen P, Hyvärinen S, Isola J, et al. The effect of age and density of the breast on the sensitivity of breast cancer diagnostic by mammography and ultasonography. Breast Cancer Res Treat 2001;67:117–23.

[11] Yaghjyan L, Colditz GA, Collins LC, Schnitt SJ, Rosner B, Vachon C, et al.

Mammographic breast density and subsequent risk of breast cancer in postmenopausal women according to tumor characteristics. JNCI J Natl Cancer Inst 2011;103:1179–89.

[12] Van Veen EM, Brentnall AR, Byers H, Harkness EF, Astley SM, Sampson S, et al. Use of single-nucleotide polymorphisms and mammographic density plus classic risk factors for breast cancer risk prediction. JAMA Oncol 2018;4:476–82.

[13] D'Orsi C, Sickles E, Mendelson E, Morris, EA et al. ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System. Reson VA, Am Coll Radiol 2013:https://www.acr.org/Clinical-Resources/Reporting-a.

[14] Tice JA, Cummings SR, Smith-Bindman R, Ichikawa L, Barlow WE, Kerlikowske K. Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model. Ann Intern Med 2008;148:337–47.

[15] Ciatto S, Houssami N, Apruzzese A, Bassetti E, Brancato B, Carozzi F, et al. Categorizing breast mammographic density: Intra- and interobserver reproducibility of BI-RADS density categories. Breast 2005;14:269–75.

[16] Sprague BL, Conant EF, Onega T, Garcia MP, Beaber EF, Herschorn SD, et al. Variation in mammographic breast density assessments among radiologists in clinical practice. Ann Intern Med 2016;165:457.

[17] Alomaim W, O'Leary D, Ryan J, Rainford L, Evanoff M, Foley S. Variability of breast density classification between US and UK radiologists. J Med Imaging Radiat Sci 2019;50:53–61.

[18] Irshad A, Leddy R, Ackerman S, Cluver A, Pavic D, Abid A, et al. Effects of changes in BI-RADS density assessment guidelines (fourth versus fifth edition) on breast density assessment: intra-and interreader agreements and density distribution. Am J Roentgenol 2016;207:1366–71..

[19] Alikhassi A, Esmaili Gourabi H, Baikpour M. Comparison of inter- and intra-observer variability of breast density assessments using the fourth and fifth editions of Breast Imaging Reporting and Data System. Eur J Radiol Open 2018;5:67–72.

[20] Gastounioti A, McCarthy AM, Pantalone L, Synnestvedt M, Kontos D, Conant EF.

Effect of mammographic screening modality on breast density assessment: digital mammography versus digital breast tomosynthesis. Radiology 2019;291:320–7.

[21] Philpotts LE. Density variation among mammographic modalities will likely impact imaging management and risk stratification. Radiology 2019;291:328–9. 42.

[22] Kang E, Lee EJ, Jang M, Kim SM, Kim Y, Chun M, et al. Reliability of computer-assisted breast density estimation: Comparison of interactive thresholding, semiautomated, and fully automated methods. AJR Am J Roentgenol 2016;207:126–34.

[23] Balleyguier C, Arfi-Rouche J, Boyer B, Gauthier E, Helin V, Loshkajian A, et al. A new automated method to evaluate 2D mammographic breast density according to BI-RADS® Atlas Fifth Edition recommendations. Eur Radiol 2019;29:3830–8.

[24] Lehman CD, Yala A, Schuster T, Dontchos B, Bahl M, Swanson K, et al. Mammographic breast density assessment using deep Learning: Clinical Implementation. Radiology 2019;290:52–8.

[25] Chan H-P, Helvie MA. Deep learning for mammographic breast density assessment and beyond. Radiology 2018;290:59–60.

[26] Wang J, Kato F, Yamashita H, Baba M, Cui Y, Li R, et al. Automatic estimation of volumetric breast density using artificial neural network-based calibration of full-field digital mammography: feasibility on japanese women with and without breast cancer. J Digit Imaging 2017;30:215–27.

[27] Puliti D, Zappa M, Giorgi Rossi P, Pierpaoli E, Manneschi G, Ambrogetti D, et al. Volumetric breast density and risk of advanced cancers after a negative screening episode: a cohort study. Breast Cancer Res 2018;20:1–7.

[28] Wanders JOP, Holland K, Veldhuis WB, Mann RM, Pijnappel RM, Peeters PHM, et al. Volumetric breast density affects performance of digital screening mammography. Breast Cancer Res Treat 2017;162:95–103.

[29] https://mypebs.eu/. My PeBs n.d.

[30] Veron L. Feasibility of breast cancer risk assessment and personalized breast screening recommendations delivery in community practice: a national prospective study (EBCC

poster). 2018.

[31] Balleyguier C. Feasibility of risk assessment and personalized breast screening recommendations delivery in community radiology practice : a national prospective study (NCT02997384), ECR oral communication, 2018.

[32] Landis SH, Murray T, Bolden S, Wingo PA. Cancer statistics, 1998. CA Cancer J Clin 1998;48:6–29.

[33] Thomassin-Naggara I, Touboul C. Women's imaging: What's new in 2019? Diagn Interv Imaging. 2019 Oct;100(10):535-536.

[34] Kerlikowske K, Ma L, Scott CG, Mahmoudzadeh AP, Jensen MR, Sprague BL, et al. Combining quantitative and qualitative breast density measures to assess breast cancer risk. Breast Cancer Res 2017;19:1–9.

[35] Trentham-Dietz A, Kerlikowske K, Stout NK, Miglioretti DL, Schechter CB, Ergun MA, et al. Tailoring breast cancer screening intervals by breast density and risk for women aged 50 Years or older: collaborative modeling of screening outcomes. Ann Intern Med 2016;165:700.

[36] Østerås BH, Martinsen ACT, Gullien R, Skaane P. Digital mammography versus breast tomosynthesis: Impact of breast density on diagnostic performance in population-based screening. Radiology 2019;293:60–8.

[37] van Engeland S, Snoeren PR, Huisman H, Boetes C, Karssemeijer N. Volumetric breast density estimation from full-field digital mammograms. IEEE Trans Med Imaging 2006;25:273–82.

[38] Brandt KR, Scott CG, Ma L, Mahmoudzadeh AP, Jensen MR, Whaley DH, et al. Comparison of clinical and automated breast density measurements: implications for risk prediction and supplemental screening. Radiology 2016;279:710–9.

[39] Youk JH, Gweon HM, Son EJ, Kim J. Automated volumetric breast density measurements in the era of the BI-RADS fifth edition: a comparison with visual assessment. AJR Am J Roentgenol 2016:1056–62.

[40] Thomassin-Naggara I, Balleyguier C, Ceugnart L, Heid P, Lenczner G, Maire A, et al. Artificial intelligence and breast screening: French Radiology Community position

paper. Diagn Interv Imaging 2019;100:553–66.

[41] Freer PE. Mammographic breast density: Impact on breast cancer risk and implications for screening. Radiographics 2015;35:302–15.

[42] Weigert J, Cavanaugh N, Ju T. Evaluating mammographer acceptance of MammoRisk software. Radiol Technol 2018;89:344–50.

[43] Eklund M, Broglio K, Yau C, Connor JT, Stover Fiscalini A, Esserman LJ. The WISDOM personalized breast cancer screening trial: simulation study to assess potential bias and analytic approaches. JNCI Cancer Spectr 2018;2:1–7.

[44] Pashayan N, Morris S, Gilbert FJ, Pharoah PDP. Cost-effectiveness and benefit-to-harm ratio of risk-stratified screening for breast cancer a life-table model. JAMA Oncol 2018;4:1504–10.

[45] Houssami N, Turner RM. Rapid review: Estimates of incremental breast cancer detection from tomosynthesis (3D-mammography) screening in women with dense breasts. The Breast 2016;30:141–5.

[46] Marinovich ML, Hunter KE, Macaskill P, Houssami N. Breast cancer screening using tomosynthesis or mammography: A meta-analysis of cancer detection and recall. J Natl Cancer Inst 2018;110:942–9.

[47] Carbonaro LA, Di Leo G, Clauser P, Trimboli RM, Verardi N, Fedeli MP, et al. Impact on the recall rate of digital breast tomosynthesis as an adjunct to digital mammography in the screening setting. A double reading experience and review of the literature. Eur J Radiol 2016;85:808–14.

[48] Bernardi D, Macaskill P, Pellegrini M, Valentini M, Fantò C, Ostillio L, et al. Breast cancer screening with tomosynthesis (3D mammography) with acquired or synthetic 2D mammography compared with 2D mammography alone (STORM-2): a population-based prospective study. Lancet Oncol 2016;17:1105–13.

[49] Mariscotti G, Durando M, Houssami N, Fasciano M, Tagliafico A, Bosco D, et al. Comparison of synthetic mammography, reconstructed from digital breast tomosynthesis, and digital mammography: evaluation of lesion conspicuity and BI-RADS assessment categories. Breast Cancer Res Treat 2017;166:765–73.

[50] Hofvind S, Hovda T, Holen ÅS, Lee CI, Albertsen J, Bjørndal H, et al. Digital breast tomosynthesis and synthetic 2D mammography versus digital mammography: Evaluation in a population-based screening program. Radiology 2018;287:787–94.

[51] Hovda T, Holen ÅS, Lång K, Albertsen JL, Bjørndal H, Brandal SHB, et al. Interval and consecutive round breast cancer after digital breast tomosynthesis and synthetic 2d mammography versus standard 2d digital mammography in breastscreen Norway. Radiology 2020;294:256–64.

[52] Zuckerman SP, Conant EF, Keller BM, Maidment ADA, Barufaldi B, Weinstein SP, et al. Implementation of synthesized two-dimensional mammography in a population-based digital breast tomosynthesis screening program. Radiology 2016;281:730–6.

[53] Aujero MP, Gavenonis SC, Benjamin R, Zhang Z, Holt JS. Clinical performance of synthesized two-dimensional mammography combined with tomosynthesis in a large screening population. Radiology 2017;283:70–6.

[54] Mandoul C, Verheyden C, Millet I, Orliac C, Pages E, Thomassin I, et al. Breast tomosynthesis: What do we know and where do we stand? Diagn Interv Imaging. 2019 Oct;100(10):537-551.

FIGURES LEGENDS

Figure 1: Study flow chart

Figure 2: Histogram and statistic table showing breast density evaluation agreement between senior and junior radiologists on digital mammography (n=311). DM: digital mammography

Horizontal line with A, B, C and D correspond to the mammographic breast density (MBD) gold standard results provided by the senior on digital mammography (DM). Colored column correspond to MBD provided by the junior radiologist on DM. If there is more than on column per horizontal category, additional column corresponds to junior radiologist's misinterpretations. For example, the blue column under A corresponds to misinterpretation by junior (type A classified as B by junior).

Figure 3: Histogram and statistic table showing breast density senior evaluation agreement between digital and synthetic mammography (n=311). DM: digital mammography; SM: synthetic mammography

Horizontal lines with A, B, C and D correspond to the mammographic breast density (MBD) gold standard results provided by the senior on synthetic mammography (SM). Colored column correspond to MBD provided by the junior radiologist on SM. If there is more than on column per horizontal category, additional column corresponds to junior radiologist's misinterpretations. For example, the red column under B corresponds to misinterpretation by junior (type B classified as A by junior).

Figure 4: Histogram and statistic table showing breast density evaluation agreement between senior radiologist and AI model (automatic) on digital mammography (n=311). DM: digital mammography

Horizontal lines with A, B, C and D correspond to the mammographic breast density (MBD) gold standard results provided by the senior on digital mammography (DM). Colored column correspond to MBD provided by AI (artificial intelligence; DenSeeMammo®) on DM. If there is more than on column per horizontal category, additional column corresponds to AI's misinterpretations. For example, the blue column under A corresponds to misinterpretation by AI (type A classified as B by AI).

Figure 5: Graph showing breast cancer risk at 5 years depending on senior and junior radiologists breast density assessment on digital mammography (n=311)

Figure 6: Graph showing breast cancer risk at 5 years depending on senior radiologist and AI (artificial intelligence; DenSeeMammo®) breast density assessment on digital mammography (n=311)

Figure 7: Graph showing breast cancer risk at 5 years depending on breast density assessment by senior radiologist on digital and synthetic mammography (n=311). DM: digital mammography; SM: synthetic mammography

Figure 8: Digital (A) and synthetic mammography (B) cranio-caudal views of a 44 y/o woman without personal or familial history of breast cancer included in our study

Breast density evaluated by senior radiologist was C on digital mammography and B on synthetic mammography. Breast cancer risk at 5 years was low (respectively 0.9% and 0.6%).