Pros and cons for breast cancer screening with tomosynthesis – a review of the literature

Julian Hans Kleinknecht¹, Anca Ileana Ciurea², Cristiana Augusta Ciortea³

Abstract
Breast cancer screening programs using mammography proved their value in detecting breast cancer at early stages and, consequently, reducing the mortality from this disease.

Due to the technological progress, the screening programs have shifted from screen-film mammography to digital mammography and nowadays digital breast tomosynthesis became the focus of breast imaging research. Using tomosynthesis in screening increases cancer detection rates and decreases recall and false-positive rates, thus improving the effectiveness of breast cancer screening programs, with positive consequences on health care costs and on patient psychology.

More long-term follow-up data must be collected for assessing absolute sensitivity and specificity of digital breast tomosynthesis, together with efforts for addressing the limitations of the method.

Keywords: digital breast tomosynthesis, breast cancer, screening mammography

Introduction
Breast cancer is a major cause of mortality and morbidity among women and represents the most common cancer diagnosed in women. Screening programs using mammography proved their value in detecting breast cancer at early stages, and consequently, reducing the mortality from this disease [1-5].

Due to the technological progress, the screening programs have shifted from screen-film mammography to full-field digital mammography (FFDM or 2D) [6]. According to most of the studies, digital mammography increases the cancer detection rates with a decrease of false positive results. Therefore, digital mammography is considered to increase the effectiveness of breast cancer screening programs and can be seen as the current standard in the developed countries of the world [7,8].

However, FFDM has two major limitations: a low sensitivity in women with dense breasts because of the masking effect of the overlying parenchyma and a low specificity because of the summation of normal parenchyma on the conventional images, which occasionally may simulate a cancer and, consequently, induces false positive recalls of the patients [1,6]. Due to the need to address these aspects, digital breast tomosynthesis (DBT or 3D mammography) became the focus of breast imaging research.

Technical basics of tomosynthesis
Tomosynthesis is a three-dimensional mammogram that uses X-rays to obtain sectional images of the breast, which are then reconstructed into a 3D volume.

Tomosynthesis is performed by moving the X-ray tube on a circular arch and making a series of low-dose exposures, from different angles, in a few seconds. Image acquisition can be done in both standard views, middle-lateral-oblique (MLO) and cranial-caudal (CC), as well as in additional views. The obtained projection images are then reconstructed into 1 mm-thick images for review. The total number of reconstructed images depends on the thickness of the compressed breast. In a 60 mm
compressed breast, for instance, there will be 60 images for the physician to review [9,10].

Subsequently, all images are used to reconstruct a synthesized 2D image, in which all the details recorded on the thin images are represented, without overlapping, as is the case in the classical 2D mammography.

The combo-mode (2D+3D) is frequently used in the clinical practice for image acquisition. Using the combo-mode, within the same examination, 2D digital mammography and tomosynthesis are performed, with the advantage for the radiologist to compare the information provided by the two modalities and to increase the accuracy of the interpretation, but also with the disadvantage for the patient due to the risk of the increased irradiation dose. The replacement of 2D mammography with the synthesized 2D images from tomosynthesis might reduce with almost 50% the irradiation dose, and this is the subject of many studies conducted in the present [1,11,12].

Breast screening in Europe and in the United States of America (USA)

In Europe, breast screening is recommended by the European Commission since 2003 to women aged 50-69 years. In 2007, in 22 states (out of 28), programs based on the European Union Commission indications were active or being organized. The population-based breast cancer screening programs among Europe differ from screening programs in the USA. In the USA, although several organizations recommend routine screening, screening practices differ because of the preferences of the medical services providers and because of the various health care plans, which may influence the accessibility to health care services. Most frequently, screening is performed opportunistically, due to a possible elevated risk of developing breast cancer, or in response to recommendations made during a routine medical consultation [13].

Screening and tomosynthesis

There are several studies comparing DBT+FFDM with FFDM in a screening setting; 15 studies (10 conducted in the USA and 5 conducted in Europe) were included in this review [14-28]. These studies varied considerably since they have been carried out in different health care systems with different screening systems and guidelines.

The inclusion criteria for the studies analyzed in this review were: retrospective and prospective studies with 1000 or more participants, which evaluated women participating in a breast cancer screening program or undergoing opportunistic mammography screening, and which analyzed FFDM alone compared to DBT alone, to DBT+FFDM and to DBT+synthesized 2D.

The exclusion criteria used in this review were: women with diagnosed and treated breast cancer; women participating in a breast cancer screening program which were recalled for further evaluation; women with symptoms suggestive for breast cancer; studies conducted before 2008, if they evaluated DBT systems in regard with technical issues, if they compared DBT with screen-film mammography or if the results were reported in other languages than English.

The studies included in this review analyzed the relative cancer detection, recall and false positive rates for DBT+FFDM and FFDM alone [1].

The studies published in the USA (Table I) are mainly retrospective studies in which geographical or historical FFDM control groups have been used for comparison [14-21], and only two are prospective trials [22,23]. In seven of the USA studies [14,15,17-20,22], single reader design was applied, in two studies [16,23] double reading of mammography images was applied and in one of the studies it was not reported whether it was single or double reading [21]. It is to be mentioned that single reading of mammograms used to be standard practice in the USA [6]. All the USA studies used two-view images (CC and MLO) for FFDM as well as for DBT.

There are particularities among the USA studies, in terms of number, age, personal history of breast cancer, risk group of the women included and also regarding the readers’ experience in breast imaging and their training in DBT interpretation, particularities that were to be expected. Another aspect worthy to be mentioned is that some studies analyzed the performance of DBT on baseline screening and some studies evaluated the longitudinal performance of DBT, which means to evaluate “whether the improved outcomes observed after initial installation of DBT screening are sustainable over time at a population level and to evaluate the effect of more than one DBT screening at the individual level” [17,18].

The five European studies included in this review (Table II) are all prospective studies [24-28]. Two studies were conducted in Scandinavia, the OTST trial in Oslo [24] and the MBTST trial in Malmö [25]. The other three trials, STORM 1 [26], STORM 2 [27] and Reggio Emilia [28] were carried out in Italy.

These trials on DBT have different study designs. In all five studies, women invited to an organized population-based mammography screening program were included, but the age groups varied among the five trials. They all applied double reading process, with and without arbitration. In two of the studies (OTST and STORM 1) the women were invited to undergo consecutively two-view FFDM and two-view DBT. One study (STORM 2) compared integrated 3D mammography (dual-acquisition 2D–3D mammography or 2D synthetic–3D mammography) with 2D mammography alone, and one study (MBTST) compared one-view DBT alone (MLO view) versus two-view FFDM and a combination of one-view DBT (MLO view) and one-view FFDM (CC view) in a sequential reading mode [1]. The Reggio Emilia trial is a two-arm test-and-treat trial where women were randomized to undergo FFDM+DBT or FFDM alone at baseline.
Advantages of tomosynthesis in screening

All trials have demonstrated a better visibility of a cancer on DBT and a significantly increase in the cancer detection rate when compared to FFDM alone, even if there are differences between the sites where the study was conducted [14,19,26-30]. The comparison between integrated 2D-3D mammography and synthetic 2D-3D mammography showed that the cancer detection rate did not differ significantly [27]. In the USA studies, the increase in overall cancer detection and invasive cancer detection was smaller than in the European studies. This may be explained by the older age of women taking part in the European screening programs, and also by the shorter screening intervals in the USA [6].

One of the two prospective USA studies compared the screening metrics in subgroups based on patient age and breast density. The conclusions were that a significantly higher cancer detection rate with DBT was reported in women 60–69 years old and cancer detection rate was found to be higher for DBT compared with FFDM for both the less dense and denser breast categories, but only being statistically significant in the less dense category [20]. Regarding breast density, similar results were obtained in one of the European trials [26].

An important finding of the USA and European studies, except for one in the USA [20] and two in Europe [28,30], is the statistically significant decrease in the recall rate, which confirms the higher specificity of combined
2D+3D mammography. In terms of screening metrics, in only one study in the USA, there was no significant difference in the recall rate between FFDM and DBT by patient decade and by breast density [20]. It is important to mention that in the USA study which showed an increase in the recall rate, the number of unnecessary recalls was reduced, which was statistically demonstrated by an increase in the positive predicted value. In the European trial using one-view DBT, which showed a relative increase in the recall rate, this was within the acceptable range as recommended by the European guidelines. Another aspect worthy to be mentioned is that the recall rates in the USA sites are higher compared to the European sites.

The European trials also showed a reduction in pre-arbitration false-positive scores when using DBT+FFDM, in single-reader and double-reader study approaches [24,29]. When compared FFDM+DBT to synthetic 2D+DBT, the analysis showed that there was no significant difference regarding false positive scores between them [31].

Since there are only few data available regarding long-term follow-up, it is not possible to evaluate absolute sensitivity and specificity of tomosynthesis [1]. However, the study published by Conant in 2016, which was the first study concentrating on long-term follow-up, reported an increase in specificity for the combination of DBT+FFDM [19].

**Challenges and limitations of tomosynthesis**

Taking into consideration all the above, DBT might be considered the future of breast cancer screening. But, nevertheless, there are challenges and limitations that must be addressed before.

**Irradiation dose**

The combo mode (FFDM+DBT) means almost doubling the irradiation of a presumed healthy woman. Consequently, there is a debate whether there is a need for 2D when using DBT because, theoretically, DBT might make 2D images dispensable. However, many radiologists prefer to include 2D images for comparison of current with prior screening examinations [32]. Beside this, some cancers are more obvious on one of the two standard views, and studies have proven that for cancer detection two-view DBT is superior to one-view tomosynthesis and the combination FFDM+DBT is superior to FFDM alone [33,34]. Therefore, experiences so far favor the combined use of 2-view FFDM with 2-view DBT in breast screening. The solution of the radiation issue might be the use of synthetic 2D images, which do not require any extra radiation exposure. The synthetic 2D images have now reached the quality that makes them an alternative to FFDM and will probably eradicate the need for FFDM in breast cancer screening with tomosynthesis [27,31,35,36]. Also, the STORM 2 trial concluded that there is no significant difference in the performance of 2D mammograms and synthetic ones regarding cancer detection rate [27].

**Microcalcifications**

A controversial issue is the ability of DBT to detect and characterize microcalcifications. A study released in 2011 came to the conclusion that calcifications can be demonstrated with equal or even greater clarity on DBT [37], whereas another study concluded that in a minority of cases, microcalcifications are scored and classified differently, which might have clinical relevance [38]. Furthermore, it has been observed that calcifications are “highlighted” on synthetic 2D images. This fact may ease the detection of calcifications in women with dense breasts [39,40], thus presenting another reason why synthetic 2D images will most probably become the future of DBT in screening.

**Lesions visible only on DBT**

A question to be answered when using tomosynthesis is how to manage the spiculated masses or architectural distortions visible only on DBT and not being detecatable on conventional imaging, even retrospectively (“tomo only lesions”) [40]. These lesions have a high chance of malignancy [41] and finding a solution to reach a positive diagnosis is imperative. Tomosynthesis-guided interventional procedures (biopsy or hook-wire localization) can solve the problems concerning occult lesions found only on DBT, as they have been proven accurate and fast [42,43].

**Overdiagnosis**

“Overdiagnosis” is a big debate in regard with the false positive interpretations causing unnecessary recalls. As for all modalities of investigations, using DBT will have that consequence, but almost all the studies in the literature have shown a decrease or a non-significant increase in the false positive rate of DBT compared to FFDM alone [24,29].

**Reading time**

Using DBT means a longer time needed for the interpretation of the mammographic images [29,30]. An experimental clinical study reported a mean reading time for 2D+3D of 77 sec. as compared with 33 sec. for 2D only [44]. This disadvantage must be weighed against the advantages of DBT and further studies will show if technological support as computer aided detection is a reliable tool to speed up reading time.

**Breast density**

One study in the USA [20] and one trial in Europe [26] reported a significantly increase in cancer detection rate for low dense breasts.

For high dense breasts, the increase in cancer detection rate was only statistically significant in the European studies, whereas the USA studies demonstrated a non-significant increase. In the evaluation of very dense breast parenchyma the use of ultrasonography for screening must be taken into consideration. Studies regarding this
aspect were undertaken and revealed controversial results, one showing increased incremental cancer detection for ultrasonography compared to DBT in women with mammographic negative dense breasts and another showing an overall better diagnostic performance of DBT compared to ultrasonography in women with dense breast parenchyma [45,46].

Conclusions
Breast cancer screening with DBT+FFDM increases the cancer detection rates and decreases the recall and false-positive rates compared with screening with FFDM alone.

Therefore, screening with DBT+FFDM increases the effectiveness of breast cancer screening programs, with positive consequences on health care costs and on patient psychology.

Nevertheless, more long-term follow-up data must be collected for assessing absolute sensitivity and specificity of digital breast tomosynthesis, together with efforts for addressing the limitations of the method.

The present article is part of the graduation license thesis entitled „Tomosynthesis in Breast Pathology”, elaborated by Julian Hans Kleinknecht in the Radiology Department of the „Iuliu Hatieganu” University of Medicine and Pharmacy.

References


