More cancers detected with combination of mammography and tomosynthesis

The interim analysis of data from the Italian Reggio Emilia Tomosynthesis (RETomo) trial has just been published [1]. This randomized trial involving nearly 30000 women was designed to compare the use of digital breast tomosynthesis (DBT) combined with standard digital mammography (DM) versus the use of DM alone for the detection of breast cancer. The interim results show that the combination of DBT and DM detects significantly greater numbers of breast cancers than DM alone.

We wanted to find out more about the RETomo trial in general and the implications of the preliminary results in particular, so we spoke to Dr. Pierpaolo Pattacini, director of Radiology at the AUSL Reggio Emilia hospital in Reggio Emilia, Italy.

Dr. Pierpaolo Pattacini,
Director of Radiology,
AUSL Reggio Emilia hospital,
Reggio Emilia, Italy.
email: Pierpaolo.Pattacini@ausl.re.it

Q *The RETomo study, whose interim results have just been published was initiated several years ago. What was the underlying rationale behind the study?*

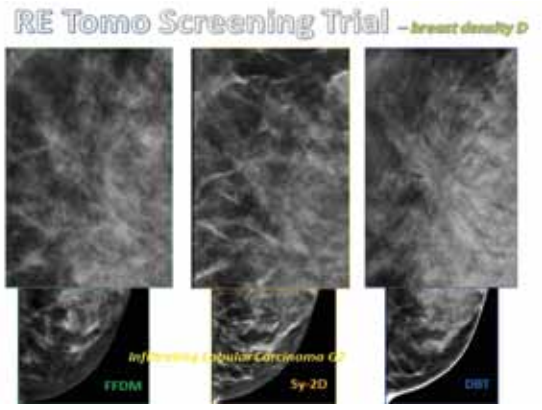
First of all, a bit of background. The radiology department of which I am head serves the province of Reggia Emilia, located in the north of Italy, with a total population in the province of 535000. Ever since 1994, we have carried out a province-wide breast cancer screening program, with an excellent participation rate of about 80%. In 2012 we acquired via a public tender 11 identical digital mammography systems supplied by GE Healthcare Systems. Three of the systems were then upgraded in 2013 to be able to carry out tomosynthesis. The availability of our new tomo and mammography systems put us in a position in which organizationally we could plan a robustly designed randomized double arm study (digital mammography (DM) alone versus DM plus digital breast tomosynthesis (DBT), using optimal technology. The aim of the study was to generate much more information, mainly about the overall health outcome of the participating women.

Q *And briefly, how was the study designed and what methodology and equipment was used?*

As I said, the study is based on a randomized double-arm design, which was created and validated by our epidemiologists and statisticians, Paolo Giorgi Rossi and Cinzia Campari, who are co-authors on the paper [1]. We used GE Essential systems throughout. Methodologically, we considered that the most important data and end-points that we would need to evaluate and compare would be the incidence of interval and advanced cancers (T2+) at the two subsequent rounds of screening in the respective two arms of the study (DM vs DM + DBT), which otherwise were absolutely identical in terms of the women being examined.

In practice, all women who agreed to take part in the study first had a 2D mammography examination. Then randomization algorithms allocated each woman to one of the arms, with only the woman herself and the technologist /radiographer being aware of which arm she was in. The randomization process was carried out by an independent Screening Co-ordination Centre (SCC) who had overall responsibility for running the study.

Women in the control arm had no further examination at this stage, whereas women in the experimental arm also had a DBT examination. All data were recorded using a structured reporting system and stored both in the hospital RIS and also in the information system of the screening organization (SCC) which also feeds the regional screening database for all the screening programs being carried out in the region. Cumulative incidences of interval and T2+ cancers are obtained from the "Reggio Emilia Cancer Registry".

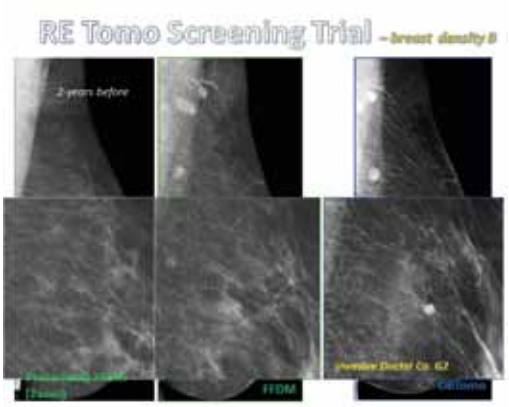


Infiltrating Lobular Carcinoma G2. Left panel, Full Field Digital Mammography (FFDM) image. Center panel, Synthetic 2D image. Right Panel, Digital Breast Tomosynthesis (DBT).

Q *Several trials to evaluate the impact of DBT using various, different study designs have already been reported. What are the features of the RETomo trial that distinguish it from the others?*

Our study has a simple and effective design that not only avoids selection bias but also minimizes any bias in the way the images are read. A randomized design is the only one allowing evaluation of any differences in long term outcomes, since women are managed either by DM alone in one arm and by the experimental protocol, DBT+DM, in the other arm. In contrast, in double testing studies women are always managed according to a combination of the results of two protocols, so the long term health outcomes cannot be attributed to one procedure or to the other.

Of course one consequence of our study design is that the DBT+DM combination means that women have a double compression of the breast, with that for tomo lasting slightly longer than that for DM. In addition the women in the DBT + DM arm have increased exposure to ionizing radiation. (In the future it is very likely that this drawback will be overcome through use of synthetic 2D images derived from the DBT data and so removing the need for a separate DM examination). The double acquisition resulted in some women declining to participate in the study. However, at the time we designed the study we considered that the alternatives, namely DBT alone or the use of an as-yet non-validated synthetic 2D, would have been less justifiable ethically.



Invasive Ductal Cancer G2. Left panel, Prior Full Field Digital Mammography (FFDM) image (two years earlier). Center panel, Full Field Digital Mammography (FFDM). Right Panel, Digital Breast Tomosynthesis (DBT) image.

Q *And now the interim report has been published; what are the key results so far and their implications?*

The increase in the cancer detection rate (DR) in the experimental arm is quite striking (we'll talk more about this later). However what I consider to be even more important than this, is the improvement in the experimental arm compared to the control arm of the positive predictive value (PPV) with the same recall rate, and without an increase in reading time in cases to be recalled. This means that the readers have greater confidence in the diagnostic interpretation, which would also benefit younger or less experienced screening readers, with an acceleration in their learning curve.

Q *When are the final results due?*

The end of the enrolment was in August 2017, so the complete dataset will be analyzed by August 2021. We'll probably also have results of the complete interim analysis published this year.

Q *The increase in detection rate in the DBT + DM arm of your study compared to that of DM alone is even higher than that reported in other studies. Any thoughts on the reason for this?*

In fact, the detection rate in the experimental arm of our study, is similar to that of other studies; the main difference is the detection rate of the control arm, which is lower than that found in the other prospective European study.

We have three possible explanations for this.

" ... the increase in the cancer detection rate in the experimental arm is quite striking .... however what I consider to be even more important..., is the improvement of the positive predictive value (PPV) with the same recall rate..." ,

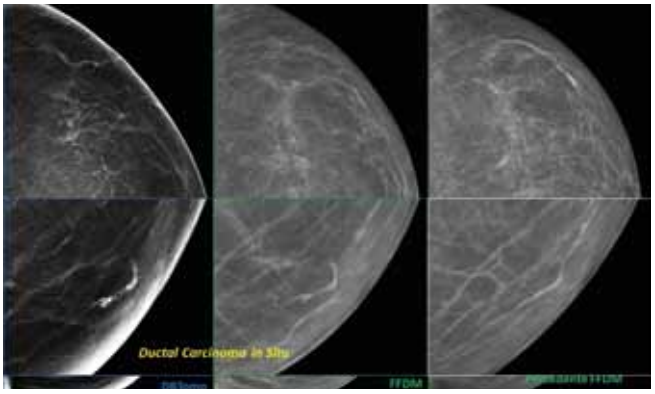
Firstly — and I don't know exactly why — we have never had a detection rate near 6 per thousand, even when we consider the entire population, including the prevalence round. We will have to examine our study cohort more closely and compare it with those of other studies. For example, in our province, we have a separate, dedicated process for hereditary-familial risk surveillance, which means that women at high risk of breast cancer are systematically excluded from the screening program. In addition, some women with other risk factors, such as receiving hormone replacement therapy, may be less likely to participate in screening programmes since they often prefer a spontaneous requested examination, in which they have direct close contact with, and feedback from the radiologist. All such factors could contribute to the inclusion in our study of women with a relatively low breast cancer risk.

The second possible explanation is that the sensitivity of detecting lesions in our control arm was lower than in other studies, i.e. we somehow missed more lesions with DM alone. This theory is however not supported by the low interval cancer rate that we have historically and regularly observed in our screening programs…

The third explanation — which I find the most plausible — is the fact that the women in our study had already undergone at least one round of screening.

Q *Likewise the recall rates were similar in the two arms of your study, in contrast to other trials where recall rates have been sometimes been higher with DBT, sometimes lower. How can this be explained?*

It's true that in Europe recall rates vary a lot across the different screening programs, so it's impossible to compare results. What is important is that in our setting the new method increased the PPV.



Ductal carcinoma in situ (DCIS). Left Panel, Digital Breast tomosynthesis (DBT). Center Panel, Full Field Digital Mammography (FFDM). Right Panel, prior FFDM (two years previously).

Q *As you mentioned earlier, the radiation dose in the DBT + DM arm is higher than that in DM alone. Can you foresee the eventual use of DBT alone, with the accompanying DM being replaced by a synthetic 2D DM image*

For sure, synthetic 2D DM is the immediate future — at least until future young screening readers and viewing systems could completely do without a 2D image….: in this respect the history of CT scans must surely have taught us something!

Q *The extended reading time necessary with DBT could be a brake on the introduction of DBT in routine screening. Could technologies such as Computer-Aided Detection help in this respect?*

We absolutely have to reduce reading time somehow. CAD could indeed be a tool for this, and we are looking forward to evaluating CAD systems. Other possible approaches to reducing the reading time are by reducing the number of tomo slices via increases in their thickness. Slices of 1 mm are probably not that useful diagnostically and certainly give rise to too many images; on average 55 per exam. My feeling is that 1 mm is too thin, and 10 mm too thick, given that the objective is to detect lesions that — to date, at least - shouldn't be larger than 10 mm.

Q *DBT is particularly useful in women with dense breasts where DBT can detect lesions that might otherwise be masked on DM alone by fibroglandular tissue. What was the density profile of the women in the RETomo study population?*

We don't routinely determine the BI-RAD Density in BI-RAD 1 breasts — we only evaluate density in recall patients — so currently I can't answer this question. In the future, we'll most likely have an automatic system for the evaluation of breast density. Then we will analyze the breast density in the entire cohort.

Q *Could the increased detection rate shown in the preliminary results of the DBT+ DM arm of your study eventually lead to an adjustment of current typical breast screening protocols, e.g. by extending the time interval between screening exams?*

The straight answer to this is no, since it's impossible to do so without going through a formal process of updating the recommendations that we have to follow, namely the EU guidelines. Nevertheless, when the long term outcomes of this study and those of other ongoing trials on tomosynthesis are published there will for sure be a movement to reconsider/re-evaluate the recommendations, .

Q *Overdiagnosis. This is currently a controversial subject, with some statisticians even going so far as to say that overdiagnosis could invalidate the basic rationale and cost-efficiency of screening mammography. What is your opinion on the overdiagnosis debate? Will the test-and-treat design of the RETomo study enable evaluation of the extent of overdiagnosis? On what time scale ?*

This is indeed a big subject and could influence the speed at which tomo is introduced to routine screening or not.
For example, the management in elderly patients of a large number of non-aggressive DCIS or Grade 1 tumors (whose detection rate could and, in fact seems to be increased through DBT screening), presents an economic and ethical problem for public health systems. In general all women want such lesions removed immediately. When this is done, the women feel that the Breast Cancer Screening Program has been effective for them. Nevertheless, in strict terms of impact on patient outcome, this probably could not be considered beneficial.
Thus, the most important future research to be carried out in tomo screening should be directed to the precise evaluation of overdiagnosis. Randomized trials, with a long follow up can do this better than other studies. In RETomo, we'll consider results of the follow up at least for 4 years: final results in late 2021.

Q *Breast tomosynthesis seems to be a very fast-growing imaging modality with the wind currently in its sails. How do you see future developments? With more and more DBT systems being installed and many different clinical trials of the technology being carried out, what, if any, co-operation is there between various European investigators in the field?*

As I said, in the future, I think all screening systems will have the tomo options; breast radiologists should at least be prepared for the next guidelines update! The most recent EU guidelines have already endorsed tomo (+ Synthetic 2D view) as an alternative to mammography so it's easy to imagine that the next guidelines, will at least confirm superiority in terms of DR, PPV, etc. But any move "*en bloc*" to tomo screening will first need a better understanding/quantification of overdiagnosis. This will inevitably require large studies, with large cohorts and long follow-up. (One such study that comes to mind, the TMIST trial, is ongoing now in US and Canada). Pooling data from ongoing studies is ethically highly desirable but I personally think that such data pooling should only involve studies with the same (or very similar) design. There's too much at stake to risk confusion by pooling incompatible data.
In this respect we are very fortunate in Italy in that the Ministry of Health has provided funds to pool data from (almost) identical studies in the North (MAITA): Piedmont, Lombardia, Tuscany, Veneto…and. of course, "RETomo".

REFERENCE

1. Pattacini P et al. Digital Mammography versus Digital Mammography Plus Tomosynthesis for Breast Cancer Screening: The Reggio Emilia Tomosynthesis Randomized Trial. Radiology. 2018 Jun 5:172119. doi: 10.1148/radiol.2018172119