Mammographic compression and diagnostic performance

By Prof. K Grimbergen & Prof. A den Heeten

One of the factors dissuading women from continuing participation in breast screening mammography programmes is the pain and discomfort caused by the compression of the breast necessary to optimize the quality of the mammographic image. Recently a new system for breast compression has been developed based on the use of pressure (i.e. the force applied to the breast divided by the contact area of the breast with the paddle).

This article reviews the question of breast compression in mammography in general and discusses the significance of two recently published articles on the relation between breast compression and the performance characteristics of mammography. If an optimal compression pressure, namely approximately 10kPa, is applied in screening mammography, then an increase of 5% in one-year sensitivity may be achievable.

Mammography is typically carried out in two situations, either as a screening tool or as a clinical/diagnostic tool. Of these two situations, the majority of investigations, — at least 150 million per year — are carried out in either an organized or non–organized setup for screening purposes [1].

The Authors

Prof. Kees Grimbergen

Professor of Medical Technology, Biomedical Engineering & Physics Department, Academic Medical Center, University of Amsterdam, The Netherlands. email: c.a.grimbergen@amc.uva.nl

&

Prof. Ard den Heeten

Professor of Radiology, Radiology Department, Academic Medical Center, University of Amsterdam, The Netherlands.

email: denheeten@gmail.com

In Western Europe, one in eight women will develop breast cancer in her lifetime, of whom more than 75 % are detected after the age of 50 [2]. Every year breast cancer kills more than 500000 women and recently it has become one of the leading causes of morbidity and mortality in low- and middle-income countries (LMICs), where 60% of the world's total of new cases are currently diagnosed [3, 4]. In LMICs the 5 years survival rate varies between 10 and 40%, in contrast to a survival rate of 80% in countries with better healthcare organization and low threshold mammographic breast cancer screening.

Since mammography — together with other factors such as the development of new therapeutic treatments — has undoubtedly played an important role in the improvement of the survival rate of breast cancer patients [5], we thought it useful to take a closer look at the future relevance of the modality and possible future innovations and improvements. In particular, in this article we focus on the recent insights in the importance of mammographic compression.

Over the last few years, there has been a significant technology innovation with the introduction of tomosynthesis, which, as a variant of mammography still requires breast compression as an essential element.

COMPRESSION IN MAMMOGRAPHY:

The term "compression" is misused in mammography. In almost every other technical discipline, such as engineering, physics and computer science, the term compression is used to describe the process involved in making something smaller in volume. In mammography the only reduction is in the thickness of the breast after deformation, but this barely relates to the actual volume of the breast itself. Thus in this context, the word "compression" should perhaps be replaced by the word "flattening", which is a better description of the clamping of the breast between two flat parallel plates, the paddle and the bucky plate. The flattened breast is at right angles to the x-ray source, thereby minimizing the differences in distance travelled by the x-rays in the breast tissues, which would not be the case in a round or oval breast without flattening.

However one important question has never been satisfactorily answered, namely at which point in time is breast flattening sufficient? Put another way, when is it optimal to stop increasing the force used to flatten the breast?

Such simple, but in practice extremely important, questions are impossible to answer without some basic physics background. Breast flattening is caused by a combination of the pressure and the flexible (but not compressible) nature of the breast. In mammography systems this flattening is a result of an applied force generated by a motor. The force will instantaneously be distributed over the contact area, and a mean pressure between the breast and the paddle will be felt by the woman

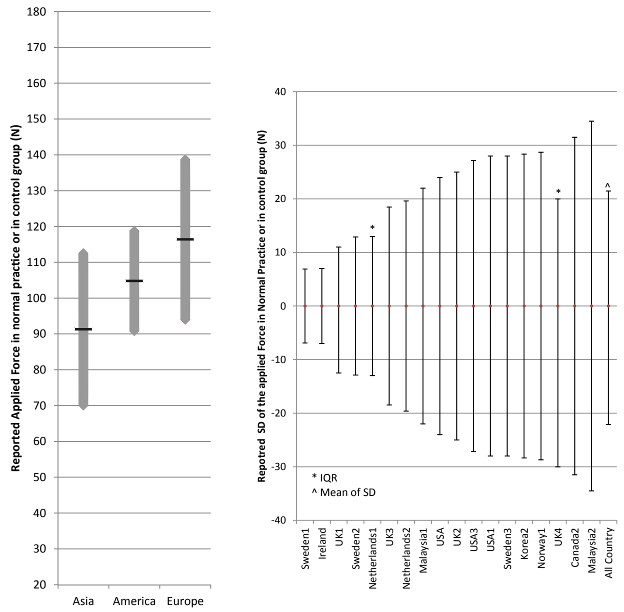

Figure 1. Mean compression forces and standard deviations in newtons applied throughout the world; left panel per continent and right panel per country. [7]

involved. At the same time an identical reaction force is generated by the bucky plate. When the motor is exposed to a counter force, the force it generates increases from 0 daN (1 decanewton (daN) is close to 1 kilogram of force) to a maximum of 20 daN [6].

The mean pressure which will ultimately be reached in manual mammography, i.e. without the intervention of a pressure sensor, is dependent on the individual properties (volume, stiffness) of the woman's breasts, her personal pain threshold and on the skill and experience of the technician (awareness, empathy, training, and guidelines).

In pressure-guided compression all four independent parameters influencing the patient's experience are grouped into one factor, and can be managed in a simple way by monitoring only this single physical parameter, namely the mean pressure. In this way, the flattening procedure stops at a reasonable point and this is — in most cases — below the pain threshold of women.

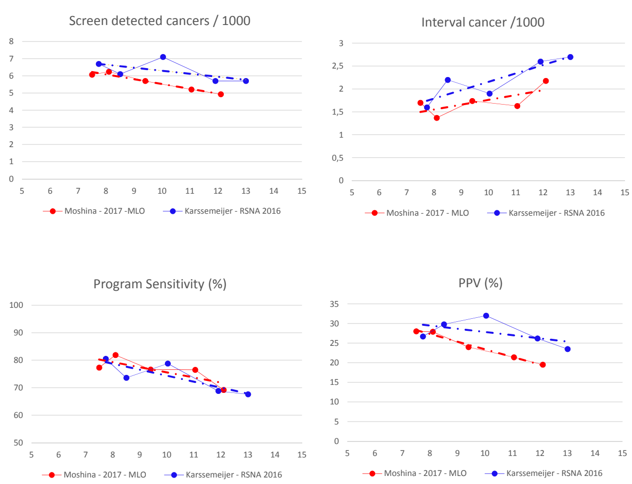

Figure 2. Comparison of the performance measures versus mean pressure in kPa of the Norwegian study and the Dutch study (the dots indicate the median values of the 5 pressure groups in both studies).

GUIDELINES

When a woman undergoing mammography asks the technician: "When will you stop compressing my breast?" the most frequent answer is : "When the skin is taut". While the breast compression is progressing, the technician generally uses her finger to probe the skin and so estimate its tautness. At some indeterminate point the technician then simply decides the compression is sufficient and stops. Other — so-called — guidelines suggest aiming for a force between 12-18 daN. In fact, the only absolutely clear guideline is that provided on page 76 of the European guidelines, namely "there is no optimal value known for the force" [6]. Figure 1 illustrates the consequences of this lack of guidelines and lack of consistency in practice which leads to a wide variation in the compression forces typically used in different countries and continents.

EFFECT OF PRESSURE ON SCREENING PERFORMANCE.

An important question arising in this respect is whether the level of pressure applied could influence the efficacy of mammography in its primary role, namely the detection of cancers. The possibility of analyzing this aspect further depends on the availability of large, accurate and continuously monitored databases, which can be found in organized breast cancer screening programs in countries with national cancer registers. A second important condition is that the mammographic images are stored in a "for processing format" so that a new generation of software can be used that opens up the possibility of datamining the large datasets generated by screening mammography. Such software is provided by the Volpara company and enables a reliable estimate to be established of the contact area of the breast on the paddle, based on the mammography image and other important parameters such as the volume of the breast. We have validated this specific method for contact area measurement [8]. Once the breast contact area has been determined the mean contact area pressure at the moment of exposure can be easily calculated simply dividing the applied force (which is stored in the DICOM header), by the estimated contact area.

THE NORWEGIAN EXPERIENCE

A first peer-reviewed paper on the relation between compression force, pressure and the performance of screening mammography, was published by the group of Solveig Hofvind of the Norwegian Cancer Registration in Oslo [9]. They investigated the relation between compression force and pressure at the time of the mammographic screening examination with early performance measures in a population-based breast cancer screening program. In this context, the term "early performance measures" is taken to cover all performance parameters except mortality (e.g. recall rate, rates of screen detected and interval breast cancers, positive predictive value of recall (PPV), screening sensitivity and specificity, and histopathologic characteristics of screen-detected and interval breast cancers). The number of mammograms available for analysis totaled 261,641 examinations with follow-up data, in

- CC & MLO
- 50-69 years
- Subsequent rounds
- GEE, Odds ratios (crude and adjusted for confounding)
- Interval cancers 2 years
- Mean Breast Volume 964 cm³ (MLO)
- Mammographic machines of different vendors.
- Cut off values pressure bins MLO:
  ≤7,5; 7,5-8,7; 8,8-10,0;10,1-12.0; ≥ 12,1
- Recall in subsequent rounds 2,4%

- MLO
- 50-75
- First rounds and subsequent rounds
- General estimated equations for confounders (GEE)
- Interval cancers 1 and 2 years
- Mean Breast Volume 973 cm³ (MLO)
- Mammographic machines only Hologic Selenia.
- Cut off values pressure bins MLO:
  ≤7,74; 7,75-9.26; 9,27-10.80; 10,81-13; ≥13,01
- Recall in subsequent rounds 2,5%

Some relevant similarities and differences between the Norwegian study and the Dutch study

93,444 subsequently screened women. The study period was 2007–2015.
The exams were categorized into three equal sized groups. For those who are not familiar with the use of force and pressure and potential confounders, the main conclusions of the paper, as expressed in the abstract, perhaps should be explained because they appear to be in mutual disagreement. The authors concluded that high compression force and low compression pressure were associated with more favorable early performance measures in the screening program, a finding which may be, for many people, a counter-intuitive result. The relation between high forces and favorable early performance measures lies directly in the tendency of technicians to apply more force to large breasts. It may be thought that low pressure is related to low force, but a relatively high force applied to an even larger contact area will result in a lower pressure. A second source of misunderstanding is the assumption that ongoing higher forces will result in a thinner breast, lower dose and better image quality. In fact, the contrary can be the case.
For these reasons it is more sensible not to focus on force but on the results based on pressure groups and compare them with results reported earlier by the group of Karssemijer in the Netherlands, [10,11].

DUTCH EXPERIENCE

The establishment of a relation between compression pressure and screening performance in the Netherlands was performed in a series of 113,464 screening exams [10]. These examinations were carried out in the former Prevention Center (now the Midden West region) because in this screening center "for processing images" that are essential for the analysis were generated, in contrast to other Dutch screening centers. These exams were categorized into five equal groups of increasing applied pressure, in such a way that each group contained 20 % of the exams. Pressure thresholds between the groups were 7.7, 9.2, 10.7 and 12.8 kPa. Measures of screening performance were then determined for the exams in each group. It was found that PPV and the cancer detection rate varied significantly within the five groups. There was a clear indication that the group with a moderate pressure (around 10 kPa) had an overall better screening performance than those in the lower and higher pressure categories. However, in such large studies important confounders could apply and should be addressed. In a follow-up study [11], the data were supplemented with an extra year's data to reach 132,801 examinations with follow-up of the screening performance parameters. In this extended study, the level of interval cancers was known so the sensitivity and specificity of the screening program could be calculated. It is important to note that a distinction was made between one-year sensitivity and two-years sensitivity. Another important point is that in both studies the statistical technique known as Generalized Estimating Equations, (GEE) was applied to correct for confounders, such as examinations of the same women, the breast volume and the density. Thus the statistical approach differs from that used by the Norwegian group, since in the Norwegian study crude and adjusted odds ratios were calculated. It should also be noted that only the MLO projections were used in the Dutch study due to historic guidelines restricting the use of CC to first round clients, and mainly patients with more dense tissue. Thus an analysis of a mix of MLOs with CCs could skew and bias the results. For this reason when comparing the results of the Norwegian study and the Dutch study we focus solely on the MLO results of the study as presented in the addendum table A 2 2.3. of the Norwegian study.
In all the important performance parameters the trends appeared to be the same [Figure 2].

DIFFERENCES AND SIMILARITIES BETWEEN THE NORWEGIAN AND DUTCH DATA.

Despite their differences, the data from these two studies seem suitable for comparison. The most striking similarity, which could have a direct significance, is the general finding of decreasing performance parameters in the higher pressure compressions. It can be seen that the level of screen-detected cancers, the positive predictive value, the program sensitivity, and interval cancer rate all have the same trends in both studies. However, the Dutch data indicate an optimal performance in the middle group 3, i.e. around 10 kPa.
It could be expected that screening performance will be negatively affected by very low pressures as can be seen in the Dutch data and also, but less obvious in the Norwegian data. There is a trend towards worsening performance parameters for the groups with the lowest pressures. In this respect it should be noted that an important parameter is missing from the Norwegian data, namely the one-year interval cancers.

ONE-YEAR INTERVAL CARCINOMA VERSUS TWO-YEARS INTERVAL CARCINOMA.

The reason that we want to stress the value of one-year program sensitivity is because we are evaluating neither the Norwegian nor the Dutch screening performance data, but simply looking for signs of the influence of compression pressures on the performance of mammography as a diagnostic test. So metrics that come close to the pure mammographic performance (and not to the performance of organized screening as such) are the most informative. There are well known proxies for mammographic performance in the monitoring and evaluation of screening programs, one of which is one-year sensitivity [12]. Here the assumption is that interval cancers detected in the first year (almost always because they

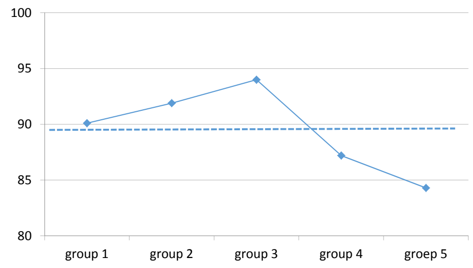

Figure 3. One year program sensitivity (%) in the Dutch study for the 5 different pressure groups. Dotted line indicates the average.

became symptomatic) were already present at the time of screening and were simply not detected at that time. So the question is: "do we see in the one-year sensitivity a similar or even more pronounced trend" — the answer is clearly yes.

BREAST VOLUME:

With the current method used for breast compression, women with the smallest breast volumes (<500mL) are clearly prone to receive the highest pressures. If in the Norwegian data we analyze the interval cancer rates plotted against volume, it can be seen that not only have these women a much higher chance of suffering from severe pain and discomfort [1], but they also have a 30% higher likelihood to be confronted with an interval carcinoma compared to women with larger breasts in group 4 and 5. A similar trend can be seen in the Dutch data.

CONCLUSION:

The many complaints from women undergoing mammography regarding pain and discomfort from breast compression together with misunderstanding and misconceptions of the mechanics of breast compression in mammography, was the main driver behind our proposal, made five years ago, of a new approach to compression [13]. We modified the mammographic compression procedure through the use of a transparent and radiolucent pressure measurement tool fitted to the compression paddle. In addition we investigated a more rational way of determining when breast compression was sufficient in the light of the set criteria, namely immobilization, sufficient image quality, and as low as reasonable radiation dose.

The research in the two papers on compression and performance discussed above, was made possible because of the existence of a datamining tool from the Volpara company that enables a reliable measure of the breast contact area duing mammography [8]. As the force value is available in the DICOM header, a retrospective calculation of the mean pressure at exposure is simple. The reported data indicate that women who received the highest pressure not only suffered from significantly more pain and discomfort, but also seemed to be subjected to a less optimal test.

Despite the fact that there are important similarities in the results of both the Norwegian and the Dutch studies, some differences cannot be explained away. The near absence in the Norwegian study of a performance decline at the low pressure end of the spectrum which was obvious in the Dutch data, deserves a closer look.

We recently discovered that mammography systems produced by different companies might differ in their force measurements by more than 10% (also within the same brands). In new studies, it is important that all mammography systems be calibrated and documented, since this can have a significant effect on the reported pressure. Another interesting point in the Norwegian data is the exclusion of women who participate in the first round (in the Netherlands such women have more dense breasts and a higher tumor detection rate), at the cost of a much lower positive predictive value.

We can conclude that the pressure used during mammographic exposure plays an important role and if the proper pressure can be reached with most patients in screening (according the Dutch data around 10 kPa), then an increase in 5% in the one-year mammographic sensitivity, might be achieved. This could be even more important than the complete effect of digitization of the breast cancer screening program in the Netherlands as a whole.

REFERENCES

1. de Groot JE, Broeders MJ, Branderhorst W, den Heeten GJ, Grimbergen CA. A novel approach to mammographic breast compression: Improved standardization and reduced discomfort by controlling pressure instead of force. Med Phys. 2013; 40: 08:1901.

2. Netherlands Cancer Registry. Comprehensive Cancer Centre the Netherlands (IKNL) (2014) http://www.dutchcancerfigures.nl/

3. WHO Position Paper on Mammography Screening. Geneva: World Health Organization; 2014. WHO Guidelines Approved by the Guidelines Review Committee.

4. Al-Sukhun S, de Lima Lopes G Jr, Gospodarowicz M, Ginsburg O, Yu PP. Global Health Initiatives of the International Oncology Community. Am Soc Clin Oncol Educ Book. 2017; 37: 395-402.

5. Sankatsing VDV, van Ravesteyn NT, Heijnsdijk EAM, Looman CWN, van Luijt PA, Fracheboud J, den Heeten GJ, Broeders MJM, de Koning HJ. The effect of population-based mammography screening in Dutch municipalities on breast cancer mortality: 20 years of follow-up. Int J Cancer. 2017 Aug 15; 141(4): 671-677.

6. European Guidelines Fourth edition. Perry et al. 2006 European Commission.

7. Oh KK, Hur J, Kim EK, Choo SS. Dosimetric evaluation of the mean glandular dose for mammography in Korean women: a preliminary report. Yonsei Med J. 2003; 44: 863.

8. Branderhorst W, de Groot JE, van Lier MGJTB, Highnam RP, den Heeten GJ, Grimbergen CA. Technical Note: Validation of two methods to determine contact area between breast and compression paddle in mammography. Med Phys. 2017 Aug;44(8):4040-4044.

9. Moshina N, Sebuødegård S, Hofvind S. Is breast compression associated with breast cancer detection and other early performance measures in a population-based breast cancer screening program? Breast Cancer Res Treat 2017; 163: 605.

10. K.Holland , I. Sechopoulos, G. den Heeten, R.M. Mann and N. Karssemeijer. "Performance of breast cancer screening depends on mammo0graphic compression"Breast Imaging-13th International Workshop of Breast Imaging (IWDM), 2014

11. N.Karssemijer, K.Holland, I.Sechopoulos, R.M. Mann, G.J. den Heeten and C.H. van Gils. "High breast compression in mammography may reduce sensitivity", Annual meeting of the Radiological Society of North America, 2016

12. National evaluation team for breast cancer screening (NETB) 1990-2011/2012, 2014 (XIII). Page 38. (downloadable https://www.erasmusmc.nl/public-health/publications-collaborations/reports/evaluatie-borstkanker/) Dutch and English.

13. Grimbergen C.A., den Heeten G.J. Pressure controlled compression of the breast in mammography. Diagnostic Imaging Europe. 2016: 32: 40-43.