A New Method for Dosimetry and Image Quality Assurance in Mammography and Breast Tomosynthesis

(including abstracts 2358373 and 2492856)

Hugo de las Heras Gala[1], Felix Schöfer[1], Britta Tiller[1], Margarita Chevalier del Río[2], Georg Zwettler[3] and Friedrich Semturs[3]

1 QUART GmbH, Kirchenweg 7, 85604, Zorneding, Germany

2 Departamento de Radiología, Facultad de Medicina. Universidad Complutense, 28040, Madrid, Spain

3 Medizinische Universität Wien, Währinger Gürtel 18-20, 1090, Wien, Austria

Abstract

Introduction and purpose: The latest developments in breast imaging technology have underlined the need for appropriate dosimetry and image evaluation methods for routine quality assurance. We are presenting QUART evaluation method (phantom and dosimeter) to address this need for digital x-ray mammography and breast tomosynthesis systems. We compare our dose measurements to the dose readings from the console, and the image quality evaluation to the known CDMAM phantom test.
Materials and Methods: The evaluation method is based on the visual detection of Landolt C's, which are constructed to be equivalent to microcalcifications inside the breast. The image quality assessment measures detectability within a wide range of background attenuation and produces a score as a function of the average glandular dose (AGD). Additionally, the phantom contains objects for a software-based evaluation of resolution and contrast-to-noise ratio, and specific features to be used for quality control of digital breast tomosynthesis. To compare the method with the CDMAM test, images were obtained with FFDM flat panel systems in a series of gradually reduced doses and evaluated following the specific protocol of each methodology. Entrance dose was measured by insertion of a small dosimeter in a specifically designed slot of the phantom.
Results: A threshold score equal or above 20 (visual test) or 2.7 (contrast-to-noise ratio) has been identified with a passed CDMAM test on acceptable values, showing the equivalence of the methods. An acceptable uncertainty is achieved with the acquisition and evaluation of 3 to5 images. The dose measurements constantly differed from the console readings by 5.0 ± 0.3 %.
Conclusion: The presented QUART phantom and dosimeter provide results and reliability equivalent to the CDMAM test. The QUART method covers a wide range of attenuation background in one shot, and is more efficient in terms of manufacture costs, operator time and evaluation effort.

Keywords: mammography, QA/QC, tomosynthesis, QUART, CDMAM.

1. Introduction

Mammography plays an important role in early detection of breast cancer due to its ability to reveal the presence of subtle lesions (masses), small details (microcalcifications) or architectural distortions of breast tissue. Due to this, the image quality provided by mammography systems is the subject of continuous assessments to ensure that breast images meet the high standards required for a correct diagnosis.
There are several approaches for the evaluation of the image quality provided by mammography systems that involve objective and subjective measurements. Objective measurements of the image quality are related with the measurement

of the modulation transfer function (MTF), noise power spectrum (NPS) and the detective quantum efficiency (DQE) [1]. The easier access to image data in digital systems has made these measurements possible as quality control procedures. The advantage of using mathematical functions is that the results are comparable and objective. However, it is important to take into account that an accurate measurement of these parameters is complex and is influenced by different system factors that limit its results.

Subjective measurements are based on the use of either clinical images or appropriate phantoms. Protocols to check image quality using clinical images have been recently reviewed [2]. Semi-quantitative phantoms have been traditionally used in quality control procedures as well as in strategies aimed to optimize the image quality/dose relationship. These phantoms usually include details simulating lesions (fibers or microcalcifications) and/or test objects that allow to perform quantitative measurements [3,4]. As an example, the phantom by the American College of Radiology (ACR) contains simulated fibres, specks and masses, of which a certain number must be visualized keeping an average glandular dose below 3 mGy [5,6]. Phantoms such as TOR(MAX) are mainly used for screen/film mammography and include test objects to measure some properties of the mammography system such as high contrast resolution or noise. These types of phantoms constitute a direct and simple means for routine image quality assessment. The growing presence of digital imaging systems has resulted in extending the use of contrast-detail phantoms to evaluate the image quality. These phantoms have been constructed based on the Rose model [7] and provide information about the contrast thresholds associated with the detectability of objects immersed in noisy backgrounds.

Reference values for the image quality of full-field digital mammography (FFDM) systems are stated in the European Guidelines for Quality Assurance in Breast Cancer Screening and Diagnosis (EPQC V4, 2006) [8,9]. These reference values are defined in terms of the contrast threshold associated with the gold discs included into the CDMAM contrast detail phantom (Artinis Medical Systems B.V., Nijmegen, The Netherlands) [10]. These guidelines have been taken as a reference to develop national protocols in several countries.

However, while excellent for science, the CDMAM phantom has received criticism regarding quality assurance. Human observer evaluation of CDMAM phantom images requires a long time and the image readouts present a high inter and intra-observer variability [11]. The use of software programs to evaluate CDMAM images requires obtaining large amounts of heavy images in order to achieve reliable results. Moreover, these results have to be processed in order to make them comparable to that from human observers [12]. In addition, its cost is high due to the complex procedure of its manufacturing. From the physical point of view, this phantom presents other drawbacks. The energy dependence of the x-ray absorption coefficient of gold (Z/A = 0.401) and the one of breast lesions is very different. Apart from that, slight variations of 1μm in the gold thickness dramatically change this energy dependence (**Figure 1**). Finally, the discs are distributed at different distances to the chest wall, so that the Heel effect must be somehow compensated before the evaluation.

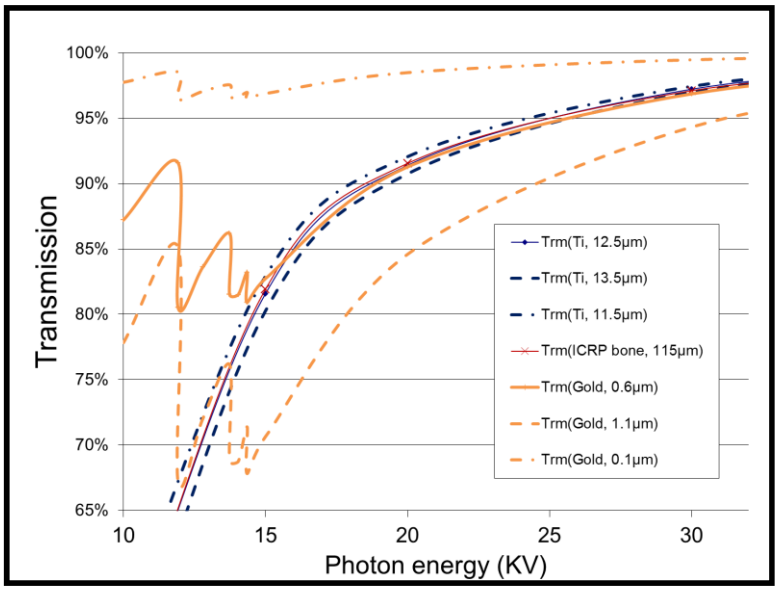


Figure 1: Transmission through 115 μm calcified biological material (cortical bone as defined by ICRP), and through thicknesses of gold and titan that provide a similar transmission. Note the larger dependence on the thickness in the case of gold.

In this work we present a semi-quantitative phantom (*QUART mam/digi EPQC*) that includes a set of intermediate contrast objects for a visual evaluation of the contrast threshold detectability, and another set of objects to determine spatial resolution via MTF and contrast-to-noise ratio (CNR). The later set can be objectively evaluated by means of a user-friendly, software-driven evaluation. The set of test objects for visual evaluation and CNR calculations are on a step wedge located close and parallel to the chest wall side (thus avoiding the need to compensate the Heel effect). The step wedge provides a wide range of attenuating thicknesses that allows assessing contrast threshold detectability for different scatter radiation conditions in a unique exposure. The software, based on the complementary measures CNR and MTF, can perform a fast and objective evaluation that can replace the visual test. Finally, the phantom can be used to evaluate digital mammography systems (FFDM and CR), as well as digital breast tomosynthesis (DBT) systems.

The purpose of the present work is to present the QUART phantom and the results from its image evaluation methodology. Acceptable and achievable thresholds are defined based on the reference values defined in the EPQC[8,9] for the contrast-detail test.

2. Materials and methods

2.1. The mam/digi EPQC phantom and the evaluation method

The phantom (QUART GmbH, Zorneding, Germany) has a semicircular shape simulating a 46 mm thick compressed breast. A radiograph of the phantom is shown in **Figure 2.** The set of test objects for visual evaluation consist of a series of rings with a gap at a position chosen at random. This type of objects is based on the so called Landolt C's, which are used in ophthalmology [13,14] since the 1950's to assess human visual detection and perception. The Landolt rings are arranged in twelve groups within a nine-step wedge made of PMMA continued by a 3-step alloy wedge, which simulate different breast thicknesses and tissue densities. Each group contains six rings, each of these having a different. As defined by Landolt, the size of the gap is always a fifth of the ring's diameter.

The contrast relation between the rings and their background (PMMA / Ti) is the inverse of the relation between microcalcifications and their background (Ca / tissue). Titan shows very similar transmission properties to those of calcified biological material for an appropriate thickness. To illustrate this, Error! Reference source not found.**1** depicts the percentage of x-ray transmission [15] through gold, titan and calcified biological material (cortical bone as defined by ICRU [15]), which as a biologically-based composition of calcium we assume to have similar x-ray attenuation properties to that of microcalcifications. The figure shows that using titan instead of gold, possible slight differences among phantoms generated during manufacture, would not introduce a noticeable effect in the evaluation of the resulting images.

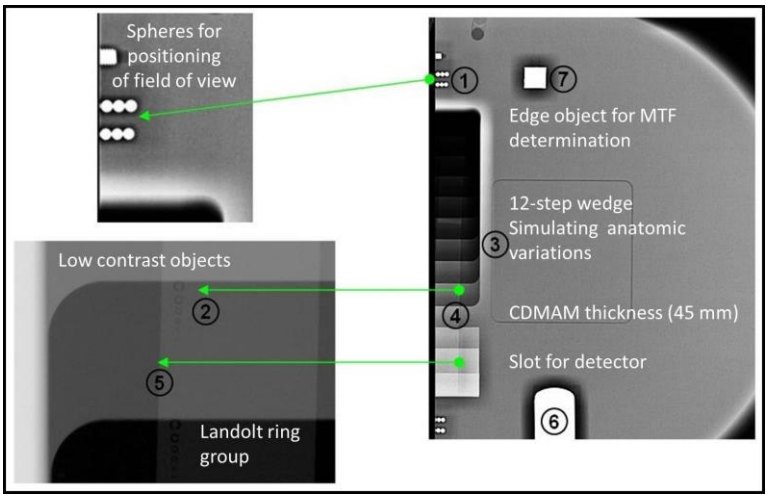


Figure 2; The features of *QUART mam/digi EPQC* phantom.

2.1.1. The evaluation method

For the visual image quality evaluation of the phantom images, the ring gaps must be found and its position within the ring must be noted (either *top*, *down*, *left* or *right*, i.e. four-alternative forced choice task) as long as a gap can be guessed. The result is then confronted with a known truth, provided with the phantom, and the number of gaps that were correctly determined serves as test indicator.
Obviously, the rings at the highest position of the step wedge (corresponding to a large PMMA thickness) show a much lower CNR than the ones at the other end. The twelve-step wedge thus covers a full range between conditions of high detectability and no detectability at all. Additionally, low contrast objects in the shape of digits are provided in each step of the wedge. The ability to distinguish the object clearly in a certain step can serve as a test indicator for low-contrast detectability.
For quick constancy checks, these visual tests can be substituted by an automated evaluation of resolution and noise. The phantom contains a brass square for the determination of the Modulation Transfer Function (MTF) and the Nyquist frequency, as well as two regions of different contrast in each wedge step to measure the contrast-to-noise ratio (CNR) at different base attenuations, which simulate different tissue densities.

2.1.2. Quality control of tomosynthesis systems

The phantom can also be used for the quality control of tomosynthesis systems. For this purpose, the titan foils are contained in two different depths inside the twelve-step wedge. The detectability limits, planar resolution and depth separation in a tomosynthesis reconstruction can be thus assessed in two separate planes using the Landolt rings at those planes. In addition to that, low-contrast resolution in each plane of the wedge can be analysed using the low-contrast digits. The wedge structure enables the check for the influence of different degrees of background attenuation.

2.2. Experimental set-up and data acquisition

We have assessed the image quality provided by a Siemens Mammomat Inspiration FDDM system (Siemens AG Healthcare Sector, Erlangen, Deutschland) at the Medical University of Vienna (Vienna, Austria) by using both CDMAM and QUART phantoms. The CDMAM images were acquired using the arrangement recommended by the phantom manufacturer (test plate with the gold discs between two 20 mm PMMA slabs). The setting for kV and target/filter determined by the automatic exposure control for a block of PMMA 50 mm thick was used. The tube load was manually adjusted in a step by step approach until reaching the acceptable limiting value for the thickness threshold for the 0.1 mm gold disc included in the CDMAM phantom, established at the EPQC V4[9], for human readout. Only this disc diameter was used in our analysis because they are considered to be in the limit of detectability for most digital or screen/film mammography systems. Sixteen images of the CDMAM phantom (which was slightly shifted for each of the images) were acquired for each of the mAs values. 16 images for each of the settings are required for performing automatic readout of the CDMAM images. The images of QUART phantom were taken using the same exposure parameters as for the series of CDMAM images, but only *three images* were analysed for each parameter setting.

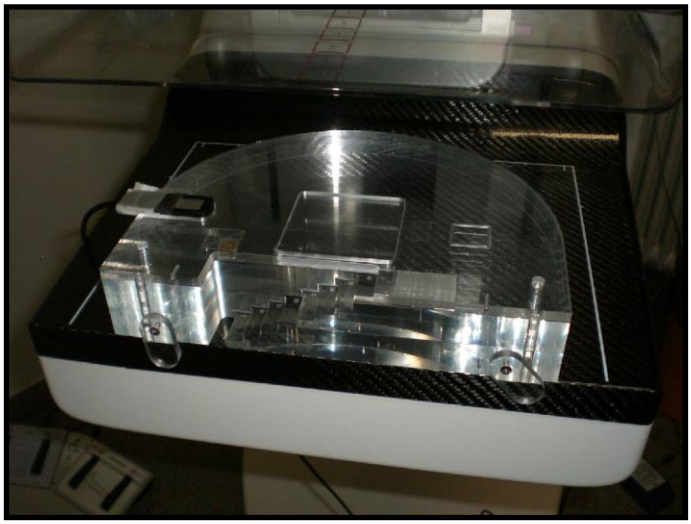


Figure 3; The QUART phantom showing the dosimeter placed in its slot.

The console readings of entrance skin exposure and average glandular dose were used as a measure of dose. These readings were checked with an ionisation chamber to be within 10 % of the actual values. However, for completeness of the QUART quality control procedure, entrance dose was additionally measured by insertion of a small dosimeter (dido2100K, QUART GmbH, Zorneding, Germany) in a specifically designed slot of the phantom (figure 3), and compared to the console readings with another system (Selenia, Hologic Inc., Bedford MA).

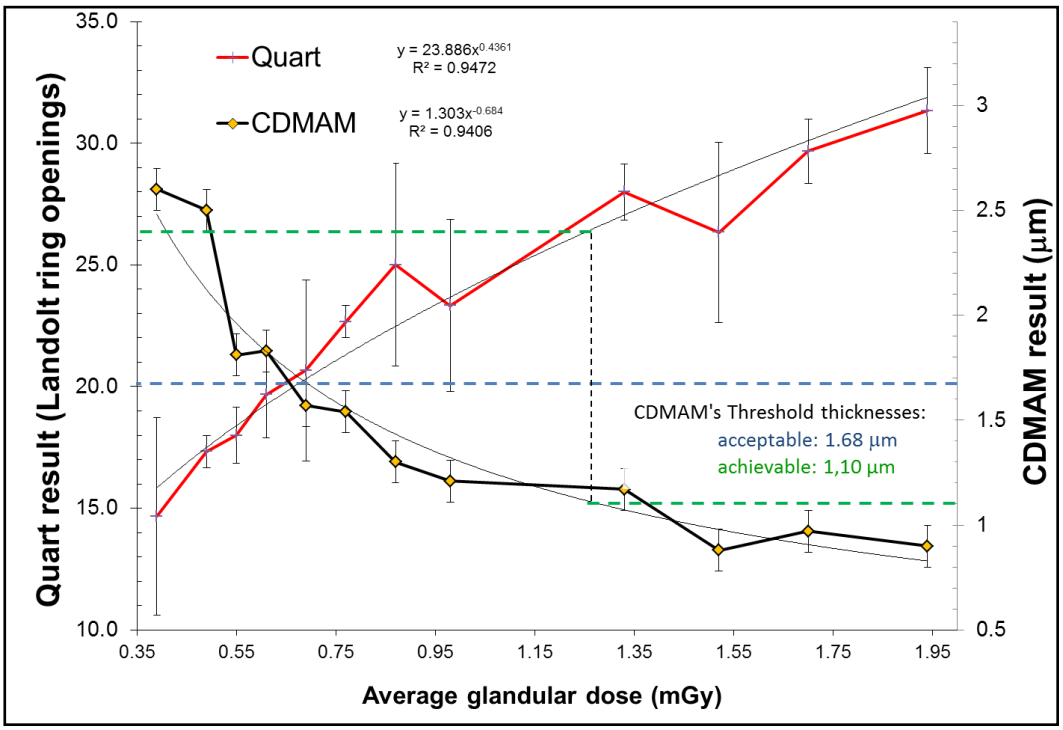
2.3. Data analysis

The CDMAM evaluation using "automatic readout" was carried out according to a combination of CDCOM Version 1.5.2 from EUREF and the CDMAM analysis software from NHS UK [16]. This software also uses "automatic readout" results to project them to "human readout" results, for which CDMAM limiting values are specified.
The QUART images were evaluated using ImageJ (release 1.43u, National Institutes of Health, USA) and following QUART guidelines. Through the evaluation of a group of three different images by three different observers, we checked that the inter-observer variability is more than twice smaller than the variability introduced by shifting the phantom between projections. Therefore, one and the same observer evaluated each of the groups of three acquired images. A 200% zoom and a narrow contrast window were used to find the Landolt rings through all steps of the wedge. For the six rings in each step, the position of the gap was determined (if visible) and then compared to the known truth shown in the manual of the phantom. The number of correct guesses in each step was noted down. The addition of the correctly guessed Landolt ring gaps from steps 1 to 7 ($LRG_{1\text{-}7}$) was used as test indicator.
Using only the first seven steps contemplates the fact that the CDMAM test determines the image quality at exposure parameters defined by the automated exposure of a 50 mm block; we thus use the data from step 4 (corresponding to 46 mm of PMMA) and its three immediate neighbours on each side. The detectability data from the other steps are useful for further tests, such as the determination of the ability of a mammography system to image structures close to the skin line, or the saturation limit of a detector, but not for the present comparison to the CDMAM results.
For the automatic evaluation of the CNR, the same images plus an additional one corresponding to the same exposure parameters, were evaluated using QUART software *MammoPro v6*. The CNR values were obtained for each one of the steps in the phantom. However, for the same reason as above, only the value corresponding to 45 mm PMMA was used in the analysis for the comparison to the CDMAM result.

3. Results and discussion

The readout results from both phantoms are shown in Figure 4. The results show that a good agreement exists between both type of results, i.e. 0.1 mm-thickness threshold and the Landolt sum ($LRG_{1\text{-}7}$ as described above). As it can be seen in Figure 4, the threshold thickness curve cuts the curve of Landolt results at the acceptable value defined in the European Protocol. From these results, a test value equal or above 20 using QUART methodology has been identified with a passed CDMAM test, whereas a value below 20 has been identified with a failure in the CDMAM test. Similarly, an achievable value equal to 26 has been determined from Figure 4 by comparison to the CDMAM achievable threshold thickness.
In all cases, the QUART uncertainty bars indicate twice the standard error in the mean ("2sem", or 95 % confidence interval) of the values obtained after the evaluation of the corresponding images. In the CDMAM results sheet, the uncertainty interval is also expressed as "2sem".

Thanks to the step wedge and the various test objects, the QUART phantom covers the analysis of the image quality associated to a wide range of breast thicknesses and tissue compositions in one shot, which to the knowledge of the authors is not the case of any other phantom. For this reason, it is required a shorter occupation of the mammography system and, in the case of human readout, a shorter evaluation time (a maximum of 5 minutes per image), as well as much less memory space inside the PACS of a hospital or clinic. The automatic evaluation software included with the QUART phantom can be used to carry out quick assurance checks without the need of a visual evaluation of images, thus saving time and reducing observer's susceptibility. As shown in Fig. 4 (bottom), this automatic evaluation also shows a very linear behaviour as a function of dose, as expected, whereas the visual evaluation methods suffer from a certain saturation effect.



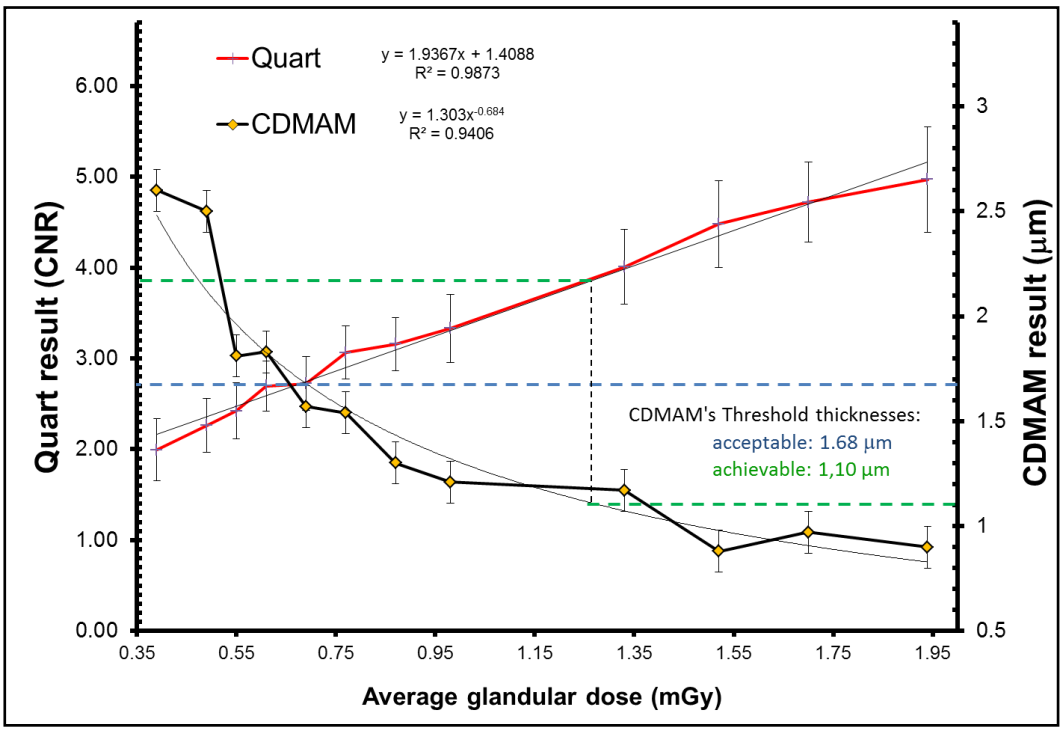


Figure 4; Image quality as a function of average glandular dose (TOP: Landolt ring evaluation, BOTTOM: Contrast-to-noise ratio). In both graphs, the left axis corresponds to QUART's result and the right axis indicates de CDMAM threshold thickness. The QUART result corresponding to the CDMAM acceptable and achievable threshold thicknesses are shown with respectively blue and green horizontal lines. Note that QUART´s result has been obtained by averaging an evaluation of three images whereas the CDMAM result has been obtained by averaging a computer evaluation of 16 images (as recommended by the manufacturer).

Finally, the entrance dose measurements provided by the detector are in good agreement with the readings of entrance skin dose (ESD) shown at the system console. However, a systematic difference of 5.0 ± 0.3 % has been detected with the Hologic system. The investigation of this small difference in the calibration of the system and the detector is in progress.

4. Conclusion

The QUART system for quality control in mammography and tomosynthesis covers a wide range of attenuation backgrounds, and it is effective in terms of manufacture costs, operator time and evaluation effort. The intuitive tests contained in the phantom, i.e. contrast-detail detection of Landolt rings and automatic evaluation of contrast-to-noise ratio, have been compared to the wide-spread CDMAM test. Agreement among all tests has been shown. Acceptable and achievable thresholds for the new phantom have been defined by comparison with the thresholds set by the CDMAM test.

Acknowledgments

The authors wish to thank Elfriede Sturm (Allgemeines Krankenhaus Wien) and Cristina Núñez (Fundación Jiménez Díaz, Madrid) for their help during data acquisition; and the team at the Johannes Gutenberg-Universität Mainz for their help during dose measurements.

Conflict of interest statement

H. de las Heras and B. Tiller work part-time as consultants for QUART GmbH, the company that manufactures the phantom we are introducing in this paper. F. Schöfer is the leader of the scientific development team at the same company. The other authors have no conflict of interest.

References

[1] International standard IEC 62220-1-2, *Medical electrical equipment –Characteristics of digital X-ray imaging devices – Part 1-2: Determination of the detective quantum efficiency – Detectors used in mammography*. Geneva: International Electrotechnical Commission, 2007.

[2] Li Y, Poulos A, McLean D, Rickard M. *A review of methods of clinical image quality evaluation in mammography.* Eur J Radiol. 2010;74(3):e122-31.

[3] Freed M, Badal A, Jennings RJ, de las Heras H, Myers KJ, Badano A. *X-ray properties of an anthropomorphic breast phantom for MRI and x-ray imaging*. Phys Med Biol 2011;56:3513-33.

[4] Carton AK, Vandenbroucke D, Struye L, et al. *Validation of MTF measurement for digital mammography quality control*. Med Phys 2005;32(6):1684-95.

[5] Hendrick RE, Bassett L, Botsco MA, et al. *Mammography Quality Control Manual*. Reston, Va: American College of Radiology, 1999.

[6] Destouet JM, Bassett LW, Yaffe MJ, Butler PF, Wilcox PA. *The ACR's Mammography Accreditation Program: Ten Years of Experience Since MQSA*. J Am Coll Radiol 2005;2(7):585-94.
http://www.acr.org/accreditation/mammography/overview/10yrs_map-msqa.aspx

[7] Rose A, *Vision: Human and Electronic*. Ed. Plenum, New York, 1972.

[8] Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L. *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition—summary document.* Ann Oncol. 2008;19(4):614-22. http://annonc.oxfordjournals.org/citmgr?gca=annonc;19/4/614

[9] Perry N, Broeders M, de Wolf C et al. (eds): *European Commission. European Guidelines for Quality Assurance in Breast Cancer Screening and Diagnosis*. 4th edition. Office for Official Publications of the European Communities, Luxembourg 2006.

[10] Bijkerk KR, Thijssen MAO, Arnoldussen TJM. *Manual CDMAM-phantom type 3.4*., University Medical Centre, Nijmegen, St Radboud, The Netherlands 2002.

[11] Fletcher-Heath L and Van Metter R, *Quantifying the performance of human and software CDMAM phantom image observers for the qualification of digital mammography system*, Proc. SPIE Medical Imaging 2005, 5745, 486–498.

[12] Young K et al. *Evaluation of software for reading images of the CDMAM test object to assess digital mammography systems*. Proc. SPIE Medical Imaging 2008, 6913, 69131C-11.

[13] Flom MC, Heath GG, Takahashi E. *Contour Interaction and Visual Resolution: Contralateral Effects*. Science 1963;142:979-80.

[14] Ludvigh E, Miller JW. *Study of Visual Acuity during the Ocular Pursuit of Moving Test Objects. I. Introduction* J Opt Soc Am. 1958;48(11):799-802.

[15] ICRU - Tissue Substitutes in Radiation Dosimetry and Measurement, Report 44 of the International Commission on Radiation Units and Measurements. Bethesda, MD 1989. See also NIST data at http://physics.nist.gov/PhysRefData/XrayMassCoef/ComTab/bone.html

[16] Young K, Alsager A, Oduko J, et al. *Evaluation of software for reading images of the CDMAM test object to assess digital mammography systems* in Proc. SPIE Medical Imaging 2008;6913:69131C. The CDCOM Software is downloadable from www.euref.org