European Journal of Radiology 117 (2019) 9–14

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Research article

Impact of using the new American College of Radiology digital mammography phantom on quality survey in modern digital mammography systems: Evidence from nationwide surveys in Taiwan

Yi-Shuan Hwang[a,b,*], Hui-Yu Tsai[c], Yu-Ying Lin[b], Ying-Lan Liao[d]

[a] Department of Medical Imaging and Intervention, Chang Gung Memorial Hospital at Linkou, Taoyuan, Taiwan
[b] Department of Medical Imaging and Radiological Sciences, Chang Gung University, Taoyuan, Taiwan
[c] Institute of Nuclear Engineering and Science, National Tsing Hua University, Hsinchu 300, Taiwan
[d] Health Physics Division, Institute of Nuclear Energy Research, Taoyuan, Taiwan

ARTICLE INFO

Keywords:
Digital mammography
Image quality
Average glandular dose
Phantom
Quality control

ABSTRACT

Objective: The purpose of this study was to investigate the impact of the new American College of Radiology (ACR) digital mammography (DM) phantom in evaluating phantom image quality (IQ) and average glandular dose (AGD) in a nationwide survey on DM systems.

Methods: On-site surveys of 239 DM units were conducted in 2017 and 2018, and comparisons were made between ACR screen-film mammography (SFM) phantom and DM phantom for accessing phantom IQ and AGD. The phantom IQ was assessed using the weighted phantom score, considering the size of each detail.

Results: When switching from SFM phantom to DM phantom, no significant difference was found in AGD ($p = 0.06$). The mean weighted phantom score was significantly higher for DM phantom than for SFM phantom in terms of fibers and specks, and so was the total weighted phantom score (DM phantom vs. SFM phantom: 8.61 ± 1.04 vs. 8.23 ± 0.77, $p < 0.0001$). The phantom IQ is thus more precise and can detect small differences when using DM phantom and investigating DM systems, especially for specks and fibers. However, the overall passing rate was lower for DM phantom (84.1%) than for SFM phantom (91.2%). This can be explained by the lower passing rate for mass (84.5%) with the DM phantom.

Conclusion: The ACR DM phantom provides better discernment to assess specks and fibers in DM systems. This study may serve as a reference for implementing a DM quality control program and when conducting large-scale surveys with the new DM phantom in the digital era.

1. Introduction

The American College of Radiology (ACR) screen-film mammography (SFM) phantom was designed earlier as a part of the ACR accreditation program for SFM units in North America and has been used for evaluating image quality (IQ) even after digital mammography (DM) systems were implemented. However, several studies have stated that the ACR SFM phantom is unsatisfactory for assessing the quality of images obtained using DM systems [1–5]. This may be related to DM systems' abilities for superior detection and morphologic characterization of microcalcifications larger than 200 μm in diameter [5,6] and superior detection of masses in the low-contrast range [7]. As stated by Gennaro et al., the ACR SFM phantom enables 68% dose reduction and 99% achievement in readings of the phantom scoring threshold for DM systems [2]. In the survey results of the quality control (QC) program in the American College of Radiology Imaging Network (ACRIN) Digital Mammographic Imaging Screening Trial (DMIST), the total failure rate was less than 1% for IQ evaluated using the SFM phantom when surveying DM systems [8]. Therefore, the ACR introduced a new DM phantom that can be used to test objects with finer gradients and smaller sizes than the SFM phantom. The new phantom has been implemented in the ACR accreditation for two-dimensional (2D) DM and

Abbreviations: ACR, American College of Radiology; ACRIN, American College of Radiology Imaging Network; AGD, average glandular dose; CR, computed radiography; DBT, digital breast tomosynthesis; DM, digital mammography; DR, digital radiography; DMIST, Digital Mammographic Imaging Screening Trial; FFDM, full-field digital mammography; IQ, image quality; QC, quality control; SFM, screen-film mammography; 2D, two-dimensional

* Corresponding author at: Department of Medical Imaging and Intervention, Chang Gung Memorial Hospital at Linkou, 5 Fushing Street, Kweishan, Taoyuan 333, Taiwan.
E-mail address: yishuan@gmail.com (Y.-S. Hwang).

https://doi.org/10.1016/j.ejrad.2019.05.014
Received 31 January 2019; Received in revised form 11 April 2019; Accepted 17 May 2019

digital breast tomosynthesis (DBT) systems [9,10].

In Taiwan, mammography systems have been included in the scope of the Medical Exposure Quality Assurance standards by the Atomic Energy Council since July 1, 2008, and all clinical facilities are required by law to perform QC procedures for mammography systems [11]. To continuously monitor the effectiveness of QC practices and the performance of mammography systems, large-scale surveys with on-site measurements have been annually conducted in Taiwan since 2008. In surveys in this decade, the ACR SFM phantom played a major role in the evaluation of phantom IQ and average glandular dose (AGD). However, all mammography systems were transited to digital systems in 2015 in Taiwan; nonetheless, using the SFM phantom to evaluate phantom IQ in a nationwide QC survey should not cause further discrimination. After the DM phantom was introduced in 2017, the ACR DM phantom was implemented together with the ACR SFM phantom in our survey. To the best of our knowledge, although the newly developed ACR DM phantom is expected to have different discrimination in phantom IQ from the ACR SFM phantom, no study has conducted IQ performance evaluation using the ACR DM phantom for DM systems in a large-scale survey to date. This study compared the performance of the mammography systems for phantom IQ and AGD evaluated using the ACR SFM phantom and ACR DM phantom based on the data of a large-scale survey and investigated the impact of using the new DM phantom in a nationwide survey for DM systems. The results of this study may serve as a reference for the future implementation of a DM QC program by using the ACR DM phantom in the digital era.

2. Materials and methods

In total, 239 DM units, representing more than 70% of the total units in Taiwan, were evaluated on-site between January 2017 and June 2018. Among the 239 systems surveyed, 186 were digital radiography (DR) systems and 53 were computed radiography (CR) systems, and the numbers and distributions of the DM systems for each manufacturer and model are listed in Table 1. Phantom IQ and AGD of DM systems with 2D acquisitions were tested using both an ACR SFM phantom and ACR DM phantom. Fig. 1 presents images of both phantoms.

Table 1
Numbers of the digital mammographic systems and the distribution of each manufacturer and model in the survey between 2017 and 2018 in Taiwan.

Manufacturer/Model	Number
DR systems	186
GE Senographe 2000D	26
GE Senographe DS	14
GE Senographe Essential	20
GE Senographe Pristina	1
Hologic Selenia (with Molybdenum target)	13
Hologic Selenia (with Tungsten target)	46
Hologic Selenia Dimensions	23
Siemens Mammomat Novation	2
Siemens Mammomat Inspiration	23
IMS Giotto Image 3D	11
IMS Giotto Class	1
IMS Giotto Tomo	1
FDR Amulet Innovality	2
Philips MicroDose SI L50	3
CR systems	53
Fuji	31
Carestream	14
Konica	7
Agfa	1
Total	239

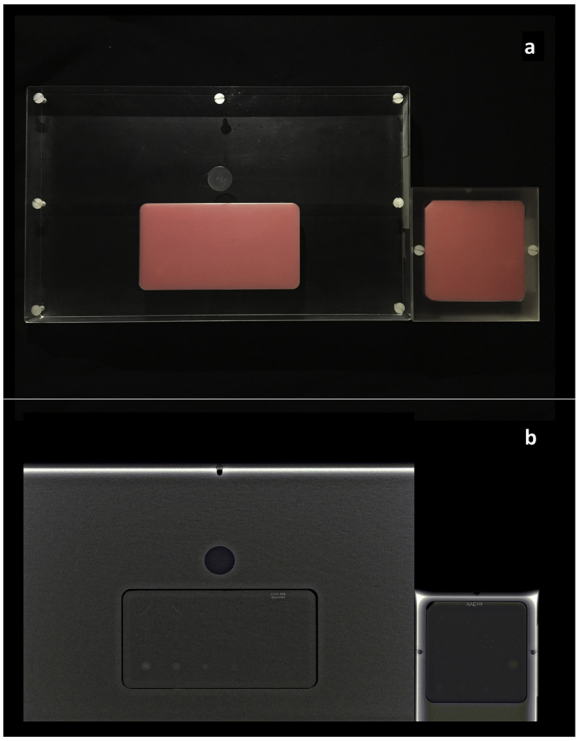

Fig. 1. Appearance (**a**) and the image (**b**) of the ACR SFM phantom (right) and ACR DM phantom (left) used in the nationwide survey.

2.1. ACR SFM phantom and ACR DM phantom

The ACR SFM phantom was designed for the ACR accreditation program of SFM systems, and it measures 10.2 × 10.8 cm^2, simulates a 4.2-cm compressed 50–50 breast, and contains six fibers, five speck groups, and five masses. The ACR DM phantom was designed to test the performance of DM systems based on the ACR DM QC manual [9,10]. It also simulates a 4.2-cm compressed 50-50 breast but has a larger dimension of 19 × 31 cm^2. The test objects in the phantom have finer gradations and smaller sizes and contain six fibers, six speck groups, and six masses.

2.2. AGD

AGD estimation in DM systems is different when using an ACR SFM phantom (based on each manufacturer's QC manual) and an ACR DM phantom (based on the ACR DM QC manual). In the QC protocols of most DM manufacturers, the SFM phantom is placed on the breast support together with the dosimeter, centered 4 cm from the chest wall edge of the image receptor, and with the sensitive level of the dosimeter at the surface of the phantom. In this study, the SFM phantom was exposed using various clinical technical factors for the averaged breast at each surveyed site. According to the protocol described in the ACR DM QC manual, breast entrance exposures were measured without the phantom, and the dosimeter was centered 4 cm from the chest wall edge of the image receptor and positioned at a height of 4.2 cm above the breast support. The exposure techniques were then manually selected for the phantom IQ evaluation of the DM phantom [9,10]. AGDs for both phantoms were calculated using the measured breast entrance exposures, and the corresponding conversion factors were calculated by Dance [9,10,12–14]. The breast entrance exposures were all measured using a 1-cc ionization chamber (Exradin A600 REF 92600, Standard Imaging Inc., Middleton, WI, USA) and an associated multimeter (Piranha model 355, RTI Electronics AB, Sweden) calibrated on mammographic X-ray beam energies. The upper limit of AGD was 3 mGy, as set by the ACR, regardless of the phantom used for estimation.

Table 2
Phantom scoring based on the ACR QC manual scoring method for the SFM phantom and the DM phantom. Specific rule for scoring mass of 0.5 point is improved for the DM phantom.

Details	1 point		0.5 point	
	SFM phantom	DM phantom	SFM phantom	DM phantom
Fiber	• full length visible (≥8 mm long) • 1 break allowed (≤width of fiber) • correct location and orientation		• at least half of length visible (≤5 and < 8 mm long) • 1 break allowed (≤width of fiber) • correct location and orientation	
Specks	• 4–6 specks visible • correct locations		• 2–3 specks visible • correct locations	
Mass	• density difference visible • border is continuous and generally circular (≥3/4 border) • correct location		• density difference visible • border is not generally circular • correct location	• density difference visible • border is not continuous or generally circular (≥1/2 and < 3/4 border) • correct location

2.3. Phantom IQ

Both the ACR SFM (CIRS model 015, CIRS, Norfolk, VA, USA) and ACR DM phantoms (Gammex model 145FFDM, Gammex Inc., Middleton, WI, USA) were applied for phantom IQ evaluations with 2D acquisitions using the clinical technique for an averaged breast at each surveyed site. Each phantom image was scored by two qualified medical physicists who attributed a score of 1, 0.5, or 0 according to their visibility for each type of detail (fiber, specks, and mass) based on the ACR scoring protocol of each phantom under optimal viewing conditions [9,10,15]. The standard phantom scores were thus obtained, and according to ACR criteria, the passing scores were the scores of four fibers, three specks, and three masses for the SFM phantom [15] and two fibers, three specks, and two masses for the DM phantom [9,10]. The scoring methods were slightly different between the ACR SFM phantom and the ACR DM phantom regarding the score of 0.5 for masses. The scoring keys for each detail for both phantoms required by the ACR are listed in Table 2. The standard phantom total score was calculated by adding the standard phantom scores of fibers, specks, and masses.

To compare the phantom IQ evaluation using the SFM and DM phantoms more directly, we transferred the standard phantom score of each detail to the weighted phantom score from the second step by using a weighted index. Weighted indexes were calculated using a modified version of Gennaro et al.'s weighting method, which is intended to increase the quantization levels of phantom scores by weighting details by their size [2]. We used the largest size of each detail of the SFM phantom as a reference size and divided the reference size by the size of each detail for both the SFM and DM phantoms, as shown in the tabulated weighted index in Table 3. The greatest weighted index matched the smallest detail compared with the reference size. After scoring each detail, a weighted phantom score of each detail could thus be obtained as the weighted index of the corresponding standard phantom score. For the standard phantom score of 0.5, the weighted phantom score was calculated as the average value between the weighted index below and above the corresponding standard phantom score of the detail. The weighted phantom total score was also calculated as the sum of the weighted phantom scores of fibers, specks, and masses. Finally, the difference between the weighted phantom score of each detail in SFM and DM phantom evaluations was calculated. It was defined as the weighted phantom score evaluated using the DM phantom subtracted by the weighted phantom score evaluated using the SFM phantom.

Table 3
Detailed sizes and the related weight indexes for the ACR SFM phantom and DM phantom. Asterisk denotes the minimum visible details required by the ACR for the SFM and DM phantoms.

Phantom			Detail					
			1	2	3	4	5	6
SFM	Fiber	Size (mm)	1.56	1.12	0.89	0.75*	0.54	0.4
		Weighted index	1.00	1.39	1.75	2.08	2.89	3.90
	Specks	Size (mm)	0.54	0.4	0.32*	0.24	0.16	–
		Weighted index	1.00	1.35	1.69	2.25	3.38	–
	Mass	Size (mm)	2	1	0.75*	0.5	0.25	–
		Weighted index	1.00	2.00	2.67	4.00	8.00	–
DM	Fiber	Size (mm)	0.89	0.75*	0.61	0.54	0.4	0.3
		Weighted index	1.75	2.08	2.56	2.89	3.90	5.20
	Specks	Size (mm)	0.33	0.28	0.23*	0.2	0.17	0.14
		Weighted index	1.64	1.93	2.35	2.70	3.18	3.86
	Mass	Size (mm)	1	0.75*	0.5	0.38	0.25	0.2
		Weighted index	2.00	2.67	4.00	5.26	8.00	10.00

2.4. Statistical analysis

All statistical analyses were performed using SPSS software (version 18.0, SPSS, Chicago, Illinois, USA). Differences in AGDs and weighted phantom scores between the SFM and DM phantoms were analyzed using the paired *t*-test and Wilcoxon-signed rank test, respectively. Differences were considered significant at $p < 0.05$.

3. Results

3.1. AGD surveys using the ACR SFM and ACR DM phantoms

The mean AGD for all 239 systems surveyed was 1.57 ± 0.51 mGy (range, 0.28–3.36 mGy) and 1.55 ± 0.51 mGy (range, 0.32–3.47 mGy) for the ACR SFM and ACR DM phantoms, respectively. The 75th percentile was estimated as 1.87 mGy from the AGD distribution evaluated using the SFM and DM phantoms. AGD distributions for the SFM and DM phantoms are presented in Fig. 2a. Only one system had AGD above the criterion of 3 mGy for both phantoms (passing rate = 99.6%).

For the comparison of exposure techniques when measuring AGD, the same target, filter, and kV were used for 76.2% of the units (182 of 239) for the exposure of the SFM and DM phantoms by using clinical automatic exposure control (AEC) techniques. However, no significant difference was observed between AGD measured using the SFM or DM phantom ($p = 0.06$).

3.2. Phantom IQ surveys using the ACR SFM and ACR DM phantoms

The mean standard phantom scores for the ACR SFM phantom were 4.4 ± 0.5, 4.0 ± 0.2, and 3.7 ± 0.4 for fibers, specks, and masses, respectively. For the ACR DM phantom, the mean scores were 3.3 ± 0.8, 4.4 ± 0.5, and 2.3 ± 0.6 for fibers, specks, and masses, respectively. The mean standard phantom total score was 12.0 ± 0.8 (range, 7.5–13.5) and 10.0 ± 1.5 (range, 5–13.5) for the SFM and DM phantoms, respectively. The distributional range of the standard phantom total score was larger for the DM phantom than for the SFM

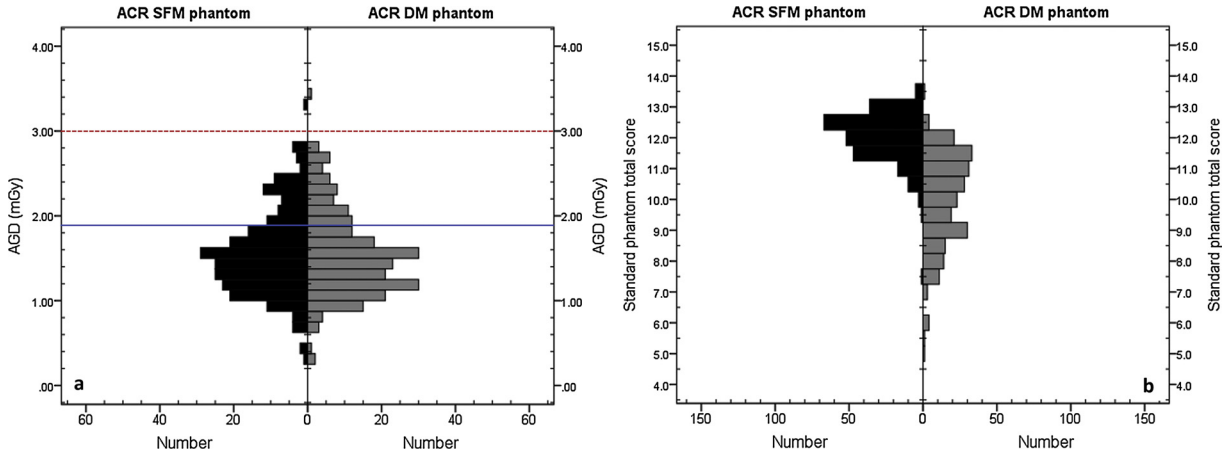

Fig. 2. Histogram of the AGD (**a**) and standard phantom total score (**b**) investigated using the ACR SFM and ACR DM phantoms in the nationwide survey. The red dashed line and blue solid line indicate the upper limit and the 75th percentile value of AGD, respectively.

phantom (Fig. 2b). According to ACR scoring criteria, the passing rate of the SFM phantom for the 239 systems was 95.4%, 100.0%, and 95.4% for fibers, specks, and masses, respectively. For the DM phantom, the passing rate was 99.2%, 100.0%, and 84.5% for fibers, specks, and masses, respectively. The overall passing rate for the standard phantom score was 91.2% and 84.1% for the SFM and DM phantoms, respectively. The passing rate reached 100% for the specks of both phantoms. However, when switching from the SFM phantom to the DM phantom, the passing rate slightly increased for the fiber scores and largely decreased from for the mass scores.

The distributions of weighted phantom scores of fibers, specks, masses, and the weighted phantom total score for the SFM and DM phantoms are presented in Fig. 3a–d. The mean weighted phantom score for fibers and specks was significantly higher for the DM phantom than for the SFM phantom (DM phantom vs. SFM phantom: 2.64 ± 0.35 vs. 2.43 ± 0.36 for fibers, $p < 0.0001$, and 2.91 ± 0.25 vs. 2.24 ± 0.10 for specks, $p < 0.0001$). The weighted phantom score was also more discriminating for the DM phantom than for the SFM phantom (2.08–2.89 and > 2.25 for fibers and specks, respectively; Fig. 3a and b). By contrast, the overall distribution of the weighted phantom score was significantly lower for the masses of the DM phantom when compared with those of the SFM phantom ($p < 0.0001$). The mean weighted phantom scores were 3.06 ± 0.70 and 3.57 ± 0.53 for the DM and SFM phantoms, respectively (Fig. 3c). However, as demonstrated in Fig. 3d, the mean weighted phantom total score was still higher for the DM phantom (8.61 ± 1.04) than for the SFM phantom (8.23 ± 0.77) ($p < 0.0001$). The distributional range of the weighted phantom total score was larger for the DM phantom (4.43–11.42) than for the SFM phantom (4.94–10.73). Differences in the weighted phantom scores between the DM and SFM phantoms are provided in Fig. 4. The results revealed that the weighted phantom score was higher or equal for 194, 238, and 88 units for fibers, specks, and masses, respectively, when using the DM phantom rather than using the SFM phantom.

4. Discussion

Both the SFM and DM phantoms simulate a 4.2-cm compressed averaged breast but with different dimensions. This difference in dimensions may result in an uneven compressed thickness, which may cause the surveyed systems to be exposed differently to these two phantoms, according to the AEC algorithm employed. In addition, the geometric setup for AGD estimations using these two phantoms was different. As demonstrated by Ng et al., the measured exposure may lead to an approximate 1.3% increase when setting the phantom next to the ion chamber, and the exposure may lead to an 8% decrease when measuring off-central axis [16]. However, in our survey, no significant difference was observed between the AGDs measured using the SFM phantom and those measured using the DM phantom. Switching from the SFM phantom to the DM phantom as well as using a different setup for the AGD measurement did not cause a significant difference in dose performance.

When comparing the passing rate for the standard phantom score investigated using these two phantoms, the passing rate for fibers was slightly increased when switching from the SFM phantom to the DM phantom in our survey, even when using the same fiber size criterion of 0.75 mm. For specks, although the visualized size criterion decreased from 0.32 mm with the SFM phantom to 0.23 mm with the DM phantom, the passing rate still reached 100% for both phantoms. High passing rates for specks may result from the higher detectability of DM systems for microcalcifications [5,17]. For masses, the passing rate largely decreased when switching from the SFM phantom to the DM phantom. Although the mass visualized size criterion remained the same (0.75 mm), the specific rules for scoring masses were improved by investigating phantom IQ using the DM phantom. The well-defined criteria for the score of 0.5 for masses when using the DM phantom caused a decrease in the mass scores. Among the 37 systems with the mass scores of < 2 using the DM phantom, 73.0% (27 of 37) were CR. This may be attributed to image nonuniformity or significant artifacts in CR images that cause the noncircular appearance of a low-contrast mass in the DM phantom. The decrease in the overall passing rate for image IQ was mainly due to the relative lower mass score obtained using the DM phantom.

The weighted phantom scores were derived from the standard phantom scores by considering the size of each detail. Because of the finer gradient between the 0.54 and 0.75 mm for fiber sizes, more scales of the weighted phantom scores between 2.08 and 2.89 exist for the DM phantom than for the SFM phantom (Fig. 3a). The higher weighted phantom score for the DM phantom may result from the smaller size visualized or finer gradient of the details. Therefore, the DM phantom provides better discrimination for digital systems with a small difference of IQ in fibers. The scoring criteria of specks were stricter when using the DM phantom than when using the SFM phantom. However, the weighted phantom scores showed higher distribution, more scales, and higher mean value for the DM phantom. As illustrated in Fig. 3b, the weighted phantom score indicated higher distribution for the specks of DM phantom. Although the passing rate reached 100% for both phantoms, the DM phantom still showed better discrimination for specks among the digital systems because of its smaller size and finer gradient design.

Fibers and specks are well-defined objects in phantoms; therefore, different observers can easily agree on the visible number of fibers and

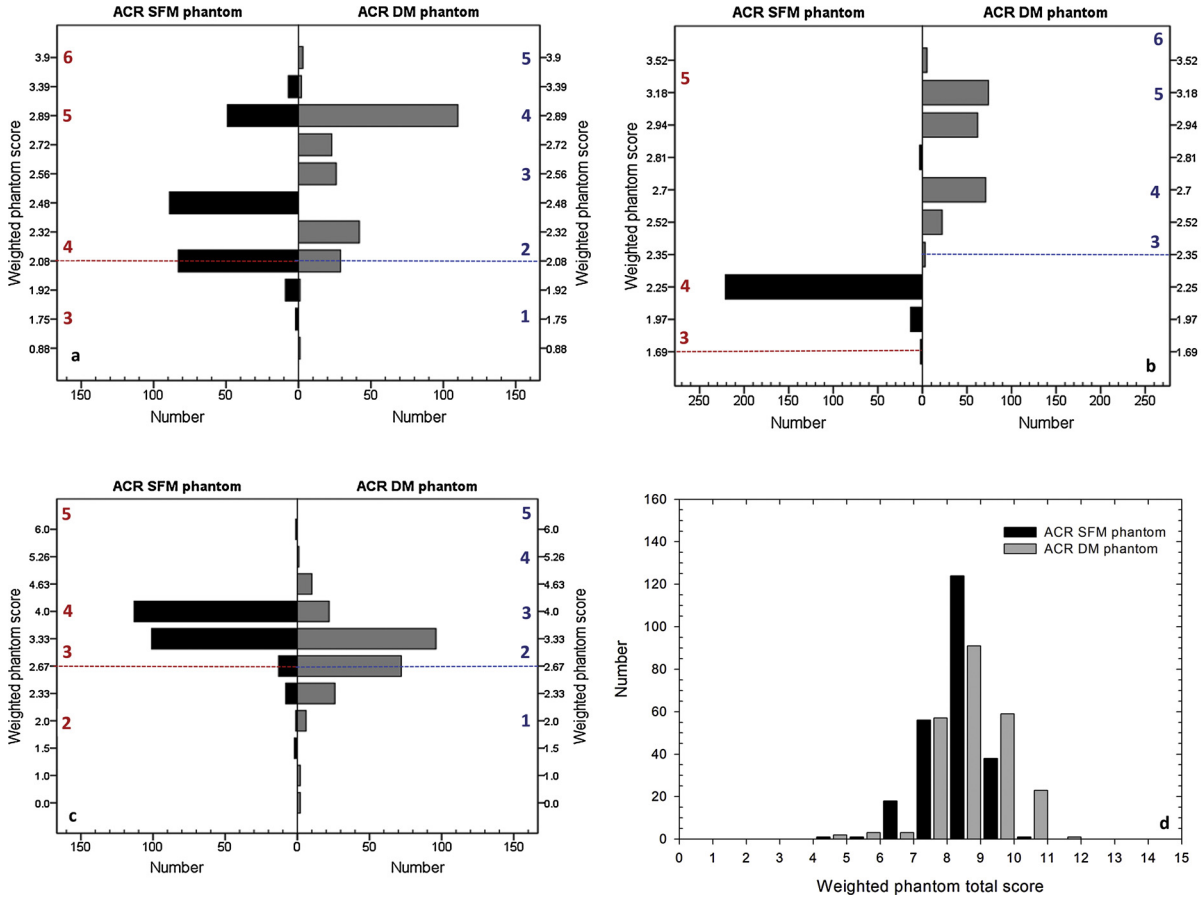

Fig. 3. Comparisons of the weighted phantom scores by using the ACR SFM phantom and ACR DM phantom. The mean weighted phantom score is significantly higher for fibers (**a**) and specks (**b**) and is significantly lower for mass (**c**) when using the DM phantom as compared to SFM phantom. The distributions of the weighted phantom total score evaluated by using these two phantoms show that the mean value and the distributional range of the weighted phantom total score are higher and larger for the DM phantom than SFM phantom (**d**). In figure **a–c**, the dashed lines illustrate the scoring criteria for each detail of each phantom. The integers in the inward of the Y-axis indicate the corresponding standard phantom score for each detail of each phantom.

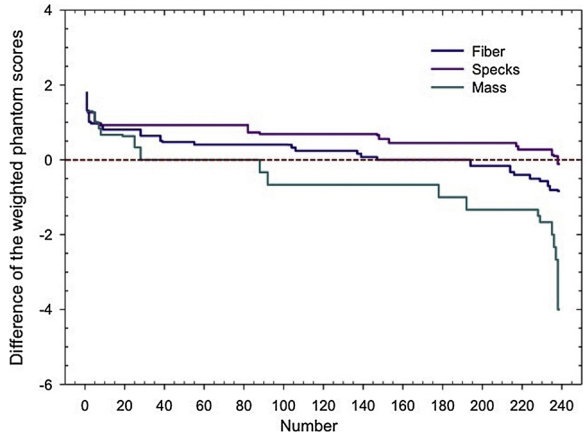

Fig. 4. Differences of the weighted phantom score of fibers, specks, and masses evaluated between the ACR DM phantom and ACR SFM phantom. The positive value of the difference between the DM phantom and SFM phantom indicates that the weighted score is higher when using the DM phantom. For the same digital system, a higher weighted phantom score estimated by the DM phantom implies that a smaller visualized size of the details was scored.

specks using the same scoring method. However, observers may have different subjective criteria for the evaluation of the visibility of masses [3]. In Taiwan, the weighted phantom scores of masses are usually between 2 and 4. However, the scales between these corresponding ranges are the same for both phantoms. This implies that in Taiwan, both the DM and SFM phantoms show comparable discrimination for masses (Fig. 3c). In addition, the well-defined criteria for the score of 0.5 for masses cause a lower mean weighted phantom mass score for the DM phantom than for the SFM phantom. Even though the mean weighted phantom score of masses is lower for the DM phantom, the weighted phantom total score for the DM phantom still shows a larger distribution and has a higher mean value than that for the SFM phantom. In Fig. 4, the positive value for the difference indicated a higher weighted phantom score for the DM phantom. For the same digital system, a higher weighted phantom score for the DM phantom implied that the smaller visualized size of the details was scored. When considering the discrimination performance for the phantom scores, compared with the SFM phantom, the DM phantom showed the highest discrimination for specks, followed by fibers. Because of the various specifications of the details and slightly different scoring methods utilized between the two phantoms, the performance of the mammography systems for phantom IQ was different when switching from the SFM phantom to the DM phantom.

Technical evolution in recent years has led to the wide usage of DBT systems, and studies have indicated that DBT improves the visibility of lesions superposed by the overlying tissue [18,19] and architectural distortions [20,21]. Meyblum et al. compared phantom scores obtained using the ACR SFM phantom between 2D full-field DM (FFDM) and DBT; they found no significant difference in the visibility of details between the 2D and DBT modes; however, when the phantom with a textured background was used, the visibility of fibers and masses was

higher for DBT than for 2D images [22]. As described by Cockmartin et al., test objects embedded in a homogeneous background are not adequate for assessing IQ between 2D FFDM and DBT systems. In their study using a designed structured phantom, Cockmartin et al. found that microcalcification detection thresholds were similar for 2D FFDM and DBT, but detection of spiculated and nonspiculated masses significantly increased for DBT than for 2D FFDM [23].

Better discrimination for fibers and specks with the ACR DM phantom investigated in this study validates better differentiation in IQ between DM systems when implemented with DM phantom. In the national quality survey in the digital era, a small quality difference in detecting architectural distortions (fibers) and microcalcifications was thus obtained for the DM phantom. Using the DM phantom, small IQ variations could also be detected in periodical QC over time. Better discrimination for smaller specks is important because DM systems are intrinsically superior in detecting microcalcifications than SFM systems [5,6]. In addition to applying the DM phantom for 2D DM systems [9], the ACR also applied the DM phantom for DBT QC in 2018 [10]. DBT should provide better visualization of architectural distortions (fibers) and masses; although the homogeneous background design of the phantom is suitable for periodic QC [24], a phantom with a uniform background may not effectively represent these superior properties of DBT systems when compared with 2D FFDM systems [22,24,25]. The impact of utilizing the ACR DM phantom, rather than using the SFM phantom, in DBT systems and 2D digital systems in evaluating phantom IQ may be similar. For a comprehensive large-scale quality survey, DM phantoms as well as phantoms with a nonhomogeneous textured background should be applied to DBT systems in the future.

5. Conclusions

This study analyzed data from nationwide on-site surveys on DM systems in Taiwan between 2017 and 2018. It is the first large-scale survey using the new ACR DM phantom. When switching from the SFM phantom to the DM phantom, no significant difference was discovered in AGD estimations between these phantoms. Phantom IQ showed better discrimination for small differences when using the DM phantom than when using the SFM phantom for DM systems, especially for specks and fibers. However, no improved discrimination was noted for assessing small changes in the IQ of masses between the systems. According to the survey results in Taiwan, the decreased passing rate for masses caused the overall passing rate to be lower with the DM phantom than with the SFM phantom. The results presented in this study can serve as reference when extending the new ACR DM phantom to the mammography QC program in the digital era.

IRB statement

This study doesn't involve any human subject or animal, and no IRB approval was required.

Conflict of interest

The authors declare that there is no conflict of interest.

Acknowledgements

The authors thank all the staffs of the mammography facilities that participated in the surveys for their help and cooperation. The authors wish to acknowledge Yu-En Wu for the assistance in on-site measurements and data analysis. This work was supported by grants from the Atomic Energy Council of Taiwan (AEC10409031L).

References

[1] M. Figl, F. Semturs, M. Kaar, R. Hoffmann, H. Kaldarar, P. Homolka, et al., Dose sensitivity of three phantoms used for quality assurance in digital mammography, Phys. Med. Biol. 58 (2013) N13–N23.
[2] G. Gennaro, L. Katz, H. Souchay, C. Alberelli, C. di Maggio, Are phantoms useful for predicting the potential of dose reduction in full-field digital mammography? Phys. Med. Biol. 50 (2005) 1851–1870.
[3] W. Huda, A.M. Sajewicz, K.M. Ogden, E.M. Scalzetti, D.R. Dance, How good is the ACR accreditation phantom for assessing image quality in digital mammography? Acad. Radiol. 9 (2002) 764–772.
[4] M.J. Yaffe, A.K. Bloomquist, G.E. Mawdsley, E.D. Pisano, R.E. Hendrick, L.L. Fajardo, et al., Quality control for digital mammography: part II recommendations from the ACRIN DMIST trial, Med. Phys. 33 (2006) 737–752.
[5] S.E. Song, B.K. Seo, A. Yie, B.K. Ku, H.Y. Kim, K.R. Cho, et al., Which phantom is better for assessing the image quality in full-field digital mammography?: American College of Radiology Accreditation phantom versus digital mammography accreditation phantom, Korean J. Radiol. 13 (2012) 776–783.
[6] K.B. Krug, H. Stützer, R. Girnus, M. Zähringer, A. Goßmann, G. Winnekendonk, et al., Image quality of digital direct flat-panel mammography versus an analog screen-film technique using a phantom model, AJR Am. J. Roentgenol. 188 (2007) 399–407.
[7] K.B. Krug, H. Stützer, R. Schröder, J. Boecker, J. Poggenborg, K. Lackner, Image quality of digital direct flat-panel mammography versus an analog screen-film technique using a low-contrast phantom, AJR Am. J. Roentgenol. 191 (2008) W80–W88.
[8] A.K. Bloomquist, M.J. Yaffe, E.D. Pisano, R.E. Hendrick, G.E. Mawdsley, S. Bright, et al., Quality control for digital mammography in the ACRIN DMIST trial: part I, Med. Phys. 33 (2006) 719–736.
[9] E.A. Berns, J.A. Baker, L.D. Barke, et al., Digital Mammography Quality Control Manual, American College of Radiology, Reston, Va, 2016.
[10] E.A. Berns, D.E. Pfeiffer, P.F. Butler, et al., Digital Mammography Quality Control Manual, American College of Radiology, Reston, Va, 2018.
[11] Y.S. Hwang, H.Y. Tsai, C.C. Chen, P.K. Tsay, H.B. Pan, G.C. Hsu, et al., Effects of quality assurance regulatory enforcement on performance of mammography systems: evidence from large-scale surveys in Taiwan, AJR Am. J. Roentgenol. 201 (2013) W307–W312.
[12] D.R. Dance, C.L. Skinner, K.C. Young, J.R. Beckett, C.J. Kotre, Additional factors for the estimation of mean glandular breast dose using the UK mammography dosimetry protocol, Phys. Med. Biol. 45 (2000) 3225–3240.
[13] D.R. Dance, K.C. Young, R.E. van Engen, Further factors for the estimation of mean glandular dose using the United Kingdom, European and IAEA breast dosimetry protocols, Phys. Med. Biol. 54 (2009) 4361–4372.
[14] D.R. Dance, Monte Carlo calculation of conversion factors for the estimation of mean glandular breast dose, Phys. Med. Biol. 35 (1990) 1211–1219.
[15] American College of Radiology, American College of Radiology (ACR) Mammography Quality Control Manual, American College of Radiology, Reston, Va, 1999.
[16] K.H. Ng, R.J. Aus, L.A. DeWerd, J.R. Vetter, Entrance skin exposure and mean glandular dose: effect of scatter and field gradient at mammography, Radiology 205 (1997) 395–398.
[17] U. Fischer, F. Baum, S. Obenauer, S. Luftner-Nagel, D. von Heyden, R. Vosshenrich, et al., Comparative study in patients with microcalcifications: full-field digital mammography vs screen-film mammography, Eur. Radiol. 12 (2002) 2679–2683.
[18] S. Vedantham, A. Karellas, G.R. Vijayaraghavan, D.B. Kopans, Digital breast tomosynthesis: state of the art, Radiology 277 (2015) 663–684.
[19] P. Elangovan, A. Mackenzie, D.R. Dance, K.C. Young, K. Wells, Lesion detectability in 2D-mammography and digital breast tomosynthesis using different targets and observers, Phys. Med. Biol. 63 (2018), https://doi.org/10.1088/1361-6560/aabd53095014–11.
[20] H.R. Peppard, B.E. Nicholson, C.M. Rochman, J.K. Merchant, R.C. Mayo III, J.A. Harvey, Digital breast tomosynthesis in the diagnostic setting: indications and clinical applications, Radiographics 35 (2015) 975–990.
[21] N. Tirada, G. Li, D. Dreizin, L. Robinson, G. Khorjekar, S. Dromi, et al., Digital breast tomosynthesis: physics, artifacts, and quality control considerations, Radiographics 39 (2019) 413–426.
[22] E. Meyblum, F. Gardavaud, T.-H. Dao, V. Fournier, P. Beaussart, F. Pigneur, et al., Breast tomosynthesis: dosimetry and image quality assessment on phantom, Diagn. Interv. Imaging 96 (2015) 931–939.
[23] L. Cockmartin, N.W. Marshall, G. Zhang, K. Lemmens, E. Shaheen, C. Van Ongeval, et al., Design and application of a structured phantom for detection performance comparison between breast tomosynthesis and digital mammography, Phys. Med. Biol. 62 (2017) 758–780.
[24] J. Sage, K.L. Fezzani, I. Fitton, L. Hadid, A. Moussier, N. Pierrat, et al., Experimental evaluation of seven quality control phantoms for digital breast tomosynthesis, Phys. Med. 57 (2019) 137–144.
[25] L. Cockmartin, N.W. Marshall, C. van Ongeval, G. Aerts, D. Stalmans, F. Zanca, et al., Comparison of digital breast tomosynthesis and 2D digital mammography using a hybrid performance test, Phys. Med. Biol. 60 (2015) 3939–3958.