Virtual Tools for the Evaluation of Breast Imaging: State-of-the Science and Future Directions

Predrag R. Bakic[1(✉)], Kyle J. Myers[2], Stephen J. Glick[2], Andrew D.A. Maidment[1], and The members of AAPM Task Group 234[1,2]

[1] Department of Radiology, University of Pennsylvania, Philadelphia, PA, USA
Predrag.Bakic@uphs.upenn.edu
[2] U.S. Food and Drug Administration, Silver Spring, MD, USA

Abstract. The beginning of this century saw the development of simulation methods for the evaluation of breast imaging, motivated by the limitations of conventional clinical trials. This has led to the formation of AAPM Task Group on Virtual Tools for the Validation of 3D/4D X-ray Breast Imaging Systems (TG234), gathering researchers from academia, industry, and government, interested in the development, testing, and adoption of these tools. TG234 is currently finalizing its report. The report has been designed as an experiential guide through the steps of simulating breast anatomy, image acquisition, image interpretation, and analysis. TG234 activities include disseminating the idea of virtual clinical trials through numerous focused conference sessions and AAPM annual meeting symposia. This paper reflects our desire to initiate wider discussion about the future directions in the development of virtual tools for the design and evaluation of novel breast imaging systems.

Keywords: Mammography · Digital breast tomosynthesis · Breast imaging simulation · Anthropomorphic phantoms · Virtual clinical trials · Model observers

1 Introduction

Modern x-ray breast imaging systems are complex, with designs based upon a multitude of modifiable parameters, all of which affect clinical performance. Conventional evaluation of these systems requires clinical imaging trials, which usually compare various image quality descriptors between repeated images of a large group of subjects acquired using different imaging systems (or different system parameters). Such studies are limited by their cost, duration, and the risk to participating subjects from repeated use of ionizing radiation.

An alternative approach based upon simulation of the complete breast imaging chain has been introduced in the last decade. The simulated components of the breast imaging chain include the object model (breast anatomy with or without simulated

The AAPM Task Group 234 roster is listed on https://www.aapm.org/org/structure/?committee_code=TG234.

A. Tingberg et al. (Eds.): IWDM 2016, LNCS 9699, pp. 518–524, 2016.
DOI: 10.1007/978-3-319-41546-8_65

abnormalities), the imaging process model (including the breast positioning, image acquisition, and post-processing or reconstruction of acquired images), and the interpretation of the image data using mathematical observer models. A generalized simulation pipeline is illustrated by Fig. 1. Such a simulation pipeline can be used to design and perform virtual clinical trials (VCTs), based upon performance metrics of specific tasks.

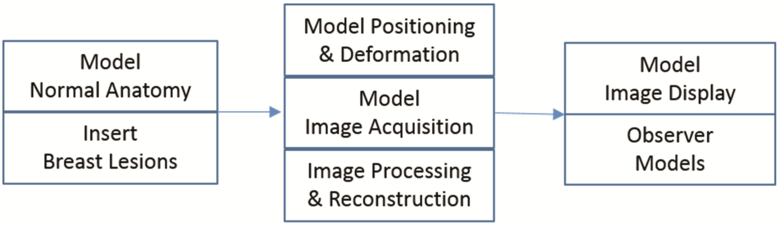


Fig. 1. Flow-chart of the breast imaging simulation for virtual evaluation.

Based upon the increasing interest in virtual evaluation of breast imaging systems, a dedicated Task Group on Virtual Tools for the Validation of 3D/4D X-ray Breast Imaging Systems (TG234) was founded by the American Association of Physicists in Medicine (AAPM) in 2012. [1] TG234 includes 24 researchers from academia, industry, and government agencies interested in the development, testing, and adoption of these tools.

The main task of TG234 is the preparation of a report, which is expected to be completed later in 2016. TG234 took an active role in disseminating the simulation approach and VCT concept throughout the breast imaging research community, by organizing focused symposia during the 2013 AAPM Annual Meeting in Indianapolis, IN [2], 2014 AAPM Annual Meeting in Austin, TX [3], and the upcoming 2016 AAPM Annual Meeting in Washington, D.C. [4], as well as by introducing a new AAPM abstract submission category related to the virtual evaluation of breast imaging. This has also motivated presentation at 2016 IWDM to provide a forum to review the task group activities and current challenges, the regulatory perspective on VCTs, and an opportunity to receive suggestions from the IWDM community about future directions of TG234.

2 State-of-the-Science of the Virtual Evaluation of Breast Imaging Systems

The TG234 report has been designed in the form of experiential guide, rather than prescriptions. It summarizes the state-of-the-science for the virtual evaluation of breast imaging systems, based upon the experience of the task group members' labs and discussions at our meetings and teleconferences. The report roughly follows the structure of the simulation pipeline, shown in Fig. 1, focusing on the steps of simulating breast anatomy, image acquisition, and interpretation.

2.1 Simulation of the Breast Anatomy (The Object Model)

The quality and realism of the simulated breast anatomy directly affects the validity of related VCTs, and at the same time critically influences the acceptance of the VCT concept in the research community. The TG234 report discusses two complementary views on phantom realism, which differ in the scope and range of the required tests of realism, namely:

1. *The task-based view* assumes that the realism depends on the specific tasks the phantom is designed for; therefore, the tests should be specific to individual evaluation tasks; and
2. *The reader-based view* assumes that the realism should be validated through radiologists' comprehensive visual assessment; thus, passing a so-called "fool-the-reader" test would warrant wide use of such assessed phantoms.

The report identifies various levels and types of realism validation, and reviews numerous examples of the realism validation from published simulation studies.

2.2 Simulation of the Breast Imaging Chain (The Imaging Acquisition Model)

Simulation of the image formation process by which the virtual phantom in object space is processed by each component of the imaging system, resulting in the detected data (synthetic images) is a necessary procedure in breast imaging VCTs. Acknowledging the complexity of the breast imaging process (which includes breast positioning and/or compression, acquisition, image pre-processing, post-processing and/or reconstruction), the TG234 report focuses primarily on the physics of projecting through computer simulated breast phantoms based upon the use of ray tracing or Monte-Carlo methods.

The effects of object discretization (i.e., phantom voxel size selection), and discrete or continuous simulation, are treated systematically, as parts of the imaging process simulation. To that end, we have introduced the simulation function, which depends on multiple parameters, some directly related to discretization and others not. The simulation function is used to compare the variation in the study endpoint introduced by different levels of discretization, against variations arising from other sources in the simulation.

2.3 Simulation of the Breast Image Interpretation

The last stage of a breast imaging VCT pipeline models the performance of a particular, relevant clinical task (e.g., lesion detection or the detection of micro calcification clusters), making use of the synthesized imaging data in either the projected data domain or following image reconstruction. This stage requires the selection and implementation of mathematical model observers. The task group report discusses possible performance tasks, choices for models for observers, and figures of merit for summarizing system performance. The report also reviews methods for estimating the uncertainty in a figure of merit, which is an essential component for hypothesis testing.

Finally, the TG234 report reviews the importance of VCTs from an industry perspective, based upon the experience of designing and utilizing the simulation methods during the development of breast imaging systems, as well as the potential use of computational modeling data in support of regulatory submissions. The U.S. Food and Drug Administration has issued draft guidance on the reporting of data from computational modeling studies in support of premarket submissions, and has made use of task-based evaluations using model observer studies in other imaging applications [5].

3 Challenges of Using Virtual Tools in Breast Imaging

Regular TG234 teleconferences and face-to-face task group meetings were instrumental in identifying the current challenges in breast imaging VCTs. These challenges have been included in the report to stimulate needed research in the field. As an illustration, here we briefly review two challenges related to the optimal selection of phantom voxel size and the training of model observers.

3.1 Optimal Size of the Phantom Voxels

Very early in the design of a VCT study, researchers are faced with the following question: *What size voxels should be used in simulation to ensure that the results are not dependent on the voxel size, or how finely should a numerical line integration routinely sample the phantom for stable integrals*? The answer is constrained by numerous factors related to various components of the breast imaging simulation flowchart (Fig. 1), including:

1. The size of the anatomical structure's detail that is being analyzed by the simulation study;
2. The trade-off between the discretization noise due to voxelizaton, and the image acquisition noise;
3. Practical limitations related to the computational time, data transfer time, and the memory storage for a given simulation platform; and
4. The accuracy required for the simulation study endpoint.
   These complex and even contradictory constraints cause the question of the optimal voxel size to persist up to the present, which will be acknowledged in the TG234 report.

3.2 Model Observer Training

Model observers are numerical algorithms that perform a task using the images resulting from the VCT simulation pipeline, which involves making decisions regarding the truth state of the virtual patient (lesion present vs. absent, for example). When the task–based performance of the model observer is determined for a large set of images, a summary figure of merit can be estimated that summarizes the quality of the imaging system for supporting that particular task.

All model observers make use of statistical properties of images, for example, the expected size or shape of the lesion, and the texture of the noise in the images. The challenge in training a model observer for a VCT study is that, for a particular simulated breast anatomy and lesion model, the size or shape of the lesion in an image will depend on the image acquisition geometry and reconstruction algorithm. Likewise, the noise texture will depend on image acquisition parameters and reconstruction method. Because VCTs are typically intended to compare multiple systems, or optimize system geometry parameters, a VCT must ensure that the model observer is well-suited for each system design or parameter combination. In other words, the model observer must be trained to be optimal for each set of system parameters, as these impact the model observer in terms of signal and noise characteristics in the resulting images. Training of the model observer involves simulation of an adequate number of subjects (breast anatomy and lesion characteristics) to ensure that the results of the VCT are stable (more training cases would not affect the VCT results). Hundreds or even thousands of simulated breasts can be used in a VCT study, in order to have confidence in the results or the investigation, allowing for adequate training of the model observer as well as data sets for testing once the model observer has been determined.

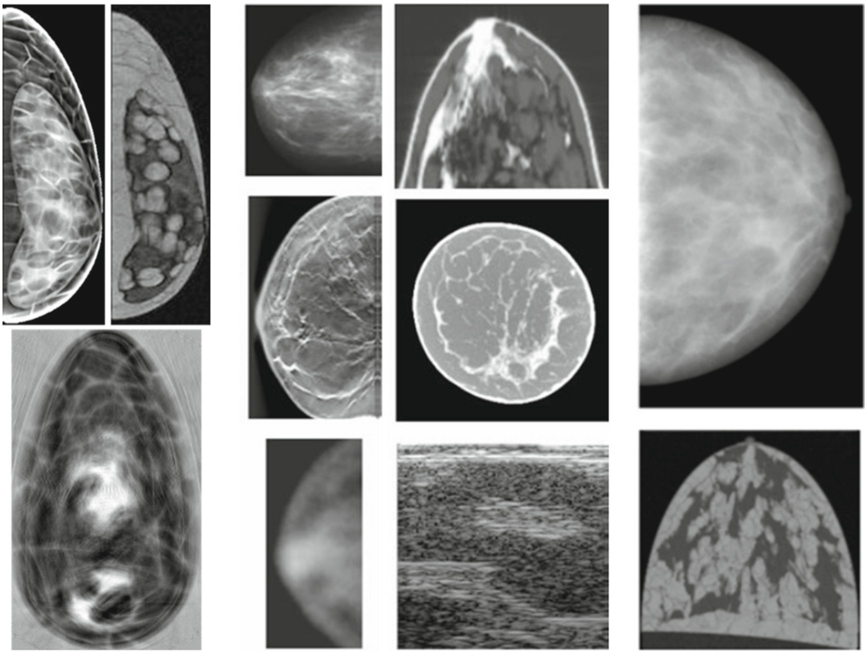
Two current challenges in the field of model observers are at the forefront of the field. The first is the development of methods for establishing that a model observer approximates ideal observer performance, which tells the investigator the best possible task performance that can be achieved given the data from the imaging system under evaluation. The second challenge is the need for recommendations on model observer designs that reliably predict human performance across a range of acquisition system parameters and reconstruction algorithms for 3D breast imaging systems so that the need for validation of new VCT studies against human data is reduced or eliminated.

4 Future Directions for the Virtual Evaluation of Breast Imaging

Following the report submission, the TG234 charge will be modified, redefining the focus of the Task Group. Anticipated future developments can be related to the extension of the VCT approach in several potential directions, including supporting other breast imaging modalities, a wider range of spatial scales, and the translation (or generalization) of the VCT concept to the imaging of other organs and diseases.

4.1 Virtual Evaluation of Other Breast Imaging Modalities

The original TG234 charge was focused on the evaluation of 3D/4D x-ray breast imaging, due to its widespread use in screening and diagnostic breast exams. Current digital breast phantoms are capable of supporting the simulation of other modalities and multi-modality studies, as the simulated breast anatomical structures are associated with various physical properties, matching the modality to be evaluated (Fig. 2). [6–8] The Penn open-source tissue simulation method builds software phantoms capable of synthesizing x-ray, MRI, and ultrasound images of the same simulated anatomy; it has been used to compare breast density estimates from mammography and ultrasound tomography. [6] Based on the segmentation of clinical breast CT data, Duke University has generated 100 phantoms capable of simulating a variety of breast imaging modalities. [7] In another recent example, a new open-source software tool was developed by Graff, to generate multimodality digital breast phantoms for use in the evaluation of both x-ray and MRI systems. [8] Of particular interest in the future, is the potential to perform virtual evaluation of systems for molecular imaging and radiomics analysis.



(a) (b) (c)

Fig. 2. Illustration of simulated multimodality images of computer breast phantoms, developed at (a) the University of Pennsylvania (modified from [6]), (b) Duke University (modified from [7]), and (c) FDA (modified from [8]).

4.2 Extending the Range of Spatial Scales in Breast Imaging Virtual Validations

Current digital breast phantoms are focused primarily on the anatomical scales visualized in clinical x-ray images of the breast. Recent progress in digital pathology and the close relationship between radiology and pathology, emphasize the need for simulating breast tissue at both radiologic and pathological scales. This introduces novel challenges to the virtual validation approach in terms of the computation speed, storage and transfer of simulated data, and the appropriate testing of phantom realism over the wide range of scales.

An example of recent efforts in the simulation of breast tissue at the radiologic and pathological scales has been reported by the researchers from the X-ray Physics Lab at the University of Pennsylvania. [9].

5 Conclusion

VCTs in breast imaging are increasingly being accepted by researchers in academia, industry, and government as an efficient preclinical approach to evaluate novel imaging system and image processing methods. This development is made evident by the growing body of published work, numerous conference sessions dedicated to VCTs, and the activities of the AAPM Task Group on Virtual Tools for 3D/4D X-ray Breast Imaging Validation (TG234).

This paper review the activities of the Task Group, current challenges related to the development and use of the virtual tools for breast imaging, and potential future directions of TG234.

Acknowledgment. The authors acknowledge the financial support and assistance from the AAPM for TG234 teleconferences and face-to-face meetings, and the Mini-Symposium on Virtual Clinical Trials held during the 2016 IWDM in Malmo, Sweden.

References

1. AAPM Task Group No. 234: Task group on virtual tools for the evaluation of new 3D/4D breast imaging systems (2012). http://www.aapm.org/org/structure/?committee_code=TG234
2. Bakic, P.R., Myers, K.J., Reiser, I., Kiarashi, N., Zeng, R.: Virtual tools for validation of x-ray breast imaging systems. Med. Phys. **40**, 390 (2013)
3. Myers, K.J., Bakic, P.R., Abbey, C.K., Kupinski, M., Mertelmeier, T.: Virtual tools for validation of x-ray breast imaging systems. Med. Phys. **41**, 446–447 (2014)
4. Myers, K.J., Bakic, P.R., Graff, C., Das, M., Avanaki, A.N.: Recent advances in virtual tools for validation of 3D/4D breast imaging systems (TG234). Med. Phys. (2016)
5. Vaishnav, J.Y., Jung, W.C., Popescu, L.M., Zeng, R., Myers, K.J.: Objective assessment of image quality and dose reduction in CT iterative reconstruction. Med. Phy. **41**, 071904 (2014)
6. Bakic, P.R., Li, C., West, E., Sak, M., Gavenonis, S.C., et al.: Comparison of 3D and 2D breast density estimation from synthetic ultrasound tomography images and digital mammograms of anthropomorphic software breast phantoms. In: Pelc, N.J., Samei, E., Nishikawa, R.M. (eds.) SPIE Medical Imaging: Physics of Medical Imaging, p. 79610Z (2011)
7. Segars, W.P., Veress, A.I., Wells, J.R., Sturgeon, G.M., Kiarashi, N., et al.: Population of 100 realistic, patient-based computerized breast phantoms for multi-modality imaging research. In: Whiting, B., Hoeschen, C., Kontos, D. (eds.) Physics of Medical Imaging: SPIE, p. 90331 (2014)
8. Graff, C.G.: A new, open-source, multi-modality digital breast phantom. In: Kontos, D., Flohr, T., Lo, J.Y. (eds.) Physics of Medical Imaging SPIE, p. 978309 (2016)
9. Bakic, P.R., Pokrajac, D.D., Batiste, R.C., Feldman, M.D., Maidment, A.D.: Simulation of breast anatomy: bridging the radiology-pathology scale gap. In: Tingberg, A., Lang, K. (eds.) International Workshop on Breast Imaging - IWDM. Springer, Malmo, Sweden (2016)