Master's Thesis Engineering Technology

2021-2022

Evidence based implementation of automated mammographic positioning quality assessment of breastcancer screening mammograms

Marthe Picard

Master of Nuclear Engineering Technology

PROBLEM STATEMENT

Breast cancer is currently the most frequent cancer in women in Europe. Through **screening**, systematic early detection can be achieved with the primary aim to reduce mortality from breast cancer [1]. Two images per breast are acquired for a mammographic examination: a cranio-caudal (CC) and a medio-lateral oblique (MLO) view [2]. Early detection of breast cancer through mammographic screening can only be achieved with **high quality mammograms,** which enable radiologists to detect the smallest abnormalities or subtle changes over time. It is the role of the radiographer to produce these high quality mammograms [3]. To assess the image quality, evaluating the positioning of the breast is a critical aspect in mammography. The positioning quality monitoring is performed **subjectively** by the radiologists during second reading and is based on a limited number of criteria. Volpara Health (Wellington, New Zealand) has developed a software tool for **automatic and objective evaluation** of the positioning quality namely **Volpara TruPGMI**. Prior to its application, the reproducibility and sensitivity of the software need to be investigated.

METHODOLOGY

- A radiologist and radiographer scored 127 screening exams with MLO and CC views of left and right breasts using 18 different positioning quality criteria provided by the Dutch expert centre for screening (LRCB, The Netherlands). These assessments were then compared to the different Volpara TruPGMI metrics [3]. PGMI stands for Perfect, Good, Moderate and Inadequate.
- By applying the software to mammograms of GE, Hologic and Siemens systems, the reproducibility and compatibility of the software was verified. The Pearson's chi square test was used to indicate significant differences. A post hoc Fisher exact test was then performed to confirm precisely where the differences were situated.
- Parameters such as compression force, compression pressure, breast volume, Volpara density grade (VDG) and dose have an influence on the quality positioning. By means of an ordinal regression it can be determined to what extent such parameters influence the overall quality of the image.

Table 1: Breast positioning criteria and associated answers for the reader study and for Volpara software

	Question		*Possible answers*	*Volpara metric*	
CC view	1	Folds present	*Yes, no*	NA	
	2	Depiction of pectoral muscle	*PGMI*	NA	
	3	Sufficient breast tissuelateral side	*Yes, no*	Tissue Cut Off	*False, True*
	4	Sufficient breast tissue medial side	*Yes ,no*	Tissue Cut Off	*False, True*
	5	Symmetry	*Yes, no*	NA	
	6	Nipple in profile	*Yes, no*	Nipple in Profile	*InProfile, NotInProfile*
	7	Overall quality	*PGMI*	Image PGMI Score	*PGMI*
	8	Retake needed	*Yes, no*	Image PGMI Score	*PGMI*
MLO view	9	Folds present	*Yes, no*	Pec Skinfold	*Present, Absent*
	9.1	If yes, where	*Pectoral, breast*		
	10	Depiction of inframammaryangle	*Visible, obscured, absent*	IMF adequacy	*VisibleWithoutSkinfold, VisibleButWithSkinfold, NotVisible*
	11	Pectoral muscle sufficiently deep	*Yes, no*	Pectoral Adequacy	*Adequate, NotAdequate (Short)*
	12	Pectoral muscle sufficiently wide	*Yes, no*	Pectoral Adequacy	*Adequate, NotAdequate (Narrow, Wide)*
	13	Symmetry	*Yes, no*	NA	
	14	Nipple in profile	*Yes, no*	Nipple in Profile	*InProfile, NotInProfile*
	15	Lower side of breast cut-off	*Yes, no*	Tissue Cut Off	*False, True*
	16	Overall quality	*PGMI*	Image PGMI score	*PGMI*
	17	Retake needed	*Yes, no*	Image PGMI score	*I versus PGM*
	18	Is there a quality problem with an image (i.e. compression, saturation, motion, noise,.. please indicate in which image and describe the problem in notes panel	*Free answer*		

RESULTS

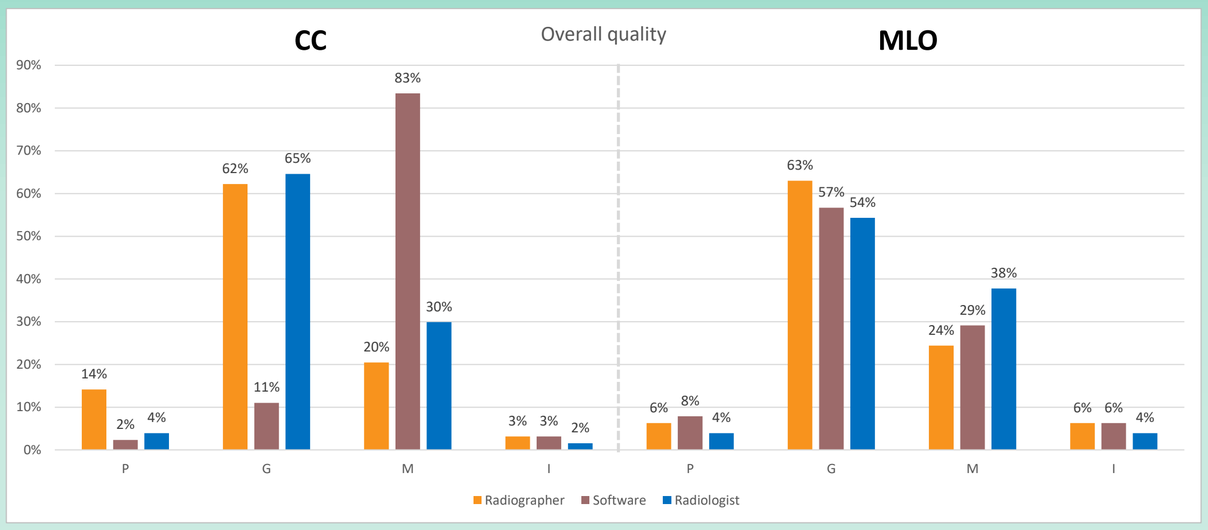


Figure 1: Bar graph showing PGMI score of radiologist, radiographer and software for CC and MLO view.

A limited number of images fall into the categories Inadequate and Perfect for CC view and MLO view. For CC view, the radiographer scored more images as Perfect than the radiologist and software. The software stated that most images fall into the Moderate category for CC view.

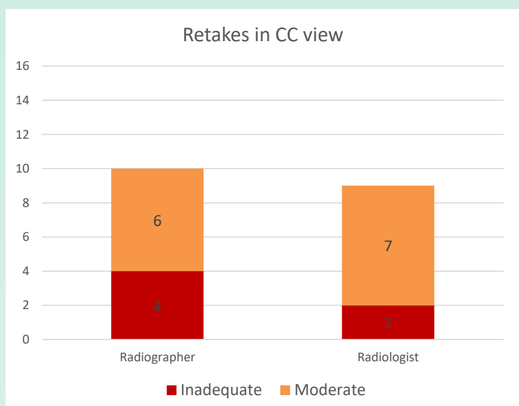
Table 2: Weighted kappa agreement for PGMI score and the binary PGM vs I score for CC and MLO view between radiologist (R1) and software (S) and between radiographer (R2)

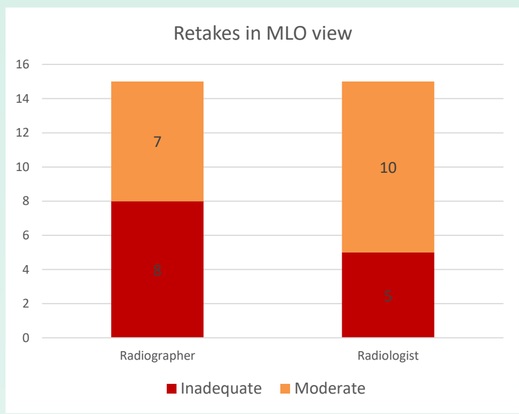
	CC		MLO	
	R1-S	*R2-S*	*R1-S*	*R2-S*
$\mathbf{Kappa}_W$	- 0.01	0.001	0.19	0.13
$\mathbf{Kappa}_W$: Binary	- 0.02	0.48	- 0.05	0.73

By combining the categories Perfect, Good and Moderate, a binary classification is obtained. The kappa agreement between radiographer (R2) and software (S) increases from a slight agreement to a moderate and substantial agreement for CC and MLO view resp. The kappa agreement decreased between radiologist (R1) and software for the binary score. Neither for CC view nor MLO view do the software and radiologist agree on the retakes.

A statistical difference was observed between the different scoring of the vendors for IMF (p =.002), MLO folds (p <.001), nipple in profile for MLO view (p =.01), pectoral width adequacy (p =.1) and for the PGMI score for CC view (p <.001).
Compression force (p <.001) and pressure (p <.001) as well as the dose (p <.001) had a significant influence on the PGMI score.

Figure 3: Histogram of the number of CC and MLO - images considered for retake with their overall quality score





CONCLUSION

The radiographer had an overall good agreement with the software compared to the radiologist who was less in line with the software. Seeing this better agreement between the radiographer who is our local teacher in screening mammograms and the software, one can conclude that the software would be useful as a continues training tool, providing immediate feedback to the radiographers. The concordance of the assessment of inadequate cases which can lead to retakes should be investigated further as there are large differences between the readers' individual scoring.
The software is not able to reproduce the outcomes for different mammographic systems. The compression force and pressure along with the dose also affect the positioning quality.
Prior to the use of automatic quality monitoring software in clinical practice, a careful evaluation and validation is needed

Supervisors / Co-supervisors / Advisors

Prof. ir. Hilde Bosmans,
Prof. dr. Brigitte Reniers,
Dr. Lesley Cockmartin

[1] N. Perry, M. Broeders, C. De Wolf, S. Törnberg, R. Holland and L. von Karsa, "Executive Summary," in European guidelines for quality assurance in breast cancer screening and diagnosis, Luxembourg, Office for Official Publications of the European Communities, 2006, pp. 5-14.
[2] Bevolkingsonderzoek.be, "Bevolkingsonderzoek," 2015. [Online]. Available: https://borstkanker.bevolkingsonderzoek.be/nl/bevolkingsonderzoek.
[3] C. van Landsveld-Verhoeven, G. J. den Heeten, J. Timmers and M. J. M. Broeders, "Mammographic positioning quality of newly trained versus experienced radiographers in the Dutch breast cancer screening programme," European Society of Radiology , p. 3322–3327, 2015.

UHASSELT

KU LEUVEN