

STUDY PROTOCOL

A systematic approach to study the effects of acquisition parameters and biological factors on computerized mammography analysis using ex vivo human tissue: A protocol description



Nicole Hernández^{1,2*}, Tomppa Pakarinen^{1,3}, Annukka Salminen^{1,2}, Santiago Laguna Castro^{4,5}, Ulla Karhunen-Enckell^{1,6}, Markus Hannula^{4,5}, Ritva Heljasvaara⁷, Jari Hyttinen^{4,5}, Katriina Joensuu^{6,8}, Otto Jokelainen^{9,10}, Arja Jukkola^{11,12}, Sanna-Maria Karppinen⁷, Auni Lindgren^{13,14}, Eero Lääperi¹, Emilia Peuhu^{15,16,17}, Taina Pihlajaniemi⁷, Renata Prunskaitė-Hyyryläinen⁷, Kirsi Rilla¹⁸, Pekka Ruusuvoori^{1,19}, Leena Latonen^{18,19}, Teemu Tolonen²⁰, Masi Valkonen¹⁹, Mira Valkonen^{1,19}, Miska Vuorlaakso^{1,21}, Said Pertuz²², Irina Rinta-Kiikka^{1,2}, Otso Arponen²³

1 Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland, **2** Department of Radiology, Tampere University Hospital, Tampere, Finland, **3** Department of Medical Physics, The Wellbeing Services County of Pirkanmaa, Tampere, Finland, **4** Computational Biophysics and Imaging Group, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland, **5** BioMediTech Unit, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland, **6** Department of Surgery, Tampere University Hospital, Tampere, Finland, **7** Faculty of Biochemistry and Molecular Medicine, University of Oulu, Oulu, Finland, **8** Department of Plastic Surgery, Tampere University Hospital, Tampere, Finland, **9** Institute of Clinical Medicine, Pathology and Forensic Medicine, University of Eastern Finland, Kuopio, Finland, **10** Department of Clinical Pathology, Diagnostic Imaging Center, Kuopio University Hospital, Kuopio, Finland, **11** Department of Oncology, Tampere University Hospital, Tampere, Finland, **12** Cancer Center, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland, **13** Department of Gynecology and Obstetrics, Kuopio University Hospital, Kuopio, Finland, **14** Institute of Clinical Medicine, University of Eastern Finland, Kuopio, Finland, **15** Institute of Biomedicine, Cancer Laboratory FICAN West, University of Turku, Turku, Finland, **16** Turku Bioscience Centre, University of Turku, Turku, Finland, **17** Åbo Akademi University, Turku, Finland, **18** Institute of Biomedicine, University of Eastern Finland, Kuopio, Finland, **19** Institute of Biomedicine, University of Turku, Turku, Finland, **20** Department of Pathology, Finlab Laboratories, Tampere, Finland, **21** Department of Musculoskeletal Surgery and Diseases, Tampere University Hospital, Tampere, Finland, **22** Department of Electrical, Electronics and Telecommunications Engineering, Universidad Industrial de Santander, Bucaramanga, Colombia, **23** Institute of Clinical Medicine, University of Eastern Finland, Kuopio, North Savo, Finland

☯ These authors contributed equally to this work.

* nicole.hernandezduran@tuni.fi

OPEN ACCESS

Citation: Hernández N, Pakarinen T, Salminen A, Castro SL, Karhunen-Enckell U, Hannula M, et al. (2025) A systematic approach to study the effects of acquisition parameters and biological factors on computerized mammography analysis using ex vivo human tissue: A protocol description. *PLoS One* 20(8): e0321658. <https://doi.org/10.1371/journal.pone.0321658>

Editor: Yuchen Qiu, University of Oklahoma, UNITED STATES OF AMERICA

Received: August 20, 2024

Accepted: July 8, 2025

Published: August 18, 2025

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0321658>

Copyright: © 2025 Hernández et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: No datasets were generated or analysed during the current study.

Abstract

Background: Mammography is the most common imaging modality for the detection of breast cancer. Artificial intelligence algorithms for mammography analysis have shown promising performance for breast cancer risk assessment and lesion detection and classification; however, these models often fail the test of external validation. The evidence points to variations in image acquisition—known as the batch effect—as a main

All relevant data from this study will be made available upon study completion.

Funding: This work is supported by the following grants, awarded to O.A.: seed funding for Health Data Science projects scheme of Tampere University, Tampere University Hospital (Project No. MJ006L), the Competitive State Research Financing of the Expert Responsibility Area of Tampere University Hospital (Project Nos. 9AC002 and T62564), the Cancer Foundation Finland and the Finnish Medical Foundation. The funders had no role in the design of this work, its execution, analyses, interpretations, or the decision to publish the article.

Competing interests: The authors have declared that no competing interests exist.

contributing factor to the lack of the models generalization and robustness. However, studies on the effects of acquisition in the mammogram have been limited due to lack of appropriate datasets. This prospective, exploratory, non-randomized study aims to study how biological and non-biological sources of heterogeneity affect the mammogram and, in turn, the computerized models for mammography analysis.

Methodology: This study will collect breast samples from 200 participants that will undergo breast resection as per clinical indications. Each sample will undergo the mammography imaging procedure several times to obtain mammograms with different combinations of imaging parameters. The resulting dataset will be used for the statistical analysis of the impact of imaging parameters in mammographic texture features and the computerized analysis of mammograms. Furthermore, biological information will be collected from the resected breast samples to study their relation to mammographic texture features.

Discussion: This study will add to the understanding of the effect of different sources of heterogeneity on mammography, ultimately aiding in the future development of robust computerized analysis models.

Introduction

Mammographic imaging and the current use of artificial intelligence

With 2.3 million new cases and 685,000 deaths in 2020, breast cancer is the most diagnosed cancer and accounts for most cancer deaths among females globally [1]. Early detection is of key importance for reducing breast cancer-related mortality [2,3], as these cases have a 5-year survival probability of approximately 96% [4]. Mammography screening programs, in which females aged 50 to 69 are invited for a biennial mammogram, are implemented across the EU [5]. Screening plays a significant role in early detection of the disease [6], as it has been shown that the successful implementation of screening programs can reduce breast cancer mortality by approximately 20% [7]. Mammography screening, however, is less effective on breasts with a higher percentage of fibroglandular tissue (FGT)—i.e., mammographically dense breasts—due to the similarity in the radiodensity between FGT and breast tumor tissue leading to reduced contrast between these tissue types [8,9]. The sensitivity of breast cancer detection decreases with increasing breast density, from 85.7% in the lowest density category to 61.0% in the highest density category [10,11]. Mammographically dense breasts (either heterogeneously dense or extremely dense) are found in about 48% of the screened population [12], and research has shown that females with higher breast density have a greater risk of developing breast cancer [11,13]. Indeed, mammographic breast cancer screening is less effective in females with mammographically dense breasts who simultaneously have the highest risk of developing breast cancer.

Computerized analysis of mammographic images by means of artificial intelligence (AI), specifically classical machine learning and deep learning (DL), is an active area of research with applications to different problems, such as breast cancer risk assessment and breast lesion detection and classification [14]. *Radiomics* is the extraction of quantitative features from medical images, with the aim of correlating them to different health conditions or outcomes [15]. AI-driven radiomics can be divided into classical machine learning approaches, which use hand-crafted radiomic features [15], and DL approaches, which implicitly build data-driven image descriptors using DL architectures [16]. Both approaches have shown potential for breast percent density classification [17,18],

breast cancer risk assessment [19,20], and mass detection [21]. Recently, AI-supported mammography readings have been suggested to reduce reader workload [22] and improve lesion detection [23]. Furthermore, radiomic biomarkers have been suggested to reflect the multiscale complexity of tissues ranging from genomic to histological (i.e., molecular and cellular) levels [24–30].

Artificial intelligence and the current understanding of the factors affecting its performance

Although AI-based mammography analysis has demonstrated great potential, the lack of access to representative datasets [31–34] and model degradation due to data distribution shift [35–39] are commonly reported challenges that are hindering the translation of AI to clinical practice. The high heterogeneity between mammography datasets due to imaging parameters is a recognized challenge for the development of robust computerized analysis tools. Research to understand the effect of these sources of heterogeneity for the goal of computerized assessment is still lacking [35,40,41].

The effect of non-biological sources of heterogeneity—imaging acquisition and the manufacturer’s pre- and post-processing—is referred to as the *batch effect*. The batch effect has been shown to introduce changes in computerized image features, which also affect their robustness and reproducibility [42–44]. This in turn is reflected in the performance of breast cancer risk assessment [37,45,46] and breast cancer classification [35] models. On the other hand, the biological properties of the breast, through the interaction with the x-ray radiation, create the radiomic features; the heterogeneity that arises from them is, technically, desired. However, as radiomic biomarkers are not initially derived from biological samples but rather from radiological images, the effect of the interplay between image acquisition and the underlying biology on the performance of AI-based radiomics is unknown [24]. Research into the role of fibroglandular and adipose tissue—from the genetic and molecular levels to cellular and tissue levels—on radiomic features is currently scarce, but it is needed to explain how radiomic features may predict the future development of cancer.

Although the amount of existing clinical mammography data is large, the retrieval of an appropriate dataset to study the sources of heterogeneity mentioned above is infeasible from existing, accessible repositories. As a source of ionizing radiation, mammographic imaging limits the number of times a person can be imaged. As a result, the acquisition parameters in clinical practice used to capture each specific mammogram are optimized for image quality and for as low a radiation dose as is reasonably achievable: this limits the range of acquisition parameters that are employed on the different breasts, which, in turn, limits the possibilities of analyzing the effect of acquisition parameters on computerized analysis of mammograms derived from *in vivo* breasts. A dataset for the study of the sources of heterogeneity would require reimaging of the breast with different acquisition settings. Moreover, the analysis of biological factors requires the collection of tissue samples. Breast tissue is not normally resected unless there is a suspicion of cancer. Previous studies have used phantoms and retrospectively collected mammograms [42,44,45], which can only partially address the association between acquisition parameters and textural features and do not allow for the systematic evaluation of the effect of biological factors on radiomic features.

Study objectives

Our hypothesis posits that variations in mammography acquisition introduce a measurable batch effect. Under this hypothesis, it is expected that the batch effect significantly affects the

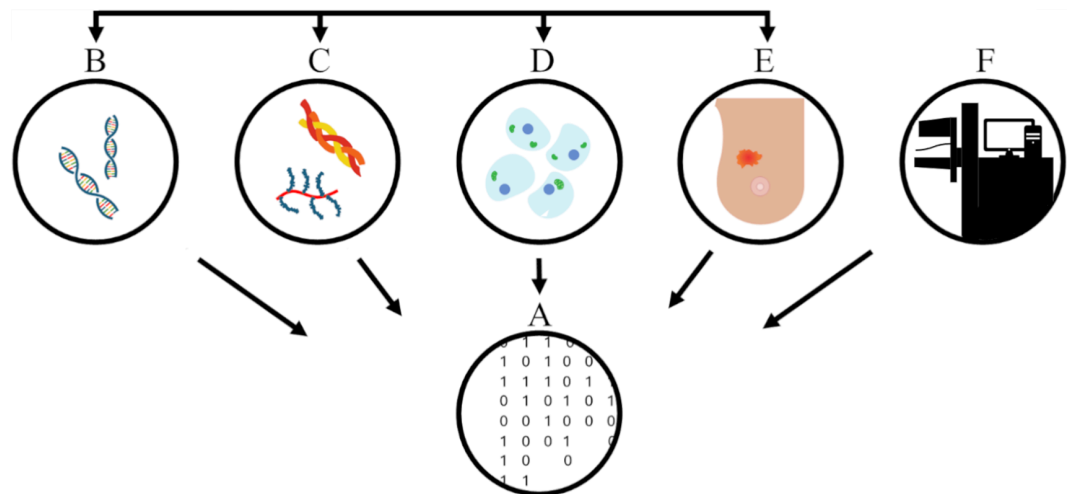


Fig 1. Factors affecting mammographic texture features. The study aims to evaluate how mammographic texture features (A) are affected by imaging parameters (F), and by genetic (B), molecular (C), cellular (D), and tissue level (E) biological factors across different scales.

<https://doi.org/10.1371/journal.pone.0321658.g001>

performance and generalizability of AI models for breast cancer detection. The characterization of this effect, as well as a better understanding of the link between biological characteristics and mammographic radiomic features, are fundamental to developing robust models (Fig 1). The study has two objectives:

1. Evaluate how imaging parameters affect the mammographic radiomic features by
 - developing a model for the batch effect in mammography images via quantification of its effect in radiomic features; and
 - assessing the impact of acquisition parameters on the performance of AI models for breast cancer detection from mammography.
2. Evaluate how biological factors across different scales affect computerized mammographic features by
 - assessing associations between different biological characteristics of the breast tissue and radiomic features.

To achieve these goals, we will use *ex vivo* human tissue—resected breasts. This dataset will be characterized by the acquisition of multiple mammograms from the resected breast tissue with different sets of imaging parameters. Such data will allow for the strict study of the interplay between acquisition parameters, breast biology, and mammographic texture features. Furthermore, we will use the acquired dataset to assess the effect of acquisition parameters and biological factors on the performance of breast cancer risk assessment and lesion detection models. In the proposed study, we will overcome the ethical constraints of the repeated use of ionizing radiation with the use of *ex vivo* human tissue for the generation of a dataset suitable for this study. This work aims to contribute to the advancement of computerized mammography analysis and to facilitate its translation into clinical practice through both radiomics and DL-based systems. Equally important, this study aims to provide insight into the associations between clinical and biological factors and

computerized radiomic features. This article primarily details the methodology and analysis pertaining to our first objective, with a focus on the study of the acquisition parameters of mammography.

Materials and methods

Methodological design

This is a prospective, exploratory, non-randomized study on adult females (≥ 18 years old) who will undergo uni- or bilateral mastectomy for clinical indications unrelated to this trial. In this protocol, we broadly define mastectomy as a surgical procedure in which fibroglandular and fatty tissue of the breast is removed. Patients will be recruited at Tampere University Hospital, Finland, the coordinating study center. Multiple study centers in Finland will analyze the resected breast samples.

The multidisciplinary research team (comprising researchers in the fields of surgery, radiology, pathology, oncology, medical physics, computer sciences, and biological sciences) enriches the methodology and scope of this study. The methodology was reviewed by the researchers' committee and approved by the Ethics Committee of the Wellbeing Services County of Pirkanmaa, Finland (study ID: R230077L), the Finnish Medicines Agency, and subsequently by the institutes participating in this study. The data collection, processing, and analyses are compliant with the European General Data Protection Regulation. Data use and sharing adhere to the data governance policies of the data controller (Tampere University Hospital). The timeline of the study is illustrated in Fig 2. It is worth noting that initial analyses of images and tissue samples will take place as soon as data have been collected from the first 10 resected samples.

Participants

Females living in the catchment area of the Wellbeing Services County of Pirkanmaa are eligible for the study if they meet all the inclusion criteria and no exclusion criteria (Table 1). Trial participants will undergo mastectomy/mastectomies based on clinical indications ranging from non-malignancy-related factors (e.g., masculinizing chest reconstruction) to risk-reducing surgery for high-risk mutation carriers and removal of cancer-affected breast(s). This study will not affect the surgical treatment predetermined for the participants as part of their standard of care treatment. Delegated surgeons will identify suitable candidates via interviews and the review of medical records. The surgeon will inform the potential participant about this study and provide a copy of the Participant Information Sheet. Informed consent will be collected prior to any protocol-specific procedure. The participant may withdraw their consent at any point.

We will recruit a maximum of 100 patients with breast cancer and 100 patients without breast cancer who are due to undergo uni- or bilateral mastectomy. Due to the exploratory nature of this investigation and the fact that insufficient prior knowledge prevents formal statistical calculations, no statistical techniques were used to determine the sample size. To allow the evaluation of the primary and secondary objectives of the study, sample size determination has been primarily based on feasibility and pragmatic considerations of the anticipated recruitment rates. We will include patients with healthy breast tissue, as we are interested in the applicability of our results to screening mammography, which encounters mainly healthy breasts.

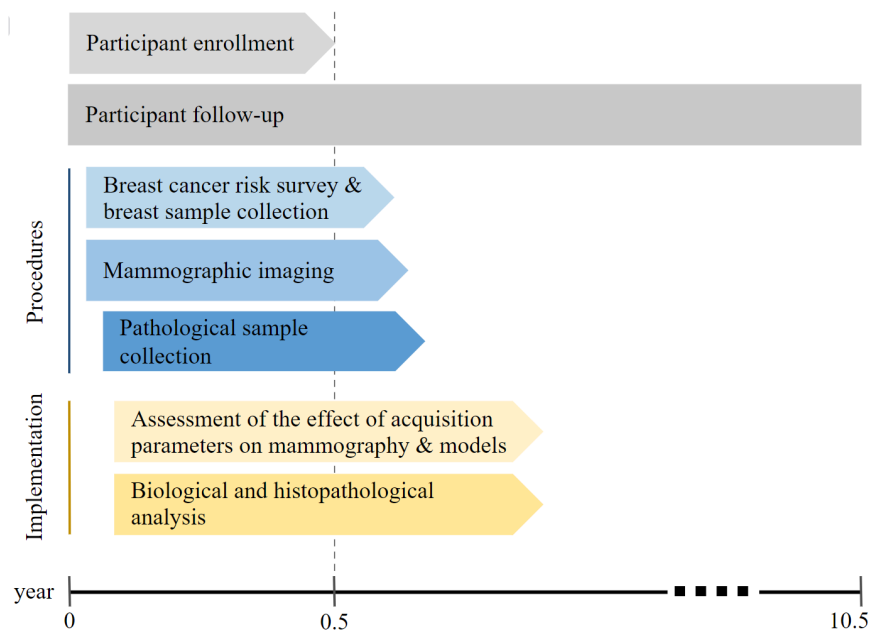


Fig 2. Timeline of the study. It starts with the recruitment of the first patient. Recruitment started on July 1st, 2024, and it is expected to end on December 15th, 2024. Participants will be followed up for 10 years. The imaging and sample collection from the resected samples starts with the first surgical intervention and ends soon after the last intervention. Imaging and biological analysis will begin after data have been obtained from the first 10 resected breasts and will continue after the imaging and sample collection have been completed. All analyses pertaining to this project are expected to be concluded by the 2-year mark after the imaging and sample collection have been finalized.

<https://doi.org/10.1371/journal.pone.0321658.g002>

Table 1. Inclusion and exclusion criteria.

Inclusion criteria	Exclusion criteria
Aged ≥ 18	Aged < 18
Biological female	Has breast implants
Ability to give informed consent	Inability to give informed consent
Clinical indication for uni- or bilateral mastectomy	

<https://doi.org/10.1371/journal.pone.0321658.t001>

Procedures

After enrolment, participants will complete a questionnaire on breast cancer risk. Subsequently, they will undergo mastectomy according to the clinical indications. We will scan the mastectomy sample with mammography and extract radiomic features used in both classical machine learning and DL. Specimens from fibroglandular tissue, fat, and, in cases of cancer, from the tumor and axillary lymph nodes will be collected. Participants will be followed-up for 10 years after enrolment to evaluate short- and long-term mortality.

Breast cancer risk survey. The trial participants will complete a questionnaire on their background information regarding comorbidities and cancer risk. The researchers will collect register-based clinical data from the healthcare records. These data will be used to determine individuals' breast cancer risk scores within 5 and 10 years using the Gail model [47] and the Tyrer–Cuzick risk calculator [48], respectively.

Mammographic imaging of the resected breast samples. Optimal mammographic acquisition parameters are contingent on several factors: the size and density of the breast,

along with patient tolerance, determine the compression thickness and force. In addition, the x-ray tube voltage peak (kVp) can be set manually or automatically via the automated exposure control (AEC), of which there are several modes. Milliampere-seconds (mAs) are controlled by the AEC but can also be specified by manual control.

Fig 3 summarizes the study pipeline and data that will be derived from the resected tissue. The resected breast samples will undergo imaging using two mammography devices prior to further procedures. Since imaging parameters are a central variable of our study, each breast sample will be imaged with different combinations of imaging parameters, yielding several images. The acquisition parameters that will be studied are the compressed breast thickness, compression force, kVp, and mAs. We will vary these parameters based on a predefined parameter matrix structured by fixed intervals of kVp and mAs; within each position of this matrix, values of breast thickness and compression force will be used according to the capacity of each specific tissue. For each breast, we will explore a range of parameters covering the clinically used distribution and overlapping with the range of parameters used for all other samples. As there is no available literature on the mammography imaging of *ex vivo* breast tissue, we will use the clinically used compression force as a guideline, but the specific values used for breast compression force and breast thickness will be decided on a breast basis, considering the limitations of each specimen. Other system-specific hardware, such as the target-filter combination and detector, will also be considered for analysis and hence retrieved. Any further information derived from the acquisition process, such as the average glandular dose, will also be retrieved and considered for analysis.

The acquired mammograms with their corresponding tags will be stored in digital imaging and communications in medicine (DICOM) format, in FOR PROCESSING (raw) and

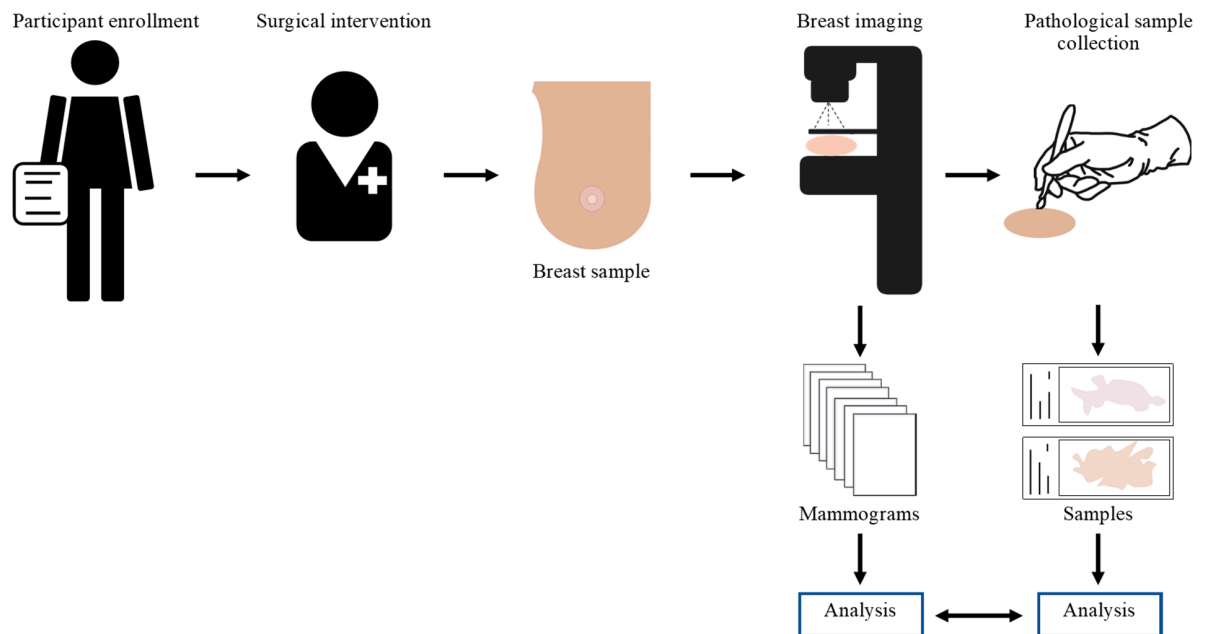


Fig 3. Study pipeline. Study pipeline and data that will be acquired from the resected breasts. Each resected breast will be scanned with two mammographic imaging machines with several predefined sets of imaging parameters, yielding several images per breast sample in both raw and postprocessed formats. Next, pathological samples will be extracted from the resected breasts for further biological analyses.

<https://doi.org/10.1371/journal.pone.0321658.g003>

FOR PRESENTATION (postprocessed) formats. Mammograms will be stored in appropriate storage devices.

Biological and histopathological samples. After the imaging of the resected breast(s), the research team will collect representative samples of fat, fibroglandular tissue and skin, and potential tumor(s) and snap-freeze them in liquid nitrogen and/or fix them in formalin for further tissue analyses.

Implementation

Our study will produce a multiparametric imaging dataset of mammograms. We will use this dataset to measure the effect of acquisition parameters on mammographic texture features, as well as for the validation of open-source AI models for breast cancer detection and risk assessment.

Measuring the effect of acquisition parameters on mammographic texture features.

We will evaluate the effect of acquisition parameters on mammograms via the assessment of their texture features, by testing both the sensitivity (in terms of statistically significant differences) and the robustness of texture features in different inter- and intra-machine scenarios. Testing for statistically significant differences between scenarios allows one to determine the magnitude of the impact of the specific parameter under observation; testing for robustness is a complementary approach measuring texture feature sensitivity to imaging parameter variation. We will study radiomic texture features that are common in mammography analysis [49].

We will follow assessment approaches found in the literature on mammography acquisition effects and texture feature robustness assessment [43–46]. The metrics used for differences and robustness assessment will include the correlation coefficient, concordance correlation coefficient, paired differences, variance, means of feature ratios, percentage coefficient variation, and the Kolmogorov–Smirnov test.

Measuring the effect of acquisition parameters at the model level. Due to rapid technological advances, we are also interested in measuring the effect that the acquisition technique might have on the performance of classical machine learning-based and DL-based models. The first step in this phase of the project is the identification of open-source algorithms published in the field of mammography analysis regarding breast cancer detection. The identified algorithms will be subjected to further development to allow for their use and external validation with our dataset; however, the nature of our dataset allows for the establishment of different validation scenarios beyond using only the images acquired with the clinically optimal acquisition settings. The possibility of testing different “versions” of a dataset acquired from the same breast cohort allows us to study the effect that the acquisition settings can have at the model level.

Furthermore, to the extent that the identified open-source models allow, we will follow the approaches presented in two previous publications that quantified the effect of model robustness to acquisition parameters and vendor using breast phantom mammography data [46] and retrospective mammography data [45]. The first approach (pipeline A) is a two-step scheme: the authors used breast phantom imaging data and self-defined assessment metrics to assess robustness, followed by the evaluation of a breast cancer risk assessment model by using groups of features with different levels of robustness. The statistical significance of the different performances obtained was assessed via a linear mixed-effects model. The second approach (pipeline B) used retrospective screening mammograms obtained with two different mammography machines. Each subject had been previously imaged with both vendors in different screening rounds. Texture feature robustness was assessed via statistics describing

equivalence, correlation, and sample distributions, and the redundancy of features was identified via clustering. To assess the effect of robustness and redundancy on the performance of a robustness assessment model, an inter-vendor scheme was constructed by training the model using one of the vendors as a reference sample and testing it on the images of the other vendor.

Pipelines A and B will be used as guidelines to assess acquisition effects at both the texture feature and the model level. The main adjustments will be the use of *ex vivo* tissue instead of a breast phantom in the case of the first study and testing intra-acquisition settings scenarios in addition to intra-machine in the case of the second study.

Biological and histopathological analysis. Biological, histopathological, and genetic analyses and imaging studies will be conducted on the tissue samples to determine their structural, biological, histopathological, and genetic features to study how biological factors across different scales correlate with mammographic texture features. The tissue analyses will include, for example, routine hematoxylin-eosin staining for tissue morphology, label-free second harmonic generation imaging or Masson trichrome staining for collagen in fibroglandular tissue, cytokeratin staining for glandular epithelium, and fluorescent lipid staining for adipocytes. Tissue stainings will be analyzed using microscopic and/or mesoscopic imaging combined with modern digital pathology methods, including AI-driven image analysis.

Ethics and dissemination

We sought a favorable opinion from the Ethics Committee of the Wellbeing Services County of Pirkanmaa (study ID: R23077L) and an organizational permission to conduct the study. The data will be pseudonymized at source. This study does not affect the clinical treatment of the participants. There will be no safety reporting on this study, since the procedures will be conducted on resected breast tissue; hence, the study does not represent risks for humans.

The study results will be published in peer-reviewed scientific journals, conference proceedings, or as part of science communication. Data use and sharing adhere to the data governance policies of the data controller (Tampere University Hospital, Finland).

Discussion

This study aims to test whether mammography imaging acquisition introduces a batch effect and its potential impact in computerized analysis. To achieve our goals, this study will use mammography images derived from resected breast tissue, which raises questions about the imaging procedure. The relative forces experienced by the components of an *ex vivo* sample differ from those of *in vivo* tissue due to the lack of attachment to the pectoralis wall in the *ex vivo* scenario. The behavior of the resected tissue during breast compression is unpredictable, as there is no available information about this procedure. We anticipate that this will limit the number of breast compression forces and thicknesses that we will be able to use and achieve during imaging. Furthermore, the discrepancies in tissue properties that pertain to the interaction of *in vivo* and *ex vivo* breast tissue with x-rays have not been studied. At the analysis level, AI-based mammographic approaches found in the literature have been designed for images acquired from *in vivo* breasts. The extent to which the use of resected breasts will restrict the direct translation of our results to an *in vivo* scenario is impossible to predict. We will assess the differences in mammograms acquired from *ex vivo* tissue at both the quantitative level, comparing their mammographic features with those of *in vivo*, retrospective counterparts (retrieved from open-source repositories), and at the qualitative level, via radiologists' visual assessment of their likeness to an *in vivo*-acquired mammogram. Furthermore, this study will enable determination of the critical features for mammograms,

the relevance of which for *in vivo* mammography settings can further be tested in prospective patient studies.

The results of this study will shed light on one of the main factors responsible for the lack of generalizability of AI-based models. Knowing the extent and ways in which acquisition parameters affect the mammographic image at the quantitative level, and ultimately the AI models based on them, will guide future approaches in addressing the batch effect as a composite problem, instead of a homogeneous issue.

Acknowledgments

We are grateful to the colleagues who have shared their knowledge and insight in preparation of this manuscript.

Author contributions

Conceptualization: Nicole Hernández, Tomppa Pakarinen, Said Pertuz, Irina Rinta-Kiikka, Otso Arponen.

Data curation: Otso Arponen.

Formal analysis: Said Pertuz, Otso Arponen.

Funding acquisition: Otso Arponen.

Investigation: Nicole Hernández, Said Pertuz, Otso Arponen.

Methodology: Nicole Hernández, Tomppa Pakarinen, Annukka Salminen, Said Pertuz, Irina Rinta-Kiikka, Otso Arponen.

Project administration: Said Pertuz, Otso Arponen.

Resources: Otso Arponen.

Software: Otso Arponen.

Supervision: Said Pertuz, Irina Rinta-Kiikka, Otso Arponen.

Validation: Otso Arponen.

Visualization: Otso Arponen.

Writing – original draft: Nicole Hernández, Said Pertuz, Otso Arponen.

Writing – review & editing: Nicole Hernández, Tomppa Pakarinen, Annukka Salminen, Santiago Laguna Castro, Ulla Karhunen-Enckell, Markus Hannula, Ritva Heljasvaara, Jari Hyttinen, Katriina Joensuu, Otto Jokelainen, Arja Jukkola, Sanna-Maria Karppinen, Auni Lindgren, Eero Lääperi, Emilia Peuhu, Taina Pihlajaniemi, Renata Prunskaitė-Hyyryläinen, Kirsi Rilla, Pekka Ruusuvoori, Leena Latonen, Teemu Tolonen, Masi Valkonen, Mira Valkonen, Miska Vuorlaakso, Said Pertuz, Irina Rinta-Kiikka, Otso Arponen.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021;71(3):209–49. <https://doi.org/10.3322/caac.21660> PMID: 33538338
2. Duffy SW, Tabár L, Yen AM-F, Dean PB, Smith RA, Jonsson H, et al. Beneficial effect of consecutive screening mammography examinations on mortality from breast cancer: a prospective study. *Radiology.* 2021;299(3):541–7. <https://doi.org/10.1148/radiol.2021203935> PMID: 33650900

3. Mayor S. Survival of women treated for early breast cancer detected by screening is same as in general population, audit shows. *BMJ: British Medical Journal*. 2008;336(7658):1398.
4. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Nikšić M, et al. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet*. 2018;391(10125):1023–75. [https://doi.org/10.1016/S0140-6736\(17\)33326-3](https://doi.org/10.1016/S0140-6736(17)33326-3) PMID: 29395269
5. Donata L, Silvia D, Crystal F, Jesus LA, Luciana N, Ciaran N, et al. Report of a European survey on the organisation of breast cancer care services - Supporting information for the European Commission initiative on breast cancer. Luxembourg: Publications Office of the European Union; 2014.
6. Tabár L, Vitak B, Chen TH-H, Yen AM-F, Cohen A, Tot T, et al. Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. *Radiology*. 2011;260(3):658–63. <https://doi.org/10.1148/radiol.11110469> PMID: 21712474
7. Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. *Lancet*. 2012;380(9855):1778–86. [https://doi.org/10.1016/S0140-6736\(12\)61611-0](https://doi.org/10.1016/S0140-6736(12)61611-0) PMID: 23117178
8. Destounis S, Johnston L, Highnam R, Arieno A, Morgan R, Chan A. Using volumetric breast density to quantify the potential masking risk of mammographic density. *AJR Am J Roentgenol*. 2017;208(1):222–7. <https://doi.org/10.2214/AJR.16.16489> PMID: 27824483
9. Hussein H, Abbas E, Keshavarzi S, Fazelzad R, Bukhanov K, Kulkarni S, et al. Supplemental breast cancer screening in women with dense breasts and negative mammography: a systematic review and meta-analysis. *Radiology*. 2023;306(3):e221785. <https://doi.org/10.1148/radiol.221785> PMID: 36719288
10. Lam PB, Vacek PM, Geller BM, Muss HB. The association of increased weight, body mass index, and tissue density with the risk of breast carcinoma in Vermont. *Cancer*. 2000;89(2):369–75. [https://doi.org/10.1002/1097-0142\(20000715\)89:2<369::aid-cnrc23>3.0.co;2-j](https://doi.org/10.1002/1097-0142(20000715)89:2<369::aid-cnrc23>3.0.co;2-j) PMID: 10918168
11. Boyd NF, Lockwood GA, Byng JW, Tritchler DL, Yaffe MJ. Mammographic densities and breast cancer risk. *Cancer Epidemiology, Biomarkers & Prevention*. 1998;7(12):1133–44.
12. Mann RM, Athanasiou A, Baltzer PAT, Camps-Herrero J, Clauser P, Fallenberg EM, et al. Breast cancer screening in women with extremely dense breasts recommendations of the European Society of Breast Imaging (EUSOBI). *Eur Radiol*. 2022;32(6):4036–45. <https://doi.org/10.1007/s00330-022-08617-6> PMID: 35258677
13. McCormack VA, dos Santos SI. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiol Biomarkers Prev*. 2006;15(6):1159–69.
14. Hernández A, Miranda DA, Pertuz S. Algorithms and methods for computerized analysis of mammography images in breast cancer risk assessment. *Computer Methods and Programs in Biomedicine*. 2021;212:106443.
15. Mayerhoefer ME, Materka A, Langs G, Häggström I, Szczypiński P, Gibbs P, et al. Introduction to radiomics. *J Nucl Med*. 2020;61(4):488–95. <https://doi.org/10.2967/jnumed.118.222893> PMID: 32060219
16. Bhowmik A, Eskreis-Winkler S. Deep learning in breast imaging. *BJR Open*. 2022;4(1):20210060. <https://doi.org/10.1259/bjro.20210060> PMID: 36105427
17. Liasis G, Pattichis C, Petroudi S. Combination of different texture features for mammographic breast density classification. In: 2012 IEEE 12th International Conference on Bioinformatics & Bioengineering (BIBE). Larnaca, Cyprus. 2012. p. 732–7.
18. Lehman CD, Yala A, Schuster T, Dontchos B, Bahl M, Swanson K, et al. Mammographic breast density assessment using deep learning: clinical implementation. *Radiology*. 2019;290(1):52–8. <https://doi.org/10.1148/radiol.2018180694> PMID: 30325282
19. Pertuz S, Sassi A, Holli-Helenius K, Kämäräinen J, Rinta-Kiikka I, Lääperi A-L, et al. Clinical evaluation of a fully-automated parenchymal analysis software for breast cancer risk assessment: a pilot study in a Finnish sample. *Eur J Radiol*. 2019;121:108710. <https://doi.org/10.1016/j.ejrad.2019.108710> PMID: 31689665
20. Li H, Giger M, Huynh BQ, Antropova NO. Deep learning in breast cancer risk assessment: evaluation of convolutional neural networks on a clinical dataset of full-field digital mammograms. *Journal of Medical Imaging*. 2017;4:041304–041304.
21. Rodriguez-Ruiz A, Lång K, Gubern-Merida A, Broeders M, Gennaro G, Clauser P, et al. Stand-alone artificial intelligence for breast cancer detection in mammography: comparison with 101 radiologists. *J Natl Cancer Inst*. 2019;111(9):916–22. <https://doi.org/10.1093/jnci/djy222> PMID: 30834436

22. Kyono T, Gilbert FJ, van der Schaar M. Improving workflow efficiency for mammography using machine learning. *J Am Coll Radiol*. 2020;17(1 Pt A):56–63. <https://doi.org/10.1016/j.jacr.2019.05.012> PMID: 31153798
23. Sun Y, Qu Y, Wang D, Li Y, Ye L, Du J, et al. Deep learning model improves radiologists' performance in detection and classification of breast lesions. *Chinese Journal of Cancer Research*. 2021;33(6):682–93.
24. Tomaszewski MR, Gillies RJ. The Biological Meaning of Radiomic Features. *Radiology*. 2021;298(3):505–16. <https://doi.org/10.1148/radiol.2021202553> PMID: 33399513
25. La Forgia D, Fanizzi A, Campobasso F, Bellotti R, Didonna V, Lorusso V, et al. Radiomic analysis in contrast-enhanced spectral mammography for predicting breast cancer histological outcome. *Diagnostics (Basel)*. 2020;10(9):708. <https://doi.org/10.3390/diagnostics10090708> PMID: 32957690
26. Ma W, Zhao Y, Ji Y, Guo X, Jian X, Liu P, et al. Breast cancer molecular subtype prediction by mammographic radiomic features. *Acad Radiol*. 2019;26(2):196–201. <https://doi.org/10.1016/j.acra.2018.01.023> PMID: 29526548
27. Son J, Lee SE, Kim E-K, Kim S. Prediction of breast cancer molecular subtypes using radiomics signatures of synthetic mammography from digital breast tomosynthesis. *Sci Rep*. 2020;10(1):21566. <https://doi.org/10.1038/s41598-020-78681-9> PMID: 33299040
28. Luo C, Zhao S, Peng C, Wang C, Hu K, Zhong X, et al. Mammography radiomics features at diagnosis and progression-free survival among patients with breast cancer. *Br J Cancer*. 2022;127(10):1886–92. <https://doi.org/10.1038/s41416-022-01958-5> PMID: 36050449
29. Davey MG, Davey MS, Ryan MRB, Lowery AJ, Kerin MJ. Radiomic differentiation of breast cancer molecular subtypes using pre-operative breast imaging—a systematic review and meta-analysis. *European Journal of Radiology*. 2021;144:109996.
30. Tamez-Peña J-G, Rodriguez-Rojas J-A, Gomez-Rueda H, Celaya-Padilla J-M, Rivera-Prieto R-A, Palacios-Corona R, et al. Radiogenomics analysis identifies correlations of digital mammography with clinical molecular signatures in breast cancer. *PLoS One*. 2018;13(3):e0193871. <https://doi.org/10.1371/journal.pone.0193871> PMID: 29596496
31. Wong DJ, Gandomkar Z, Wu W-J, Zhang G, Gao W, He X, et al. Artificial intelligence and convolution neural networks assessing mammographic images: a narrative literature review. *J Med Radiat Sci*. 2020;67(2):134–42. <https://doi.org/10.1002/jmrs.385> PMID: 32134206
32. Anderson AW, Marinovich ML, Houssami N, Lowry KP, Elmore JG, Buist DSM, et al. Independent external validation of artificial intelligence algorithms for automated interpretation of screening mammography: a systematic review. *J Am Coll Radiol*. 2022;19(2 Pt A):259–73. <https://doi.org/10.1016/j.jacr.2021.11.008> PMID: 35065909
33. Ammar MB, Ayeche FL, Ksantini R, Mahjoubi H. Pre-trained deep convolutional neural network architectures for breast cancer diagnosis in mammography: Current state-of-the-art. In: 2023 International Conference on Innovations in Intelligent Systems and Applications (INISTA). 2023. p. 1–7.
34. Ou WC, Polat D, Dogan BE. Deep learning in breast radiology: current progress and future directions. *Eur Radiol*. 2021;31(7):4872–85. <https://doi.org/10.1007/s00330-020-07640-9> PMID: 33449174
35. Wang X, Liang G, Zhang Y, Blanton H, Bessinger Z, Jacobs N. Inconsistent performance of deep learning models on mammogram classification. *J Am Coll Radiol*. 2020;17(6):796–803. <https://doi.org/10.1016/j.jacr.2020.01.006> PMID: 32068005
36. de Vries CF, Colosimo SJ, Staff RT, Dymiter JA, Yearsley J, Dinneen D, et al. Impact of different mammography systems on artificial intelligence performance in breast cancer screening. *Radiol Artif Intell*. 2023;5(3):e220146. <https://doi.org/10.1148/ryai.220146> PMID: 37293340
37. Pertuz S, Sassi A, Arponen O, Holli-Helenius K, Laaperi A-L, Rinta-Kiikka I. Do mammographic systems affect the performance of computerized parenchymal analysis? *Annu Int Conf IEEE Eng Med Biol Soc*. 2019;2019:4863–6. <https://doi.org/10.1109/EMBC.2019.8856948> PMID: 31946950
38. Hsu W, Hippe DS, Nakhaei N, Wang PC, Zhu B, Siu N. External validation of an ensemble model for automated mammography interpretation by artificial intelligence. *JAMA Network Open*. 2022;5(11):e2242343. <https://doi.org/10.1001/jamanetworkopen.2022.42343>
39. Condon JJJ, Trinh V, Hall KA, Reintals M, Holmes AS, Oakden-Rayner L, et al. Impact of transfer learning using local data on performance of a deep learning model for screening mammography. *Radiol Artif Intell*. 2024;6(4):e230383. <https://doi.org/10.1148/ryai.230383> PMID: 38717291
40. Thrall JH, Li X, Li Q, Cruz C, Do S, Dreyer K. Artificial intelligence and machine learning in radiology: opportunities, challenges, pitfalls, and criteria for success. *Journal of the American College of Radiology*. 2018;15(3):504–8.

41. Africano G, Arponen O, Rinta-Kiikka I, Pertuz S. Transfer learning for the generalization of artificial intelligence in breast cancer detection: a case-control study. *Acta Radiol.* 2024;65(4):334–40. <https://doi.org/10.1177/02841851231218960> PMID: 38115699
42. Acciavatti RJ, Cohen EA, Maghsoudi OH, Gastouniotti A, Pantalone L, Hsieh MK, et al. Robust radiomic feature selection in digital mammography: understanding the effect of imaging acquisition physics using phantom and clinical data analysis. In: Imaging M, editor. *Computer-aided diagnosis*, Houston, USA: Texas; 2020. p. 203–11.
43. Mendel KR, Li H, Lan L, Cahill CM, Rael V, Abe H, et al. Quantitative texture analysis: robustness of radiomics across two digital mammography manufacturers' systems. *Journal of Medical Imaging.* 2018;5:011002–011002.
44. Wang Y, Keller BM, Zheng Y, Acciavatti RJ, Gee JC, Maidment AD, et al. A phantom study for assessing the effect of different digital detectors on mammographic texture features. In: *Breast Imaging: 11th International Workshop*, Philadelphia, PA, USA, 2012. Springer; 2012. p. 604–10.
45. Acciavatti RJ, Cohen EA, Maghsoudi OH, Gastouniotti A, Pantalone L, Hsieh M-K, et al. Incorporating robustness to imaging physics into radiomic feature selection for breast cancer risk estimation. *Cancers (Basel).* 2021;13(21):5497. <https://doi.org/10.3390/cancers13215497> PMID: 34771660
46. Robinson K, Li H, Lan L, Schacht D, Giger M. Radiomics robustness assessment and classification evaluation: a two-stage method demonstrated on multivendor FFD. *Medical Physics.* 2019;46:2145–56.
47. Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst.* 1989;81(24):1879–86. <https://doi.org/10.1093/jnci/81.24.1879> PMID: 2593165
48. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Stat Med.* 2004;23(7):1111–30. <https://doi.org/10.1002/sim.1668> PMID: 15057881
49. Manduca A, Carston MJ, Heine JJ, Scott CG, Pankratz VS, Brandt KR, et al. Texture features from mammographic images and risk of breast cancer. *Cancer Epidemiol Biomarkers Prev.* 2009;18(3):837–45. <https://doi.org/10.1158/1055-9965.EPI-08-0631> PMID: 19258482