














Effectiveness of Mammography Screening on Breast Cancer Mortality – A Study Protocol for Emulation of Target Trials Using German Health Claims Data

Malte Braitmaier ¹, Bianca Kollhorst ¹, Miriam Heinig ², Ingo Langner ², Jonas Czwikla ³, Franziska Heinze ³, Laura Buschmann ⁴, Heike Minnerup ⁴, Xabiér García-Albéniz ^{5,6}, Hans-Werner Hense ⁴, André Karch ⁴, Hajo Zeeb ^{7,8}, Ulrike Haug ^{2,8}, Vanessa Didelez ^{1,9}

¹Department of Biometry and Data Management, Leibniz Institute for Prevention Research and Epidemiology – BIPS, Bremen, Germany; ²Department of Clinical Epidemiology, Leibniz Institute for Prevention Research and Epidemiology – BIPS, Bremen, Germany; ³SOCIUM Research Center on Inequality and Social Policy, University of Bremen, Bremen, Germany; ⁴Institute for Epidemiology and Social Medicine, Faculty of Medicine, Westfälische Wilhelms University of Münster, Münster, Germany; ⁵Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA, USA; ⁶RTI Health Solutions, Barcelona, Spain; ⁷Department of Prevention and Evaluation, Leibniz Institute for Prevention Research and Epidemiology – BIPS, Bremen, Germany; ⁸Faculty of Human and Health Sciences, University of Bremen, Bremen, Germany; ⁹Faculty of Mathematics and Computer Science, University of Bremen, Bremen, Germany

Correspondence: Vanessa Didelez, Leibniz Institute for Prevention Research and Epidemiology – BIPS, Department of Biometry and Data Management, Achterstraße 30, Bremen, 28359, Germany, Tel +49-421-56939, Fax +49-421-56941, Email didelez@leibniz-bips.de

Background: The efficacy of mammography screening in reducing breast cancer mortality has been demonstrated in randomized trials. However, treatment options - and hence prognosis – for advanced tumor stages as well as mammography techniques have considerably improved since completion of these trials. Consequently, the effectiveness of mammography screening under current conditions is unclear and controversial. The German mammography screening program (MSP), an organized population-based screening program, was gradually introduced between 2005 and 2008 and achieved nation-wide coverage in 2009.

Objective: We describe in detail a study protocol for investigating the effectiveness of the German MSP in reducing breast cancer mortality in women aged 50 to 69 years based on health claims data. Specifically, the proposed study aims at estimating per-protocol effects of several screening strategies on cumulative breast cancer mortality. The first analysis will be conducted once 10-year follow-up data are available.

Methods and Analysis: We will use claims data from five statutory health insurance providers in Germany, covering approximately 37.6 million individuals. To estimate the effectiveness of the MSP, hypothetical target trials will be emulated across time, an approach that has been demonstrated to minimize design-related biases. Specifically, the primary contrast will be in terms of the cumulative breast cancer mortality comparing the screening strategies of “never screen” versus “regular screening as intended by the MSP”.

Ethics and Dissemination: In Germany, the utilization of data from health insurances for scientific research is regulated by the Code of Social Law. All involved health insurance providers as well as the responsible authorities approved the use of the health claims data for this study. The Ethics Committee of the University of Bremen determined that studies based on claims data are exempt from institutional review. The findings of the proposed study will be published in peer-reviewed journals.

Keywords: emulated target trial, cancer screening, effectiveness, claims data, mammography

Introduction

Background

Mammography screening aims at reducing breast cancer mortality through early diagnosis of asymptomatic, early-stage cancers.¹ The prognosis of breast cancer is considerably better when diagnosed at an early stage.^{2–4} Several randomized

clinical trials were conducted in the second half of the last century demonstrating a reduction in breast cancer mortality due to screening with mammography.^{5,6}

In parallel to screening efforts increasing world-wide, novel treatment options for women with advanced breast cancer stages have also been introduced over the last two decades leading to improved survival rates, particularly for advanced stages without distant metastases.^{7,8} Consequently, the reduction in breast cancer mortality due to screening might be lower if trials were conducted nowadays. However, mammography techniques have also improved such that the present sensitivity of imaging techniques might have resulted in greater mortality reductions as compared to the earlier trials.⁹ Given that RCTs comparing mammography screening against no screening nowadays are no longer ethical, the analysis of observational data is the only option to obtain insights into the effectiveness of mammography screening under current conditions. A few large observational studies have been conducted on the effectiveness of mammography screening and indicated a reduction in breast cancer mortality in screened women.^{10–12} To date, however, there has been no observational study on this research question using the novel principle of target trial emulation, which specifically aims to minimize common time-related and other biases.

In Germany, an organized mammography screening program (MSP) was introduced from 2005 to 2008, achieving nation-wide coverage in 2009. All women aged 50 to 69 years, with German residency, are centrally invited biennially by mail to attend screening at one of the 94–95 certified mammography screening units.¹ Participation rates in the German MSP are around 50% per screening round and 83% of women in a survey said they had participated at least once over a 10-year time frame.^{13,14} Information on whether and when invitations were issued is not available due to data protection reasons; screening attendance, however, can be identified using specific health insurance claims codes.

Objectives

The proposed observational study is part of a larger research effort commissioned by the German Federal Office for Radiation protection to evaluate whether mammography screening is beneficial in Germany. Within this research effort, our proposed study will estimate the effects of different screening strategies in the German mammography screening program on breast cancer mortality over a 10-year follow-up in women aged 50 to 69 at baseline. Specifically, three screening strategies will be compared: 1) Never screening 2) Screening at least at baseline, with free choice whether to undergo screening afterwards 3) Screening at baseline and then regularly every two years.

Two primary research questions will be addressed:

Research question 1: Does participation in the German MSP reduce breast cancer mortality in the population of all eligible women?

Research question 2: Does participation in the German MSP reduce breast cancer mortality in the subgroup of screening-affine women?

While question 1 addresses the ideal situation that all those eligible participate, question 2 is relevant as it concerns those women who are most likely to participate in the MSP. Both are regarded as primary research questions. The follow-up time of 10 years refers to the time point when the first analysis will be conducted. Re-analysis based on extended follow-up is planned.

For each research question we will consider the following two contrasts:

- Primary contrast: Strategy 1 (never screened) versus strategy 3 (regular screening).
- Secondary contrast: Strategy 1 (never screened) versus strategy 2 (screening at baseline).

Specifically, in view of competing events, we will assess the total effect¹⁵ of the strategies in the primary analysis. The primary contrast reflects the original intention of the MSP and is relevant to the individual women who can decide whether to participate and adhere to the program or not. The secondary contrast addresses the effect of offering the MSP under the reality of imperfect adherence; it is therefore also relevant to public health decision makers.

Materials and Methods

Description of Data Sources

The German Pharmacoepidemiological Research Database (GePaRD) and the BARMER data warehouse (DWH) will be the main data sources for this study. GePaRD comprises health claims data from four German statutory health insurance (SHI) providers, with data on approximately 25 million individuals who have been insured with one of the participating SHI providers since 2004 or later.¹⁶ The BARMER DWH covers approximately 12.6 million individuals who were insured with BARMER between 2006 and 2017.¹⁷ In Germany, health insurance is mandatory, with 87% of the population being insured with an SHI provider (11% of the population are insured with a private insurance provider and further government schemes exist, eg for soldiers or refugees).¹⁸ The health claims data contain basic demographic information, codes for outpatient drug prescriptions, outpatient physician contacts, in- and outpatient operations, procedures, and diagnoses. Outpatient procedures and diagnoses are coded on a quarterly basis, while exact dates are available for inpatient codes and outpatient services. Reimbursed drugs are identified based on Anatomical Therapeutic Chemical (ATC) codes, diagnoses are identified based on International Classification of Diseases, tenth revision, German modification (ICD-10-GM) codes and procedures and services based on Operation and Procedure classification (OPS) codes and Uniform Assessment Standard (EBM) codes.

All regions of Germany are represented in the data from the involved SHIs. GePaRD and BARMER data will be analyzed separately for reasons of data protection. Similar data from further health insurance providers might be added to increase sample size if and when their use for this project will be approved.

Data starting in 2004 for GePaRD and 2006 for BARMER will be used for this study. For the first analysis, data up to and including 2018 will be used. The follow-up will be extended as soon as further data years are available.

Study Design

To address the research questions, we use a target trial emulation approach.^{19,20} While any observational study might suffer from bias due to uncontrolled confounding, awareness has recently increased for biases (often time-related) due to deviation from basic principles of study design. These latter, “self-inflicted” biases, can be avoided or reduced by emulating, as best as possible, the design of a hypothetical randomized trial that would ideally answer the research question (the target trial).²¹ For our proposed mammography study, the protocol of the hypothetical target trial and its emulation with health claims data are described in [Table 1](#). Multiple consecutive trials will be emulated, with one trial starting on the first day of each calendar quarter, to make full use of the longitudinal database. At the core of target trial emulation is the alignment of eligibility checks, assignment to treatment strategies, and start of follow-up at time-zero, ie baseline of each trial. A lack of such alignment is likely to entail erroneous conclusions.^{21,22} Therefore, eligibility criteria will be assessed at the baseline of each emulated trial. As a woman may qualify for multiple trials (starting in different quarters), her individual data will be copied (or “cloned”) and included in every trial for which she is eligible²³ (see [Figure S1](#)). Furthermore, at each baseline, a woman’s data might fit with more than one screening strategy. Again, information from this woman will be copied and one clone will be assigned to each screening strategy she fits. Hence one person can contribute to several trials and within one trial to several screening strategies. This cloning approach reduces time-related biases,^{23,24} while maximizing statistical efficiency.²² Data from all emulated trials and all clones within each trial will be pooled and analyzed jointly. The respective analysis dataset will then contain information on m clones across all trials originating from n women (ie $m \geq n$). Randomization is emulated by adjustment for confounding (more details are given below and in [Supplement 1](#)). Each of the emulated trials has its own baseline, defined as the first day of the calendar quarter of trial start. Pre-baseline covariates are based on information before this day, while follow-up and outcome variables are based on information starting with this day, again ensuring alignment at time zero.

Eligibility Criteria

Individuals must satisfy the eligibility criteria listed in the emulated trial column in [Table 1](#). Eligibility will be assessed at the baseline of each emulated trial.

Table 1 Tabular Study Protocol for the Ideal Target Trial and the Approximation by Our Emulated Trial

Component	Target Trial	Emulated Trial
Aim	To estimate the effect, if any, of different mammography-based screening strategies on breast cancer mortality in the German population aged 50–69.	Same
Eligibility	To be eligible, women must: <ul style="list-style-type: none"> • Be 50 to 69 years old. • Have no history of breast cancer, carcinoma in situ of the breast or unspecified lumps in the breast. • Be naive to screening or diagnostic mammography and other imaging of the breast (in order to avoid selection of the study population according to prior screening history). • Be permanently living in Germany. 	To be eligible, women must: <ul style="list-style-type: none"> • Not have missing information on sex, age, and region of residency. • Be 50 to 69 years old. • Be continuously insured for the 3 years before trial start. • Have no coded diagnosis of breast cancer, carcinoma in situ of the breast or unspecified lumps in the breast) ever before baseline. • Have no coded screening or diagnostic mammography or other imaging of the breast within 3 years before baseline. • Be permanently living in Germany. • For research question 2: have had at least one of the following preventive services coded during 3 years before trial start: screening colonoscopy, pap test or breast examination, health check-up 35, fecal occult blood test, influenza vaccine, skin cancer screening (in order to identify screening-affine women).
Screening strategies	<ol style="list-style-type: none"> 1. Never undergo screening. 2. Screening at least at baseline. 3. Regular screening (two-year intervals). <p>Women are retained under their strategy if they receive a breast cancer diagnosis or if they stop regular screening (strategy 3) at age 70 or older. Receiving a screening mammogram will be considered non-adherence for strategy 1. Under all strategies, diagnostic mammograms are allowed when clinically indicated.</p>	Same
Assignment to study arms	Randomly to one study arm. Randomization is unblinded.	Women are assigned to screening strategies based on observed screening behavior in baseline quarter. We assume random assignment within the levels of the baseline covariates described in the Supplement.
Follow-up	Start: Treatment assignment. End: Death, loss to follow-up or end of study period at 2018, whichever occurs first.	Same, except start is the first day of the quarter of trial start. Length of follow-up is 10 years.
Outcome	Death from breast cancer.	Same (as determined either by the cause of death algorithm or by record linkage).
Causal contrast	Per protocol (PP) effect.	Observational analogue of PP effect. Adjustment for baseline and time-varying post-baseline confounding is necessary.
Statistical analysis	Women are artificially censored when they deviate from their assigned screening strategy as follows: <ul style="list-style-type: none"> • No screening: Censored when a screening mammography occurs. • Screening at baseline: No censoring based on screening participation after baseline quarter. • Regular screening: Censored when no subsequent screening mammography was coded, unless the woman turned 70 or received a breast cancer diagnosis by the tenth quarter after the last screening mammography. <p>The analysis is adjusted for non-adherence using baseline and post-baseline variables (eg via inverse probability weighting).</p>	Same, except that data from each eligible woman receiving screening in the baseline quarter is cloned and assigned to screening strategies 2–3. Randomization will be emulated via adjustment for baseline confounders. Bias due to artificial censoring will be adjusted for using post-baseline confounders.

Note: The emulated trial is purely observational and exposure to the screening strategies is therefore based on the observed (appropriately censored) participation profiles. Sequential trials are emulated in each calendar quarter from 2009 to 2016, with each trial applying the eligibility criteria at its respective baseline.

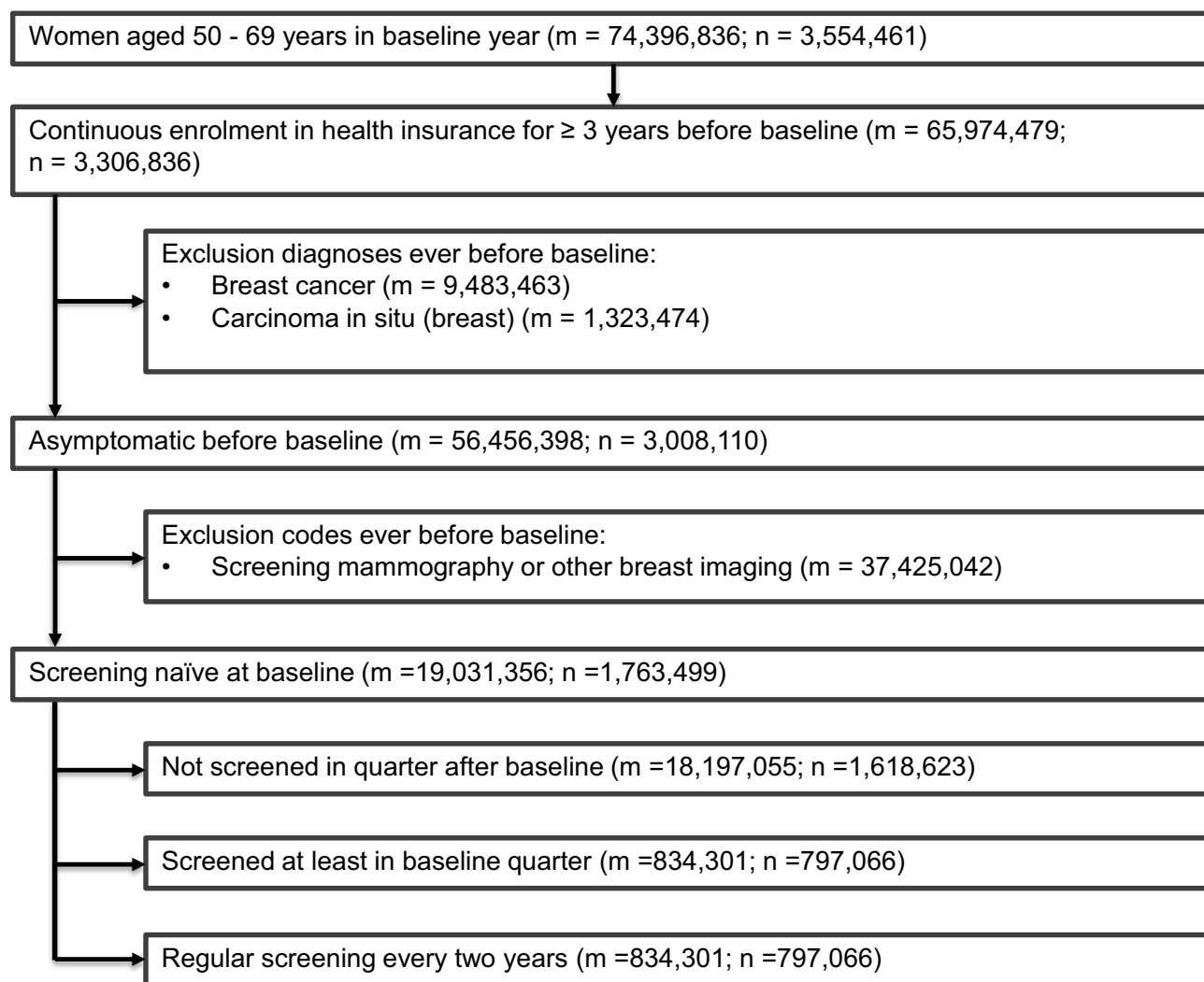


Figure 1 Flow chart of subject disposition. m refers to all clones across all emulated trials while n refers to women. GePaRD data from 2004–2016 was used; no information on the study outcome was available at the time of analysis. Clones were assigned to screening strategies as illustrated in the Supplement. Active screening strategies have identical sample sizes, since they only differ in the screening sustained over time and, therefore, in the censoring process. For the estimation of the subgroup effect in women with at least one other preventive service during three years before baseline, a further decrease in sample size of 32.6% was observed (screening naïve at cohort entry - m: 12,819,058, n: 1,503,094).

Sample Size

An overview of sample sizes (using GePaRD data from 2004 to 2016) is given in the flow chart in [Figure 1](#).

Screening Strategies, Cloning and Artificial Censoring

The mammography screening strategies to be compared will be:

No screening: Under this strategy a woman never undergoes screening (control strategy).

Screening at baseline: Under this strategy, a woman undergoes screening in the baseline quarter and may or may not attend further screenings afterwards.

Regular screening: Under this strategy, a woman undergoes screening in the baseline quarter and in regular two-year intervals thereafter as long as she is in the age range of screening.

Assignment of a woman's data to a strategy has to be done carefully using only baseline information, ie without "looking into the future". Thus, as explained above, her data will be cloned and one clone will be assigned to each strategy with which the woman's behavior is consistent in the baseline quarter, resulting in multiple clones (see [Supplement Figure S1](#) for illustration, and reference²⁰ for a methodological introduction). Note that women who died

in the baseline quarter without starting screening are consistent with and therefore assigned to all strategies. This avoids accumulation of early breast cancer deaths in the no screening strategy and, thereby, avoids bias. The same applies to women who received a breast cancer diagnosis in the baseline quarter without starting screening. Women who undergo screening within the baseline quarter are only cloned into the active screening strategies, since their observed screening behavior at baseline is not consistent with the control strategy. Furthermore, they are cloned into both active screening strategies, which avoids immortal time bias.²¹

Clones will be artificially censored at the start of the first calendar quarter during which their observed screening behavior deviates from the assigned treatment strategy ([Supplement Figures S2–S4](#)). Artificial censoring describes the analyst's decision to ignore any future data for this subject, just as if it were missing.²⁵ For example, a woman's data will be censored in the “never screened” strategy at the time when she receives a screening, and a woman's data will be censored from the “regular screening” strategy at the time when she misses a regular screening. Regular screenings are defined as screenings taking place between the fifth and tenth quarters (ie 12th to 30th month) after the quarter of the previous screening.

Artificial censoring can introduce selection bias if (time-varying) factors influence both deviation from the assigned strategy and the outcome. For instance, a woman may start taking hormone replacement therapy (HRT) and, due to the increased breast cancer risk associated with this medication, also be advised to start regular mammography screening. At that point, she deviates from the “no screening” strategy and would be artificially censored at the time of her first screening mammogram in that strategy. Thus, artificial censoring will be more likely for women using HRT than for women not using HRT, so that censoring might induce selection bias. However, this bias is avoided by weighting with the inverse probability of censoring taking HRT (and other relevant time-varying factors) into account,²⁶ note that this is equivalent to adjusting for time-varying confounding. Women will not be artificially censored under any strategy after receiving a breast cancer diagnosis or after turning 70. No artificial censoring occurs for the “screening at baseline” strategy as any behavior, regular, irregular or lack of further screening after baseline, is compatible with this strategy. Note that within all strategies, diagnostic mammography may take place at any time, as required or indicated, and does not lead to artificial censoring as we aim to assess the added benefit of the MSP.

Exposure

While invitation to screening is not captured in the data, utilization of screening mammography can be identified via a unique EBM code and, thus, is distinguishable from utilization of diagnostic mammography. Women who have not attended screening yet or never attend screening will be included in the control strategy.

Outcome

Information on cause of death is not recorded in health claims data. For the majority of the study population, breast cancer deaths will therefore be identified via an algorithm that uses available information in claims data in the year of death. The algorithm has been developed in a sample for which both claims data and the official cause of death were directly linked. The initial version of the algorithm, described by Langner et al,²⁷ showed a sensitivity of 91.3% and a specificity of 97.4%, and is currently being further optimized, eg by also considering information on cancer treatment. For study participants living in the federal states of North Rhine-Westphalia, Bavaria, and Lower Saxony official cause of death records will be directly available by linkage to the cancer registry of the respective federal state.

Covariates

Covariates for confounder adjustment were selected following subject matter knowledge and considerations about the causal relationships between covariates, exposure and outcome. An illustration of relevant causal patterns is given in [Figure S5](#) in the Supplement. Details on how confounding as a potential source of bias is considered and how relevant covariates are captured in the data are provided in [Supplements 4](#) and [5](#), and a preliminary list of covariates is given in [Table S1](#) in [Supplement 6](#). The list of covariates used in the final analysis will be finalized before data on the study outcome becomes available. All baseline covariates that can vary over time will be re-assessed at each time point, ie on a quarterly basis. All of these updated covariate values will be used to estimate inverse probability weights for artificial

censoring and competing events censoring (the latter only applies to sensitivity analyses). We will apply the usual model diagnostics and carry out balance checks. In addition to these covariates, further variables will be assessed to describe the study cohort.

Missing Data

Individuals with missing core demographic information (ie age, sex, and region of residency) will be excluded from the study. We expect this to be a negligible proportion of women.

We assume that prescriptions, diagnoses, and procedures not coded in our data did not take place. Since no information other than codes from the databases is available, this assumption cannot be verified. Over-the-counter prescriptions and medical services that are not reimbursed by health insurance providers are not coded in our database.

Loss to Follow-Up

Loss to follow-up may occur due to interruption of continuous enrolment, or end of insurance coverage. Interruptions in insurance coverage are very rare in Germany, particularly in the age group relevant for this study.²⁸ We therefore assume that loss to follow-up is neither related to screening participation nor to the risk of breast cancer. Women are censored at loss to follow-up.

Addressing Potential Sources of Bias

As explained under “Study design”, the target trial emulation principle, combined with cloning and artificial censoring, ensures the alignment at time zero and thus mitigates many typical design-related biases in observational studies. Under “Covariates” we further address how information in the claims data can be used to adjust for confounding. A systematic overview and further explanation is provided in [Supplement 4](#), addressing the topics “Confounding”, “Healthy screenee bias”, “Competing events”, “Time-related biases”, “Misclassification” and “Identifying assumptions”.

Primary Analysis

The primary analysis will consist of an estimation of the per-protocol effect of the screening strategies on breast cancer mortality, both in the overall population (research question 1) and in a subgroup of screening affine women (research question 2). One trial will be emulated for each calendar quarter from 2009 to 2016 with one woman possibly contributing to multiple screening strategies per trial. This means that the baseline of the first trial is January 1st, 2009 and follow-up extends until end of data availability. The baseline of the second trial is April 1st, 2009 and follow-up extends until end of data availability. Thus, one trial is emulated per quarter, until the last emulated trial starts on October 1st, 2016. Data from all these emulated trials will be pooled and analyzed jointly. Clones will be artificially censored as described above and reweighted with suitable inverse probability weights. In our main analysis, we estimate the total effect on breast cancer mortality, ie the effect when death from other causes is not eliminated.¹⁵ Adjusted cumulative incidence functions (CIF) will be estimated using a pooled logistic regression (for details on the statistical methods used, see [Supplement 1](#)).

The comparison of the effect of screening strategies will be done in terms of differences in CIF, ie the effect will be observed at each point of follow-up. For the comparison of strategy 1 (never screened) with strategy 3 (regular screening), the above standardization to the empirical distribution of baseline confounders will use the confounder distribution of the entire study population, ie we estimate the average treatment effect (ATE). For the comparison of strategy 1 with strategy 2 (screening at least at baseline), on the other hand, the confounder distribution among treated women will be used, ie we will estimate the average treatment effect on the treated (ATT). The ATT will be more informative to answer the health policy question of whether offering the screening given imperfect adherence (ie contrast between strategies 1 and 2) and given the confounder distribution in women who decide to undergo screening is effective in lowering breast cancer mortality. Confidence intervals will be based on a person-level bootstrap to account for cloning. For a more detailed description of the statistical methods used here, see references.^{15,19,29} The analyses may use a random sample of controls only (ie from the never screened strategy), if computationally prohibitive otherwise.

Furthermore, alternative adjustment methods may be used to adjust for baseline confounding instead of the above-described standardization if the bootstrap sampling becomes computationally prohibitive.

Sub-Group and Secondary Analyses

While primary research question 1 refers to the entire study population, primary research question 2 will assess the effect of screening in the sub-group of screening affine women. These are defined as having attended at least one of the following preventive services during the three years preceding baseline: pap test or breast examination (identified via a single claim code), health check-up after age 35, skin cancer screening, screening colonoscopy, fecal occult blood test, influenza vaccination. By choosing a more restricted and homogeneous study population for primary research question 2, we aim at minimizing residual confounding while being aware that the effect within this special group may be different than in the larger population.³⁰

As a secondary analysis, stratification by calendar year at baseline will be carried out in order to account for the implementation phase of the MSP. We will group all clones from emulated trials with baseline before or in 2011 in one stratum and all others in another stratum. The choice of the cut-off year 2011 is based on baseline characteristics in preliminary analyses (data not shown) and results in one stratum with highly variable age structure (implementation phase until 2011) and one stratum with more homogeneous age structure.

Furthermore, a restricted analysis without women who have a coded family history of breast cancer will be conducted, in order to obtain a subpopulation excluding high-risk individuals.

Sensitivity Analyses

The main analysis assesses the total effect of screening on breast cancer mortality, which encompasses the effect of screening on breast cancer mortality mediated by death due to other causes. This amounts to estimating the event-specific CIF for breast cancer death as event of interest.¹⁵ An estimation of the direct effect, ie under a hypothetical intervention which eliminates all competing events (ie death from other causes), will be conducted in a sensitivity analysis.¹⁵ A comparison with the main analysis will help assess the impact of competing events on any conclusions. Furthermore, the models from the primary analysis will be re-fitted, but with all-cause mortality as outcome variable.

Further sensitivity analyses, such as quantitative bias analysis regarding family history of breast cancer (see [Supplement 4](#)), will be added to the analysis. Results of all secondary and sensitivity analyses will be interpreted in an exploratory way.

Discussion

We propose a design for an observational study in Germany that aims to investigate whether mammography screening reduces breast cancer mortality. The first data analysis will be conducted once 10-year follow-up data are available; extension of follow-up is planned. To the best of our knowledge, there is currently no other study using the principle of target trial emulation to address this research question. García-Albéniz et al used target trial emulation to investigate the continuation of screening mammography after age 70.³¹ There have been other large observational studies investigating screening mammography and breast cancer mortality that were conceptually different from our proposed study. Furthermore, these studies did not have individual-level confounder information, nor an unscreened control group and they did not employ a per-protocol design.^{10,11} The key contribution of our proposed study will be an up-to-date assessment of mammography screening effectiveness in a real-world German population, complementing evidence from earlier randomized trials in other countries. For many reasons, we do not expect exact agreement of our results with those from previous RCTs (see eg Groenwold for a discussion³²), but instead aim at complementing past studies with the best possible evidence currently available on whether screening mammography affects breast cancer mortality in the German population. The chosen screening strategies will inform individual women's choices as well as policy makers. The proposed study design carefully accounts for potential sources of bias and ambiguity. In particular, we also conduct analyses restricted to screening affine women, which is expected to minimize healthy-screenee bias and thus leads to a high internal validity. Moreover, the large size of our database constitutes a clear strength of our proposed study.

While our study focuses on the effectiveness of mammography screening on breast cancer mortality, we are fully aware that mammography screening also has harmful effects. Overdiagnoses are considered to be one of the major harms of mammography screening including subsequent treatment of overdiagnosed cases.^{33,34} Mammography screening programs have been implemented in many countries because it is assumed that the benefits of mammography on breast cancer mortality outweigh these harms, but there is an ongoing debate with some scientists questioning this.^{33,35,36} Given that a very long follow-up of up to 30 years would be required to address overdiagnoses,³⁷ our study cannot contribute to the debate on overdiagnoses. By focusing on the question whether there is a benefit of mammography screening under current conditions in Germany, our study addresses one part of the evaluation required by law. According to German law, any screening method to detect non-communicable diseases entailing exposure to radiation must be assessed both regarding the ability to detect the disease at an early state and thereby improve prognosis and regarding the harm to benefit ratio (§84 of the German Radiation Protection Law, “Strahlenschutzgesetz/StrlSchG”).

While the design of our study addresses several sources of bias, it is still limited by other issues of observational analyses. While we have carefully considered all plausible sources of confounding as detailed in [Supplement 4](#), some unmeasured (baseline or time-dependent) confounding that cannot be mitigated with information based on claims codes cannot be ruled out. In particular, the role of family history of breast cancer, which is only partly observed, will therefore be assessed in quantitative bias analyses. Additionally, for a part of the study population the official cause of death is not available and will instead be identified based on an algorithm that has previously been validated through data linkage. The former version of this algorithm has already shown high sensitivity and specificity and is currently being further optimized.²⁷ Furthermore, imprecision in the date of some codes is present in the data, as outpatient diagnoses are coded on a quarterly basis. We mitigated this by processing all data on a quarterly basis. This, however, introduces the limitation of potential residual time-related biases in quarterly trial emulation. With these limitations in mind, we are confident that our study represents the best analysis currently possible on the effectiveness of the mammography screening program in Germany.

Given data availability, we expect to publish the results of this study by the end of 2024.

Abbreviations

ATC, anatomical therapeutic chemical; ATE, average treatment effect; ATT, average treatment effect on the treated; CIF, cumulative incidence function; DWH, data warehouse; EBM, uniform assessment standard [German: Einheitlicher Bewertungsmaßstab]; GePaRD, German Pharmacoepidemiological Research Database; HRT, hormone replacement therapy; ICD-10-GM, international classification of diseases, tenth revision, German modification; MSP, mammography screening program; OPS, operation and procedure classification [German: Operationen- und Prozedurenschlüssel]; SHI, statutory health insurance.

Ethics Approval and Informed Consent

Pseudonymized health care data may be used for scientific purposes without individual informed consent under German law if the data protection of individuals is not compromised or the public interest of the research project outweighs the data protection interest of individuals (see §75 SGB X). All SHI providers approved the use of their data for this study. Approval of the use of GePaRD data has also been given by the German Federal Office for Social Security and the Senator for Health, Women and Consumer Protection in Bremen. All published results will display aggregated data only, so that identification of individuals will be impossible.

Data Sharing Statement

In Germany, use of personal data is protected by the Federal Data Protection Act and particularly the use of claims data for research is regulated by the Code of Social Law. Researchers must apply for a project-specific permit from the statutory health insurance providers which then need an approval from their governing authorities. The use of the data on which this publication is based was only allowed for BIPS employees within the framework of the specified project and limited to a pre-defined time span. Researchers who want to access the data on which this publication is based need to ask for new approval by the statutory health insurance providers DAK-Gesundheit (service@dak.de), die Techniker

(service@tk.de), hkk Krankenkasse (info@hkk.de) and AOK Bremen/Bremerhaven (info@hb.aok.de) which upon granting approval will ask their respective authorities for approval. Please contact gepard@leibniz-bips.de for help with this process.

Acknowledgments

We are grateful to Klaus Giersiepen for his valuable contributions to this study. Furthermore, we thank Anja Gabbert and Inga Schaffer for the programming of the GePaRD flow chart. We thank all participating health insurance providers for their cooperation. Specifically, we thank the AOK Bremen/Bremerhaven, DAK-Gesundheit, hkk Krankenkasse, and die Techniker (TK) for making available their data for the flow chart shown in this study protocol. A preliminary version of this study protocol was published in a project status report on the institutional website of the study sponsor: https://doris.bfs.de/jspui/bitstream/urn:nbn:de:0221-2022030131647/3/BfS_2022_3617S42402.pdf.

Author Contributions

All authors made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; took part in drafting the article or revising it critically for important intellectual content; agreed to submit to the current journal; gave final approval of the version to be published; and agree to be accountable for all aspects of the work.

Funding

The project was/is funded by the Federal Office for Radiation Protection from funds of the Federal Ministry for the Environment, Nature Conservation and Nuclear Safety, the Federal Ministry of Health and the Kooperationsgemeinschaft Mammographie (supported by the National Association of Statutory Health Insurance Funds and the National Association of Statutory Health Insurance Physicians), grant number: 3617S42402 and 3621S42410.

Disclosure

The authors declare no competing interests.

References

- Biesheuvel C, Weigel S, Heindel W. Mammography screening: evidence, history and current practice in Germany and other European countries. *Breast Care*. 2011;6(2):104–109. doi:10.1159/000327493
- John S, Broggio J. Cancer survival in England: national estimates for patients followed up to 2017. Newport: Office for National Statistics; 2019.
- Sant M, Allemanni C, Capocaccia R, et al. Stage at diagnosis is a key explanation of differences in breast cancer survival across Europe. *Int J Cancer*. 2003;106(3):416–422. doi:10.1002/ijc.11226
- Erdmann F, Spix C, Katalinic A, et al. Krebs in Deutschland für 2017/2018; 2021. Available from: <https://edocrkide/handle/176904/9042>. Accessed August 24, 2022.
- Nelson HD, Fu R, Cantor A, Pappas M, Daeges M, Humphrey L. Effectiveness of breast cancer screening: systematic review and meta-analysis to update the 2009 U.S. preventive services task force recommendation. *Ann Intern Med*. 2016;164(4):244–255. doi:10.7326/M15-0969
- Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. *Br J Cancer*. 2013;108(11):2205–2240. doi:10.1038/bjc.2013.177
- Guarneri V, Conte PF. The curability of breast cancer and the treatment of advanced disease. *Eur J Nucl Med Mol Imaging*. 2004;31(Suppl 1):S149–S161. doi:10.1007/s00259-004-1538-5
- Jansen L, Holleczeck B, Kraywinkel K, et al. Divergent patterns and trends in breast cancer incidence, mortality and survival among older women in Germany and the United States. *Cancers*. 2020;12(9):2419. doi:10.3390/cancers12092419
- Chiarelli AM, Edwards SA, Prummel MV, et al. Digital compared with screen-film mammography: performance measures in concurrent cohorts within an organized breast screening program. *Radiology*. 2013;268(3):684–693. doi:10.1148/radiol.13122567
- Duffy SW, Tabar L, Yen AM, et al. Mammography screening reduces rates of advanced and fatal breast cancers: results in 549,091 women. *Cancer*. 2020;126(13):2971–2979. doi:10.1002/cncr.32859
- Coldman A, Phillips N, Wilson C, et al. Pan-Canadian study of mammography screening and mortality from breast cancer. *J Natl Cancer Inst*. 2014;106(11). doi:10.1093/jnci/dju261
- Lauby-Secretan B, Loomis D, Straif K. Breast-cancer screening - viewpoint of the IARC working group. *N Engl J Med*. 2015;373(15):1479. doi:10.1056/NEJMc1508733
- Kääb-Sanyal V, Hand E. Jahresbericht evaluation 2019 deutsches mammographie-screening-programm. Kooperationsgemeinschaft Mammographie; 2021.
- Schmucker C, Zok K. Informierte Teilnahme an Früherkennungsuntersuchungen: Ergebnisse einer Befragung unter GKV-Versicherten [Informed participation in screening examinations: Results of a survey among individuals enrolled in statutory health insurance]. *Versorgungsreport Früherkennung*. 2019:31–78. doi:10.32745/9783954664023-2

15. Young JG, Stensrud MJ, Tchetgen EJT, Hernan MA. A causal framework for classical statistical estimands in failure-time settings with competing events. *Stat Med.* 2020;39(8):1199–1236. doi:10.1002/sim.8471
16. Haug U, Schink T. German pharmacoepidemiological research database (GePaRD). In: Sturkenboom M, Schink T, editors. *Databases for Pharmacoepidemiological Research*. Springer; 2021:119–124.
17. Czwikla J, Urbschat I, Kieschke J, Schussler F, Langner I, Hoffmann F. Assessing and explaining geographic variations in mammography screening participation and breast cancer incidence. *Front Oncol.* 2019;9:909. doi:10.3389/fonc.2019.00909
18. Busse R, Blumel M, Knieps F, Barnighausen T. Statutory health insurance in Germany: a health system shaped by 135 years of solidarity, self-governance, and competition. *Lancet.* 2017;390(10097):882–897. doi:10.1016/S0140-6736(17)31280-1
19. Hernan MA, Robins JM. *Causal Inference - What if*. Chapman & Hall/CRC; 2020.
20. Hernan MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol.* 2016;183(8):758–764. doi:10.1093/aje/kwv254
21. Hernan MA, Sauer BC, Hernandez-Diaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol.* 2016;79:70–75. doi:10.1016/j.jclinepi.2016.04.014
22. Garcia-Albeniz X, Hsu J, Hernan MA. The value of explicitly emulating a target trial when using real world evidence: an application to colorectal cancer screening. *Eur J Epidemiol.* 2017;32(6):495–500. doi:10.1007/s10654-017-0287-2
23. Zhao SS, Lyu H, Yoshida K. Versatility of the clone-censor-weight approach: response to “trial emulation in the presence of immortal-time bias”. *Int J Epidemiol.* 2021;50(2):694–695. doi:10.1093/ije/dyaa223
24. Maringe C, Benitez Majano S, Exarchakou A, et al. Reflection on modern methods: trial emulation in the presence of immortal-time bias. Assessing the benefit of major surgery for elderly lung cancer patients using observational data. *Int J Epidemiol.* 2020;49(5):1719–1729. doi:10.1093/ije/dyaa057
25. Joffe MM. Administrative and artificial censoring in censored regression models. *Stat Med.* 2001;20(15):2287–2304. doi:10.1002/sim.850
26. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology.* 2000;11(5):550–560. doi:10.1097/00001648-200009000-00011
27. Langner I, Ohlmeier C, Haug U, Hense HW, Czwikla J, Zeeb H. Implementation of an algorithm for the identification of breast cancer deaths in German health insurance claims data: a validation study based on a record linkage with administrative mortality data. *BMJ Open.* 2019;9(7):e026834. doi:10.1136/bmjopen-2018-026834
28. Pigeot I, Ahrens W. Establishment of a pharmacoepidemiological database in Germany: methodological potential, scientific value and practical limitations. *Pharmacoepidemiol Drug Saf.* 2008;17(3):215–223. doi:10.1002/pds.1545
29. Dickerman BA, Garcia-Albeniz X, Logan RW, Denaxas S, Hernan MA. Avoidable flaws in observational analyses: an application to statins and cancer. *Nat Med.* 2019;25(10):1601–1606. doi:10.1038/s41591-019-0597-x
30. Schneeweiss S, Patrick AR, Sturmer T, et al. Increasing levels of restriction in pharmacoepidemiologic database studies of elderly and comparison with randomized trial results. *Med Care.* 2007;45(10Suppl 2):S131–S42. doi:10.1097/MLR.0b013e318070c08e
31. Garcia-Albeniz X, Hernan MA, Hsu J. Continuation of annual screening mammography and breast cancer mortality in women older than 70 years. *Ann Intern Med.* 2020;173(3):247. doi:10.7326/L20-0827
32. Groenwold RHH. Trial emulation and real-world evidence. *JAMA Netw Open.* 2021;4(3):e213845. doi:10.1001/jamanetworkopen.2021.3845
33. Loberg M, Lousdal ML, Bretthauer M, Kalager M. Benefits and harms of mammography screening. *Breast Cancer Res.* 2015;17:63. doi:10.1186/s13058-015-0525-z
34. Baum M. Harms from breast cancer screening outweigh benefits if death caused by treatment is included. *BMJ.* 2013;346:f385. doi:10.1136/bmj.f385
35. Herrmann C, Vounatsou P, Thurlimann B, Probst-Hensch N, Rothermundt C, Ess S. Impact of mammography screening programmes on breast cancer mortality in Switzerland, a country with different regional screening policies. *BMJ Open.* 2018;8(3):e017806. doi:10.1136/bmjopen-2017-017806
36. Biller-Andorno N, Juni P. Abolishing mammography screening programs? A view from the Swiss Medical Board. *N Engl J Med.* 2014;370(21):1965–1967. doi:10.1056/NEJMp1401875
37. Lee CI, Etzioni R. Missteps in current estimates of cancer overdiagnosis. *Acad Radiol.* 2017;24(2):226–229. doi:10.1016/j.acra.2016.05.020

Clinical Epidemiology

Dovepress

Publish your work in this journal

Clinical Epidemiology is an international, peer-reviewed, open access, online journal focusing on disease and drug epidemiology, identification of risk factors and screening procedures to develop optimal preventative initiatives and programs. Specific topics include: diagnosis, prognosis, treatment, screening, prevention, risk factor modification, systematic reviews, risk & safety of medical interventions, epidemiology & biostatistical methods, and evaluation of guidelines, translational medicine, health policies & economic evaluations. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use.

Submit your manuscript here: <https://www.dovepress.com/clinical-epidemiology-journal>