IQWiG Reports – Commission No. S21-01

Review of age limits in the mammography screening programme[1]

Extract

[1] Translation of Chapters 1 to 6 of the final report S21-01 *Überprüfung der Altersgrenzen im Mammografie-Screening-Programmem* (Version 1.1; Status: 16 August 2022 [German original], 25 November 2022 [English translation]). Please note: This document was translated by an external translator and is provided as a service by IQWiG to English-language readers. However, solely the German original text is absolutely authoritative and legally binding.

Publishing details

Publisher
Institute for Quality and Efficiency in Health Care

Topic
Review of age limits in the mammography screening programme

Commissioning agency
Federal Joint Committee

Commission awarded on
22 April 2021

Internal Commission No.
S21-01

Address of publisher
Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Im Mediapark 8
50670 Köln
Germany

Phone: +49 221 35685-0
Fax: +49 221 35685-1
E-mail: berichte@iqwig.de
Internet: www.iqwig.de

This report was prepared in collaboration with external experts.

The responsibility for the contents of the report lies solely with IQWiG.

According to §139b (3) No. 2 of Social Code Book (SGB) V, Statutory Health Insurance, external experts who are involved in the Institute's research commissions must disclose "all connections to interest groups and contract organizations, particularly in the pharmaceutical and medical devices industries, including details on the type and amount of any remuneration received". The Institute received the completed *Form for disclosure of potential conflicts of interest* from each external expert. The information provided was reviewed by a Committee of the Institute specifically established to assess conflicts of interests. The information on conflicts of interest provided by the external experts and external reviewers is presented in Chapter A10 of the full report. No conflicts of interest were detected that could endanger professional independence with regard to the work on the present commission.

External experts (modelling)

- Lára Rún Hallson, UMIT Private University for Health Sciences, Medical Informatics and Technology, Hall in Tirol
- Beate Jahn, UMIT Private University for Health Sciences, Medical Informatics and Technology, Hall in Tirol
- Felicitas Kühne, UMIT Private University for Health Sciences, Medical Informatics and Technology, Hall in Tirol
- Nikolai Mühlberger, UMIT Private University for Health Sciences, Medical Informatics and Technology, Hall in Tirol
- Uwe Siebert, UMIT Private University for Health Sciences, Medical Informatics and Technology, Hall in Tirol
- Gaby Sroczynski, UMIT Private University for Health Sciences, Medical Informatics and Technology, Hall in Tirol

IQWiG thanks the external experts for their collaboration in the project.

Involvement of affected women

Women affected by mammography screening were consulted during the preparation of the report. Barbara Cornelissen, Jutta Fiedler and 3 other women took part in the interview. IQWiG would like to thank them for their participation in the discussion about their experiences. They were not involved in the actual report preparation.

IQWiG employees

- Konstanze Angelescu
- Catharina Brockhaus
- Moritz Felsch
- Wolfram Groß
- Tatjana Hermanns
- Heike Kölsch
- Martina Markes
- Stefan Sauerland
- Mareike Störchel

Keywords

Mass Screening, Mammography, Breast Neoplasms, Benefit Assessment, Systematic Review

Key statement

Research question

The aims of the present investigation are to assess the benefit of mammography screening for breast cancer versus no screening (or breast palpation alone)

- in women aged 45 to 49 years with no signs of breast cancer and no specifically increased risk of breast cancer (Question 1) and

- in women aged 70 years and older with no signs of breast cancer and no specifically increased risk of breast cancer (Question 2)

with regard to patient-relevant outcomes.

Conclusion

Question 1: Age group 45 to 49 years

For breast cancer-specific mortality, the data provide a hint of a benefit of mammography screening versus no screening in women aged 45 to 49 years. With regard to all-cause mortality, there was no statistically significant result. The result on all-cause mortality is no argument against a benefit of mammography screening, because all-cause mortality is considerably influenced by other causes of death. It is assumed that the overall effect of reduced breast cancer mortality through screening was too small to have an impact on all-cause mortality.

However, mammography screening leads to negative consequences (indication of harm) through false-positive screening results. In addition, overdiagnosis occurs (hint of harm). However, the extent of harm is not so great that it outweighs the mortality advantage. With regard to mastectomies, the data provide no hint of a benefit or harm from mammography screening. No data were available for adverse events and health-related quality of life, so no hint of a benefit or harm was shown for these outcomes. However, the effect of screening on the rate of adverse events and on health-related quality of life is likely to be essentially captured by the outcome of overdiagnosis.

In summary, the data provide a hint of a benefit of mammography screening for women aged 45 to 49 years versus no screening; the benefit of mammography screening thus outweighs harm. However, an individual assessment and weighing of benefits and harms is still essential in view of the very small mortality advantage in this age group. Therefore, all conditions must be fulfilled to enable women to make an informed decision.

Question 2: Age group 70 years and older

For breast cancer-specific mortality, the data provide a hint of a benefit of mammography screening versus no screening in women aged 70 to 74 years. This is based on the results for this age group, the transferability of effects from adjacent younger age groups, and the confirmatory results of the modelling performed. With regard to all-cause mortality, there was

no statistically significant result. This is no argument against a benefit of mammography screening, because with increasing age, breast cancer accounts for a smaller and smaller proportion of all deaths, and all-cause mortality is greatly influenced by competing causes of death. It is therefore assumed that the overall effect of reduced breast cancer mortality through screening was too small to have an impact on all-cause mortality.

Mammography screening leads to overdiagnosis and to negative consequences of false-positive screening results (in each case a hint of harm in women aged 70 and older). It is assumed that the expected mortality advantage outweighs the expected harm. No data were available for mastectomies, adverse events and health-related quality of life, so no hint of a benefit or harm was shown for these outcomes.

In summary, the data provide a hint of a benefit of mammography screening versus no screening in women aged 70 to 74 years; the benefit of mammography screening thus outweighs harm. However, an individual assessment and weighing of benefits and harms, considering the individual health status and life expectancy, is still essential in view of the very small mortality advantage, harm observed from overdiagnosis and the consequences of false-positive results, as well as the uncertainties regarding the quantification of effects. Therefore, all conditions must be fulfilled to enable women to make an informed decision.

An ongoing randomized controlled trial (AgeX) is expected to provide meaningful data for both questions of this report in the near future. As soon as the study results are available, it should therefore be examined whether the results are in line with the recommendations to extend the age limits.

Mammography screening for women aged 75 and older currently offers neither a benefit nor a potential versus no screening due to a lack of informative data. The possible benefit with regard to breast cancer-specific mortality, which is, however, quite questionable in this age group, is likely to be countered by harm expected from overdiagnosis and consequences of false-positive screening results.

Table of contents

List of tables

Page

List of abbreviations

Abbreviation	Meaning
BfS	Bundesamt für Strahlenschutz (Federal Office for Radiation Protection)
CI	confidence interval
CNBSS	Canadian National Breast Screening Study
DCIS	ductal carcinoma in situ
ECIBC	European Commission Initiative for Breast Cancer
G-BA	Gemeinsamer Bundesausschuss (Federal Joint Committee)
HIP	Health Insurance Plan of Greater New York
ICD-10	International Statistical Classification of Diseases and Related Health Problems, Revision 10
IDR	incidence density quotient
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (Institute for Quality and Efficiency in Health Care)
MMST	Malmö Mammographic Screening Trial
OR	odds ratio
QALY	quality-adjusted life years
RCT	randomized controlled trial
RR	relative risk
SR	systematic review

1 Background

Breast cancer or mammary carcinoma (ICD-10[2] C50) refers to a malignant neoplasm of the mammary gland. The International Agency for Research on Cancer (IARC) estimated that in 2020, breast cancer was the most common cancer among all new cancers (excluding non-melanoma skin cancer) in women worldwide and the leading cause of cancer death [1]. According to data from the Robert Koch Institute (RKI), this is also true for women in Germany [2]. Known factors influencing the risk of developing breast cancer include age, hormones (e.g., age at first menstrual period, number of births, age at first birth, hormone replacement therapy during or after menopause), previous breast cancer, and benign breast lesions [2]. Breast density has also been described as an independent risk factor. In addition, higher breast density is associated with a decrease in the sensitivity and specificity of mammography [3]. In addition to general factors, women with a family history or genetic predisposition for breast cancer have a specific increased risk of developing breast cancer [4].

The mammography screening programme is a scheme for the early detection of breast cancer in women in Germany. All women in a certain age group are offered a breast x-ray examination at set intervals. If the screening results are abnormal, further procedures are performed for clarification. If an ultrasound examination or a repeated mammography is not sufficient to exclude the suspicion of cancer, the removal of a tissue sample (biopsy) from the breast is recommended [5]. The aim of mammography screening is to improve the chances of a cure by detecting the cancer at an early stage, while less burdensome treatment options are possible due to the earlier tumour stage [3]. At the same time, screening can also lead to false-positive results and thus unnecessary biopsies [6,7]. A proportion of breast cancer diagnoses are overdiagnoses, that is, diagnoses of cancerous changes that would not have been detected without screening and would not have caused symptoms. Accordingly, the treatment that follows this overdiagnosis is unnecessary. Like overdiagnosis, the diagnosis of ductal carcinoma in situ (DCIS) after mammography [8] can also result in unnecessary treatment. DCIS, which has a highly variable potential for malignancy, is considered to be a precancerous breast lesion. However, whether DCIS actually develops into invasive breast cancer cannot be reliably predicted [3].

To ensure the best possible treatment, the current guideline recommendation is that patients with abnormal biopsies should receive further therapy at certified breast centres [3]. If the malignancy of the specimen is pathologically confirmed, besides surgery, there are various therapeutic options such as radiation, hormone therapy, and chemotherapy. In this context, for all non-advanced breast cancer, complete removal of the tumour is the basis of therapy [3].

In Germany, mammography screening was included in the guideline of the Federal Joint Committee (G-BA) on the early detection of cancer, based on European guidelines [9]. In 2005, the national screening programme started [3]. In Germany, every woman between the ages of

[2]International Statistical Classification of Diseases and Related Health Problems, Revision 10.

50 and 69 years is currently invited to participate in mammography screening every 2 years, unless she has objected to the invitation [10]. According to the current guidelines for breast cancer screening and diagnosis by the European Commission Initiative for Breast Cancer (ECIBC), mammography screening is also recommended for women aged 45 to 49 years and 70 to 74 years with no signs and an average risk of breast cancer. Screening is not recommended for women younger than 45 years, and there is currently no recommendation for or against screening for women older than 74 years [11]. In contrast, non-European guidelines, such as the guideline of the Canadian Task Force on Preventive Health Care, do not recommend mammography screening for the age group of women between 40 and 49 years, while, as in the European guideline, women between 70 and 74 years are recommended to be screened every 2 to 3 years [6].

Whether and to what extent women between 45 and 49 years of age and women aged 70 years and older can also benefit from mammography screening is the question addressed in the present benefit assessment.

2 Research question

The aims of the present investigation are to assess the benefit of mammography screening for breast cancer versus no screening (or breast palpation alone)

- in women aged 45 to 49 years with no signs of breast cancer and no specifically increased risk of breast cancer (Question 1) and
- in women aged 70 years and older with no signs of breast cancer and no specifically increased risk of breast cancer (Question 2)

with regard to patient-relevant outcomes.

3 Methods

The target population of the benefit assessment for Question 1 was women aged 45 to 49 years with no signs of breast cancer and no specifically increased risk of breast cancer. The target population of the benefit assessment for Question 2 was women aged 70 years and older with no signs of breast cancer and no specifically increased risk of breast cancer. The test intervention for both questions was mammography screening. The control intervention was no screening or breast palpation alone.

The following patient-relevant outcomes were considered in the assessment:

- Mortality (e.g. all-cause mortality and breast cancer-specific mortality)
- Morbidity (e.g. breast amputations, consequences of false screening results, overdiagnosis)
- Health-related quality of life

- Adverse events

Randomized controlled trials (RCTs) were included in the benefit assessment. Studies that were designated as randomized by the authors, but were considered quasi-randomized studies in the context of the present assessment, were also included. There was no restriction with regard to study duration.

For Question 1 (women 45 to 49 years of age), studies were included even if less than 80% of the women corresponded to the relevant age group. Other age ranges were considered if the data referred to an age group overlapping with the above age group (e.g., 40 to 49 years), because it was sufficiently plausible that such findings would be applicable to the target population of women 45 to 49 years of age. For Question 2, the problem of overlapping age ranges did not exist, as there were no data that referred to women both under and over 70 years of age. If within a question, no usable data on an outcome were available at all for the relevant age group or the corresponding 10-year age group, data from a subgroup with a broader age range were used and issues of transferability were considered accordingly.

In parallel with the preparation of the protocol (report plan), a search for systematic reviews was conducted in the MEDLINE (also includes the Cochrane Database of Systematic Reviews) and HTA databases, as well as on the websites of the National Institute for Health and Care Excellence (NICE) and the Agency for Healthcare Research and Quality (AHRQ). The search was limited to articles published from 2011 onwards.

It was checked whether at least 1 high-quality and current systematic review was eligible where information retrieval could be used as a basis for the assessment (hereinafter: basic SR).

As this was the case, a supplementary search for studies for the period not covered by the basic SR was performed in a second step.

The systematic literature search for studies was performed in MEDLINE, Embase, and the Cochrane Central Register of Controlled Trials.

The following sources of information and search techniques were also used: study registries, documents sent by the G-BA, and screening of reference lists.

Relevant studies were selected by 2 persons independently from one another. Any discrepancies were resolved by discussion. Data were extracted into standardized tables. To assess the qualitative certainty of results, criteria across outcomes and outcome-specific criteria for the risk of bias were assessed, and the risk of bias was rated as high or low in each case. The results of the individual studies were organized according to outcomes and described.

In addition to the comparison of the results of the individual studies, meta-analyses and sensitivity analyses were to be conducted and effect modifiers investigated, provided that the methodological prerequisites had been met.

For each outcome, a conclusion was drawn on the evidence for (greater) benefit and (greater) harm, with 4 levels of certainty of conclusions: proof (highest certainty of conclusions), indication (moderate certainty of conclusions), hint (lowest certainty of conclusions), or neither of the above 3. The latter is the case if no data are available or the available data do not permit classification into one of the 3 other categories. In that case, the conclusion "There is no hint of (greater) benefit or (greater) harm" was drawn.

Subsequently, an assessment of benefit and harm was carried out across outcomes.

If no hint of a benefit or harm could be derived, a conclusion was drawn on the potential of the intervention.

In addition to the benefit assessment, a modelling study was conducted by an external expert group. In the modelling study, the following outcomes were calculated, among others: the lifetime risk of dying from detected invasive breast cancer, the remaining life expectancy, the quality-of-life-adjusted remaining life expectancy in quality-adjusted life years (QALY), and the lifetime risk of overdiagnosis (see Section A4, p. 19 of the full report). The modelling results were incorporated into the report as supplemental information. Details are provided in Section 5. The methods for this modelling study (and its results) are presented in Section A4 of the full report.

4 Results

4.1 Information retrieval results

One systematic review was assessed as a basic SR and considered for the purpose of identifying primary studies.

Information retrieval identified 9 studies relevant to the research question, including 8 randomized and 1 quasi-randomized controlled trial (Table 1). One ongoing study was identified. Planned studies were not identified. Search strategies for bibliographic databases and study registries are provided in the Appendix. The last search was performed on 11 March 2022.

Table 1: Study pool of the benefit assessment

Study	Documents available	
	Full-text publication (in scientific journals)	**Registry entry / Results report from study registries**
AgeX Pilot	yes [12]	yes [13] / no
CNBSS	yes [14-22]	no / no
Edinburgh	yes [23-25]	no / no
Gothenburg	yes [26-28]	no / no
HIP	yes [29-35]	no / no
Malmö	yes [36-40]	no / no
Stockholm	yes [41-44]	no / no
Swedish Two-County	yes [45-55]	yes [56] / no
UK Age	yes [57-64]	yes [65] / no
CNBSS: Canadian National Breast Screening Study; HIP: Health Insurance Plan of Greater New York		

4.2 Characteristics of studies included in the assessment

All 9 eligible studies included women aged 45 to 49 years (Question 1), even though none of the studies covered the exact age range of the question. Only 3 studies were identified that included women aged 70 years and older (Question 2). These 3 studies referred to women up to a maximum of 74 years at the time of randomization.

The AgeX Pilot study [12] is a multicentre cluster RCT involving 6 breast cancer units in the United Kingdom. Between June 2009 and May 2010, the pilot study recruited women aged 47 to 49 years and 71 to 73 years. Thus, the study is used for both Question 1 and Question 2. The screening period lasted only 12 months, as it was essentially a feasibility study (for the currently ongoing AgeX study, see Section 4.6). Accordingly, no mortality data were collected and there was no follow-up beyond the screening period. In the screening group, 47- to 49-year-old women and 70- to 73-year-old women each underwent 1 screening in addition to the national screening programme offered between 50 and 70 years of age.

The Canadian National Breast Screening Study (CNBSS) [17] is a multicentre RCT with study centres in Canada. It started in 1980 and includes 2 age groups (women 40 to 49 years and women 50 to 59 years), where data from the younger age group can be used for Question 1. Before randomization, women underwent a clinical breast examination and received breast self-examination guidance. Women in the intervention group were invited for annual screening. The duration of the screening period (the term here and hereafter refers to the period during which one study group is screened and the other is not) was up to 5 years, but the women recruited later participated in only 3 or 4 rounds of screening. Women in the control group received breast self-examination guidance and were contacted annually.

The Edinburgh study [24] is a cluster RCT that started in 1979. A large proportion of the city's GP practices consented to participate. Practices were randomly assigned to the screening or

control group. Women in the age group 45 to 64 years registered with each practice were assigned to the groups according to the allocation by the practice. Women in the intervention group were screened every 2 years for a total of 7 years. Women in the control group were not screened.

The Health Insurance Plan of Greater New York (HIP) study [34] is by far the oldest of the studies included. The HIP study started in New York in 1963, 13 years earlier than the second oldest study included. By 1966, over 61,000 women aged 40 to 64 years had been recruited and randomly assigned to either annual screening for 4 rounds of screening or to a no-screening group.

Four of the studies included were conducted in Sweden: the Gothenburg Study [26], the Malmö Study [37], the Stockholm Study [42], and the Swedish Two-County Study [51].

The Gothenburg Study started in 1982 and included about 52,000 women aged 39 to 59 years, of whom about half were under 50 years. The screening period for this age group (39 to 49 years) was 7 years. During this time, women in the intervention group were offered screening every 18 months, for a total of 5 rounds of screening.

The Malmö Study comprises 2 study populations recruited according to the same study protocol: The first Mammographic Screening Trial (MMST 1) started in 1976 and included women who were between 45 and 70 years of age at recruitment. Age grouping in one of the analyses was based on exact age at randomization (rather than year of birth), so separate data were available for 70-year-old women, allowing the study to be used for both Question 1 and Question 2. In MMST 2, which started in 1978, women aged 44 to 49 years were recruited. The screening period lasted a mean of 8.8 years, based on the entire study population (MMST 1 and 2). Women underwent up to 9 (MMST 1) or a mean of 5 (MMST 2) screenings during the period, with a screening interval between 1.5 and 2 years. The 70-year-old women underwent a maximum of 6 screenings. Separate data on screening parameters for the two age groups relevant to this report were not available.

The Stockholm study, which started in 1981, recruited women who were between 40 and 64 years old at baseline. Although the study was described as randomized in the associated publications, because allocation to the screening and control groups was based on the day of birth, it was classified as quasi-randomized. Women in the intervention group were screened twice with an interval of 2 years, which was only once more than women in the control group, who underwent a final screening.

The cluster RCT Swedish Two-County Study recruited women in Kopparberg (Dalarna) and Östergötland from 1977 onwards and evaluated results from women aged 40 to 74 years. The study is thus used for both Question 1 and Question 2. The screening period lasted between 6 and 8 years, and during that time, the 40- to 49-year-old women underwent 4 rounds of screening and the 70- to 74-year-old women underwent 2 rounds of screening. The scheduled

screening interval was 2 or 3 years, depending on age. For women from 40 to 49 years of age, the mean screening interval was 24 months, and for women 50 years of age and older, 33 months; separate data on women of 70 years of age and older were not available.

The UK Age Study [60] included women aged 39 to 41 years at randomization. The study started in 1991 and included 23 study sites in the United Kingdom. The screening period lasted until the age of 48, ranging from 7 to 9 years. Thus, this is the only one of the studies included whose study population remained below 50 years of age throughout the screening period. Women in the intervention group were invited for annual mammography. Women in the control group received standard care without screening. With over 160,000 randomized women, UK Age is the largest study included.

In all studies included (except the AgeX Pilot study), breast cancer-specific mortality was considered the primary outcome. The inclusion criteria for all studies comprised only the age and the place of residence of the women (in the HIP study, instead of the place of residence, membership in the HIP). Exclusion criteria were only mentioned in 3 studies; besides previous breast cancer (CNBSS, Edinburgh and Gothenburg study), they comprised a current pregnancy as well as a mammogram in the last year (CNBSS). Furthermore, the studies differed with regard to the screening strategy: although all women in the screening group underwent mammography screening, the number of screening rounds (1 to 8) and the screening interval, i.e., the interval between screening rounds varied (12 to 24 months).

In 3 studies (Gothenburg, Stockholm, Swedish Two-County), women from the control group underwent screening after the end of the screening period as part of the study. In another 4 studies, the national screening programme started soon after the end of the screening period (AgeX Pilot, Edinburgh, Malmö, UK Age), so that women from both groups subsequently received a screening offer, sometimes restricted to certain age groups. In 1 study (AgeX Pilot), women aged 71 to 73 years underwent 1 screening following the national screening programme. In the remaining 2 studies, no information was available on whether and, if yes, when screening of all women occurred after the end of the screening period (CNBSS, HIP); at least in the HIP study, however, it can be assumed that screening of all women did not follow initially.

Data characterizing the study population were available from only 1 of the 9 studies (CNBSS). In this study, selected prognostic characteristics (e.g., reproductive status, smoking status) were presented in addition to demographic information (e.g., marital status, occupation). A family history, one of the main prognostic characteristics, was present in 9493 women (37.6%) in the intervention group and 9652 women (38.3%) in the control group.

4.3 Overview of patient-relevant outcomes

Data on patient-relevant outcomes could be extracted from 9 studies. Table 2 shows an overview of the usable data on patient-relevant outcomes from the studies included. For mastectomies, usable data were only available for Question 1. No usable data on adverse events and health-related quality of life were available from any study.

With the exception of the UK Age Study, which included women between 39 and 41 years of age, all studies contained results broken down by age subgroups, mostly summarized into 10-year age groups (e.g., 40 to 49 years), each based on age at the time of randomization.

Age subgroups used for Question 1

Data on precisely the relevant age group of 45- to 49-year-olds were available only sporadically. However, it seemed plausible with sufficient certainty that results from slightly younger women would be applicable to the actual target population. Data were therefore used from all age subgroups that included the 39- to 49-year age range (e.g., 39-49 years, 39-41 years, 45-49 years, 47-49 years).

Age subgroups used for Question 2

The AgeX Pilot study provided data on the 71- to 73-year-old age group. From the Malmö study, data on the age subgroup of over 69-year-olds could be used. It can be assumed that these were exclusively 70-year-old women (see Section 4.2). From the Swedish Two-County Study, data on the age subgroup of 70- to 74-year-olds were used.

Table 2: Matrix of patient-relevant outcomes

Study	Outcomes						
	Mortality		Morbidity				HrQoL
	All-cause mortality	Breast cancer-specific mortality	Mastectomies	Consequences of false-positive screening results	Adverse events	Overdiagnosis	Health-related quality of life (instrument)
Question 1							
AgeX Pilot	–	–	–	●	–	–	–
CNBSS	●	●	●	●	–	●	–
Edinburgh	–	●	–	–	–	–	–
Gothenburg	●	●	–	●	–	●	–
HIP	–	●	–	●	–	●	–
Malmö	–	●	–	–	–	●	–
Stockholm	–	●	–	●	–	–	–
Swedish Two-County	●	●	–	–	–	●	–
UK Age	●	●	–	●	–	●	–
Question 2							
AgeX Pilot	–	–	–	●	–	–	–
Malmö	–	●	–	–	–	–	–
Swedish Two-County	●	●	–	–	–	●	–

●: Data available and usable.
–: Data not available and not usable.

CNBSS: Canadian National Breast Screening Study; HIP: Health Insurance Plan of Greater New York; HrQoL: health-related quality of life

4.4 Assessment of the risk of bias of the results

The risk of bias of the studies included was rated as low across outcomes for 1 study (UK Age) and as high for the remaining 8 studies. For the 3 studies that were included for both questions, there were no differences in the assessment of the risk of bias between Question 1 and 2.

In 7 of 8 studies with a high risk of bias, it was unclear whether allocation to the intervention and control groups was concealed. At this point, special reference should be made to the CNBSS study. In this study, a clinical breast examination was performed before randomization. Compared with other studies, each of which diagnosed fewer cases of advanced breast cancer

in the screening group than in the control group, it is notable that there were markedly more cases of advanced breast cancer in the screening arm of the CNBSS than in the control group [66]. Therefore, some authors in the literature have critically questioned whether women with a palpable tumour in the clinical examination may have been more likely to be assigned to the screening group [66-68]. If this were true, the results of the CNBSS with regard to all outcomes might be strongly biased against screening and could only be used with restrictions. The preferential allocation of women with palpable tumours to the screening group would, among other things, lead to an overestimation of the total number of breast cancer cases in this group compared with the control group. Consequently, no valid estimate would be possible with regard to overdiagnosis, because not only would more breast cancer cases be identified in the intervention group by screening, but additional cases would result from the arbitrary assignment of women with abnormal palpable findings. This would potentially overestimate the screening effect for the outcome of overdiagnosis. Furthermore, since the prognosis of advanced breast cancer is less favourable than that of early breast cancer, the increased number of women with advanced breast cancer in the screening group would be expected to result in increased mortality compared with the control group. Consequently, there would be an underestimation of the screening effect. However, whether such arbitrary allocation and associated bias actually occurred remains unknown on the basis of the data.

In 4 studies (Edinburgh, Gothenburg, HIP, Swedish Two-County), it was also unclear whether the randomization sequence was adequately generated. In the Edinburgh study, cluster randomization led to group differences in social status [69] and thus to apparently unequal groups, which was also reflected in a marked difference in all-cause mortality (related to the total study population of 45- to 64-year-olds, not presented in the report for this reason). A relevant bias in the results on breast cancer-specific mortality was therefore also possible; this was taken into account in sensitivity analyses.

For one other study (Stockholm), the generation of the randomization sequence was not adequate because assignment to groups was based on date of birth. The Stockholm study was therefore considered quasi-randomized.

The high risk of bias across outcomes directly translates to the risk of outcome-specific bias, so that all outcomes reported in these 8 studies have a high risk of outcome-specific bias.

Thus, the outcome-specific risk-of-bias criteria were assessed only for the UK Age study. Outcome-specific risk of bias for the results on all-cause mortality, breast cancer mortality, and overdiagnosis was assessed as low. For the results on consequences of false-positive screening results (biopsies with benign findings), on the other hand, it was assessed as high because the blinding of the outcome assessors and the adequate implementation of the intention-to-treat (ITT) principle were unclear due to insufficient reporting.

4.5 Results on patient-relevant outcomes

4.5.1 Question 1: age group 45 to 49 years

4.5.1.1 Results on mortality

Results on all-cause mortality

For all-cause mortality, usable data separated by age subgroups were available from 1 study with a high (UK Age) and 3 studies with a moderate qualitative certainty of results (CNBSS, Gothenburg, Swedish Two-County). Data from 10-year age groups (40-49 and 39-49 years) were used from 3 studies (CNBSS, Gothenburg, Swedish Two-County), and data from the entire study population, women aged 39 to 41 years, were used from the fourth study (UK Age). Results referred to follow-up periods between 7.9 and 22.8 years since randomization, with data available at multiple times in 2 studies (CNBSS, UK Age). Data from 7.9 to 13 years of follow up were included in the pooled estimate.

Neither the study with a high qualitative certainty of results at 3 different times nor the pooled estimate from all 4 studies showed a statistically significant effect (incidence density ratio [IDR] 0.98; 95% confidence interval [95% CI]: [0.90; 1.08]; p = 0.602).

There was no effect modification on all-cause mortality for the characteristics “screening interval” and “screening period duration”. No data were available to examine effect modification by the characteristics “age” and “breast density”.

Thus, there is no hint of a benefit or harm of mammography screening for all-cause mortality.

Results on breast cancer-specific mortality

For breast cancer-specific mortality, usable data were available from all 8 studies, including 1 study (UK Age) with a high, 6 with a moderate (CNBSS, Edinburgh, Gothenburg, HIP, Malmö, Swedish Two-County), and 1 with a low qualitative certainty of results (Stockholm). Data from the following age groups were used: the 10-year age groups (40-49 and 39-49 years) from 3 studies (CNBSS, Stockholm, Swedish Two-County), the 5-year age group (45-49 years) from 1 study (Edinburgh), both the 10- and 5-year age groups from 2 studies (Gothenburg, HIP), as well as the age groups of 44 to 49 years from 1 study (Malmö) and 39 to 41 years from 1 study (entire study population, UK Age). From the Gothenburg and HIP studies, data from the 5-year age group were primarily used, whereas data from the 10-year age group, which actually do not fit the report question so well, were used to test the overall impact of a broader definition of the age group limit in a sensitivity analysis. Results of follow-up periods between 4 and 24 years since randomization were presented.

Meta-analyses were performed at 4 different times, summarizing results at approximately 5, 10, 15, and 20 years after randomization. At about 10 years, data from all 8 studies could be included; at the other times, data could be included only from 4 or 5 studies.

Neither the results of the one study with a high certainty of results (UK Age) included in the meta-analyses at the 4 times of analysis, nor the respective pooled estimate, showed a statistically significant difference. In the meta-analysis at approximately 10 years – the time at which most of the data were available – statistical significance was narrowly missed (IDR 0.85; 95% CI: [0.73; 1.002]; p = 0.052).

In the UK Age study, data were available at 2 different times (9 years and 10.7 years), both of which matched the time of approximately 10 years and from which the data at the later time (10.7 years) were selected for the main analysis. The data at 9 years after randomization (also with a high certainty of results) showed a statistically significant effect (IDR 0.75; 95% CI: [0.58; 0.96]). A sensitivity analysis was performed to examine the extent to which it makes a difference as to which of the two times in the UK Age Study that fits this analysis is included in the meta-analysis. In contrast to the main analysis, this sensitivity analysis yielded a statistically significant result (IDR 0.83; 95% CI: [0.70; 0.99]; p = 0.041). There is already an indication of an effect of the statistically significant result of the single study with a high qualitative certainty of results. The pooled estimate from all studies confirms this effect. Therefore, proof of an effect in favour of mammography screening is derived in the sensitivity analysis.

According to the study authors, a possible explanation for the fact that the UK Age Study showed an effect at 9 years, but no longer at 10.7 years, is that while there was a survival advantage for women with Stage 1 and 2 tumours in both the short and long term, for women with Stage 3 tumours, this advantage was only shown within the first 10 years. From this, they concluded that screening may prolong survival in women with more aggressive tumours, but does not prevent death from breast cancer, thus diluting the effect. This explanation seems plausible, and prolongation of survival should also be considered an advantage.

Therefore, when the main analysis at 10 years (which only narrowly misses statistical significance) and the sensitivity analysis with the 9-year data from the UK Age Study (proof of an effect) are considered together, the data provide a hint of a benefit from mammography screening.

The fact that there was no statistically significant effect at the other times (5, 15, and 20 years) does not contradict the results at 10 years. This is because at the other times, data were available from only 4 and 5 studies, with correspondingly lower precision of the pooled effect estimate. Furthermore, an effect is most likely to be expected at about 10 and 15 years, based on the fact that it takes at least 9 to 12 years for an effect of screening to become apparent [70], and on the fact that any effect that may be present is expected to be diluted with increasing time after the end of the screening period.

Further sensitivity analyses did not contradict these results.

No effect modification was shown for the characteristics “screening interval” and “screening period duration”. No data were available to investigate effect modification by the characteristics “age” and “breast density”.

Overall, the results on breast cancer-specific mortality thus provide a hint of a benefit of mammography screening.

Overall assessment of mortality

For all-cause mortality, there was no hint of a benefit or harm from mammography screening, but for breast cancer-specific mortality, there was a hint of a benefit. Since causes of death other than breast cancer are much more common in this age group – only 12.7% of deaths in women aged 45 to 54 years are attributable to breast cancer (ICD-10 code C50) [71] – the result for all-cause mortality is no argument against a benefit of mammography screening.

4.5.1.2 Results on mastectomies

Mammography screening aims to reduce mortality from breast cancer by detecting and treating breast cancer cases at an earlier stage. In addition, starting treatment at an earlier stage of the disease is associated with the expectation that therapy can be less invasive without jeopardizing the success of treatment. An advantage for breast cancer patients may therefore result from the fact that breast-conserving surgery can be performed more frequently without having a detrimental effect on mortality. However, a lower proportion of mastectomies among all breast cancer surgeries in the screening group can only be considered an advantage if the absolute rate of mastectomies in relation to all randomized women is also lower in the screening group than in the control group.

For breast cancer surgery, usable data were available from 1 study with a moderate qualitative certainty of results (CNBSS). Results were available for 40- to 49-year-olds. Breast-conserving surgery and mastectomies were reported in the study. Results after 7 years were used. The results after the first round of screening are presented only as supplemental information. This is because a relevantly higher number of mastectomies can be expected at this time, since screening also detects more breast cancers.

As expected, after the first round of screening there was a statistically significant difference to the disadvantage of screening with respect to total surgeries (odds ratio [OR] 1.58; 95% CI: [1.15; 2.18]; p = 0.004) as well as mastectomies (OR 1.67; 95% CI: [1.10; 2.52]; p = 0.014). The proportion of mastectomies among all breast cancer surgeries was not statistically significantly different between the groups (61.2% vs. 58.1%; OR 1.14; 95% CI: [0.60; 2.18]; p = 0.709).

After 7 years, the total number of surgeries was statistically significantly higher in the screening group than in the control group (OR 1.33; 95% CI: [1.15; 1.54]; p < 0.001). However, there was no statistically significant difference in mastectomies between groups (OR 1.17; 95% CI: [0.94; 1.45]; p = 0.158). Similarly, the proportion of mastectomies among all breast cancer surgeries

was not statistically significantly different between groups (44.1% vs. 50.2%; OR 0.78; 95% CI: [0.58; 1.05]; p = 0.116).

Thus, the higher number of surgeries in the screening group at 7 years was not associated with a reduced absolute rate of mastectomies, i.e., a discernibly lower invasiveness of surgeries. Based on this, the data provide no hint of a benefit or harm from mammography screening.

4.5.1.3 Results on consequences of false-positive screening results

For the outcome "consequences of false-positive screening results", data were used for screening participants who had a positive screening result and for whom the suspicion of breast cancer was not confirmed in the subsequent invasive diagnostic procedures. For this outcome, both data on purely diagnostic interventional clarification and on surgical therapeutic interventions were relevant if the diagnostic and therapeutic intervention for breast tissue of unclear histological classification (malignant or benign) could not be clearly separated. Results were available on biopsies with benign findings (AgeX Pilot, CNBSS, HIP, Malmö, Stockholm, UK Age) and surgery with benign findings (AgeX Pilot, Gothenburg). By definition, data on this outcome refer only to screened women. A group comparison is therefore not applicable, even though diagnostic and therapeutic interventions with benign findings may also occur in the control group.

Usable data on the consequences of false-positive screening results were available from 6 studies, 5 of which had a moderate certainty of results (AgeX Pilot, CNBSS, HIP, Malmö, UK Age) and 1 of which had a low certainty of results (Stockholm). Data from the following age groups were used: the 10-year age groups (40-49 years; CNBSS, HIP), the 5-year age groups (45-49 years; Malmö, Stockholm) as well as the age groups of 47 to 49 years (entire study population, AgeX Pilot) and 39 to 41 years (UK Age). Events were presented differently in the studies: in 1 study (HIP), only data on the proportion of biopsies with benign findings out of all biopsies were available, without specifying the reference period. The results of the remaining 5 studies referred either to the number per screening year (up to a maximum of 5 years; CNBSS) or per screening round (up to a maximum of 5 rounds; Malmö, Stockholm, UK Age) or cumulatively to the entire follow-up period (AgeX Pilot: 12 months; Malmö: 10 to 15.5 years). Therefore, no overall summary estimate is provided for this outcome, but a range of effect estimates from the individual studies [minimum; maximum].

The proportion of women in the first screening round or year who had a biopsy only because of a false-positive result in screening ranged from 0.16% to 3.4%. In subsequent screening rounds or years, the proportions per round or year were somewhat lower, ranging from 0.07% to 1.8%. The proportion of biopsies with benign findings among all biopsies was 90% (HIP study).

The proportion of screening participants who underwent surgery only because of a false-positive screening result ranged from 0.15% (AgeX Pilot) to 0.27% (Gothenburg) in the first screening round and from 0.03% to 0.13% (Gothenburg) in subsequent rounds. Among these,

the proportion of surgeries with benign findings among all breast cancer surgeries ranged from 16% (4/25) to 65% (32/49) in the Gothenburg study, and the proportion of surgical procedures with benign findings was 23% (19/84) in the AgeX Pilot study. It is to be expected that with current diagnostics, fewer surgeries with benign findings occur than in the 1980s, when the Gothenburg study was performed. Therefore, the results of the AgeX Pilot study are more reflective of the current status.

Thus, for consequences of false-positive results, based on the results on biopsies with benign findings and surgeries with benign findings, overall there is an indication of harm from mammography screening.

4.5.1.4 Results on adverse events

No data were available on this outcome.

4.5.1.5 Results on overdiagnosis

For overdiagnosis (based on breast cancer incidence), data were available from 6 studies, 1 with a high (UK Age) and 5 with a moderate qualitative certainty of results (CNBSS, Gothenburg, HIP, Malmö, Swedish Two-County). Data from 10-year age groups were available from 5 studies (CNBSS, HIP, Swedish Two-County: 40-49 years; Gothenburg: 39-49 years; Malmö: 45-54 years). In 2 studies, additional data on 5-year age groups were available (Gothenburg, HIP: 45-49 years). Data on 39- to 41-year-olds were available from 1 study (UK Age).

The methods used for this assessment to calculate overdiagnosis were only meaningfully applicable in 2 studies (CNBSS, HIP). This is because this calculation is based on a comparison of the breast cancer incidence of screened women with a control group without screening, i.e., in which only symptomatic breast cancer cases were recorded. In each of the other 4 studies there was a final screening in the control group or screening of all women after the end of the screening period. Once a control group is also screened, this group will include pre- and asymptomatic breast cancer cases. In such cases, sufficiently valid results on overdiagnosis are therefore not to be expected on the basis of the methods used for the present assessment.

Overdiagnosis related to women invited for screening

In the CNBSS study, the risk of overdiagnosis related to women invited for screening was 0.41%. At 0.04% (age group 40-49 years) and 0.01% (age group 45-49 years), the values in the HIP study were markedly lower.

If the limited validity of the calculation method in the other 4 studies is disregarded, the risk for the Malmö study would be 0.49% (after screening of controls) and for the Swedish Two-County Study, it would be 0.25% (after screening of controls). These two values are between those of the CNBSS and HIP studies.

As in the Malmö and Swedish Two-County studies, an approximate calculation of the risk of overdiagnosis was not possible for the Gothenburg and UK Age studies, as the number of breast

cancer diagnoses in the control group was higher than in the screening group at the analysis times reported. However, the authors of the UK Age study themselves reported a value of 0.2% in relation to the women invited for screening [57]. The calculation included test sensitivity and sojourn time (duration of preclinical detectability of asymptomatic breast cancer), which was estimated using interval carcinomas. The reported value of 0.2% lies in the range of the values calculated in this report for the other studies included.

It should be noted that it cannot be excluded that the value of 0.41% determined for the CNBSS study is overestimated. As described in Sections 4.2 and 4.4, a clinical breast examination was performed in all women prior to randomization. During the course of the study, a striking cluster of advanced breast cancers was observed in the screening group, which presumably had already been palpable at the initial examination. Some authors therefore suspected that women with abnormal or unclear findings might have been preferentially assigned to the screening group [67]. If this were true, it would influence the overall number of breast cancer diagnoses in both the screening group and the control group, which would lead to an overestimation of the risk of overdiagnosis.

In the overall view of the results of all studies, however, this does not result in a limitation, since, at 0.01% to 0.41%, the values for the overdiagnosis risk in relation to the women invited for screening in the age group between 45 and 49 years point in one direction in all studies. A meta-analytical summary of the results was not performed because the proportions were markedly different.

Overdiagnosis related to women diagnosed with breast cancer during the screening period

In the CNBSS study, the risk of overdiagnosis among women who received a breast cancer diagnosis during the screening period was 31.61%. At 8.06% (age group 40-49 years) and 2.63% (age group 45-49 years), the values in the HIP study were markedly lower.

If the limited validity of the calculation method in the other 4 studies were disregarded, the risk would be 12.98% for the Malmö study (after screening of controls) and 21.00% for the Swedish Two-County Study (after screening of controls). These two values lie between those of the CNBSS and HIP studies.

For the Gothenburg and UK Age studies, an approximate calculation of the risk of overdiagnosis in women who had received a breast cancer diagnosis during the screening period was also not possible, since at the analysis times reported, the number of breast cancer diagnoses was higher in the control group than in the screening group. However, the authors of the UK Age study report a value of 8.5% [57]. The calculation included test sensitivity and sojourn time, which was estimated using interval carcinomas. The reported value of 8.5% lies in the range of the values calculated in this report for the other studies included.

As already noted in the upper section, it cannot be excluded that the calculated value of the CNBSS overestimates the risk of overdiagnosis.

In the overall view of the results of all studies, however, this does not result in a limitation, since, at 2.63% and 31.61%, the values for the risk of overdiagnosis in women in the age group between 45 and 49 years who had received a breast cancer diagnosis during the screening period point in one direction in the studies. A meta-analytic summary of the results was not performed because the proportions were markedly different.

Summarized assessment for overdiagnosis

After invasive diagnostics to confirm the diagnosis, breast cancer diagnoses are usually followed by treatment. Both the further diagnostic procedures and the subsequent treatment carry a risk of adverse effects and complications. Since overdiagnosis represents the diagnosis of a disease that would never have become apparent and would not have caused any symptoms without screening, the subsequent interventions, together with their consequences, are superfluous. Even if in some cases less invasive therapeutic approaches are used today (e.g. more frequent breast-conserving surgeries instead of mastectomies), overtreatment remains a detriment.

Overall, there is therefore a hint of harm from mammography screening versus no screening with respect to the outcome of overdiagnosis in women between 45 and 49 years of age.

4.5.1.6 Results on health-related quality of life

No data were available on this outcome.

4.5.2 Question 2: age group from 70 years

4.5.2.1 Results on mortality

Results on all-cause mortality

For all-cause mortality, usable data with a moderate qualitative certainty of results were available from 1 study (Swedish Two-County) at 7.9 years after randomization.

There was no statistically significant effect (relative risk [RR] 0.98; 95% CI: [0.93; 1.03]; p = 0.384).

No suitable data were available to examine effect modifiers.

Thus, for all-cause mortality, there is no hint of a benefit or harm of mammography screening in women aged 70 years and older.

Results on breast cancer-specific mortality

For breast cancer-specific mortality, usable data with a moderate qualitative certainty of results were available from both studies. Data were available at 13.6 years after randomization from

the Malmö study and at 7.9, 10.8, and 20 years from the Swedish Two-County Study. Because of the different data cuts, only the results at 13.6 and 10.8 years were summarized meta-analytically. In the pooled estimate from both studies, there was no statistically significant difference at this point (IDR 0.91; 95% CI: [0.61; 1.36]; p = 0.655). Here, the result of the pooled estimate is driven by the Swedish Two-County Study because of the approximately 20-fold higher weighting. The results at the other 2 times in the Swedish Two-County Study also showed no statistically significant difference, but considerably more favourable effect estimates (7.9 years: IDR 0.77; 95% CI: [0.47; 1.27]; p = 0.3; 20 years: RR 0.76; 95% CI: [0.53; 1.08]; p = 0.113). Why the mean data cut yielded a different result than the consistent results before and after is unclear. All results are used for the further weighing of benefits and harms.

For interpretation purposes it must be taken into account that in the Swedish Two-County Study, the sample size for the 70 to 74-year-old age group is considerably lower than for the middle-aged groups (e.g., only half as large as that for 60-69 year-old women). This results in a correspondingly lower precision of the estimate. However, the relative effect in women aged 70-74 years, for example, when looking at 20-year results (IDR 0.76, 95% CI: [0.53; 1.08]), is consistent with that in the middle-aged groups (50-59 years: RR 0.86; 95% CI: [0.68; 0.97]; 60-69 years: RR 0.67; 95% CI: [0.54; 0.83], see, for example, the meta-analyses in [72]). The results of the study suggest that the effect might be somewhat smaller in women between 70 and 74 years of age than in the middle-aged segment. However, there is no reason to assume that the effects differ strongly, that there is no effect, or even that there is a reversal of effects.

The two studies, in this case predominantly the Swedish Two-County Study, are the best evidence available on this sub-question. However, when interpreting them it must be taken into account that screening started more than 40 years ago. Since then, the remaining life expectancy of 70-year-old women (from 1986/1988 to 2018/2020) has increased by an average of 3.5 years (from 13.5 to 17.0 years) [73], i.e. by a good 25%. The consistent effect described above is further supported by the plausible assumption that 65- to 69-year-old women in the 1980s are equivalent to today's 70- to 74-year-old women in health status and life expectancy. Thus the effect of screening on mortality among the then 60- to 69-year-old women can essentially be transferred to today's 70- to 74-year-olds.

However, the age of the studies, which is partly due to the need for very long follow-up periods, also leads to interpretation problems, as the health care situation has changed. Influencing factors that could have an impact on the prognosis of the disease and possibly also on the effect of screening are primarily further developments in breast cancer diagnostics and therapy (see e.g. [74,75]). With today's technology, more and different tumours are being found, and better health care through less invasive surgical techniques, modern chemotherapy and radiation techniques, psycho-oncology, and palliative care should contribute to a better prognosis in early and late stages. However, it is unknown how these improvements in health care affect mammography screening outcomes individually and, more importantly, collectively. Incidentally, the changes would affect all age groups, so would by no means only call screening in older age groups into question.

Overall, there are no reasons not to use the two studies as a basis for assessment and recommendation.

Finally, the assumption of a positive screening effect on breast cancer-specific mortality in 70- to 74-year-olds is also supported by our own modelling results. The modelling results on breast cancer-specific mortality appear robust because at the lower age limit, they are consistent with the results from the RCTs. To this extent, this robustness can also be assumed for the upper age limit.

In the overall view of the data and arguments, there is a hint of a benefit of mammography screening in women aged 70 to 74 years for breast cancer-specific mortality as a basis for expansion of the current screening programme.

In contrast to the age group up to 74 years, no data are available for even higher age groups (≥ 75 years). Currently – entirely without corresponding data – it is not assumed that the screening effect is transferable to breast cancer-specific mortality in even higher age groups. For women aged 75 years and older, there is therefore no hint of a benefit or harm of mammography screening in this respect.

Summary for mortality

For all-cause mortality, there was no hint of a benefit or harm from mammography screening, but for breast cancer-specific mortality, there was a hint of a benefit in women 70 to 74 years of age. Since breast cancer accounts for an increasingly smaller proportion of all deaths with increasing age [76], the result on all-cause mortality is no argument against a benefit with regard to mammography screening.

4.5.2.2 Results on mastectomies

No usable data are available on this outcome.

4.5.2.3 Results on consequences of false-positive screening results

For the outcome of consequences of false-positive screening results, data were used for screening participants who had a positive screening result and for whom the suspicion of breast cancer was not confirmed in the subsequent invasive diagnostic procedures. For this outcome, both data on purely diagnostic interventional clarification and on surgical therapeutic interventions were relevant if the diagnostic and therapeutic intervention for breast tissue of unclear histological classification (malignant or benign) could not be clearly separated. By definition, data on this outcome refer only to screened women. A group comparison is thus not applicable, even though diagnostic and therapeutic interventions with benign findings may also occur in the control group.

Usable data on the consequences of false-positive results were available from 1 study with a moderate qualitative certainty of results (AgeX Pilot). The results referred to the age group

from 71 to 73 years. Data on needle biopsies with benign findings and surgeries with benign findings after 1 round of screening could be extracted from the study.

The proportion of screening participants who underwent needle biopsy only because of a false-positive screening result was reported to be 0.66% in this study. The proportion of screening participants who underwent surgery, also only because of a false-positive screening result, was 0.03%. The proportion of surgeries with benign findings was 3% (2/72) of all surgeries.

Thus, based on the results on needle biopsies with benign findings and surgeries with benign finding, for the outcome of consequences of false-positive results there is a hint of harm from mammography screening in women aged 70 years and older.

4.5.2.4 Results on adverse events

No data are available on this outcome.

4.5.2.5 Results on overdiagnosis

Data on overdiagnosis (based on breast cancer incidence) were available from only 1 study with a moderate qualitative certainty of results (Swedish Two-County Study). The age group included women aged 70 to 74 years.

However, the methods used for this assessment to calculate overdiagnosis were not applicable to this study. This is because the calculation is based on a comparison of the breast cancer incidence of screened women with a control group without screening, i.e., in which only symptomatic breast cancer cases were recorded. However, in the Swedish Two-County Study, final screening was performed in the control group. Once a control group is also screened, pre-symptomatic and asymptomatic breast cancer cases will be included in this group, resulting in a higher total number of breast cancer cases in the control group, thus underestimating the true effect. Nevertheless, these data from the Swedish Two-County Study were used in the analysis to give an idea of the effect size.

Overdiagnosis in relation to women invited for screening

If the limited validity of the calculation method is disregarded for the Swedish Two-County Study, the risk of overdiagnosis in relation to all women invited for screening is 0.48% (after screening controls).

Overdiagnosis related to women diagnosed with breast cancer during the screening period

If the limited validity of the calculation method is disregarded for the Swedish Two-County Study, the risk of overdiagnosis in relation to women diagnosed with breast cancer during the screening period is 19.45% (after screening controls).

Summarized assessment of overdiagnosis

Invasive diagnostic procedures confirming the diagnosis of breast cancer are followed by treatment in almost all cases. Both diagnostic procedures and subsequent treatment carry a risk of adverse effects and complications. Since overdiagnosis represents the diagnosis of a disease that would never have become apparent and would not have caused any symptoms without screening, the subsequent interventions along with their consequences are superfluous. Even if in some cases more gentle therapeutic approaches are used today (e.g., more frequent breast-conserving surgery instead of mastectomies), overtreatment remains a detriment.

Even though the certainty of the results of the Swedish Two-County Study with regard to the calculation of overdiagnosis is limited due to the use of data from the control group after the final screening (thus an underestimation of the effect must be assumed), the results show that overdiagnosis occurs in breast cancer screening in this age group.

Thus, the data provide a hint of harm from mammography screening versus no screening with respect to the outcome of overdiagnosis in women aged 70 years and older.

4.5.2.6 Results on health-related quality of life

No data on this outcome were available.

4.6 Summarized assessment of results

Evidence map

The following Table 3 shows the evidence map in relation to patient-relevant outcomes.

Table 3:Evidence map in relation to patient-relevant outcomes

	Mortality		Morbidity				HrQoL
	All-cause mortality	Breast-cancer specific mortality	Mastectomies	Consequences of false-positive screening results	Adverse events	Overdiagnosis	Health-related quality of life
Question 1: 45–49 years	⇔	⇗	⇔	⇓	–	⇘	–
Question 2: ≥ 70 years	⇔	⇗[a]	–	⇘	–	⇘	–

⇗: hint of a benefit
⇔: no hint, indication, or proof; homogeneous result
⇘: hint of harm
⇓: indication of harm
–: no data reported

a. Limited to women aged 70 to 74 years

HrQoL: health-related quality of life

Assessment of the extent of unpublished data

The available information does not indicate publication bias.

Ongoing AgeX study

An ongoing cluster randomized trial in the United Kingdom was identified with 4.4 million women recruited at ages 47 to 49 and 71 to 73 years (AgeX [77]; the previous AgeX pilot study has been completed and was included in this report).

The ongoing AgeX study, which completed recruitment in spring 2020, is expected to provide further meaningful evidence on both questions of the report in the mid-2020s. It is the largest randomized trial to date of the benefit of screening versus no screening and, apart from the UK Age Study, the only known study investigating the extension of age limits of an already established breast cancer screening programme. The study also meets the question of the present report better than the previously studies included with regard to the age groups considered, although the screening interval of 3 years investigated in the study deviates from that currently used in Germany.

In the comments on the preliminary report, different opinions were provided on the extent to which the study can be expected to produce meaningful results for the German context. This question was therefore also discussed during the oral debate. Among other things, it was stated that the results could only be applied to a limited extent due to the different screening interval.

However, the evidence available to date on the possible influence of the screening interval (examination of a possible effect modification, see Section 4.5.1.1; modelling, see Section A4 of the full report; ECIBC analyses comparing screening intervals, see [78]) so far provide no reason to assume markedly different effects on breast cancer-specific mortality based on the screening interval.

As soon as the study results are available, it should therefore be examined whether they confirm the previous recommendations to extend the age limits in mammography screening.

Weighing of benefits and harms

All screening causes harm through false screening results and overdiagnosis. Screening is only justified if the harm is more than outweighed by the benefit. In addition, when weighing benefits and harms, it must be taken into account that the results for the different outcomes must be weighted differently.

Question 1

The following Table 4 shows an overview of all patient-relevant outcomes of Question 1.

Table 4: Question 1 – Overview of the results for all patient-relevant outcomes used for a weighing of benefits and harms (2-page table)

Patient-relevant outcomes	Results	Basic risk[a] per 10 000 women	Risk[b] per 10 000 invited women	Absolute effect per 10 000 invited women [95% CI]	Comments
Mortality					
All-cause mortality after 7.9 to 13 years of follow up	IDR 0.98; 95% CI: [0.90; 1.08]; p = 0.602	273	268	5 [−22; 27]	It is not proven that mammography screening decreases or increases all-cause mortality.
Breast-cancer specific mortality after approx. 10 years	IDR 0.85; 95% CI: [0.73; 1.002]; p = 0.053	35	30	5 [0; 9]	Without mammography screening, 35 out of 10,000 women die of breast cancer. With mammography screening, 30 out of 10,000 women die of breast cancer. Mammography screening prevents about 5 in 10,000 women from dying of breast cancer within about 10 years. Looking at the sensitivity analysis, 35 women die of breast cancer without mammography screening and 29 out of 10,000 women die with mammography screening. Mammography screening thus prevents about 6 out of 10,000 women from dying of breast cancer within about 10 years.
Morbidity					
Mastectomies	OR 1.17[c]; 95% CI: [0.94; 1.45][c]; p = 0.158[c]	62	72	−10 [−28; 4]	It is not proven that mammography screening decreases or increases the risk of mastectomy. The impact of mammography screening on the proportion of mastectomies among all breast surgeries is not proven.
Consequences of false-positive screening results: biopsies with benign finding	See Table 29 of the full report	-	-	7 to 340	7 to 340 out of 10,000 women undergo invasive diagnostic clarification per year or per screening round, with subsequent benign findings.

Table 4: Question 1 – Overview of the results for all patient-relevant outcomes used for a weighing of benefits and harms (2-page table)

Patient-relevant outcomes	Results	Basic risk[a] per 10 000 women	Risk[b] per 10 000 invited women	Absolute effect per 10 000 invited women [95% CI]	Comments
Consequences of false-positive screening results: surgery with benign finding	See Table 30 of the full report	-	-	15[d]	15 of 10,000 women undergo surgery after first screening round, with a subsequent benign finding[d].
Adverse events	No data reported	-	-	-	-
Overdiagnosis	Range [minimum; maximum] of point estimates from individual studies for overdiagnosis risk relative to women invited for screening: 0.01% to 0.41%[e]	-	-	1 to 41	Screening detects breast cancer in 1 to 41 of 10,000 screened women that would not have caused symptoms during their remaining life. These women are subjected to diagnostic and therapeutic procedures that are unnecessary and, in some cases, associated with complications. The overdiagnosis risk calculated from the individual studies in relation to the women diagnosed with breast cancer during the screening period ranges from 2.63% to 31.61%[e]
HrQoL					
HrQoL	No data reported	-	-	-	-

a. Median risk of the control group
b. Median risk of the intervention group
c. Based on the study CNBSS 7 years after start of study
d. Based on the results of the study AgeX Pilot
e. Based on the results of the studies CNBSS and HIP

CI: confidence interval; CNBSS: Canadian National Breast Screening Study; HIP: Health Insurance Plan of Greater New York; HrQoL: health-related quality of life; IDR: incidence density ratio; OR: odds ratio

Mammography screening prevents about 5 out of 10 000 women from dying of breast cancer within about 10 years. However, based on the study results, it cannot be shown that screening also affects all-cause mortality. It is assumed that, due to competing causes of death, the reduction in breast cancer mortality was not large enough to statistically show an effect on all-cause mortality. Overall, the result on all-cause mortality is no argument against a hint of a benefit of mammography screening with respect to breast cancer-specific mortality.

It is not proven that mammography screening reduces or increases the risk of needing a mastectomy. An impact of mammography screening on the proportion of mastectomies among all surgeries is also not proven.

In the case of false-positive screening results, women screened suffer harm from being informed about a suspicious finding, from the subsequent diagnostic clarification, and from any resulting complications. 7 to 340 women out of 10,000 invited to screening undergo invasive diagnostic clarification per year or per screening round, with subsequent benign findings. In addition, 15 of 10,000 women invited to screening undergo surgery after one round of screening with subsequent benign findings.

The risk of overdiagnosis relative to women with a breast cancer diagnosis during the screening period ranged from 2.63% (in the HIP study) to 31.61% (in the CNBSS). Relative to women invited for screening, the overdiagnosis risk ranged from 0.01% (in the HIP study) to 0.41% (in the CNBSS) for the same period. The studies thus found that an estimated 1 to 41 of every 10,000 women in the age group investigated who were invited to mammography screening would have received a diagnosis of breast cancer that, however, would not have caused any symptoms during their remaining life.

No data were available on adverse events of mammography screening. Possible adverse events associated with mammography screening are mainly due to the subsequent treatments, e.g. infection and anaesthesia risks of surgery as well as adverse effects of any chemotherapy, radio- and/or hormonal therapy. Adverse events due to diagnosis and treatment of cancer should be considered as harm from screening if they are due to overdiagnosis; otherwise, (early) diagnosis and subsequent treatment (with the related adverse events) represent a goal of screening and should therefore not be considered as harm from screening. It can therefore be assumed that the effect of screening on the rate of adverse events is mainly represented by the outcome of overdiagnosis. In addition, adverse events can also be expected among women who eventually receive a negative screening result. First, the psychological distress reported by some women, especially anxiety, until notification of the negative screening result [79] should be taken into account here, although this is generally likely to last only a few days. Second, pain from breast compression during mammography can also be considered an adverse event. Although pain was not recorded in the present studies, other studies showed that many women experience the examination as painful, although the pain is usually mild [80]. In this regard, a disadvantage of mammography screening should be considered proven.

No data were available on health-related quality of life. It can be assumed that being informed about an abnormal result affects the health-related quality of life of screening participants. Since this effect is likely to be only short-term in the case of false-positive results, a relevant impairment is to be expected primarily in screening participants with true-positive results. Impairment of health-related quality of life due to diagnosis and treatment of cancer should be considered as harm from screening if it is due to overdiagnosis; otherwise, (early) diagnosis and subsequent treatment (including the associated impairment of health-related quality of life) are a goal of screening and should therefore not be considered as harm from screening. The effect of screening on health-related quality of life is therefore likely to be mainly reflected by the outcome of overdiagnosis.

Whether mammography screening over several years, including follow-up diagnostics, is permissible in this age group under German radiation protection law is to be determined in the near future by the Federal Ministry for the Environment, Nature Conservation, Nuclear Safety and Consumer Protection (BMUV; formerly Federal Ministry for the Environment, Nature Conservation and Nuclear Safety) in the form of a legal ordinance. An older analysis by the Federal Office for Radiation Protection (BfS) noted that mammography screening before the age of 46 should be viewed very critically [81].

Overall, across outcomes the data provide a hint of a benefit of mammography screening in 45- to 49-year-old women. This is because the harm identified is not so large that it outweighs the mortality advantage. However, given the comparatively small mortality advantage in this age group, an individual assessment and weighing of benefits and harms is still essential.

Question 2: age group 70 years and older

A tabular comparison of benefits and harms for Question 2 was omitted because the data on benefits and harms were not sufficient to present absolute effects.

The available data from 3 studies with about 25,000 participants analysed refer to women up to a maximum age of 74 years at the time of randomization. Data on mortality were available for about 18,000 women, too few for a sufficiently precise estimate of the effect. The conclusion on the evidence available is therefore additionally based on data from the middle-aged groups, on considerations concerning the applicability of the results, and on the modelling results (see Chapter 5 and Section A4 of the full report), so that overall, it can be assumed that screening also reduces breast cancer-specific mortality in the age group 70-74 years. However, these data do not allow an estimation of absolute effects.

Furthermore, it cannot be shown on the basis of the study results that screening also has an effect on all-cause mortality. It is assumed that – to an even greater extent than in women between 45 and 49 years – due to other causes of death, the reduction in breast cancer mortality was not large enough to be able to show statistically an effect on all-cause mortality. Overall, the result on all-cause mortality is no argument against a hint of a benefit of mammography screening with regard to breast cancer-specific mortality.

With regard to women aged 70 years and older, special attention must be paid to the risk of overdiagnosis, which increases with age and is caused by competing risks [3]. This is especially true for women with a lower remaining life expectancy due to serious comorbidities. Although data on overdiagnosis are available from 1 RCT and already provide a hint of harm, further data would be needed to reliably assess the extent of overdiagnosis in this age group. It is possible that the AgeX study will help to close this data gap.

Data on the consequences of false-positive results are available from 1 study on the age group of women between 71 and 73 years. These data also suggest a hint of harm in the age group 70 years and older.

No data were available on mastectomies, adverse events and health-related quality of life. As was already the case in the age group of 45-49 year old women, it can be assumed that the effect of screening on adverse events and health-related quality of life in the age group 70 years and older is largely represented by the outcome of overdiagnosis.

The BfS is currently assessing whether under the German radiation protection law, mammography screening over several years, including follow-up diagnostics, is also permissible for this age group.

Overall, the data provide a hint of a benefit of mammography screening in 70- to 74-year-old women across all outcomes. This is because it is assumed that the expected mortality advantage outweighs the expected harm. However, in view of the very small mortality advantage expected in this age group, the harm observed from overdiagnosis and consequences of false-positive results, as well as the uncertainties regarding the quantification of the effects (especially breast cancer-specific mortality and overdiagnosis), an individual assessment and weighing of benefit and harms, considering individual health status and remaining life expectancy, is still essential.

Assessment of potential

No data could be identified for women aged 75 years and older in this assessment. It is also at least questionable whether the effects of mammography screening on breast cancer-specific mortality in 50- to 69-year-old women can also be applied to this even older population. The review of the systematic reviews revealed only 1 study that specifically dealt with screening in the age group of 75 years and older. This comparative observational study with screened compared to unscreened women aged 80 years and older [38] tends to support the assumption that, at least from this age, the harm from overdiagnosis and consequences of false-positive results outweighs the benefit. Therefore, on the basis of the available evidence, no potential of mammography screening for the age group 75 years and older could be identified.

5 Classification of the assessment result

Informative value of results

When interpreting the studies, some aspects have to be taken into account that influence the informative value of the results. The main ones are explained below.

Study designs

This report assesses the benefits and harms of extending the age limits of mammography screening. The core question is therefore whether screening that starts earlier or is carried out for longer has an additional benefit compared to the established screening programme (screening every 2 years between 50 and 69 years). In contrast, the studies available so far have mainly aimed to show the benefit of screening over no screening. Such studies were considered to be a sufficient approximation to the actual research question of the report. Nevertheless, the questions investigated in the studies included vary in their similarity to the questions investigated in the present report. Studies where final screening was offered to the control group after the screening period (Gothenburg, Stockholm, Swedish Two-County) or in which a general screening programme started after the screening period (Gothenburg in women ≥50, Edinburgh, Malmö, UK Age; AgeX Pilot plays no role here due to missing mortality data) better represent the question of earlier vs. later screening (Question 1) than those studies in which no final screening and no transition to a general screening programme took place (CNBSS; data missing for the HIP study, but it can be assumed that initially there was no follow-up screening for all women). The influence of the different study designs on the interpretation and applicability of the results differs for the outcomes considered. This is explained below.

Breast cancer-specific mortality

Breast cancer-specific mortality was analysed differently in the studies. Either only those breast cancer cases detected within the screening period were included in the analysis (evaluation model, see Section A1.2 of the full report) or all breast cancer cases were considered, including those only occurring after the end of the screening period (follow-up model, see Section A1.2 of the full report). With regard to Question 1, studies in which general screening of all women begins after the screening period and in which breast cancer-specific mortality is analysed using the follow-up model best reflect the question asked in this report. With this type of analysis, however, a dilution effect is to be expected, as the breast cancer-specific mortality rates in both groups converge through the later screening. Studies with such a design are the UK Age Study and the Gothenburg Study.

Both studies contain additional analyses according to the evaluation model. If one compares the results of the different analysis types within these two studies, however, no marked differences can be observed. This suggests that the type of analysis used in each case did not have a major influence on the results.

Overdiagnosis

Overdiagnosis could only be meaningfully calculated in 2 studies (CNBSS, HIP) according to the methods used for this assessment. However, the results on overdiagnosis from these two studies should be interpreted with caution for other reasons (see Section 4.5.1.5). In addition, the relevant literature also notes that aggregated data often overestimate overdiagnosis [39]. Even if the data available make it difficult to draw reliable conclusions on the extent of overdiagnosis, there is no question that screening inevitably leads to overdiagnosis [3].

According to Duffy et al [40], overdiagnoses result from 2 different disease entities. One is potentially progressing tumours, which, however, occur in women who die for other reasons before the tumour becomes symptomatic. The other consists of non-progressive or very slowly progressing tumours. A substantial proportion of DCIS cases probably belong to these tumour types. The two entities described play a different role in the two age groups considered. For example, it is to be expected that only very few women under 50 die of other causes before the tumour becomes symptomatic. This form of overdiagnosis therefore contributes very little to the overall cases of overdiagnosis in this age group, so that a large part of the cases of overdiagnoses can be expected to consist of non-progressing or slowly progressing tumours. With regard to Question 1, it can therefore be expected that bringing forward the start of screening by 5 years will not result in many additional cases of overdiagnosis, because in most cases only the time of diagnosis (and thus also the time of overdiagnosis) is brought forward. The modelling results support this expectation: with the established screening strategy (screening every 2 years between 50 and 69 years), a lifetime risk of overdiagnosis of 6.3% related to all participating women is calculated, which increases only moderately to 7.2% when screening is extended to the 45- to 49-year-old age group (see Table 11 in Section A4 of the full report).

In contrast, for older women (Question 2) it is estimated that the risk of overdiagnosis increases markedly with screening [41], mainly due to competing causes of death. However, this is not reflected in the modelling results: a lifetime risk of overdiagnosis of 6.3% to 6.6% is calculated for an extension of screening beyond 69 years, depending on the screening interval and the end of screening (compared to 6.3% for the reference strategy), which is a marginal difference at best compared to the established screening programme. A possible explanation for the fact that in modelling, the risk of overdiagnosis is not greater in older women than in younger ones could be a different frequency of DCIS cases in the age groups. However, the authors of the modelling report also point out that, due to the underlying empirical data, there is great uncertainty with regard to the outcome of overdiagnosis.

Age subgroups investigated

For Question 1, only few data are available that refer exactly to the age range of 45- to 49-year-olds under investigation, so that data from age subgroups that included women under 45 years of age (e.g. age subgroup 39-49 years) were also used for the assessment. Under the premise that the effect of mammography screening on mortality decreases with younger age, the

inclusion of "too young" women in the age subgroup considered tends to result in a conservative estimate of the screening effect. However, during the screening period, some of the women who are in the age range searched for (45-49 years) at the beginning of the screening period grow into the age at which the screening programme begins in Germany. This means that the effect in women who are between 40 and 49 years old at the start of the study is partly due to women whose breast cancer was not detected until after the age of 50 and would thus have also been found with the established screening programme. In terms of mortality, this tends to lead to an overestimation of the effect of screening in the 45-49 age group. It is not possible to answer which of the two opposing effects has a stronger impact here or whether they cancel each other out.

For Question 2, the potential problem of larger and overlapping age ranges did not exist, as the available data referred to the age group 70-74 years, which is completely within the range of the target population of this question.

Study age and applicability

Only the AgeX Pilot study, with a recruitment period between 2009 and 2010, can still be considered a current study, but it provides only limited information due to its nature as a pilot study. The study start date for most of the other studies included was in the 1970s and 1980s. Exceptions are the HIP study, which started in 1963 and thus even 13 years earlier than the second oldest study included – at that time, adjuvant chemotherapy for breast cancer had not even become established – and the UK Age study, which started in 1991. Except for the AgeX Pilot study, which does not include results on mortality and overdiagnosis, all studies included were thus conducted in the context of a different health care situation and under different social circumstances. It makes sense that the applicability of the study results to today's situation decreases with the increasing age of the study. As already explained in Section 4.5.2.1, today's higher life expectancy contributes to the fact that the effect of screening on breast cancer-specific mortality is likely to be more evident today, at least among 70- to 74-year-olds. However, it is not known how the further developments in breast cancer diagnostics and therapy individually and in their entirety affect the effects of screening. Despite these limitations, the studies included here represent the best available evidence, and overall there are no reasons not to use them as a basis for assessment and recommendation. Nevertheless, other data sources should be considered, especially the modelling performed, as this source better reflects the current life expectancy and health care situation.

Classification of the modelling results

The modelling carried out as part of this assessment (see Section A4 of the full report) brings together data from various sources and makes predictions beyond the period covered by the studies available to date. Modelling alone is insufficient to answer the question of benefits and harms. However, modelling is based on RCTs and reliable epidemiological data and thus represents a solid basis for assessment and is therefore suitable to greatly support the conclusions based on the RCTs analysed. The strength of modelling also lies in the fact that it

can support the weighing of benefits and harms by showing how benefits and harms are related to each other and in which direction and to what extent benefits and harms shift when certain screening parameters are changed.

A key finding of modelling is that survival is potentially prolonged the longer and more often mammography screening is performed, but that at the same time, beyond a certain screening intensity, the incremental harm-benefit ratio deteriorates again. (The incremental harm-benefit ratio expresses what additional harm must be accepted in order to achieve an additional unit of benefit).

According to the modelling results, the strategy with the largest incremental QALY gain versus the reference strategy is screening from 45 to 74 years every 2 years (3.5 per 100 women screened; Table 10 in Section A4 of the full report), which supports the conclusions based on the studies included. The modelling results also fit the meta-analysis conducted: raising the upper age limit (screening up to age 74 every 2 years) versus decreasing the lower age limit (screening from age 45 every 2 years) in the base case analysis yields less than half the benefit in terms of lifetime gained versus the reference strategy (screening from age 50 to 69 every 2 years), (2.4 versus 6.0 years per 100 women, Table 10 in Section A4 of the full report). In the modelling of QALYs, the better benefit-harm ratio becomes even more apparent at the lower versus the upper age limit, because here the incremental QALY gain differs by about a factor of about 5 (3.1 versus 0.6 per 100 women; Table 10 in Section A4). However, there are also modelled outcomes for which more additional harm is to be expected in the 5-year interval at the lower age limit than at the upper age limit, e.g. with regard to the rate of false-positive screening results (see Table 11 in Section A4 of the full report).

In addition to empirical data, modelling is based on several (more or less reliable) assumptions and thus represents a partially simplified reflection of reality. Different results between meta-analysis and modelling therefore do not necessarily indicate a contradiction. The fact that modelling also calculates a survival advantage when screening is extended beyond the age of 70 (Question 2), while the available study results on women between 70 and 74 showed no statistically significant difference, can most likely be explained by the fact that the actual absolute survival advantage was not large enough to achieve sufficient statistical precision with the studies available so far.

However, it should be noted that the consequences of the different screening strategies in the modelling study are essentially based on test accuracy. On the basis of this, the stage shift resulting from the changes in the screening programme parameters is modelled with the inclusion of numerous other input parameters. In combination with stage-specific survival, this manifests itself in a survival advantage of the more intensive strategies. On the basis of modelling alone, it is therefore not possible to say with absolute certainty how well the calculated effect reflects reality in women aged 70 and over. Overall, however, modelling supports the conclusion, based on the studies included, that the benefit in terms of breast cancer-

specific mortality outweighs the harm from overdiagnosis and consequences of false-positive screening results.

Findings from observational studies

For Question 1, the studies included already provided a sufficient basis for assessment, so that the additional consideration of observational studies does not play a relevant role here. The situation is different for Question 2. In view of the sparse evidence based on RCTs, the question arises here whether sufficient evidence for conclusions on the benefits and harms of mammography screening can be obtained with the additional inclusion of observational studies. However, based on the review of systematic reviews in the context of information retrieval (see Section A7.1 of the full report), no observational studies could be identified that refer to screening in 70- to 79-year-old women; only one observational study comparing screened versus unscreened women aged 80 years and older was found ([82], see Section 4.6).

Several observational studies were submitted during the hearing on the preliminary report. These data essentially confirm the results of the RCTs for the age groups of 50-69 years currently included in regular screening (see Section 4.5.2.1). However, they do not provide any robust information on screening in 70- to 74-year-olds. A detailed discussion of the submitted studies is provided in Section A5.3.1 of the full report.

Considerations on the screening interval

If mammography screening were extended to 45- to 49-year-old or 70- to 74-year-old women, it would be necessary to determine whether the screening interval of 2 years established for 50- to 69-year-olds would be retained or adjusted. In most screening programmes in countries of the European Union, mammography screening is offered every 2 years, also in relation to the age groups 45-49 and 70-74 years, if included (as of 2016) [86]. However, there is still a need for research on the question of which screening interval is optimal, especially for younger women [3]. This question cannot be answered satisfactorily on the basis of the studies included. The screening intervals in the studies were 1 to 2 years. In Question 1, no influence of the screening interval on the outcomes examined could be shown when considering studies with a screening interval of ≤ 1.5 years or > 1.5 years. In Question 2, a corresponding analysis was not possible due to the lack of evidence available. However, modelling could help in this regard.

Question 1

In the modelling study (see Section A4 of the full report), the impact of different screening intervals (in each case keeping the 2-year screening interval in 50- to 69-year-olds) on certain outcomes was investigated in comparison to the established screening strategy – i.e. to screening every 2 years between 50 and 69 years. The results suggest that for women who start screening at age 45, the average gain in lifetime per woman could be greater with screening once per year between 45 and 49 years of age than with screening every 2 years. However, the absolute difference between the two strategies is small (38.77 years versus 38.76 years of

remaining life expectancy per woman; 38.70 years of remaining life expectancy is calculated with the established screening strategy; see Table 10 in Section A4).

According to the modelling study, the lifetime risk of overdiagnosis changes only moderately from 6.3% per participating woman to 7.2% (for both screening intervals studied, i.e. for 1 and 2 years) due to the start of screening 5 years earlier. However, false-positive screening results become more frequent, from 86.7 events per 100 participating women in the reference strategy to 147.4 (for a 1-year interval) and 123.1 events (for a 2-year interval) per 100 women (see Table 11 in Section A4 of the full report). Given the rather small difference in remaining life expectancy between 1- and 2-year intervals, the change in false-positive results is the most important and supports preference of a 2-year interval. The incremental harm-benefit ratios also indicate a more favourable outcome with a 2-year interval.

Question 2

Modelling also considered the impact of different screening intervals (2 years or 3 years) beyond 69 years of age on the outcomes studied. The absolute difference between the two strategies is negligible (38.71 years versus 38.73 years remaining life expectancy per woman after age 45; with the established screening strategy, 38.70 years remaining life expectancy is calculated; see Table 10 in Section A4). In view of the comparatively short time span between 70 and 74 years, a substantial influence of the screening interval would not be expected, because with a short interval the women receive only 1 screening examination more than with the longer interval. However, as with the extension of the lower age limit, false-positive screening results become more frequent, namely from 86.7 events per 100 participating women in the reference strategy to 104.8 (if the interval of 2 years is continued) or 92.9 events per 100 women (if the interval is extended to 3 years in over 70-year-olds, see Table 11 in Section A4 of the full report). Given the small difference in remaining life expectancy between the 2-year and 3-year intervals, the change in false positives is more important and may support the extension of the interval from 70 years of age. However, organizational reasons could be cited against this. In addition, a change to a longer interval might not find sufficient acceptance [78].

Perspective of women affected

In discussions during the preparation of the present report with women affected, these women reported potential reasons for not participating in the screening. These included pain during the examination and an atmosphere that was perceived as very impersonal during the entire appointment. According to the women, the latter was mainly because there was almost no communication with them from the staff carrying out the mammography and the women felt left alone. The women also found the transmission of findings by post stressful. Such or similar problems are known from the literature [79,87].

The points of criticism raised by the women interviewed could indeed be taken into account in the context of a screening. This could also contribute to increasing the participation rate of currently about 50% in Germany [70]. Because the individual weighing of benefits and harms is crucial in mammography screening (especially at the lower and upper age limits),

communication and counselling should be optimally organized. The decision aid distributed to all eligible women in Germany can only represent supportive information here [88].

Public health perspective

The World Health Organisation (WHO) only makes a (conditional) recommendation for mammography screening for women aged 40 to 49 years or 70 to 75 years if resources are sufficient [89]. In contrast, if resources are limited, it is against screening in these age groups [89]. With regard to the question as to whether sufficient resources are available in the German health care context for an extension of the age limits, medical and para-medical personnel resources should be considered as possible limiting factors in addition to monetary resources.

Need for research on other screening methods

In a comment on the preliminary report, sonography and magnetic resonance imaging are addressed as possible alternative or supplementary screening methods, especially for younger women or women with higher breast density. As an advantage over mammography, the lack of radiation exposure and, with regard to sonography, the better suitability for high breast density are emphasized. In guidelines, however, the use of these screening methods is not generally recommended at present, with reference to the evidence base [3,90] and a need for research is seen in this regard [3]. In addition, tomosynthesis, another mammography-based procedure that could be suitable as a screening procedure, is currently being investigated (e.g. [91]).

6 Conclusion

Question 1: Age group 45 to 49 years.

For breast cancer-specific mortality, the data provide a hint of a benefit of mammography screening versus no screening in women aged 45 to 49 years. With regard to all-cause mortality, there was no statistically significant result. The result on all-cause mortality is no argument against a benefit of mammography screening, because all-cause mortality is considerably influenced by other causes of death. It is assumed that the overall effect of reduced breast cancer mortality through screening was too small to have an impact on all-cause mortality.

However, mammography screening leads to negative consequences (indication of harm) through false-positive screening results. In addition, overdiagnosis occurs (hint of harm). However, the extent of harm is not so great that it outweighs the mortality advantage. With regard to mastectomies, the data provide no hint of a benefit or harm from mammography screening. No data were available for adverse events and health-related quality of life, so no hint of a benefit or harm was shown for these outcomes. However, the effect of screening on the rate of adverse events and on health-related quality of life is likely to be essentially captured by the outcome of overdiagnosis.

In summary, the data provide a hint of a benefit of mammography screening for women aged 45 to 49 years versus no screening; the benefit of mammography screening thus outweighs harm. However, an individual assessment and weighing of benefits and harms is still essential

in view of the very small mortality advantage in this age group. Therefore, all conditions must be fulfilled to enable women to make an informed decision.

Question 2: Age group 70 years and older

For breast cancer-specific mortality, the data provide a hint of a benefit of mammography screening versus no screening in women aged 70 to 74 years. This is based on the results for this age group, the transferability of effects from adjacent younger age groups, and the confirmatory results of the modelling performed. With regard to all-cause mortality, there was no statistically significant result. This is no argument against a benefit of mammography screening, because with increasing age, breast cancer accounts for a smaller and smaller proportion of all deaths, and all-cause mortality is greatly influenced by competing causes of death. It is therefore assumed that the overall effect of reduced breast cancer mortality through screening was too small to have an impact on all-cause mortality.

Mammography screening leads to overdiagnosis and to negative consequences of false-positive screening results (in each case a hint of harm in women aged 70 and older). It is assumed that the expected mortality advantage outweighs the expected harm. No data were available for mastectomies, adverse events and health-related quality of life, so no hint of a benefit or harm was shown for these outcomes.

In summary, the data provide a hint of a benefit of mammography screening versus no screening in women aged 70 to 74 years; the benefit of mammography screening thus outweighs harm. However, an individual assessment and weighing of benefits and harms, considering the individual health status and life expectancy, is still essential in view of the very small mortality advantage, harm observed from overdiagnosis and the consequences of false-positive results, as well as the uncertainties regarding the quantification of effects. Therefore, all conditions must be fulfilled to enable women to make an informed decision.

An ongoing randomized controlled trial (AgeX) is expected to provide meaningful data for both questions of this report in the near future. As soon as the study results are available, it should therefore be examined whether the results are in line with the recommendations to extend the age limits.

Mammography screening for women aged 75 and older currently offers neither a benefit nor a potential versus no screening due to a lack of informative data. The possible benefit with regard to breast cancer-specific mortality, which is, however, quite questionable in this age group, is likely to be countered by harm expected from overdiagnosis and consequences of false-positive screening results.

References for English extract

Please see full final report for full reference list.

1. Sung H, Ferlay J, Siegel RL 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): 209-249. https://dx.doi.org/10.3322/caac.21660.

2. Robert Koch Institut, Gesellschaft der epidemiologischen Krebsregister in Deutschland. Krebs in Deutschland für 2015/2016 [online]. 2020 [Accessed: 12.04.2021]. URL: https://www.krebsdaten.de/Krebs/DE/Content/Publikationen/Krebs_in_Deutschland/kid_2019/krebs_in_deutschland_2019.pdf?__blob=publicationFile.

3. Leitlinienprogrammem Onkologie (Deutsche Krebsgesellschaft DK, AWMF). Interdisziplinäre S3-Leitlinie für die Früherkennung, Diagnostik, Therapie und Nachsorge des Mammakarzinoms [online]. 2020 [Accessed: 01.04.2021]. URL: https://www.leitlinienprogrammem-onkologie.de/fileadmin/user_upload/Downloads/Leitlinien/Mammakarzinom_4_0/Version_4.3/LL_Mammakarzinom_Langversion_4.3.pdf.

4. Loibl S, Poortmans P, Morrow M et al. Breast cancer. Lancet 2021; 397(10286): 1750-1769. https://dx.doi.org/10.1016/S0140-6736(20)32381-3.

5. Gemeinsamer Bundesausschuss. Mammographie-Screening; eine Entscheidungshilfe [online]. 2017 [Accessed: 31.03.2021]. URL: https://www.g-ba.de/downloads/17-98-2232/2015-11-13_Merkblatt-Mammographie_bf.pdf?

6. Klarenbach S, Sims-Jones N, Lewin G et al. Recommendations on screening for breast cancer in women aged 40-74 years who are not at increased risk for breast cancer. CMAJ 2018; 190(49): E1441-E1451. https://dx.doi.org/10.1503/cmaj.180463.

7. Nelson HD, Pappas M, Cantor A et al. Harms of Breast Cancer Screening: Systematic Review to Update the 2009 U.S. Preventive Services Task Force Recommendation. Ann Intern Med 2016; 164(4): 256-267. https://dx.doi.org/10.7326/m15-0970.

8. Badve SS, Gökmen-Polar Y. Ductal carcinoma in situ of breast: update 2019. Pathology 2019; 51(6): 563-569. https://dx.doi.org/10.1016/j.pathol.2019.07.005.

9. Perry N, Broeders M, de Wolf C et al. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition--summary document. Ann Oncol 2008; 19(4): 614-622. https://dx.doi.org/10.1093/annonc/mdm481.

10. Gemeinsamer Bundesausschuss. Richtlinie des Gemeinsamen Bundesausschusses über die Früherkennung von Krebserkrankungen (Krebsfrüherkennungs-Richtlinie/KFE-RL) [online]. 2020 [Accessed: 01.04.2021]. URL: https://www.g-ba.de/downloads/62-492-2238/KFE-RL_2020-06-18_iK-2020-08-28.pdf.

11. Schünemann HJ, Lerda D, Quinn C et al. Breast Cancer Screening and Diagnosis: A Synopsis of the European Breast Guidelines. Ann Intern Med 2020; 172(1): 46-56. https://dx.doi.org/10.7326/m19-2125.

12. Moser K, Sellars S, Wheaton M et al. Extending the age range for breast screening in England: pilot study to assess the feasibility and acceptability of randomization. J Med Screen 2011; 18(2): 96-102. https://dx.doi.org/10.1258/jms.2011.011065.

13. University of Oxford. Pilot Study: Age Extension of NHS Breast Screening Programmeme [online]. 2010 [Accessed: 24.08.2021]. URL: https://ClinicalTrials.gov/show/NCT00890864.

14. Baines CJ, To T, Miller AB. Revised estimates of overdiagnosis from the Canadian National Breast Screening Study. Prev Med 2016; 90: 66-71. https://dx.doi.org/10.1016/j.ypmed.2016.06.033.

15. Miller AB. The costs and benefits of breast cancer screening. Am J Prev Med 1993; 9(3): 175-180.

16. Miller AB, Baines CJ, To T et al. Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years. CMAJ 1992; 147(10): 1459-1476.

17. Miller AB, Howe GR, Wall C. The National Study of Breast Cancer Screening Protocol for a Canadian Randomized Controlled trial of screening for breast cancer in women. Clin Invest Med 1981; 4(3-4): 227-258.

18. Miller AB, To T, Baines CJ et al. The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years. Ann Intern Med 2002; 137(5 Part 1): 305-312. https://dx.doi.org/10.7326/0003-4819-137-5_part_1-200209030-00005.

19. Miller AB, Wall C, Baines CJ et al. Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial. BMJ 2014; 348: g366. https://dx.doi.org/10.1136/bmj.g366.

20. Narod SA, Sun P, Wall C et al. Impact of screening mammography on mortality from breast cancer before age 60 in women 40 to 49 years of age. Curr Oncol 2014; 21(5): 217-221. https://dx.doi.org/10.3747/co.21.2067.

21. Shaevitch D, Taghipour S, Miller AB et al. Tumour size distribution of invasive breast cancers and the sensitivity of screening methods in the Canadian National Breast Screening Study. Journal of Cancer Research &Therapeutics 2017; 13(3): 562-569. https://dx.doi.org/10.4103/0973-1482.174539.

22. Taghipour S, Caudrelier LN, Miller AB et al. Using Simulation to Model and Validate Invasive Breast Cancer Progression in Women in the Study and Control Groups of the Canadian National Breast Screening Studies I and II. Med Decis Making 2017; 37(2): 212-223. https://dx.doi.org/10.1177/0272989x16660711.

23. Alexander FE, Anderson TJ, Brown HK et al. 14 years of follow-up from the Edinburgh randomised trial of breast-cancer screening. Lancet 1999; 353(9168): 1903-1908. https://dx.doi.org/10.1016/s0140-6736(98)07413-3.

24. Roberts MM, Alexander FE, Anderson TJ et al. The Edinburgh randomised trial of screening for breast cancer: description of method. Br J Cancer 1984; 50(1): 1-6. https://dx.doi.org/10.1038/bjc.1984.132.

25. Roberts MM, Alexander FE, Anderson TJ et al. Edinburgh trial of screening for breast cancer: mortality at seven years. Lancet 1990; 335(8684): 241-246. https://dx.doi.org/10.1016/0140-6736(90)90066-e.

26. Bjurstam N, Björneld L, Duffy SW et al. The Gothenburg breast screening trial: first results on mortality, incidence, and mode of detection for women ages 39-49 years at randomization. Cancer 1997; 80(11): 2091-2099.

27. Bjurstam N, Björneld L, Warwick J et al. The Gothenburg Breast Screening Trial. Cancer 2003; 97(10): 2387-2396. https://dx.doi.org/10.1002/cncr.11361.

28. Bjurstam NG, Bjorneld LM, Duffy SW. Updated results of the Gothenburg Trial of Mammographic Screening. Cancer 2016; 122(12): 1832-1835. https://dx.doi.org/10.1002/cncr.29975.

29. Aron JL, Prorok PC. An analysis of the mortality effect in a breast cancer screening study. Int J Epidemiol 1986; 15(1): 36-43. https://dx.doi.org/10.1093/ije/15.1.36.

30. Chu KC, Smart CR, Tarone RE. Analysis of breast cancer mortality and stage distribution by age for the Health Insurance Plan clinical trial. J Natl Cancer Inst 1988; 80(14): 1125-1132. https://dx.doi.org/10.1093/jnci/80.14.1125.

31. Habbema JD, van Oortmarssen GJ, van Putten DJ et al. Age-specific reduction in breast cancer mortality by screening: an analysis of the results of the Health Insurance Plan of Greater New York study. J Natl Cancer Inst 1986; 77(2): 317-320.

32. Shapiro S. Evidence on screening for breast cancer from a randomized trial. Cancer 1977; 39(6 Suppl): 2772-2782. https://dx.doi.org/10.1002/1097-0142(197706)39:6<2772::aid-cncr2820390665>3.0.co;2-k.

33. Shapiro S. Periodic screening for breast cancer: the HIP Randomized Controlled Trial. Health Insurance Plan. J Natl Cancer Inst Monogr 1997; (22): 27-30. https://dx.doi.org/10.1093/jncimono/1997.22.27.

34. Shapiro S, Strax P, Venet L et al. Proceedings: Changes in 5-year breast cancer mortality in a breast cancer screening programme. Proc Natl Cancer Conf 1972; 7: 663-678.

35. Shapiro S, Venet W, Strax P et al. Ten- to fourteen-year effect of screening on breast cancer mortality. J Natl Cancer Inst 1982; 69(2): 349-355.

36. Andersson I. Radiographic screening for breast carcinoma. I. Programme and primary findings in 45--69 year old women. Acta Radiol Diagn (Stockh) 1981; 22(2): 185-194. https://dx.doi.org/10.1177/028418518102200213.

37. Andersson I, Aspegren K, Janzon L et al. Mammographic screening and mortality from breast cancer: the Malmö mammographic screening trial. BMJ 1988; 297(6654): 943-948. https://dx.doi.org/10.1136/bmj.297.6654.943.

38. Andersson I, Janzon L. Reduced breast cancer mortality in women under age 50: updated results from the Malmö Mammographic Screening Programme. J Natl Cancer Inst Monogr 1997; (22): 63-67. https://dx.doi.org/10.1093/jncimono/1997.22.63.

39. Nyström L, Andersson I, Bjurstam N et al. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. Lancet 2002; 359(9310): 909-919. https://dx.doi.org/10.1016/s0140-6736(02)08020-0.

40. Zackrisson S, Andersson I, Janzon L et al. Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. BMJ 2006; 332(7543): 689-692. https://dx.doi.org/10.1136/bmj.38764.572569.7C.

41. Frisell J, Eklund G, Hellstrom L et al. Randomized study of mammography screening--preliminary report on mortality in the Stockholm trial. Breast Cancer Res Treat 1991; 18(1): 49-56. https://dx.doi.org/10.1007/BF01975443.

42. Frisell J, Glas U, Hellström L et al. Randomized mammographic screening for breast cancer in Stockholm. Design, first round results and comparisons. Breast Cancer Res Treat 1986; 8(1): 45-54. https://dx.doi.org/10.1007/bf01805924.

43. Frisell J, Lidbrink E. The Stockholm Mammographic Screening Trial: Risks and benefits in age group 40-49 years. J Natl Cancer Inst Monogr 1997; (22): 49-51. https://dx.doi.org/10.1093/jncimono/1997.22.49.

44. Frisell J, Lidbrink E, Hellström L et al. Followup after 11 years--update of mortality results in the Stockholm mammographic screening trial. Breast Cancer Res Treat 1997; 45(3): 263-270. https://dx.doi.org/10.1023/a:1005872617944.

45. Duffy SW, Tabar L, Fagerberg G et al. Breast screening, prognostic factors and survival--results from the Swedish two county study. Br J Cancer 1991; 64(6): 1133-1138. https://dx.doi.org/10.1038/bjc.1991.477.

46. Duffy SW, Tabar L, Vitak B et al. The Swedish Two-County Trial of mammographic screening: cluster randomisation and outcome evaluation. Ann Oncol 2003; 14(8): 1196-1198. https://dx.doi.org/10.1093/annonc/mdg322.

47. Tabar L, Chen TH, Yen AM et al. Effect of Mammography Screening on Mortality by Histological Grade. Cancer Epidemiology, Biomarkers &Prevention 2018; 27(2): 154-157. https://dx.doi.org/10.1158/1055-9965.Epi-17-0487.

48. Tabar L, Duffy SW, Yen MF et al. All-cause mortality among breast cancer patients in a screening trial: support for breast cancer mortality as an outcome. J Med Screen 2002; 9(4): 159-162. https://dx.doi.org/10.1136/jms.9.4.159.

49. Tabar L, Fagerberg CJ, Gad A et al. Reduction in mortality from breast cancer after mass screening with mammography. Randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare. Lancet 1985; 1(8433): 829-832. https://dx.doi.org/10.1016/s0140-6736(85)92204-4.

50. Tabar L, Fagerberg G, Chen HH et al. Efficacy of breast cancer screening by age. New results from the Swedish Two-County Trial. Cancer 1995; 75(10): 2507-2517. https://dx.doi.org/10.1002/1097-0142(19950515)75:10<2507::aid-cncr2820751017>3.0.co;2-h.

51. Tabar L, Fagerberg G, Duffy SW et al. The Swedish two county trial of mammographic screening for breast cancer: recent results and calculation of benefit. J Epidemiol Community Health 1989; 43(2): 107-114. https://dx.doi.org/10.1136/jech.43.2.107.

52. Tabar L, Fagerberg G, Duffy SW et al. Update of the Swedish two-county programme of mammographic screening for breast cancer. Radiol Clin North Am 1992; 30(1): 187-210.

53. Tabár L, Vitak B, Chen HH et al. The Swedish Two-County Trial twenty years later. Updated mortality results and new insights from long-term follow-up. Radiol Clin North Am 2000; 38(4): 625-651. https://dx.doi.org/10.1016/s0033-8389(05)70191-3.

54. Tabár L, Vitak B, Chen TH et al. Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. Radiology 2011; 260(3): 658-663. https://dx.doi.org/10.1148/radiol.11110469.

55. Yen AM, Duffy SW, Chen TH et al. Long-term incidence of breast cancer by trial arm in one county of the Swedish Two-County Trial of mammographic screening. Cancer 2012; 118(23): 5728-5732. https://dx.doi.org/10.1002/cncr.27580.

56. Dalarna County Council Sweden. The Swedish Two-County Trial of Mammography Screening (WE) [online]. 2018 [Accessed: 24.08.2021]. URL: https://ClinicalTrials.gov/show/NCT03217539.

57. Duffy S, Vulkan D, Cuckle H et al. Annual mammographic screening to reduce breast cancer mortality in women from age 40 years: long-term follow-up of the UK Age RCT. Health Technol Assess 2020; 24(55): 1-24. https://dx.doi.org/10.3310/hta24550.

58. Duffy SW, Vulkan D, Cuckle H et al. Effect of mammographic screening from age 40 years on breast cancer mortality (UK Age trial): final results of a randomised, controlled trial. Lancet Oncol 2020; 21(9): 1165-1172. https://dx.doi.org/10.1016/s1470-2045(20)30398-3.

59. Johns LE, Moss SM, Age Trial Management G. False-positive results in the randomized controlled trial of mammographic screening from age 40 ("Age" trial). Cancer Epidemiol Biomarkers Prev 2010; 19(11): 2758-2764. https://dx.doi.org/10.1158/1055-9965.EPI-10-0623.

60. Moss S. A trial to study the effect on breast cancer mortality of annual mammographic screening in women starting at age 40. Trial Steering Group. J Med Screen 1999; 6(3): 144-148. https://dx.doi.org/10.1136/jms.6.3.144.

61. Moss S, Thomas I, Evans A et al. Randomised controlled trial of mammographic screening in women from age 40: results of screening in the first 10 years. Br J Cancer 2005; 92(5): 949-954. https://dx.doi.org/10.1038/sj.bjc.6602396.

62. Moss S, Waller M, Anderson TJ et al. Randomised controlled trial of mammographic screening in women from age 40: predicted mortality based on surrogate outcome measures. Br J Cancer 2005; 92(5): 955-960. https://dx.doi.org/10.1038/sj.bjc.6602395.

63. Moss SM, Cuckle H, Evans A et al. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years'follow-up: a randomised controlled trial. Lancet 2006; 368(9552): 2053-2060. https://dx.doi.org/10.1016/s0140-6736(06)69834-6.

64. Moss SM, Wale C, Smith R et al. Effect of mammographic screening from age 40 years on breast cancer mortality in the UK Age trial at 17 years'follow-up: a randomised controlled trial. Lancet Oncol 2015; 16(9): 1123-1132. https://dx.doi.org/10.1016/s1470-2045(15)00128-x.

65. Medical Research Council. UKCCCR trial to study the effect on breast cancer mortality of annual mammographic screening starting at age 40 [online]. 2020 [Accessed: 24.08.2021]. URL: https://www.isrctn.com/ISRCTN24647151.

66. Tarone RE. The excess of patients with advanced breast cancer in young women screened with mammography in the Canadian National Breast Screening Study. Cancer 1995; 75(4): 997-1003. https://dx.doi.org/10.1002/1097-0142(19950215)75:4<997::aid-cncr2820750415>3.0.co;2-m.

67. Kopans DB. The Canadian National Breast Screening Studies are compromised and their results are unreliable. They should not factor into decisions about breast cancer screening. Breast Cancer Res Treat 2017; 165(1): 9-15. https://dx.doi.org/10.1007/s10549-017-4302-9.

68. Yaffe MJ, Seely JM, Gordon PB et al. The randomized trial of mammography screening that was not-A cautionary tale. J Med Screen 2022; 29(1): 7-11. https://dx.doi.org/10.1177/09691413211059461.

69. Alexander FE, Anderson TJ, Brown HK et al. The Edinburgh randomised trial of breast cancer screening: results after 10 years of follow-up. Br J Cancer 1994; 70(3): 542-548. https://dx.doi.org/10.1038/bjc.1994.342.

70. Deutsches Mammographie-Screening-Programmem. Jahresbericht Evaluation 2019 [online]. 2021 [Accessed: 03.12.2021]. URL: https://fachservice.mammo-programmem.de/download/evaluationsberichte/neu_KOOPMAMMO_Jahresbericht_Eval_2019_20211112_web-Einzelseite.pdf.

71. Statistisches Bundesamt. Tabelle 2.1.2: Die drei häufigsten Todesursachen nach Geschlecht und Altersgruppen [Gesundheitliche Lage der Männer in Deutschland, 2014] [online]. [Accessed: 22.11.2021]. URL: https://www.gbe-bund.de/gbe/abrechnung.prc_abr_test_logon?p_uid=gast&p_aid=0&p_knoten=FID&p_sprache=D&p_suchstring=19740.

72. Nelson HD, Cantor A, Humphrey L et al. Screening for Breast Cancer: A Systematic Review to Update the 2009 U.S. Preventive Services Task Force Recommendation [online]. 2016 [Accessed: 11.06.2021]. URL: https://www.ncbi.nlm.nih.gov/books/n/es124/pdf/.

73. Gesundheitsberichterstattung des Bundes. Durchschnittliche Lebenserwartung im Alter von ... Jahren je Person. Gliederungsmerkmale: Zeitraum, Region, Alter, Geschlecht [online]. 2022 [Accessed: 14.06.2022]. URL: https://www.gbe-bund.de/gbe/!pkg_olap_tables.prc_set_page?p_uid=gast&p_aid=52621804&p_sprache=D&p_help=2&p_indnr=524&p_ansnr=27245048&p_version=3&D.001=1000001&D.003=43.

74. Ades F, Tryfonidis K, Zardavas D. The past and future of breast cancer treatment-from the papyrus to individualised treatment approaches. Ecancermedicalscience 2017; 11: 746. https://dx.doi.org/10.3332/ecancer.2017.746.

75. Gøtzsche PC, Jørgensen KJ. Screening for breast cancer with mammography. Cochrane Database Syst Rev 2013; 2013(6): Cd001877. https://dx.doi.org/10.1002/14651858.CD001877.pub5.

76. Stang A, Jockel KH. The Impact of Cancer Screening on All-Cause Mortality. Dtsch Arztebl Int 2018; 115(29-30): 481-486. https://dx.doi.org/10.3238/arztebl.2018.0481.

77. University of Oxford. Nationwide cluster-randomised trial of extending the NHS breast screening age range in England [online]. 2020 [Accessed: 24.08.2021]. URL: https://www.isrctn.com/ISRCTN33292440.

78. European Commission Initiative on Breast Cancer. European guidelines on breast cancer screening and diagnosis [online]. 2021 [Accessed: 28.12.2021]. URL: https://healthcare-quality.jrc.ec.europa.eu/ecibc/european-breast-cancer-guidelines.

79. Tormann D. Mammographie-Screening zwischen Sicherheitsbedürfnis und nachhaltiger Verunsicherung [online]. 2016 [Accessed: 13.01.2022]. URL: https://dgpfg.de/blog/gyne-072016-mammographie-screening-zwischen-sicherheitsbeduerfnis-und-nachhaltiger-verunsicherung/.

80. Armstrong K, Moye E, Williams S et al. Screening mammography in women 40 to 49 years of age: a systematic review for the American College of Physicians. Ann Intern Med 2007; 146(7): 516-526. https://dx.doi.org/10.7326/0003-4819-146-7-200704030-00008.

81. Nekolla EA, Griebel J, Brix G. Strahlenrisiko infolge von Mammographie-Screening-V ntersuchungen für Frauen unter 50 Jahren. Z Med Phys 2008; 18(3): 170-179. https://dx.doi.org/10.1016/j.zemedi.2007.12.004.

82. Schonberg MA, Silliman RA, Marcantonio ER. Weighing the benefits and burdens of mammography screening among women age 80 years or older. J Clin Oncol 2009; 27(11): 1774-1780. https://dx.doi.org/10.1200/JCO.2008.19.9877.

83. Chaltiel D, Hill C. Estimations of overdiagnosis in breast cancer screening vary between 0% and over 50%: why? BMJ Open 2021; 11(6): e046353. https://dx.doi.org/10.1136/bmjopen-2020-046353.

84. Duffy SW, Agbaje O, Tabar L et al. Overdiagnosis and overtreatment of breast cancer: estimates of overdiagnosis from two trials of mammographic screening for breast cancer. Breast Cancer Res 2005; 7(6): 258-265. https://dx.doi.org/10.1186/bcr1354.

85. Ryser MD, Lange J, Inoue LYT et al. Estimation of Breast Cancer Overdiagnosis in a U.S. Breast Screening Cohort. Ann Intern Med 2022. https://dx.doi.org/10.7326/M21-3577.

86. Basu P, Ponti A, Anttila A et al. Status of implementation and organization of cancer screening in The European Union Member States-Summary results from the second European screening report. Int J Cancer 2018; 142(1): 44-56. https://dx.doi.org/10.1002/ijc.31043.

87. Dierks ML, Schmacke N. Mammographie-Screening und informierte Entscheidung: mehr Fragen als Antworten. In: Böcken J, Braun B, Meierjürgen R (Ed). Gesundheitsmonitor 2014; Bürgerorientierung im Gesundheitswesen; Kooperationsprojekt der Bertelsmann Stiftung und der BARMER GEK. Gütersloh: Bertelsmann Stiftung; 2014. p. 55-91.

88. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Mammographie-Screening - Eine Entscheidungshilfe - Programmem zur Früherkennung von Brustkrebs für Frauen zwischen 50 und 69 Jahren [online]. 2016 [Accessed: 13.01.2022]. URL: https://www.iqwig.de/download/p14-03_entscheidungshilfe_mammographie.pdf.

89. World Health Organisation. WHO position paper on mammography screening [online]. 2014 [Accessed: 28.12.2021]. URL: https://apps.who.int/iris/handle/10665/137339.

90. Siu AL, Force USPST. Screening for Breast Cancer: U.S. Preventive Services Task Force Recommendation Statement. Ann Intern Med 2016; 164(4): 279-296. https://dx.doi.org/10.7326/M15-2886.

91. Heindel W, Weigel S, Gerss J et al. Digital breast tomosynthesis plus synthesised mammography versus digital screening mammography for the detection of invasive breast cancer (TOSYMA): a multicentre, open-label, randomised, controlled, superiority trial. Lancet Oncol 2022; 23(5): 601-611. https://dx.doi.org/10.1016/S1470-2045(22)00194-2.

92. Lefebvre C, Glanville J, Briscoe S et al. Cochrane Handbook for Systematic Reviews of Interventions; Version 6.2; Technical Supplement to Chapter 4: Searching for and selecting studies [online]. 2021 [Accessed: 27.05.2021]. URL: https://training.cochrane.org/handbook/version-6.1/chapter-4-tech-suppl.

93. Wong SS, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455.

The full report (German version) is published under

https://www.iqwig.de/en/projects/s21-01.html

Appendix A – Search strategies

A.1 – Searches in bibliographic databases

Search for systematic reviews

1. PubMed

Search interface: NLM

#	Searches
1	mammography screening
2	systematic [sb]
3	#1 and #2
4	#3 AND 2011:2021[DP]

2. Health Technology Assessment Database

Search interface: INAHTA

#	Searches
1	mammography
2	screening
3	#2 AND #1 2011 to 2021

Search for primary studies

1. MEDLINE

Search interface: Ovid

- Ovid MEDLINE(R) 1946 to March 10, 2022

The following filters were adopted:

- RCT: Lefebvre [92] – Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2008 revision)

#	Searches
1	exp Breast Neoplasms/
2	(breast* adj3 (cancer* or carcinoma* or tumour* or tumour*)).ti,ab.
3	or/1-2
4	Mass Screening/
5	Early Detection of Cancer/
6	screening*.mp.
7	or/4-6
8	exp Mammography/
9	mammogr*.ti,ab.

#	**Searches**
10	or/8-9
11	and/3,7,10
12	randomized controlled trial.pt.
13	controlled clinical trial.pt.
14	(randomized or placebo or randomly or trial or groups).ab.
15	drug therapy.fs.
16	or/12-15
17	16 not (exp animals/ not humans.sh.)
18	and/11,17
19	18 not (comment or editorial).pt.
20	19 and (english or german).lg.
21	20 and 20160401:3000.(dt).

Search interface: Ovid

- Ovid MEDLINE(R) Epub Ahead of Print and In-Process, In-Data-Review & Other Non-Indexed Citations March 10, 2022

#	**Searches**
1	(breast* adj3 (cancer* or carcinoma* or tumour* or tumour*)).ti,ab.
2	screening*.mp.
3	mammogr*.ti,ab.
4	and/1-3
5	(clinical trial* or random* or placebo).ti,ab.
6	trial.ti.
7	or/5-6
8	and/4,7
9	8 not (comment or editorial).pt.
10	9 and (english or german).lg.
11	10 and 20160401:3000.(dt).

2. Embase

Search interface: Ovid

- Embase 1974 to 2022 March 10

The following filters were adopted:

- RCT: Wong [93] – Strategy minimizing difference between sensitivity and specificity

#	**Searches**
1	exp Breast Cancer/
2	(breast* adj3 (cancer* or carcinoma* or tumour* or tumour*)).ti,ab.
3	or/1-2
4	Cancer Screening/
5	screening*.mp.
6	or/4-5
7	exp Mammography/
8	mammogr*.ti,ab.
9	or/7-8
10	(random* or double-blind*).tw.
11	placebo*.mp.
12	or/10-11
13	12 not (exp animal/ not exp human/)
14	and/3,6,9,13
15	14 not (Conference Abstract or Conference Review or Editorial).pt.
16	15 and (english or german).lg.
17	16 and 20160401:3000.(dc).
18	17 not medline.cr.

3. The Cochrane Library

Search interface: Wiley

- Cochrane Central Register of Controlled Trials: Issue 2 of 12, February 2022

#	**Searches**
#1	[mh "Breast Neoplasms"]
#2	(breast* NEAR/3 (cancer* or carcinoma* or tumour* or tumour*)):ti,ab
#3	#1 or #2
#4	[mh ^"Mass Screening"]
#5	[mh ^"Early Detection of Cancer"]
#6	screening*:ti,ab
#7	#4 or #5 or #6
#8	[mh "Mammography"]

#	Searches
#9	mammogr*:ti,ab
#10	#8 or #9
#11	#3 and #7 and #10
#12	#11 not (*clinicaltrial*gov* or *who*trialsearch* or *clinicaltrialsregister*eu* or *anzctr*org*au* or *trialregister*nl* or *irct*ir* or *isrctn* or *controlled*trials*com* or *drks*de*):so
#13	#12 not ((language next (afr or ara or aze or bos or bul or car or cat or chi or cze or dan or dut or es or est or fin or fre or gre or heb or hrv or hun or ice or ira or ita or jpn or ko or kor or lit or nor or peo or per or pol or por or pt or rom or rum or rus or slo or slv or spa or srp or swe or tha or tur or ukr or urd or uzb)) not (language near/2 (en or eng or english or ger or german or mul or unknown)))
#14	#13 with Cochrane Library publication date from Apr 2016 to present, in Trials

A.2 – Searches in study registries

1. ClinicalTrials.gov

Provider: U.S. National Institutes of Health

- URL: http://www.clinicaltrials.gov
- Type of search: Expert Search

Search strategy
mammography AND screening AND breast

2. International Clinical Trials Registry Platform Search Portal

Provider: World Health Organization

- URL: https://trialsearch.who.int
- Type of search: Standard Search

Search strategy
mammogr* AND screen* AND breast OR mammography AND screening AND breast