Überprüfung der Altersgrenzen im Mammografie-Screening-Programm

Vorbericht (vorläufige Nutzenbewertung)

Auftrag: S21-01
Version: 1.0
Stand: 16.02.2022

Impressum

Herausgeber
Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema
Überprüfung der Altersgrenzen im Mammografie-Screening-Programm

Auftraggeber
Gemeinsamer Bundesausschuss

Datum des Auftrags
22.04.2021

Interne Auftragsnummer
S21-01

Anschrift des Herausgebers
Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Im Mediapark 8
50670 Köln

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

Dieser Bericht wurde unter Beteiligung externer Sachverständiger erstellt. Externe Sachverständige, die wissenschaftliche Forschungsaufträge für das Institut bearbeiten, haben gemäß § 139b Abs. 3 Satz 2 Sozialgesetzbuch – Fünftes Buch – Gesetzliche Krankenversicherung „alle Beziehungen zu Interessenverbänden, Auftragsinstituten, insbesondere der pharmazeutischen Industrie und der Medizinprodukteindustrie, einschließlich Art und Höhe von Zuwendungen“ offenzulegen. Das Institut hat von jedem der Sachverständigen ein ausgefülltes „Formblatt zur Offenlegung von Beziehungen“ erhalten. Die Angaben wurden von dem speziell für die Beurteilung der Interessenkonflikte eingerichteten Gremium des Instituts bewertet. Es wurden keine Interessenkonflikte festgestellt, die die fachliche Unabhängigkeit im Hinblick auf eine Bearbeitung des vorliegenden Auftrags gefährden.

Dieser Vorbericht ist eine vorläufige Nutzenbewertung. Er wird zur Anhörung gestellt und es können schriftliche Stellungnahmen eingereicht werden. Das Ende der Stellungnahmefrist wird auf der Website des IQWiG (www.iqwig.de) bekannt gegeben. Dort sind auch die notwendigen Formblätter zu finden. Stellungnahmen können alle interessierten Personen, Institutionen und Gesellschaften abgeben. Gegebenenfalls wird eine wissenschaftliche Erörterung zur Klärung unklarer Aspekte aus den schriftlichen Stellungnahmen durchgeführt. Die Anhörung kann zu Änderungen und / oder Ergänzungen des Berichts führen.

Schlagwörter

Reihenuntersuchung, Mammografie, Mammatumoren – Mensch, Nutzenbewertung, Systematische Übersicht

Keywords

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

Kernaussage

Fragestellung

Die Ziele der vorliegenden Untersuchung sind die Nutzenbewertung eines Screenings auf Brustkrebs mittels Mammografie im Vergleich zu keinem Screening (oder der alleinigen Palpation der Brust)

- bei Frauen im Alter zwischen 45 und 49 Jahren ohne Verdacht auf Brustkrebs und ohne spezifisch erhöhtes Brustkrebsrisiko (Fragestellung 1) sowie
- bei Frauen ab einem Alter von 70 Jahren ohne Verdacht auf Brustkrebs und ohne spezifisch erhöhtes Brustkrebsrisiko (Fragestellung 2)

hinsichtlich patientenrelevanter Endpunkte.

Fazit

Altersgruppe 45 bis 49 Jahre

Für die brustkrebsspezifische Mortalität liegt ein Anhaltspunkt für einen Nutzen des Mammografie-Screenings im Vergleich zu keinem Screening bei Frauen zwischen 45 und 49 Jahren vor. In Hinblick auf die Gesamtmortalität zeigte sich kein statistisch signifikantes Ergebnis. Das Ergebnis zur Gesamtmortalität spricht nicht gegen einen Nutzen bezüglich des Mammografie-Screenings, denn die Gesamtmortalität wird wesentlich durch das Auftreten konkurrierender Todesursachen beeinflusst. Es wird davon ausgegangen, dass der Effekt der verringerten Brustkrebsmortalität durch das Screening insgesamt zu klein war um sich auf die Gesamtmortalität auszuwirken. Das Mammografie-Screening führt jedoch über falsch-positive Screeningbefunde zu negativen Konsequenzen (Hinweis auf einen Schaden). Auch kommt es zu Überdiagnosen (Anhaltspunkt für einen Schaden). In Bezug auf Mastektomien ergab sich kein Anhaltspunkt für einen Nutzen oder Schaden des Mammografie-Screenings. Für unerwünschte Ereignisse und gesundheitsbezogene Lebensqualität lagen keine Daten vor, sodass sich auch hier jeweils kein Anhaltspunkt für einen Nutzen oder Schaden ergab. Der Effekt des Screenings auf die Rate unerwünschter Ereignisse und auf die gesundheitsbezogene Lebensqualität dürfte jedoch im Wesentlichen durch den Endpunkt Überdiagnosen erfasst sein.

Zusammenfassend wird ein Anhaltspunkt für einen Nutzen des Mammografie-Screenings für Frauen zwischen 45 und 49 Jahren gegenüber keinem Screening festgestellt und somit, dass der Nutzen des Mammografie-Screenings den Schaden überwiegt. Eine individuelle Bewertung und Abwägung bleibt aber angesichts des sehr kleinen Mortalitätsvorteils in dieser Altersgruppe unerlässlich. Deshalb müssen alle Voraussetzungen getroffen werden, um den Frauen eine informierte Entscheidung zu ermöglichen.

Altersgruppe ab 70 Jahren

Ein Einfluss des Mammografie-Screenings auf die Mortalität im Vergleich zu keinem Screening ist bei Frauen ab 70 Jahren nicht belegt (kein Anhaltspunkt für einen Nutzen oder Schaden).

Das Mammografie-Screening führt zu Überdiagnosen und zu negativen Konsequenzen aus falsch-positiven Screeningbefunden (jeweils ein Anhaltspunkt für einen Schaden). Für Mastektomien, unerwünschte Ereignisse und die gesundheitsbezogene Lebensqualität lagen keine Daten vor, sodass sich hier jeweils kein Anhaltspunkt für einen Nutzen oder Schaden ergab.

Der endpunktübergreifende Nutzen bzw. Schaden eines Mammografie-Screenings bei Frauen ab 70 Jahren bleibt daher wegen unzureichender Daten insgesamt unklar.

Jedoch lässt sich für die Altersgruppe der 70- bis 74-jährigen Frauen ein Potenzial erkennen. Eine Erprobungsstudie zu dieser Fragestellung erscheint dennoch verzichtbar, weil eine laufende randomisierte kontrollierte Studie (AgeX) in absehbarer Zeit aussagekräftige Daten erwarten lässt.

Inhaltsverzeichnis

Seite

Kernaussage ... **iv**
Tabellenverzeichnis ... **x**
Abbildungsverzeichnis ... **xii**
Abkürzungsverzeichnis ... **xiii**
1 Hintergrund ... **1**
2 Fragestellung ... **3**
3 Methoden ... **4**
4 Ergebnisse ... **6**
4.1 Ergebnisse der Informationsbeschaffung ... **6**
4.2 Charakteristika der in die Bewertung eingeschlossenen Studien ... **6**
4.3 Übersicht der patientenrelevanten Endpunkte ... **8**
4.4 Bewertung des Verzerrungspotenzials der Ergebnisse ... **10**
4.5 Ergebnisse zu patientenrelevanten Endpunkten ... **12**
4.5.1 Fragestellung 1: Altersgruppe 45 bis 49 Jahre ... 12
4.5.1.1 Ergebnisse zur Mortalität ... 12
4.5.1.1.1 Ergebnisse zur Gesamtmortalität ... 12
4.5.1.1.2 Ergebnisse zur brustkrebsspezifischen Mortalität ... 12
4.5.1.1.3 Zusammenfassende Betrachtung für die Mortalität ... 14
4.5.1.2 Ergebnisse zu Mastektomien ... 14
4.5.1.3 Ergebnisse zu Konsequenzen aus falsch-positiven Screeningbefunden ... 15
4.5.1.4 Ergebnisse zu unerwünschten Ereignissen ... 16
4.5.1.5 Ergebnisse zu Überdiagnosen ... 16
4.5.1.6 Ergebnisse zur gesundheitsbezogenen Lebensqualität ... 19
4.5.2 Fragestellung 2: Altersgruppe ab 70 Jahre ... 19
4.5.2.1 Ergebnisse zur Mortalität ... 19
4.5.2.1.1 Ergebnisse zur Gesamtmortalität ... 19
4.5.2.1.2 Ergebnisse zur brustkrebsspezifischen Mortalität ... 19
4.5.2.1.3 Zusammenfassende Betrachtung für die Mortalität ... 19
4.5.2.2 Ergebnisse zu Mastektomien ... 19
4.5.2.3 Ergebnisse zu Konsequenzen aus falsch-positiven Screeningbefunden ... 20
4.5.2.4 Ergebnisse zu unerwünschten Ereignissen ... 20
4.5.2.5 Ergebnisse zu Überdiagnosen ... 20
4.5.2.6 Ergebnisse zur gesundheitsbezogenen Lebensqualität ... 21

4.6 Zusammenfassende Bewertung der Ergebnisse ... 22
5 Einordnung des Arbeitsergebnisses ... 29
6 Fazit ... 36
Details des Berichts ... 37
A1 Projektverlauf ... 37
A1.1 Zeitlicher Verlauf des Projekts ... 37
A1.2 Spezifizierungen und Änderungen im Projektverlauf ... 37
A2 Methodik gemäß Berichtsplan 1.0 ... 40
A2.1 Kriterien für den Einschluss von Studien in die Untersuchung ... 40
A2.1.1 Population ... 40
A2.1.2 Prüf- und Vergleichsintervention ... 40
A2.1.3 Darstellung der Einzelstudien ... 40
A2.1.4 Patientenrelevante Endpunkte ... 41
A2.1.5 Studientypen ... 42
A2.1.6 Studiendauer ... 42
A2.1.7 Publikationssprache ... 42
A2.1.8 Tabellarische Darstellung der Kriterien für den Studieneinschluss ... 42
A2.1.9 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen ... 43
A2.2 Informationsbeschaffung ... 43
A2.2.1 Fokussierte Informationsbeschaffung von systematischen Übersichten ... 43
A2.2.2 Umfassende Informationsbeschaffung von Studien ... 44
A2.2.3 Selektion relevanter Studien aus der umfassenden Informationsbeschaffung ... 45
A2.3 Informationsbewertung und -synthese ... 46
A2.3.1 Bewertung des Verzerrungspotenzials der Ergebnisse ... 46
A2.3.2 Metaanalysen ... 46
A2.3.3 Sensitivitätsanalysen ... 47
A2.3.4 Subgruppenmerkmale und andere Effektmodifikatoren ... 48
A2.3.5 Aussagen zur Beleglage ... 48
A2.4 Modellierung ... 49
A3 Details der Ergebnisse ... 51
A3.1 Informationsbeschaffung ... 51
A3.1.1 Fokussierte Informationsbeschaffung nach systematischen Übersichten ... 51
A3.1.2 Umfassende Informationsbeschaffung ... 51
A3.1.2.1 Primäre Informationsquellen ... 51
A3.1.2.1.1 Bibliografische Datenbanken ... 51
A3.1.2.1.2 Studienregister ... 53
A3.1.2.2 Weitere Informationsquellen und Suchtechniken ... 53

A3.1.2.2.1 Durch den G-BA übermittelte Dokumente ... 53
A3.1.2.2.2 Anwendung weiterer Suchtechniken ... 54
A3.1.2.2.3 Autorenanfragen ... 54
A3.1.2.2.4 Zusätzliche relevante Studien bzw. Dokumente ... 54
A3.1.3 Resultierender Studienpool ... 55
A3.1.4 Studien ohne berichtete Ergebnisse ... 55
A3.2 Charakteristika der in die Bewertung eingeschlossenen Studien ... 57
A3.2.1 Fragestellung 1 – Studiendesign und Studienpopulationen ... 57
A3.2.2 Fragestellung 2 – Studiendesign und Studienpopulationen ... 66
A3.2.3 Bewertung endpunktübergreifender Kriterien des Verzerrungspotenzials ... 67
A3.3 Fragestellung 1 – Altersgruppe 45 bis 49 Jahre ... 68
A3.3.1 Mortalität ... 68
A3.3.1.1 Gesamtmortalität ... 68
A3.3.1.2 Brustkrebsspezifische Mortalität ... 72
A3.3.2 Mastektomien ... 80
A3.3.3 Konsequenzen aus falsch-positiven Screeningbefunden ... 83
A3.3.4 Unerwünschte Ereignisse ... 86
A3.3.5 Überdiagnosen ... 86
A3.3.6 Gesundheitsbezogene Lebensqualität ... 90
A3.4 Fragestellung 2 – Altersgruppe ab 70 Jahren ... 90
A3.4.1 Mortalität ... 90
A3.4.1.1 Gesamtmortalität ... 90
A3.4.1.2 Brustkrebsspezifische Mortalität ... 92
A3.4.2 Mastektomien ... 94
A3.4.3 Konsequenzen aus falsch-positiven Screeningbefunden ... 94
A3.4.4 Unerwünschte Ereignisse ... 95
A3.4.5 Überdiagnosen ... 95
A3.4.6 Gesundheitsbezogene Lebensqualität ... 98
A4 Bericht der externen Expertengruppe zur Modellierung ... 99
A5 Kommentare ... 192
A5.1 Bericht im Vergleich zu anderen systematischen Übersichten ... 192
A5.2 Bericht im Vergleich zu internationalen Leitlinien ... 192
A5.3 Kritische Reflexion des Vorgehens ... 194
A5.3.1 Betrachtete Studientypen ... 194
A6 Literatur ... 195
A7 Studienlisten ... 205
A7.1 Liste der gesichteten systematischen Übersichten ... 205

A7.2 Liste der ausgeschlossenen Publikationen aus der fokussierten Informationsbeschaffung 205
A7.3 Liste der ausgeschlossenen Publikationen aus der bibliografischen Recherche mit Ausschlussgründen 206
A7.4 Liste der ausgeschlossenen Dokumente aus den durch den G-BA übermittelten Dokumenten mit Ausschlussgründen 208
A8 Bewertung der Qualität der berücksichtigten systematischen Übersichten aus der fokussierten Recherche 210
A9 Suchstrategien 211
A9.1 Bibliografische Datenbanken 211
A9.2 Studienregister 214

Tabellenverzeichnis

Seite

Tabelle 1: Matrix der patientenrelevanten Endpunkte 10
Tabelle 2: Landkarte der Beleglage in Bezug auf die patientenrelevanten Endpunkte 22
Tabelle 3: Fragestellung 1: Überblick und Erläuterung der Ergebnisse für alle patientenrelevanten Endpunkte für eine Nutzen-Schaden-Abwägung 24
Tabelle 4: Übersicht über die Kriterien für den Studieneinschluss 43
Tabelle 5: Regelhaft abgeleitete Aussagesicherheiten für verschiedene Evidenzsituationen beim Vorliegen von Studien derselben qualitativen Ergebnissicherheit 49
Tabelle 6: Berücksichtigte systematische Übersichten 51
Tabelle 7: In Studienregistern identifizierte relevante Studien bzw. Dokumente 53
Tabelle 8: In vom G-BA übermittelten Dokumenten identifizierte relevante Studien bzw. Dokumente 54
Tabelle 9: Durch Anwendung weiterer Suchtechniken identifizierte relevante Studien bzw. Dokumente 54
Tabelle 10: Zusätzlich identifizierte relevante Studien bzw. Dokumente 55
Tabelle 11: Studienpool der Nutzenbewertung 55
Tabelle 12: In der Informationsbeschaffung identifizierte Studien ohne berichtete Ergebnisse 56
Tabelle 13: Fragestellung 1 – Charakterisierung der eingeschlossenen Studien 57
Tabelle 14: Fragestellung 1 – Charakterisierung der Interventionen in den eingeschlossenen Studien 59
Tabelle 15: Fragestellung 1 – Ein- / Ausschlusskriterien für die zu rekrutierenden Frauen 62
Tabelle 16: Fragestellung 1 – Charakterisierung Teilnehmerfluss innerhalb der Studien 63
Tabelle 17: Fragestellung 2 – Charakterisierung der eingeschlossenen Studien 66
Tabelle 18: Fragestellung 2 – Charakterisierung der Interventionen in den eingeschlossenen Studien 66
Tabelle 19: Fragestellung 2 – Ein- / Ausschlusskriterien für die zu rekrutierenden Frauen 67
Tabelle 20: Fragestellung 2 – Charakterisierung Teilnehmerfluss innerhalb der Studien 67
Tabelle 21: Endpunktübergreifendes Verzerrungspotenzial 68
Tabelle 22: Fragestellung 1 – endpunktspezifisches Verzerrungspotenzial: Gesamtmortalität 69
Tabelle 23: Fragestellung 1 – Ergebnisse: Gesamtmortalität 70
Tabelle 24: Fragestellung 1 – endpunktspezifisches Verzerrungspotenzial: brustkrebsspezifische Mortalität 73
Tabelle 25: Fragestellung 1 – Ergebnisse: brustkrebsspezifische Mortalität 74
Tabelle 26: Fragestellung 1 – endpunktspezifisches Verzerrungspotenzial: Mastektomien 81
Tabelle 27: Fragestellung 1 – Ergebnisse: Brustkrebs-Operationen einschließlich Mastektomien 82

Tabelle 28: Fragestellung 1 – endpunktspezifisches Verzerrungspotenzial: Konsequenzen aus falsch-positiven Screeningbefunden ... 83
Tabelle 29: Fragestellung 1 – Ergebnisse: Konsequenzen aus falsch-positiven Screeningbefunden: Biopsien mit benignem Befund ... 84
Tabelle 30: Fragestellung 1 – Ergebnisse: Konsequenzen aus falsch-positiven Screeningbefunden: Operationen mit benignem Befund ... 85
Tabelle 31: Fragestellung 1 – endpunktspezifisches Verzerrungspotenzial: Überdiagnosen ... 86
Tabelle 32: Fragestellung 1 – Ergebnisse: Überdiagnosen ... 87
Tabelle 33: Fragestellung 2 – endpunktspezifisches Verzerrungspotenzial: Gesamtmortalität ... 90
Tabelle 34: Fragestellung 2 – Ergebnisse: Gesamtmortalität ... 91
Tabelle 35: Fragestellung 2 – endpunktspezifisches Verzerrungspotenzial: brustkrebsspezifische Mortalität ... 92
Tabelle 36: Fragestellung 2 – Ergebnisse: brustkrebsspezifische Mortalität ... 93
Tabelle 37: Fragestellung 2 – endpunktspezifisches Verzerrungspotenzial: Konsequenzen aus falsch-positiven Screeningbefunden ... 94
Tabelle 38: Fragestellung 2 – Ergebnisse: Konsequenzen aus falsch-positiven Screeningbefunden: chirurgische Biopsien mit benignem Befund ... 95
Tabelle 39: Fragestellung 2 – endpunktspezifisches Verzerrungspotenzial: Überdiagnosen) ... 96
Tabelle 40: Fragestellung 2 – Ergebnisse: Überdiagnosen ... 97
Tabelle 41: Bewertung der Informationsbeschaffung zur systematischen Übersicht Canelo-Aybar 2021 ... 210

Abbildungsverzeichnis

Seite

Abbildung 1: Ergebnis der umfassenden Informationsbeschaffung aus den bibliografischen Datenbanken und der Studienselektion 52
Abbildung 2: Fragestellung 1 – Gesamtmortalität 72
Abbildung 3: Fragestellung 1 – brustkrebsspezifische Mortalität, 5 Jahre nach Randomisierung 78
Abbildung 4: Fragestellung 1 – brustkrebsspezifische Mortalität, 10 Jahre nach Randomisierung 78
Abbildung 5: Fragestellung 1 – brustkrebsspezifische Mortalität, 15 Jahre nach Randomisierung 78
Abbildung 6: Fragestellung 1 – brustkrebsspezifische Mortalität, 20 Jahre nach Randomisierung 79
Abbildung 7: Fragestellung 1 – brustkrebsspezifische Mortalität, 10 Jahre nach Randomisierung, Sensitivitätsanalyse mit 9-Jahres-Daten der UK Age 79
Abbildung 8: Fragestellung 1 – brustkrebsspezifische Mortalität, 10 Jahre nach Randomisierung, Sensitivitätsanalyse ohne Daten der Edinburgh-Studie 80
Abbildung 9: Fragestellung 1 – brustkrebsspezifische Mortalität, 10 Jahre nach Randomisierung, Sensitivitätsanalyse mit Daten der 10-Jahres-Altersgruppe der Göteborg-Studie 80
Abbildung 10: Fragestellung 2 – brustkrebsspezifische Mortalität, 10 Jahre nach Randomisierung 94

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AHRQ	Agency for Healthcare Research and Quality
BMU	Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit
CNBSS	Canadian National Breast Screening Study
DCIS	duktales Carcinoma in situ
ECIBC	European Commission Initiative on Breast Cancer
G-BA	Gemeinsamer Bundesausschuss
HIP	Health Insurance Plan of Greater New York
IARC	International Agency for Research on Cancer (Internationale Agentur für Krebsforschung)
ICD-10	International Statistical Classification of Diseases and Related Health Problems, Revision 10 (Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme, 10. Revision)
IDR	Inzidenzdichtequotient
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
KI	Konfidenzintervall
MMST	Malmö Mammographic Screening Trial
NICE	National Institute for Health and Care Excellence
OP	Operation
OR	Odds Ratio
QALY	Quality-adjusted life years (qualitätsadjustierte Lebensjahre)
RCT	Randomized controlled Trial (randomisierte kontrollierte Studie)
RKI	Robert Koch-Institut
RR	relatives Risiko
SÜ	systematische Übersichtsarbeit
WHO	Weltgesundheitsorganisation (World Health Organization)

1 Hintergrund

Brustkrebs oder Mammakarzinom (ICD-10 C50) bezeichnet eine bösartige Neubildung der Brustdrüse. Die Internationale Agentur für Krebsforschung (IARC) schätzte, dass im Jahr 2020 weltweit bei Frauen Brustkrebs die häufigste Krebserkrankung unter allen Krebsneuerkrankungen (ausgenommen nicht melanotischer Hautkrebs) war und die häufigste Krebstodesursache darstellt [1]. Dies trifft laut Daten des Robert Koch-Instituts (RKI) zufolge auch für Frauen in Deutschland zu [2]. Bekannte Faktoren, die das Erkrankungsrisiko für Brustkrebs beeinflussen, sind das Alter, Hormone (z. B. Alter bei der 1. Regelblutung, Anzahl der Geburten, Alter bei der 1. Geburt, Hormonersatztherapie in oder nach den Wechseljahren), vorherige Brustkrebserkrankung und gutartige Brustveränderungen [2]. Auch die Brustdichte wird als unabhängiger Risikofaktor beschrieben. Darüber hinaus geht eine höhere Brustdichte mit der Abnahme der Sensitivität und Spezifität der Mammografie einher [3]. Neben den allgemeinen Faktoren haben Frauen, bei denen eine familiäre Belastung oder eine genetische Disposition für Brustkrebs vorliegt, ein spezifisch erhöhtes Risiko an Brustkrebs zu erkranken [4].

Das Mammografie-Screening-Programm ist ein Vorhaben zur Früherkennung von Brustkrebs bei Frauen in Deutschland. Hierbei wird allen Frauen einer bestimmten Altersgruppe in festgelegten Abständen eine Untersuchung der Brust mittels Röntgen angeboten. Bei einem auffälligen Screeningbefund erfolgen weitere Untersuchungen zur Abklärung. Reicht eine Ultraschalluntersuchung oder eine erneute Mammografie nicht aus, um den Krebsverdacht auszuschließen, wird die Entnahme einer Gewebeprobe (Biopsie) aus der Brust empfohlen [5]. Das Ziel des Mammografie-Screenings ist es, durch ein frühzeitiges Erkennen der Krebserkrankung die Heilungschancen zu verbessern, während weniger belastende Therapieoptionen aufgrund des früheren Tumorstadiums möglich sind [3]. Gleichzeitig kann das Screening auch zu falsch-positiven Befunden und somit zu unnötigen Biopsien führen [6,7]. Bei einem Teil der Brustkrebsdiagnosen handelt es sich um Überdiagnosen. In diesem Fall sind dies Diagnosen krebsartiger Veränderungen, die ohne Screening nicht entdeckt und auffällig geworden wären. Die sich dieser Überdiagnose anschließende Behandlung ist dementsprechend eine unnötige Behandlung. Wie Überdiagnosen kann auch die Diagnose eines duktalen Carcinoma in situ (DCIS) infolge einer Mammografie [8] unnötige Behandlungen zur Folge haben. Das DCIS, welches ein sehr unterschiedliches Malignitätspotenzial aufweist, gilt als eine Brustkrebsvorstufe. Ob sich aus einem DCIS tatsächlich ein invasiver Brustkrebs entwickelt, kann aber nicht zuverlässig vorhergesagt werden [3].

Um die bestmögliche Behandlung zu gewährleisten, soll nach aktueller Leitlinienempfehlung die weiterführende Therapie der Patientinnen mit auffälligen Biopsien in zertifizierten Brustzentren erfolgen [3]. Sollte die Malignität der Probe pathologisch bestätigt werden, gibt es neben einer Operation verschiedene Therapiemöglichkeiten wie Bestrahlung, Hormon- und Chemotherapie. Dabei ist für alle nicht fortgeschrittenen Brustkrebskarzinome die vollständige Entfernung des Tumors Basis der Therapie [3].

In Deutschland wurde das Mammografie-Screening auf der Grundlage europäischer Leitlinien [9] in die Richtlinie des Gemeinsamen Bundesausschusses (G-BA) über die Früherkennung von Krebserkrankungen mitaufgenommen. Im Jahr 2005 startete das nationale Screeningprogramm [3]. In Deutschland wird aktuell jede Frau im Alter zwischen 50 und 69 Jahren – sofern sie der Einladung nicht widersprochen hat – alle 2 Jahre zur Teilnahme am Mammografie-Screening eingeladen [10]. Nach den aktuellen Leitlinien der European Commission Initiative for Breast Cancer Screening and Diagnosis wird ein Mammografie-Screening zusätzlich für Frauen im Alter zwischen 45 und 49 Jahren, bei denen kein Verdacht auf Brustkrebs besteht und die ein durchschnittliches Brustkrebsrisiko aufweisen, empfohlen, ebenso wie für Frauen zwischen 70 und 74 Jahren. Für Frauen unter 45 Jahren wird kein Screening empfohlen, für Frauen über 74 Jahren gibt es aktuell keine Empfehlung für oder gegen ein Screening [11]. In außereuropäischen Leitlinien, wie der Leitlinie der Canadian Task Force on Preventive Health Care, wird hingegen für die Altersgruppe der Frauen zwischen 40 und 49 Jahren kein Mammografie-Screening empfohlen, Frauen im Alter zwischen 70 und 74 Jahren wird ebenso wie in der europäischen Leitlinie ein Screening alle 2 bis 3 Jahre empfohlen [6].

Ob und in welchem Maße auch Frauen zwischen 45 und 49 Jahren beziehungsweise Frauen, die 70 Jahre und älter sind, von einem Mammografie-Screening profitieren können, ist die Fragestellung der vorliegenden Nutzenbewertung.

2 Fragestellung

Die Ziele der vorliegenden Untersuchung sind die Nutzenbewertung eines Screenings auf Brustkrebs mittels Mammografie im Vergleich zu keinem Screening (oder der alleinigen Palpation der Brust)

- bei Frauen im Alter zwischen 45 und 49 Jahren ohne Verdacht auf Brustkrebs und ohne spezifisch erhöhtes Brustkrebsrisiko (Fragestellung 1) sowie
- bei Frauen ab einem Alter von 70 Jahren ohne Verdacht auf Brustkrebs und ohne spezifisch erhöhtes Brustkrebsrisiko (Fragestellung 2)

hinsichtlich patientenrelevanter Endpunkte.

3 Methoden

Die Zielpopulation der Nutzenbewertung für die Fragestellung 1 bildeten Frauen im Alter zwischen 45 und 49 Jahren, bei denen kein Verdacht auf Brustkrebs bestand und kein spezifisch erhöhtes Brustkrebsrisiko vorlag. Die Zielpopulation der Nutzenbewertung für die Fragestellung 2 bildeten Frauen im Alter von 70 Jahren und mehr, bei denen kein Verdacht auf Brustkrebs bestand und kein spezifisch erhöhtes Brustkrebsrisiko vorlag. Die Prüfintervention bildete für beide Fragestellungen das Mammografie-Screening. Als Vergleichsintervention galt kein Screening oder die alleinige Palpation der Brust.

Für die Untersuchung wurden folgende patientenrelevante Endpunkte betrachtet:

- Mortalität (zum Beispiel Gesamtmortalität und brustkrebsspezifische Mortalität)
- Morbidität (zum Beispiel Brustamputationen, Konsequenzen aus falschen Screeningbefunden, Überdiagnosen)
- gesundheitsbezogene Lebensqualität
- unerwünschte Ereignisse

Es wurden randomisierte kontrollierte Studien (RCTs) in die Nutzenbewertung eingeschlossen. Studien, die von den Autorinnen und Autoren als randomisiert bezeichnet, jedoch im Rahmen der vorliegenden Bewertung als quasirandomisierte Studien aufgefasst wurden, wurden ebenfalls eingeschlossen. Hinsichtlich der Studiendauer bestand keine Einschränkung.

Für Fragestellung 1 (Frauen von 45 bis 49 Jahren) wurden Studien auch dann eingeschlossen, wenn weniger als 80 % der Frauen der relevanten Altersgruppe entsprachen bzw. es wurde auch auf Daten von Subgruppen zurückgegriffen, die sich auf die entsprechende 10-Jahres-Altersgruppe (z. B. 40 bis 49 Jahre) bezogen, weil hinreichend sicher plausibel war, dass die in einer solchen Studie bzw. in einer solchen Subgruppe gewonnenen Erkenntnisse auf die Zielpopulation der Frauen von 45 bis 49 Jahre anwendbar sind. Für Fragestellung 2 bestand das Problem der überlappenden Altersbereiche nicht, weil es keine Daten gab, die sich auf Frauen sowohl unter als auch über 70 Jahren bezogen. Waren in einer Fragestellung zu einem Endpunkt gar keine verwertbaren Daten der relevanten Altersgruppe oder der entsprechenden 10-Jahres-Altersgruppe verfügbar, wurde auf Daten einer Subgruppe mit einem breiteren Altersbereich zurückgegriffen und Fragen der Übertragbarkeit entsprechend berücksichtigt.

Parallel zur Erstellung des Berichtsplans erfolgte eine Suche nach systematischen Übersichten in den Datenbanken MEDLINE (umfasst auch die Cochrane Database of Systematic Reviews) und HTA Database sowie auf den Websites des National Institute for Health and Care Excellence (NICE) und der Agency for Healthcare Research and Quality (AHRQ). Die Suche wurde auf die Publikationsjahre ab 2011 eingeschränkt.

Es wurde geprüft, ob mindestens 1 hochwertige und aktuelle systematische Übersicht infrage kommt, deren Informationsbeschaffung als Grundlage für die Bewertung verwendet werden kann (im Folgenden: Basis-SÜ).

Da dies der Fall war, erfolgte in einem 2. Schritt eine ergänzende Suche nach Studien für den Zeitraum, der nicht durch die Basis-SÜ abgedeckt war.

Die systematische Literaturrecherche nach Studien wurde in den Datenbanken MEDLINE, Embase und Cochrane Central Register of Controlled Trials durchgeführt.

Darüber hinaus wurden folgende Informationsquellen und Suchtechniken berücksichtigt: Studienregister, vom G-BA übermittelte Dokumente und die Sichtung von Referenzlisten.

Die Selektion relevanter Studien erfolgte von 2 Personen unabhängig voneinander. Diskrepanzen wurden durch Diskussion zwischen beiden aufgelöst. Die Datenextraktion erfolgte in standardisierte Tabellen. Zur Einschätzung der qualitativen Ergebnissicherheit wurden endpunktübergreifende und endpunktspezifische Kriterien des Verzerrungspotenzials bewertet und das Verzerrungspotenzial jeweils als niedrig oder hoch eingestuft. Die Ergebnisse der einzelnen Studien wurden nach Endpunkten geordnet beschrieben.

Über die Gegenüberstellung der Ergebnisse der Einzelstudien hinaus wurden Metaanalysen und Sensitivitätsanalysen durchgeführt sowie Effektmodifikatoren untersucht, sofern die methodischen Voraussetzungen erfüllt waren.

Für jeden Endpunkt wurde eine Aussage zur Beleglage des (höheren) Nutzens und (höheren) Schadens in 4 Abstufungen bezüglich der jeweiligen Aussagesicherheit getroffen: Es lag entweder ein Beleg (höchste Aussagesicherheit), ein Hinweis (mittlere Aussagesicherheit), ein Anhaltspunkt (schwächste Aussagesicherheit) oder keine dieser 3 Situationen vor. Der letzte Fall trat ein, wenn keine Daten vorlagen oder die vorliegenden Daten keine der 3 übrigen Aussagen zuließen. In diesem Fall wurde die Aussage „Es liegt kein Anhaltspunkt für einen (höheren) Nutzen oder (höheren) Schaden vor“ getroffen.

Abschließend erfolgte eine endpunktübergreifende Bewertung des Nutzens und Schadens.

Falls kein Anhaltspunkt für einen Nutzen oder Schaden ableitbar war, wurde eine Aussage zum Potenzial der Behandlungsmethode getroffen.

Zusätzlich zur Nutzenbewertung wurde von einer externen Expertengruppe eine Modellierungsstudie erstellt. Die Ergebnisse der Modellierung gingen unterstützend in den Bericht ein. Details hierzu finden sich in Abschnitt 5. Die Methoden für diese Modellierung (und ihre Ergebnisse) werden im Abschnitt A4 dargestellt.

4 Ergebnisse

4.1 Ergebnisse der Informationsbeschaffung

Es wurde 1 systematische Übersicht als Basis-SÜ bewertet und zum Zweck der Identifizierung von Primärstudien berücksichtigt.

Die Informationsbeschaffung ergab 8 für die Fragestellung relevante Studien, davon 7 randomisierte und 1 quasirandomisierte kontrollierte Studie. Es wurde 1 laufende Studie identifiziert. Geplante Studien wurden nicht identifiziert. Die letzte Suche fand am 17.05.2021 statt.

4.2 Charakteristika der in die Bewertung eingeschlossenen Studien

Alle 8 eingeschlossenen Studien schlossen Frauen im Alter von 45 bis 49 Jahren (Fragestellung 1) ein, auch wenn keine Studie hierbei den exakten Altersbereich der Fragestellung abbildet. Lediglich 2 Studien konnten identifiziert werden, in denen Frauen ab 70 Jahren (Fragestellung 2) eingeschlossen wurden.

Die Canadian National Breast Screening Study (**CNBSS**) [12] ist eine multizentrische RCT mit Studienzentren in Kanada. Sie startete 1980 und umfasst 2 Altersgruppen (Frauen von 40 bis 49 Jahren und Frauen von 50 bis 59 Jahren), von denen Daten aus der jüngeren Altersgruppe für die Fragestellung 1 herangezogen werden können. Vor Randomisierung erhielten die Frauen eine klinische Untersuchung der Brust und eine Anleitung zur Selbstuntersuchung der Brust. Die Frauen der Interventionsgruppe wurden jährlich zum Screening eingeladen. Die Dauer der Screeningphase (der Begriff bezieht sich hier und im Folgenden auf die Zeitspanne, während der die eine Studiengruppe gescreent wird und die andere nicht) betrug bis zu 5 Jahre, die später rekrutierten Frauen nahmen jedoch nur an 3 oder 4 Screeningrunden teil. Die Frauen in der Vergleichsgruppe erhielten eine Anleitung zur Selbstuntersuchung der Brust und wurden jährlich angeschrieben.

Die **Edinburgh**-Studie [13] ist eine Cluster-RCT, die 1979 startete. Ein Großteil der Hausarztpraxen der Stadt willigte in die Teilnahme ein. Die Praxen wurden randomisiert der Screening- oder der Kontrollgruppe zugeteilt. Die in der jeweiligen Praxis registrierten Frauen der Altersgruppe 45 bis 64 Jahre wurden entsprechend der Zuteilung der Praxis den Gruppen zugeteilt. Die Frauen der Interventionsgruppe wurden alle 2 Jahre über insgesamt 7 Jahre gescreent. Die Frauen in der Vergleichsgruppe wurden nicht gescreent.

Die Health-Insurance-Plan-of-Greater-New-York (**HIP**)-Studie [14] ist mit Abstand die älteste der eingeschlossenen Studien. Die HIP-Studie startete 1963 in New York und damit 13 Jahre früher als die zweitälteste eingeschlossene Studie. Bis 1966 wurden über 61 000 Frauen im Alter von 40 bis 64 Jahren rekrutiert und entweder einem jährlichen Screening für 4 Screeningrunden oder einer Gruppe ohne Screening randomisiert zugeteilt.

4 der eingeschlossenen Studien wurden in Schweden durchgeführt: die Göteborg-Studie [15], die Malmö-Studie [16], die Stockholm-Studie [17] und die Swedish-Two-County-Studie [18].

Die **Göteborg**-Studie startete 1982 in Göteborg und schloss etwa 52 000 Frauen im Alter von 39 bis 59 Jahren ein, davon war etwa die Hälfte der Frauen unter 50 Jahre. Die Screeningphase für diese Altersgruppe (39 bis 49 Jahre) betrug 7 Jahre. In dieser Zeit wurde den Frauen der Interventionsgruppe alle 18 Monate ein Screening angeboten, also insgesamt 5 Screening-runden.

Die **Malmö**-Studie umfasst 2 Studienkollektive, die nach demselben Studienprotokoll rekrutiert wurden: Die erste Mammographic Screening Trial (MMST 1) startete 1976 und schloss Frauen ein, die bei Rekrutierung zwischen 45 und 70 Jahre alt waren. Die Einteilung in Altersgruppen basierte in einer der Auswertungen auf dem exakten Alter bei Randomisierung (anstatt auf dem Geburtsjahrgang), daher waren separate Daten für 70-jährige Frauen verfügbar, sodass die Studie sowohl für Fragestellung 1 als auch für Fragestellung 2 herangezogen werden konnte. In der MMST 2, die 1978 startete, wurden Frauen zwischen 44 und 49 Jahren rekrutiert. Die Screeningphase dauerte im Mittel 8,8 Jahre, bezogen auf die gesamte Studienpopulation (MMST 1 und 2). Die Frauen erhielten in der Zeit bis zu 9 (MMST 1) bzw. im Mittel 5 (MMST 2) Screenings bei einem Screeningintervall zwischen 1,5 und 2 Jahren. Die 70-jährigen Frauen erhielten maximal 6 Screenings. Separate Angaben zu den Screeningparametern für die beiden für den vorliegenden Bericht relevanten Altersgruppen waren nicht verfügbar.

Die 1981 gestartete **Stockholm**-Studie rekrutierte Frauen, die zu Studienbeginn zwischen 40 und 64 Jahre alt waren. Die Studie wurde in den zugehörigen Publikationen zwar als randomisiert bezeichnet, da die Zuteilung zur Screening- und Vergleichsgruppe jedoch anhand des Tages der Geburt erfolgte, wurde sie als quasirandomisiert klassifiziert. Die Frauen in der Interventionsgruppe wurden 2-mal mit einem Abstand von 2 Jahren gescreent und damit nur 1-mal mehr als die Frauen der Kontrollgruppe, die ein Abschlussscreening erhielten.

Die Cluster-RCT **Swedish Two County** rekrutierte ab 1977 Frauen in Kopparberg (Dalarna) und Östergötland und wertete Ergebnisse von Frauen zwischen 40 und 74 Jahren aus. Die Studie wird somit sowohl für die Fragestellung 1 als auch für Fragestellung 2 herangezogen. Die Screeningphase dauerte zwischen 6 und 8 Jahren, und währenddessen erhielten die 40- bis 49-jährigen Frauen 4, die 70- bis 74-jährigen 2 Screeningrunden. Das geplante Screening-intervall betrug je nach Alter 2 oder 3 Jahre. Bei den Frauen von 40 bis 49 Jahren ergab sich im Mittel ein Screeningintervall von 24 Monaten, bei Frauen ab 50 Jahren von 33 Monaten; separate Angaben zu Frauen ab 70 Jahren waren nicht verfügbar.

In der **UK-Age**-Studie [19] wurden Frauen im Alter von 39 bis 41 Jahren bei Randomisierung eingeschlossen. Die Studie startete 1991 und umfasste 23 Studienzentren im Vereinigten Königreich. Die Screeningphase dauerte bis zum 48. Lebensjahr und damit zwischen 7 und 9 Jahre. Damit ist dies die einzige der 8 eingeschlossenen Studien, deren Studienpopulation über die gesamte Screeningphase unter 50 Jahren blieb. Die Frauen in der Interventionsgruppe

wurden jährlich zur Mammografie eingeladen. Die Frauen in der Vergleichsgruppe erhielten die Grundversorgung ohne Screening. Mit über 160 000 randomisierten Frauen ist die UK Age die größte eingeschlossene Studie.

In allen eingeschlossenen Studien wurde die brustkrebsspezifische Mortalität als primärer Endpunkt betrachtet. Die Einschlusskriterien für die Studien umfassten lediglich das Alter und den Wohnort der Frauen (in der HIP-Studie anstelle des Wohnortes die Mitgliedschaft beim HIP), Ausschlusskriterien wurden nur in 3 Studien genannt und umfassten neben einem früheren Mammakarzinom (CNBSS, Edinburgh- und Göteborg-Studie) eine bestehende Schwangerschaft sowie eine Mammografie im letzten Jahr (CNBSS). Im Weiteren unterschieden sich die Studien hinsichtlich der Screeningstrategie: Zwar erhielten alle Frauen in der Screeninggruppe ein Mammografie-Screening, jedoch variierte die Anzahl der Screeningrunden (2 bis 8) sowie das Screeningintervall, also der Abstand zwischen den Screeningrunden (12 bis 24 Monate). In 3 Studien (Göteborg, Stockholm, Swedish Two County) erhielten die Frauen aus der Vergleichsgruppe im Rahmen der Studie ein Screening nach Abschluss der Screeningphase. In weiteren 3 Studien startete bald nach Ende der Screeningphase das nationale Screeningprogramm (Edinburgh, Malmö, UK Age), sodass anschließend die Frauen beider Gruppen – zum Teil beschränkt auf bestimmte Altersgruppen – ein Screeningangebot erhielten. In den übrigen 2 Studien lagen keine Angaben dazu vor, ob und ggf. wann nach Ende der Screeningphase ein Screening aller Frauen erfolgte (CNBSS, HIP); zumindest bei der HIP-Studie kann jedoch vermutet werden, dass sich zunächst kein Screening aller Frauen anschloss.

Nur aus 1 der 8 Studien (CNBSS) lagen Angaben zur Charakterisierung der Studienpopulation vor. In dieser Studie wurden neben demografischen Angaben (z. B. Familienstand, berufliche Tätigkeit) auch ausgewählte prognostische Merkmale dargestellt (z. B. Reproduktionsstatus, Raucherstatus). Eine familiäre Vorbelastung, eines der wesentlichen prognostischen Merkmale, wiesen 9493 Frauen (37,6 %) der Interventionsgruppe bzw. 9652 Frauen (38,3 %) der Vergleichsgruppe auf.

4.3 Übersicht der patientenrelevanten Endpunkte

Aus 8 Studien konnten Daten zu patientenrelevanten Endpunkten extrahiert werden. Tabelle 1 zeigt die Übersicht der verwertbaren Daten zu patientenrelevanten Endpunkten aus den eingeschlossenen Studien. Zu Mastektomien lagen nur zu Fragestellung 1 verwertbare Daten vor. Zu unerwünschten Ereignissen und zur gesundheitsbezogenen Lebensqualität lagen aus keiner Studie verwertbare Daten vor.

Bis auf die UK-Age-Studie, die Frauen zwischen 39 und 41 Jahren einschloss, enthielten alle Studien nach Alterssubgruppen aufgeschlüsselte Ergebnisse, meist zusammengefasst zu 10-Jahres-Altersgruppen (z. B. 40 bis 49 Jahre), jeweils bezogen auf das Alter zum Zeitpunkt der Randomisierung.

Herangezogene Alterssubgruppen für Fragestellung 1

Daten zu genau der relevanten Altersgruppe der 45- bis 49-Jährigen lagen nur vereinzelt vor. Es erschien jedoch hinreichend sicher plausibel, dass die Ergebnisse von etwas jüngeren Frauen auf die eigentliche Zielpopulation anwendbar sind. Es wurden daher Daten aller Alterssubgruppen herangezogen, die den Altersbereich 39 bis 49 Jahre beinhalteten (z. B. 39–49 Jahre, 39–41 Jahre, 45–49 Jahre).

Herangezogene Alterssubgruppen für Fragestellung 2

Aus der Studie Malmö konnten Daten zur Alterssubgruppe der über 69-Jährigen herangezogen werden. Es kann davon ausgegangen werden, dass es sich dabei ausschließlich um 70-jährige Frauen handelte (siehe Abschnitt 4.2). Aus der Studie Swedish Two County wurden Daten zur Alterssubgruppe der 70- bis 74-Jährigen herangezogen. Zusätzlich wurden aus dieser Studie in Bezug auf 1 Endpunkt (Konsequenzen aus falsch-positiven Befunden) Daten der gesamten Studienpopulation (40 bis 74 Jahre) dargestellt, da zu diesem Endpunkt in dieser Fragestellung andernfalls gar keine Daten herangezogen worden wären. Fragen hinsichtlich Übertragbarkeit wurden entsprechend berücksichtigt (siehe Abschnitt 4.5.2.3).

Tabelle 1: Matrix der patientenrelevanten Endpunkte

Studie	Endpunkte						
	Mortalität		Morbidität				LQ
	Gesamtmortalität	Brustkrebsspezifische Mortalität	Mastektomien	Konsequenzen aus falsch-positiven Screeningbefunden	Unerwünschte Ereignisse	Überdiagnosen	Gesundheitsbezogene Lebensqualität (Instrument)
Fragestellung 1							
CNBSS	●	●	●	●	–	●	–
Edinburgh	–	●	–	–	–	–	–
Göteborg	●	●	–	●	–	●	–
HIP	–	●	–	●	–	●	–
Malmö	–	●	–	–	–	●	–
Stockholm	–	●	–	●	–	–	–
Swedish Two County	●	●	–	–	–	●	–
UK Age	●	●	–	●	–	●	–
Fragestellung 2							
Malmö	–	●	–	–	–	–	–
Swedish Two County	●	●	–	●	–	●	–

●: Daten vorhanden und verwertbar
–: Daten nicht vorhanden oder nicht verwertbar

CNBSS: Canadian National Breast Screening Study; HIP: Health Insurance Plan of Greater New York; LQ: gesundheitsbezogene Lebensqualität

4.4 Bewertung des Verzerrungspotenzials der Ergebnisse

Das Verzerrungspotenzial der eingeschlossenen Studien wurde endpunktübergreifend für 1 Studie (UK Age) als niedrig und für die übrigen 7 Studien als hoch eingestuft. Bei den 2 Studien, die für beide Fragestellungen herangezogen wurden, ergaben sich keine Unterschiede in der Einschätzung des Verzerrungspotenzials je nach Fragestellung.

Bei 6 von 7 Studien mit hohem Verzerrungspotenzial war unklar, ob die Zuteilung zur Interventions- und Kontrollgruppe verdeckt erfolgte. An dieser Stelle ist besonders auf die Studie CNBSS hinzuweisen. In dieser Studie erfolgte vor Randomisierung eine klinische Untersuchung der Brust. Im Vergleich zu anderen Studien, die in der Screeninggruppe jeweils weniger Fälle von fortgeschrittenem Brustkrebs als in den Kontrollgruppe diagnostizierten, fällt

auf, dass im Screeningarm der CNBSS deutlich mehr Fälle von fortgeschrittenem Brustkrebs als in der Kontrollgruppe auftraten [20]. In der Literatur wurde daher von einigen Autorinnen und Autoren kritisch hinterfragt, ob Frauen mit tastbarem Tumor in der klinischen Untersuchung möglicherweise eher der Screeninggruppe zugeordnet wurden [20,21]. Träfe dies zu, wären die Ergebnisse der CNBSS in Bezug auf alle Endpunkte möglicherweise stark zuungunsten des Screenings verzerrt und nur unter Einschränkungen heranzuziehen. Die bevorzugte Zuordnung von Frauen mit tastbarem Tumor zur Screeninggruppe würde unter anderem dazu führen, dass in dieser Gruppe die Gesamtzahl der Brustkrebsfälle im Vergleich zur Kontrollgruppe überschätzt wird. In der Folge wäre keine valide Einschätzung hinsichtlich der Überdiagnosen möglich, weil nicht nur durch das Screening mehr Brustkrebsfälle in der Interventionsgruppe identifiziert werden, sondern zusätzliche Fälle aus der willkürlichen Zuordnung von Frauen mit auffälligen Tastbefunden resultierten. Damit würde der Screeningeffekt für den Endpunkt Überdiagnosen potenziell überschätzt. Da zudem die Prognose eines fortgeschrittenen Brustkrebses ungünstiger ist als die eines frühen Brustkrebs, wäre zu erwarten, dass die erhöhte Anzahl an Frauen mit fortgeschrittenem Brustkrebs in der Screeninggruppe zu einer erhöhten Mortalität im Vergleich zur Kontrollgruppe führt. In der Konsequenz käme es zu einer Unterschätzung des Screeningeffekts. Ob eine solche willkürliche Zuordnung und damit einhergehende Verzerrung tatsächlich erfolgte, bleibt jedoch anhand der Angaben unbekannt.

Bei 4 Studien (Edinburgh, Göteborg, HIP, Swedish Two County) war zusätzlich unklar, ob die Randomisierungssequenz adäquat erzeugt wurde. Bei der Studie Edinburgh führte die Clusterrandomisierung zu Gruppenunterschieden im sozialen Status [22] und damit zu offenbar ungleichen Gruppen, was sich auch in einer deutlich unterschiedlichen Gesamtmortalität zeigte (bezogen auf die gesamte Studienpopulation der 45- bis 64-Jährigen, bereits aus diesem Grund im Bericht nicht dargestellt). Eine relevante Verzerrung der Ergebnisse zur brustkrebsspezifischen Mortalität war daher ebenfalls möglich; dies wurde im Rahmen von Sensitivitätsanalysen berücksichtigt.

Bei 1 weiteren Studie – Stockholm – war die Erzeugung der Randomisierungssequenz nicht adäquat, da die Zuteilung zu den Gruppen anhand des Geburtsdatums erfolgte. Die Stockholm-Studie wurde daher als quasirandomisiert aufgefasst.

Das endpunktübergreifend hohe Verzerrungspotenzial schlägt sich direkt auf das endpunktspezifische Verzerrungspotenzial nieder, sodass alle in diesen 7 Studien berichteten Ergebnisse ein hohes endpunktspezifisches Verzerrungspotenzial aufweisen.

Die Kriterien des endpunktspezifischen Verzerrungspotenzials wurden somit nur für die Studie UK Age bewertet. Das endpunktspezifische Verzerrungspotenzial für die Ergebnisse zu Gesamtmortalität, Brustkrebsmortalität und Überdiagnosen wurde als niedrig bewertet. Für die Ergebnisse zu Konsequenzen aus falsch-positiven Screeningbefunden (Biopsien mit benignem Befund) wurde es hingegen als hoch bewertet, da die Verblindung der Endpunkterheber und

die adäquate Umsetzung des Intention to treat (ITT)-Prinzips wegen unzureichender Beschreibung unklar waren.

4.5 Ergebnisse zu patientenrelevanten Endpunkten

4.5.1 Fragestellung 1: Altersgruppe 45 bis 49 Jahre

4.5.1.1 Ergebnisse zur Mortalität

4.5.1.1.1 Ergebnisse zur Gesamtmortalität

Zur Gesamtmortalität lagen verwertbare, nach Alterssubgruppen getrennte Daten vor aus 1 Studie mit hoher (UK Age) und 3 Studien mit mäßiger qualitativer Ergebnissicherheit (CNBSS, Göteborg, Swedish Two County). Aus 3 Studien (CNBSS, Göteborg, Swedish Two County) wurden Daten von 10-Jahres-Altersgruppen (40–49 bzw. 39–49 Jahre) herangezogen, aus der 4. Studie (UK Age) Daten der gesamten Studienpopulation, Frauen zwischen 39 und 41 Jahren. Die Ergebnisse bezogen sich auf Beobachtungsdauern zwischen 7,9 und 22,8 Jahren seit Randomisierung, wobei in 2 Studien (CNBSS, UK Age) Daten zu mehreren Zeitpunkten verfügbar waren. In den gepoolten Schätzer gingen Daten von 7,9 bis 13 Jahren Beobachtungsdauer ein.

Weder in der Studie mit hoher qualitativer Ergebnissicherheit zu 3 verschiedenen Zeitpunkten noch im gepoolten Schätzer aus allen 4 Studien zeigte sich ein statistisch signifikanter Effekt (Inzidenzdichtequotient [IDR] 0,98; 95 %-Konfidenzintervall [95 %-KI]: [0,90; 1,08]; p = 0,602).

Für die Merkmale Screeningintervall und Screeningphasendauer zeigte sich keine Effektmodifikation auf die Gesamtmortalität. Zur Untersuchung einer Effektmodifikation durch die Merkmale Alter und Brustdichte lagen keine Daten vor.

Somit ergibt sich für die Gesamtmortalität kein Anhaltspunkt für einen Nutzen oder Schaden des Mammografie-Screenings.

4.5.1.1.2 Ergebnisse zur brustkrebsspezifischen Mortalität

Zur brustkrebsspezifischen Mortalität lagen verwertbare Daten aus allen 8 Studien vor, davon aus 1 Studie (UK Age) mit hoher, aus 6 mit mäßiger (CNBSS, Edinburgh, Göteborg, HIP, Malmö, Swedish Two County) und aus 1 mit geringer qualitativer Ergebnissicherheit (Stockholm). Aus 3 Studien (CNBSS, Stockholm, Swedish Two County) wurden Daten allein von 10-Jahres-Altersgruppen (40–49 bzw. 39–49 Jahre) herangezogen, aus 1 Studie (Edinburgh) alle in der 5-Jahres-Altersgruppe (45–49 Jahre), aus 2 Studien sowohl der 10- als auch der 5-Jahres-Altersgruppe (Göteborg, HIP), aus 1 Studie von 44- bis 49-Jährigen (Malmö) und aus 1 Studie die gesamte Studienpopulation der 39- bis 41-Jährigen (UK Age). Aus der Göteborg- und der HIP-Studie wurden primär die Daten der 5-Jahres-Altersgruppe herangezogen, während die Daten der 10-Jahres-Altersgruppe, die eigentlich weniger gut zur Berichtsfragestellung passen, dazu verwendet wurden, um in einer Sensitivitätsanalyse zu

prüfen, welchen Einfluss ein breiteres Fassen der Altersgruppengrenze insgesamt hat. Dargestellt wurden Ergebnisse von Beobachtungsdauern zwischen 4 und 24 Jahren seit Randomisierung.

Es wurden Metaanalysen zu 4 verschiedenen Zeitpunkten durchgeführt und darin jeweils Ergebnisse zu ca. 5, 10, 15 und 20 Jahren nach Randomisierung zusammengefasst. Zum Zeitpunkt ca. 10 Jahre konnten Daten aus allen 8 Studien einbezogen werden, zu den anderen Zeitpunkten jeweils nur aus 4 bzw. 5 Studien.

Weder in den Ergebnissen der 1 Studie mit hoher Ergebnissicherheit (UK Age), die in die Metaanalysen zu den 4 Zeitpunkten eingingen, noch im jeweiligen gepoolten Schätzer ergab sich ein statistisch signifikanter Unterschied. In der Metaanalyse zum Zeitpunkt ca. 10 Jahre – der Zeitpunkt, zu dem die meisten Daten vorlagen – wird die statistische Signifikanz knapp verfehlt (IDR 0,85; 95 %-KI: [0,73; 1,002]; p = 0,052).

In der Studie UK Age lagen Daten zu 2 verschiedenen Messzeitpunkten vor (9 Jahre und 10,7 Jahre), die beide zum Zeitpunkt ca. 10 Jahre passten und von denen die Daten zum späteren Zeitpunkt (10,7 Jahre) für die Hauptanalyse ausgewählt wurden. Die Daten zum Zeitpunkt 9 Jahre nach Randomisierung (ebenfalls hohe Ergebnissicherheit) zeigten einen statistisch signifikanten Effekt (IDR 0,75; 95 %-KI: [0,58; 0,96]). Im Rahmen einer Sensitivitätsanalyse wurde überprüft, inwieweit es einen Unterschied macht, welcher der beiden zu dieser Analyse passenden Zeitpunkte der UK-Age-Studie in die Metaanalyse eingeht. Diese Sensitivitätsanalyse ergab – im Unterschied zur Hauptanalyse – ein statistisch signifikantes Ergebnis (IDR 0,83; 95 %-KI: [0,70; 0,99]; p = 0,041). Aus dem statistisch signifikanten Ergebnis der Einzelstudie mit hoher qualitativer Ergebnissicherheit ergibt sich bereits ein Hinweis auf einen Effekt. Der gepoolte Schätzer aus allen Studien bestätigt diesen Effekt. Daher wird in der Sensitivitätsanalyse ein Beleg für einen Effekt zugunsten des Mammografie-Screenings abgeleitet.

Eine mögliche Erklärung dafür, dass sich in der Studie UK Age ein Effekt zu 9 Jahren zeigte, jedoch nicht mehr zu 10,7 Jahren, sehen die Autorinnen und Autoren der Studie darin, dass sich zwar für Frauen mit Tumoren im Stadium 1 und 2 sowohl kurzfristig als auch langfristig ein Überlebensvorteil zeigte, für Frauen mit Tumoren im Stadium 3 jedoch nur innerhalb der ersten 10 Jahre. Daraus schlossen sie, dass das Screening bei aggressiveren Tumoren zwar das Überleben verlängert, den Tod durch Brustkrebs jedoch nicht verhindern kann und so der Effekt wieder verdünnt wird. Diese Erklärung erscheint plausibel, und eine Verlängerung des Überlebens ist ebenfalls als Vorteil zu werten.

In der gemeinsamen Betrachtung der Hauptanalyse zu 10 Jahren (die nur knapp die statistische Signifikanz verfehlt) und der Sensitivitätsanalyse mit den 9-Jahres-Daten der UK-Age-Studie (Beleg für einen Effekt) ergibt sich daher ein Anhaltspunkt für einen Nutzen des Mammografie-Screenings.

Dass sich zu den anderen Zeitpunkten (5, 15 und 20 Jahre) kein statistisch signifikanter Effekt zeigte, steht nicht im Widerspruch zu den Ergebnissen zu 10 Jahren. Denn zu den anderen Zeitpunkten lagen Daten aus nur 4 bzw. 5 Studien vor, mit entsprechend geringerer Präzision des gepoolten Effektschätzers. Außerdem ist ein Effekt am ehesten zu den Zeitpunkten ca. 10 und 15 Jahre erwartbar, ausgehend davon, dass es mindestens 9 bis 12 Jahre dauert, bis sich ein Effekt des Screenings abzeichnen kann [23], sowie davon, dass ein ggf. vorhandener Effekt nach Ende der Screeningphase mit zunehmender Zeit erwartbar wieder verdünnt wird.

Weitere Sensitivitätsanalysen widersprachen diesen Ergebnissen nicht.

Für die Merkmale Screeningintervall und Screeningphasendauer zeigte sich keine Effektmodifikation. Zur Untersuchung einer Effektmodifikation durch die Merkmale Alter und Brustdichte lagen keine Daten vor.

Insgesamt ergibt sich somit aus den Ergebnissen zur brustkrebsspezifischen Mortalität ein Anhaltspunkt für einen Nutzen des Mammografie-Screenings.

4.5.1.1.3 Zusammenfassende Betrachtung für die Mortalität

Für die Gesamtmortalität zeigte sich kein Anhaltspunkt für einen Nutzen oder Schaden des Mammografie-Screenings, für die brustkrebsspezifische Mortalität hingegen ein Anhaltspunkt für einen Nutzen. Da andere Todesursachen als Brustkrebs in dieser Altersgruppe viel häufiger sind – lediglich 12,7 % der Todesfälle bei Frauen in einem Alter von 45 bis 54 Jahren sind zurückzuführen auf Brustkrebs (ICD-10-Code C50) [24] –, spricht das Ergebnis zur Gesamtmortalität nicht gegen einen Nutzen bezüglich des Mammografie-Screenings. Es ergibt sich daher für den Endpunkt Mortalität ein Anhaltspunkt für einen Nutzen des Mammografie-Screenings.

4.5.1.2 Ergebnisse zu Mastektomien

Das Mammografie-Screening zielt darauf ab, die Mortalität durch Brustkrebs zu verringern, indem Brustkrebsfälle bereits in einem früheren Stadium entdeckt und behandelt werden. Darüber hinaus ist mit dem Behandlungsbeginn in einem früheren Krankheitsstadium die Erwartung verbunden, dass die Therapie weniger invasiv ausfallen kann, ohne damit den Behandlungserfolg zu gefährden. Ein Vorteil für Brustkrebs-Patientinnen kann sich daher daraus ergeben, dass häufiger brusterhaltend operiert werden kann, ohne dass sich dies nachteilig auf die Mortalität auswirkt. Ein geringerer Anteil Mastektomien an allen Brustkrebs-Operationen in der Screeninggruppe ist jedoch nur dann als Vorteil zu werten, wenn auch die absolute Rate der Mastektomien bezogen auf alle randomisierten Frauen in der Screeninggruppe geringer ist als in der Kontrollgruppe.

Zu Brustkrebs-Operationen lagen verwertbare Daten aus 1 Studie mit mäßiger qualitativer Ergebnissicherheit vor (CNBSS). Verfügbar waren Ergebnisse von 40- bis 49-Jährigen. In der Studie wurden brusterhaltende Operationen (OPs) und Mastektomien berichtet. Herangezogen wurden die Ergebnisse nach 7 Jahren. Lediglich ergänzend dargestellt werden die Ergebnisse

nach der 1. Screeningrunde. Denn zu diesem Zeitpunkt ist eine relevant höhere Anzahl Mastektomien erwartbar, da durch das Screening auch mehr Brustkrebsfälle entdeckt werden.

Wie zu erwarten, zeigte sich nach der 1. Screeningrunde sowohl hinsichtlich OPs gesamt (Odds Ratio [OR] 1,58; 95 %-KI: [1,15; 2,18]; p = 0,004) als auch hinsichtlich Mastektomien (OR 1,67; 95 %-KI: [1,10; 2,52]; p = 0,014) ein statistisch signifikanter Unterschied zuungunsten des Screenings. Der Anteil der Mastektomien an allen Brustkrebs-Operationen war dabei zwischen den Gruppen nicht statistisch signifikant unterschiedlich (61,2 % vs. 58,1 %; OR 1,14; 95 %-KI: [0,60; 2,18]; p = 0,709).

Nach 7 Jahren war zwar die Anzahl der OPs insgesamt in der Screeninggruppe statistisch signifikant höher als in der Kontrollgruppe (OR 1,33; 95 %-KI: [1,15; 1,54]; p < 0,001). Bei den Mastektomien zeigte sich jedoch kein statistisch signifikanter Unterschied zwischen den Gruppen (OR 1,17; 95 %-KI: [0,94; 1,45]; p = 0,158). Auch der Anteil der Mastektomien an allen Brustkrebs-Operationen war zwischen den Gruppen nicht statistisch signifikant unterschiedlich (44,1 % vs. 50,2 %; OR 0,78; 95 %-KI: [0,58; 1,05]; p = 0,116).

Die höhere Anzahl OPs in der Screeninggruppe nach 7 Jahren ging also nicht einher mit einer reduzierten absoluten Rate der Mastektomien, also einer erkennbar geringeren Invasivität der OPs. Basierend darauf ergibt sich kein Anhaltspunkt für einen Nutzen oder Schaden des Mammografie-Screenings.

4.5.1.3 Ergebnisse zu Konsequenzen aus falsch-positiven Screeningbefunden

Für den Endpunkt Konsequenzen aus falsch-positiven Screeningbefunden wurden Angaben für Screeningteilnehmerinnen herangezogen, die einen positiven Screeningbefund hatten und für die sich in der nachfolgenden invasiven Abklärungsdiagnostik der Verdacht auf Brustkrebs nicht bestätigt hat. Für den Endpunkt wurden sowohl Daten zur rein diagnostisch interventionellen Abklärung herangezogen als auch Daten zu chirurgisch therapeutischen Eingriffen, wenn die Behandlung und Diagnostik von Brustgewebe unklarer Dignität nicht eindeutig voneinander zu trennen waren. Verfügbar waren Ergebnisse zu Biopsien mit benignem Befund (CNBSS, HIP, Malmö, Stockholm, UK Age) sowie OPs mit benignem Befund (Göteborg). Daten zu diesem Endpunkt beziehen sich definitionsgemäß nur auf die gescreenten Frauen. Ein Gruppenvergleich entfällt damit, auch wenn diagnostische und therapeutische Eingriffe mit benignem Befund auch in der Kontrollgruppe auftreten können.

Zu Konsequenzen aus falsch-positiven Screeningbefunden lagen verwertbare Daten aus 5 Studien vor, davon aus 4 Studien Ergebnisse mit mäßiger (CNBSS, HIP, Malmö, UK Age) und aus 1 mit geringer qualitativer Ergebnissicherheit (Stockholm). Es wurden Daten von 10-Jahres-Altersgruppen (40–49 Jahre; CNBSS, HIP), 5-Jahres-Altersgruppen (45–49 Jahre; Malmö, Stockholm) bzw. der gesamten Studienpopulation der 39- bis 41-Jährigen (UK Age) herangezogen. Die Darstellung der Ereignisse erfolgte in den Studien unterschiedlich: In 1 Studie (HIP) waren lediglich Daten zum Anteil der Biopsien mit benignem Befund an allen Biopsien verfügbar, ohne Angabe des Bezugszeitraums. Die Ergebnisse der übrigen 4 Studien

bezogen sich entweder auf die Anzahl je Screeningjahr (bis maximal 5 Jahre; CNBSS) oder je Screeningrunde (bis maximal 5 Runden; Malmö, Stockholm, UK Age) oder kumulativ auf den gesamten Beobachtungszeitraum (10 bis 15,5 Jahre; Malmö). Daher wurde für diesen Endpunkt keine zusammenfassende Gesamtschätzung angegeben, sondern nachfolgend eine Spanne der Effektschätzer aus den einzelnen Studien [Minimum; Maximum].

Der Anteil der Frauen in der 1. Screeningrunde bzw. im 1. Jahr, bei denen eine Biopsie nur durch das falsch-positive Ergebnis im Screening notwendig geworden war, reichte von 0,16 % bis 3,4 %. In den nachfolgenden Screeningrunden bzw. Jahren lagen die Anteile je Runde bzw. Jahr mit 0,07 % bis 1,8 % etwas niedriger. Der Anteil der Biopsien mit benignem Befund an allen Biopsien betrug 90 % (HIP-Studie).

Der Anteil der Screeningteilnehmerinnen, bei denen eine Operation nur durch das falsch-positive Ergebnis im Screening notwendig wurde, lag in der 1. Screeningrunde bei 0,27 %, in den nachfolgenden Runden zwischen 0,03 % und 0,13 %. Der Anteil von Operationen mit benignem Befund an allen Brustkrebs-Operationen lag dabei zwischen 16 % (4/25) und 65 % (32/49). Es ist allerdings zu erwarten, dass mit aktueller Diagnostik wesentlich weniger Operationen mit benignem Befund vorkommen, sodass nicht von einer Übertragbarkeit auf die heutige Situation ausgegangen werden kann.

Somit ergibt sich für Konsequenzen aus falsch-positiven Befunden – basierend auf den Ergebnissen zu Biopsien mit benignem Befund – insgesamt ein Hinweis auf einen Schaden des Mammografie-Screenings.

4.5.1.4 Ergebnisse zu unerwünschten Ereignissen

Zu diesem Endpunkt lagen keine Daten vor.

4.5.1.5 Ergebnisse zu Überdiagnosen

Zu den Überdiagnosen (basierend auf Brustkrebsinzidenzen) lagen Daten aus 6 Studien vor, davon aus 1 Studie mit hoher (UK Age) und aus 5 mit mäßiger qualitativer Ergebnissicherheit (CNBSS, Göteborg, HIP, Malmö, Swedish Two County). Aus 5 Studien lagen Daten von 10-Jahres-Altersgruppen vor (CNBSS, HIP, Swedish Two County: 40–49 Jahre; Göteborg: 39–49 Jahre; Malmö: 45–54 Jahre). In 2 Studien lagen zusätzlich Daten zu 5-Jahres-Altersgruppen vor (Göteborg, HIP: 45–49 Jahre). Aus 1 Studie (UK Age) lagen Daten zu 39- bis 41-Jährigen vor.

Die für diese Bewertung zugrunde gelegte Methodik zur Berechnung von Überdiagnosen war nur in 2 Studien sinnvoll anwendbar (CNBSS, HIP). Denn diese Berechnung beruht auf einem Vergleich der Brustkrebsinzidenz der gescreenten Frauen mit einer Kontrollgruppe ohne Screening, in der also nur symptomatische Brustkrebsfälle erfasst wurden. In den anderen 4 Studien erfolgte jeweils ein Abschlussscreening in der Kontrollgruppe bzw. ein Screening aller Frauen nach Ende der Screeningphase. Sobald auch eine Kontrollgruppe gescreent wird, werden in dieser Gruppe prä- und asymptomatische Brustkrebsfälle mit enthalten sein. Daher

lässt die hier angewendete Methodik in solchen Fällen keine ausreichend validen Ergebnisse der Überdiagnosen erwarten.

Überdiagnosen bezogen auf die zum Screening eingeladenen Frauen

In der Studie CNBSS betrug das Risiko für eine Überdiagnose bezogen auf die zum Screening eingeladenen Frauen 0,41 %. Mit 0,04 % (Altersgruppe 40–49 Jahre) bzw. 0,01 % (Altersgruppe 45–49 Jahre) lagen die Werte in der Studie HIP deutlich niedriger.

Wird die bei den anderen 4 Studien eingeschränkte Validität der Berechnungsmethode außer Acht gelassen, würde sich für die Studie Malmö ein Risiko von 0,49 % (nach dem Screening der Kontrollen) und für die Studie Swedish Two County ein Risiko von 0,25 % (nach dem Screening der Kontrollen) ergeben. Diese beiden Werte liegen in der Größenordnung zwischen denen der Studien CNBSS und HIP.

Für die Studien Göteborg und UK Age war auch eine näherungsweise Berechnung des Risikos für Überdiagnosen wie bei den Studien Malmö und Swedish Two County nicht möglich, da zu den berichteten Auswertungszeitpunkten die Anzahl der Brustkrebsdiagnosen in der Kontrollgruppe höher war als in der Screeninggruppe. Allerdings haben die Autorinnen und Autoren der UK-Age-Studie selbst einen Wert von 0,2 % bezogen auf die zum Screening eingeladenen Frauen berichtet [25]. In die Berechnung flossen dabei die Testsensitivität und die Sojourn Time (Dauer der präklinischen Detektierbarkeit des asymptomatischen Brustkrebses) ein, die mittels der Intervallkarzinome geschätzt wurde. Der berichtete Wert von 0,2 % liegt zwischen den in diesem Bericht berechneten Werten der anderen eingeschlossenen Studien.

Anzumerken ist, dass nicht ausgeschlossen werden kann, dass der für die Studie CNBSS ermittelte Wert von 0,41 % überschätzt ist. Wie in Abschnitt 4.2 und 4.4 beschrieben, erfolgte bei allen Frauen vor Randomisierung eine klinische Untersuchung der Brust. Im Verlauf der Studie wurde eine auffällige Häufung von fortgeschrittenen Brustkrebsfällen in der Screeninggruppe beobachtet, die vermutlich bereits bei der Eingangsuntersuchung tastbar gewesen waren. Es wurde daher von einigen Autorinnen und Autoren der Verdacht geäußert, dass Frauen mit auffälligem oder unklarem Befund möglicherweise bevorzugt der Screeninggruppe zugeordnet wurden [21]. Träfe dies zu, wäre hiervon sowohl in der Screeninggruppe als auch in der Kontrollgruppe die Anzahl von Brustkrebsdiagnosen insgesamt beeinflusst, was zu einer Überschätzung des Risikos für eine Überdiagnose führen würde.

In der Gesamtschau der Ergebnisse aller Studien ergibt sich hieraus aber keine Einschränkung, da die Werte für das Überdiagnoserisiko bezogen auf die zum Screening eingeladenen Frauen in der Altersgruppe zwischen 45 und 49 Jahren in allen Studien mit 0,01 % bis 0,41 % in eine Richtung zeigen. Auf eine metaanalytische Zusammenfassung der Ergebnisse wurde verzichtet, da die Anteile deutlich unterschiedlich sind.

Überdiagnosen bezogen auf die Frauen mit Brustkrebsdiagnose während der Screeningphase

In der Studie CNBSS lag das Risiko für eine Überdiagnose bei Frauen, die eine Brustkrebsdiagnose während der Screeningphase erhalten hatten, bei 31,61 %. Mit 8,06 % (Altersgruppe 40–49 Jahre) bzw. 2,63 % (Altersgruppe 45–49 Jahre) lagen die Werte in der Studie HIP deutlich niedriger.

Wird die bei den anderen 4 Studien eingeschränkte Validität der Berechnungsmethode außer Acht gelassen, würde sich für die Studie Malmö ein Risiko von 12,98 % (nach dem Screening der Kontrollen) und für die Studie Swedish Two County von 21,00 % (nach dem Screening der Kontrollen) ergeben. Diese beiden Werte liegen in der Größenordnung zwischen denen der Studien CNBSS und HIP.

Für die Studien Göteborg und UK Age war auch eine näherungsweise Berechnung des Risikos für Überdiagnosen bei Frauen, die eine Brustkrebsdiagnose während der Screeningphase erhalten hatten, nicht möglich, da zu den berichteten Auswertungszeitpunkten die Anzahl der Brustkrebsdiagnosen in der Kontrollgruppe höher war als in der Screeninggruppe. Allerdings nennen die Autorinnen und Autoren der UK-Age-Studie selbst einen Wert von 8,5 % [25]. In die Berechnung flossen dabei die Testsensitivität und die Sojourn Time ein, die mittels der Intervallkarzinome geschätzt wurde. Der berichtete Wert von 8,5 % liegt zwischen den in diesem Bericht berechneten Werten der anderen eingeschlossenen Studien.

Wie bereits im oberen Abschnitt angemerkt, kann nicht ausgeschlossen werden, dass der berechnete Wert der CNBSS das Risiko einer Überdiagnose überschätzt.

In der Gesamtschau der Ergebnisse aller Studien ergibt sich hieraus aber keine Einschränkung, da die Werte für das Überdiagnoserisiko bei Frauen in der Altersgruppe zwischen 45 und 49 Jahren, die eine Brustkrebsdiagnose während der Screeningphase erhalten hatten, in den Studien mit 2,63 % und 31,61 % in eine Richtung zeigen. Auf eine metaanalytische Zusammenfassung der Ergebnisse wurde verzichtet, da die Anteile deutlich unterschiedlich sind.

Zusammenfassende Betrachtung für die Überdiagnosen

Nach der invasiven Diagnostik zur Diagnosesicherung erfolgt bei nahezu allen Brustkrebsdiagnosen eine medikamentöse, operative und / oder Strahlentherapie. Sowohl die weitergeführte Diagnostik als auch die anschließende Behandlung bergen ein Risiko von unerwünschten Nebenwirkungen und Komplikationen. Da eine Überdiagnose die Diagnose einer Erkrankung darstellt, die sich ohne ein Screening nie bemerkbar gemacht und keine Beschwerden ausgelöst hätte, sind die nachfolgenden Interventionen samt ihrer Folgen überflüssig.

Insgesamt ergibt sich somit ein Anhaltspunkt für einen Schaden des Mammografie-Screenings im Vergleich zu keinem Screening bezüglich des Endpunkts Überdiagnosen bei Frauen zwischen 45 und 49 Jahren.

4.5.1.6 Ergebnisse zur gesundheitsbezogenen Lebensqualität

Zu diesem Endpunkt lagen keine Daten vor.

4.5.2 Fragestellung 2: Altersgruppe ab 70 Jahre

4.5.2.1 Ergebnisse zur Mortalität

4.5.2.1.1 Ergebnisse zur Gesamtmortalität

Zur Gesamtmortalität lagen verwertbare Daten mit mäßiger qualitativer Ergebnissicherheit aus 1 Studie (Swedish Two County) zum Zeitpunkt 7,9 Jahre nach Randomisierung vor.

Es zeigte sich kein statistisch signifikanter Effekt (RR 0,98; 95 %-KI: [0,93; 1,03]; p = 0,384).

Zur Untersuchung von Effektmodifikatoren lagen keine geeigneten Daten vor.

Somit ergibt sich für die Gesamtmortalität kein Anhaltspunkt für einen Nutzen oder Schaden des Mammografie-Screenings bei Frauen ab 70 Jahren.

4.5.2.1.2 Ergebnisse zur brustkrebsspezifischen Mortalität

Zur brustkrebsspezifischen Mortalität lagen verwertbare Daten mit mäßiger qualitativer Ergebnissicherheit aus beiden Studien vor. Aus der Studie Malmö waren Daten verfügbar zu 13,6 Jahren nach Randomisierung, aus der Studie Swedish Two County zu 7,9, 10,8 und 20 Jahren. Die Ergebnisse zu 13,6 und 10,8 Jahren wurden metaanalytisch zusammengefasst.

Weder im gepoolten Schätzer aus beiden Studien (IDR 0,91; 95 %-KI: [0,61; 1,36]; p = 0,655). noch zu den weiteren 2 Zeitpunkten in der Studie Swedish Two County ergab sich ein statistisch signifikanter Unterschied.

Zur Untersuchung von Effektmodifikatoren lagen keine geeigneten Daten vor.

Somit ergibt sich für die brustkrebsspezifische Mortalität kein Anhaltspunkt für einen Nutzen oder Schaden des Mammografie-Screenings bei Frauen ab 70 Jahren.

4.5.2.1.3 Zusammenfassende Betrachtung für die Mortalität

Weder für die Gesamtmortalität noch für die brustkrebsspezifische Mortalität ergab sich ein Anhaltspunkt für einen Nutzen oder Schaden des Mammografie-Screenings bei Frauen ab 70 Jahren, daher auch nicht insgesamt für die Mortalität.

4.5.2.2 Ergebnisse zu Mastektomien

Zu diesem Endpunkt lagen keine verwertbaren Daten vor.

4.5.2.3 Ergebnisse zu Konsequenzen aus falsch-positiven Screeningbefunden

Für den Endpunkt Konsequenzen aus falsch-positiven Screeningbefunden wurden Angaben für Screeningteilnehmerinnen herangezogen, die einen positiven Screeningbefund hatten und für die sich in der nachfolgenden invasiven Abklärungsdiagnostik der Verdacht auf Brustkrebs nicht bestätigt hat. Für den Endpunkt waren sowohl Daten zur rein diagnostisch interventionellen Abklärung als auch zu chirurgisch therapeutischen Eingriffen relevant, wenn die Behandlung und Diagnostik von Brustgewebe unklarer Dignität nicht eindeutig voneinander zu trennen waren. Daten zu diesem Endpunkt beziehen sich definitionsgemäß nur auf die gescreenten Frauen. Ein Gruppenvergleich entfällt damit, auch wenn diagnostische und therapeutische Eingriffe mit benignem Befund auch in der Kontrollgruppe auftreten können.

Die einzigen verfügbaren Daten zu diesem Endpunkt für diese Fragestellung waren solche zu chirurgischen Biopsien mit benignem Befund und bezogen sich auf die gesamte Studienpopulation der 40- bis 74-Jährigen aus der Studie Swedish Two County.

Der Anteil der Frauen in der 1. Screeningrunde, bei denen eine chirurgische Biopsie nur durch das falsch-positive Ergebnis im Screening notwendig geworden war, wurde mit 0,6 % angegeben. In den nachfolgenden Screeningrunden wurde ein Anteil von im Mittel 0,1 % berichtet. Die Bezugsgröße blieb dabei unklar.

Zwar liegen keine Daten aufgeschlüsselt nach Alterssubgruppen vor, aber es gibt keinen Grund anzunehmen, dass in der Altersgruppe der 70- bis 74-Jährigen keine chirurgischen Biopsien mit benignem Befund vorkommen, auch wenn sich möglicherweise die Häufigkeit zwischen den Altersgruppen unterscheidet. Dies gilt auch für andere diagnostische und therapeutische Eingriffe, die unter diesen Endpunkt fallen, für die jedoch nur Daten zu anderen Alterssubgruppen vorliegen (siehe Abschnitt 4.5.1.3).

Somit ergibt sich auf Basis der Ergebnisse und der aufgeführten Überlegungen für Konsequenzen aus falsch-positiven Befunden ein Anhaltspunkt für einen Schaden des Mammografie-Screenings bei Frauen ab 70 Jahren.

4.5.2.4 Ergebnisse zu unerwünschten Ereignissen

Zu diesem Endpunkt lagen keine Daten vor.

4.5.2.5 Ergebnisse zu Überdiagnosen

Zu den Überdiagnosen (basierend auf Brustkrebsinzidenzen) lagen Daten aus nur 1 Studie mit mäßiger qualitativer Ergebnissicherheit (Swedish Two County) vor. Die Altersgruppe umfasst Frauen von 70 bis 74 Jahren.

Die für diese Bewertung zugrunde gelegte Methodik zur Berechnung von Überdiagnosen war jedoch für diese Studie nicht anwendbar. Denn diese Berechnung beruht auf einem Vergleich der Brustkrebsinzidenz der gescreenten Frauen mit einer Kontrollgruppe *ohne* Screening, in der also nur symptomatische Brustkrebsfälle erfasst wurden. In der Swedish-Two-County-Studie

erfolgte aber ein Abschlussscreening in der Kontrollgruppe. Sobald auch eine Kontrollgruppe gescreent wird, werden in dieser Gruppe prä- und asymptomatische Brustkrebsfälle mit enthalten sein, wodurch die Gesamtzahl der Brustkrebsfälle in der Kontrollgruppe höher ausfällt, der wahre Effekt also unterschätzt wird. Trotzdem wurden diese Daten der Studie Swedish Two County für die näherungsweise durchgeführte Analyse herangezogen, um einen Eindruck von der Größenordnung des Effekts zu erhalten.

Überdiagnosen bezogen auf die zum Screening eingeladenen Frauen

Wird bei der Swedish-Two-County-Studie die eingeschränkte Validität der Berechnungsmethode außer Acht gelassen, ergibt sich ein Risiko für Überdiagnosen bezogen auf alle zum Screening eingeladenen Frauen von 0,48 % (nach dem Screening der Kontrollen).

Überdiagnosen bezogen auf die Frauen mit Brustkrebsdiagnose während der Screeningphase

Wird bei der Swedish-Two-County-Studie die eingeschränkte Validität der Berechnungsmethode außer Acht gelassen, ergibt sich ein Risiko für Überdiagnosen bezogen auf Frauen, die während der Screeningphase eine Brustkrebsdiagnose erhalten haben, von 19,45 % (nach dem Screening der Kontrollen).

Zusammenfassende Betrachtung für die Überdiagnosen

Nach der invasiven Diagnostik zur Diagnosesicherung erfolgt bei nahezu allen Brustkrebsdiagnosen eine medikamentöse, operative und / oder Strahlentherapie. Sowohl die weitergeführte Diagnostik als auch die anschließende Behandlung bergen ein Risiko von unerwünschten Nebenwirkungen und Komplikationen. Da eine Überdiagnose die Diagnose einer Erkrankung darstellt, die sich ohne ein Screening nie bemerkbar gemacht und keine Beschwerden ausgelöst hätte, sind die nachfolgenden Interventionen samt ihrer Folgen überflüssig.

Auch wenn die Sicherheit der Ergebnisse der Studie Swedish Two County hinsichtlich der Berechnung der Überdiagnosen aufgrund der Verwendung der Daten der Kontrollgruppe nach dem Abschlussscreening eingeschränkt ist und somit von einer Unterschätzung des Effekts ausgegangen werden muss, zeigen die Ergebnisse, dass beim Brustkrebsscreening in dieser Altersgruppe Überdiagnosen auftreten.

Es ergibt sich somit ein Anhaltspunkt für einen Schaden des Mammografie-Screenings im Vergleich zu keinem Screening bezüglich des Endpunkts Überdiagnosen bei Frauen ab 70 Jahren.

4.5.2.6 Ergebnisse zur gesundheitsbezogenen Lebensqualität

Zu diesem Endpunkt lagen keine Daten vor.

4.6 Zusammenfassende Bewertung der Ergebnisse

Landkarte der Beleglage

Die folgende Tabelle 2 zeigt die Landkarte der Beleglage in Bezug auf die patientenrelevanten Endpunkte.

Tabelle 2: Landkarte der Beleglage in Bezug auf die patientenrelevanten Endpunkte

	Mortalität	Morbidität				LQ
	Gesamtmortalität und brustkrebsspezifische Mortalität	Mastektomien	Konsequenzen aus falsch-positiven Screeningbefunden	Unerwünschte Ereignisse	Überdiagnosen	Gesundheitsbezogene Lebensqualität
Fragestellung 1: 45–49 Jahre	⇗[a]	⇔	⇓	–	⇘	–
Fragestellung 2: ≥ 70 Jahre	⇔	–	⇘	–	⇘	–

⇗: Anhaltspunkt für einen Nutzen
⇔: kein Anhaltspunkt, Hinweis oder Beleg, homogenes Ergebnis
⇘: Anhaltspunkt für einen Schaden
⇓: Hinweis auf einen Schaden
–: keine Daten berichtet

a. Basierend auf einem Anhaltspunkt für einen Nutzen bei der brustkrebsspezifischen Mortalität. Die Ergebnisse zur Gesamtmortalität stehen nicht im Widerspruch dazu, auch wenn sich ein Einfluss auf die Gesamtmortalität nicht belegen ließ.

LQ: gesundheitsbezogene Lebensqualität

Bewertung des Umfangs unpublizierter Daten

Die vorliegenden Informationen deuten nicht auf einen Publication Bias hin.

Es wurde eine in Großbritannien laufende clusterrandomisierte Studie mit 4,4 Millionen im Alter von 47 bis 49 bzw. von 71 bis 73 Jahren rekrutierten Frauen identifiziert (AgeX [26]). Diese Studie, deren Rekrutierung im Frühjahr 2020 beendet wurde, lässt Mitte der 2020er Jahre weitere aussagekräftige Erkenntnisse zu beiden Fragestellungen des Berichts erwarten. Sie ist die bisher größte randomisierte Studie zum Nutzen eines Screenings im Vergleich zu keinem Screening und neben der UK-Age-Studie die einzige bekannte Studie, die sich mit einer Ausweitung von Altersgrenzen eines bereits etablierten Brustkrebs-Screeningprogramms befasst. Die Studie trifft auch hinsichtlich der betrachteten Altersgruppen die Fragestellung des vorliegenden Berichts besser als die bisher eingeschlossenen Studien. Das in der Studie untersuchte Screeningintervall von 3 Jahren weicht von dem in Deutschland derzeit üblichen ab. Dies wird bei der Interpretation der Ergebnisse und ihrer Übertragung auf den deutschen

Versorgungskontext zu berücksichtigen sein, lässt aber keine grundsätzlichen Probleme erwarten.

Nutzen-Schaden-Abwägung

Jedes Screening schadet durch falsche Screeningbefunde und Überdiagnosen. Ein Screening ist nur gerechtfertigt, wenn der Schaden durch den Nutzen mehr als aufgewogen wird. Bei der Nutzen-Schaden-Abwägung ist darüber hinaus zu berücksichtigen, dass die Ergebnisse für die verschiedenen Endpunkte unterschiedlich gewichtet werden müssen.

Fragestellung 1

Die folgende Tabelle 3 zeigt einen Überblick und Erläuterungen für alle patientenrelevanten Endpunkte der Fragestellung 1.

Tabelle 3: Fragestellung 1: Überblick und Erläuterung der Ergebnisse für alle patientenrelevanten Endpunkte für eine Nutzen-Schaden-Abwägung (mehrseitige Tabelle)

Patientenrelevanter Endpunkt	**Ergebnisse**	**Basisrisiko[a] pro 10 000 Frauen**	**Risiko[b] pro 10 000 eingeladene Frauen**	**Absoluter Effekt pro 10 000 eingeladene Frauen [95 %-KI]**	**Erläuterung**
Mortalität					
Gesamtmortalität nach 7,9 bis 13 Jahren Beobachtungsdauer	IDR 0,98; 95 %-KI: [0,90; 1,08]; p = 0,602	273	268	5 [−22; 27]	Es ist nicht belegt, dass das Mammografie-Screening die Gesamtmortalität senkt oder erhöht.
brustkrebsspezifische Mortalität zu ca. 10 Jahren	IDR 0,85; 95 %-KI: [0,73; 1,002]; p = 0,053	35	30	5 [0; 9]	Ohne Mammografie-Screening sterben 35 von 10 000 Frauen an Brustkrebs. Mit Mammografie-Screening sterben 30 von 10 000 Frauen an Brustkrebs. Das Mammografie-Screening bewahrt ca. 5 von 10 000 Frauen innerhalb von etwa 10 Jahren davor, an Brustkrebs zu versterben. Bei Betrachtung der Sensitivitätsanalyse sterben ohne Mammografie-Screening 35 und mit Mammografie-Screening 29 von 10 000 Frauen an Brustkrebs. Das Mammografie-Screening bewahrt somit ca. 6 von 10 000 Frauen innerhalb von etwa 10 Jahren davor, an Brustkrebs zu versterben.
Morbidität					
Mastektomien	OR 1,17[c]; 95 %-KI: [0,94; 1,45][c]; p = 0,158[c]	62	72	−10 [−28; 4]	Es ist nicht belegt, dass ein Mammografie-Screening das Risiko einer Mastektomie senkt oder erhöht. Auch ein Einfluss des Mammografie-Screenings auf den Anteil der Mastektomien an allen Brust-OPs ist nicht belegt.
Konsequenzen aus falsch-positiven Screening-befunden: Biopsien mit benignem Befund	siehe Tabelle 29	-	-	7 bis 340	7 bis 340 Frauen von 10 000 erhalten je Jahr bzw. je Screeningrunde eine invasive Abklärungsdiagnostik bei anschließendem gutartigem Befund.

Tabelle 3: Fragestellung 1: Überblick und Erläuterung der Ergebnisse für alle patientenrelevanten Endpunkte für eine Nutzen-Schaden-Abwägung (mehrseitige Tabelle)

Patientenrelevanter Endpunkt	**Ergebnisse**	**Basisrisiko[a] pro 10 000 Frauen**	**Risiko[b] pro 10 000 eingeladene Frauen**	**Absoluter Effekt pro 10 000 eingeladene Frauen [95 %-KI]**	**Erläuterung**
unerwünschte Ereignisse	keine Daten berichtet	-	-	-	-
Überdiagnosen	Spanne [Minimum; Maximum] der Punktschätzungen der einzelnen Studien für das Überdiagnoserisiko bezogen auf die zum Screening eingeladenen Frauen: 0,01 % bis 0,41 %[d]	-	-	1 bis 41	Durch das Screening wird bei 1 bis 41 von 10 000 Screeningteilnehmerinnen Brustkrebs festgestellt, der im Verlauf der restlichen Lebenszeit der Frau keine Beschwerden verursacht hätte. Diese Frauen werden diagnostischen und therapeutischen Prozeduren unterzogen, die nicht erforderlich und zum Teil mit Komplikationen behaftet sind. Das aus den einzelnen Studien berechnete Überdiagnoserisiko bezogen auf die während der Screeningphase mit Brustkrebs diagnostizierten Frauen liegt zwischen 2,63 % und 31,61 %[d].
LQ					
gesundheitsbezogene LQ	keine Daten berichtet	-	-	-	-

a. medianes Risiko der Kontrollgruppe
b. medianes Risiko der Interventionsgruppe
c. basierend auf der Studie CNBSS 7 Jahre nach Studienbeginn
d. basierend auf den Ergebnissen der Studien CNBSS und HIP

CNBSS: Canadian National Breast Screening Study; HIP: Health Insurance Plan of Greater New York; IDR: Inzidenzdichtequotient; KI: Konfidenzintervall; LQ: Lebensqualität; OP: Operation; OR: Odds Ratio

Das Mammografie-Screening bewahrt ca. 5 von 10 000 Frauen innerhalb von ca. 10 Jahren davor, an Brustkrebs zu versterben. Auf Basis der Studienergebnisse lässt sich jedoch nicht zeigen, dass sich das Screening auch auf die Gesamtmortalität auswirkt. Es wird davon ausgegangen, dass aufgrund konkurrierender Todesursachen die Verringerung der Brustkrebsmortalität nicht groß genug war, um auch einen Effekt auf die Gesamtmortalität statistisch nachweisen zu können. Insgesamt spricht das Ergebnis zur Gesamtmortalität nicht gegen einen Nutzen bezüglich des Mammografie-Screenings, und es ergibt sich für den Endpunkt Mortalität ein Anhaltspunkt für einen Nutzen des Mammografie-Screenings.

Es ist nicht belegt, dass ein Mammografie-Screening das Risiko, dass eine Mastektomie durchgeführt werden muss, senkt oder erhöht. Auch ein Einfluss des Mammografie-Screenings auf den Anteil der Mastektomien an allen OPs ist nicht belegt.

Bei falsch-positiven Screeningbefunden erleiden Personen einen Schaden durch die Mitteilung eines besorgniserregenden Befunds durch die sich anschließende Abklärungsdiagnostik und durch die damit einhergehenden Komplikationen. 7 bis 340 zum Screening eingeladene Frauen von 10 000 erhalten je Jahr bzw. je Screeningrunde eine invasive Abklärungsdiagnostik bei anschließendem gutartigem Befund.

Das Überdiagnoserisiko bezogen auf die Frauen mit einer Brustkrebsdiagnose während der Screeningphase lag zwischen 2,63 % (in der HIP-Studie) und 31,61 % (in der CNBSS). Bezogen auf die zum Screening eingeladenen Frauen betrug das Überdiagnoserisiko zwischen 0,01 % (in der HIP-Studie) und 0,41 % (in der CNBSS) für denselben Zeitraum. Die Studien ergaben damit, dass geschätzt 1 bis 41 von 10 000 zum Mammografie-Screening eingeladene Frauen der untersuchten Altersgruppe die Diagnose Brustkrebs erhalten hätten, der im Verlauf der restlichen Lebenszeit aber keine Beschwerden verursacht hätte.

Zu unerwünschten Ereignissen des Mammografie-Screenings lagen keine Daten vor. Mögliche unerwünschte Ereignisse im Zusammenhang mit dem Mammografie-Screening sind vor allem bedingt durch die nachfolgenden Therapien, z. B. Infektions- und Narkoserisiken der OP sowie Nebenwirkungen einer Chemo-, Radio- und / oder hormonellen Therapie. Unerwünschte Ereignisse durch Diagnose und Behandlung einer Krebserkrankung sind dann als ein durch das Screening verursachten Schaden zu werten, wenn diese auf eine Überdiagnose zurückzuführen sind; ansonsten stellen die (frühe) Diagnose und die nachfolgende Behandlung (auch mit den damit einhergehenden unerwünschten Ereignissen) ein Ziel des Screenings dar und sollten deshalb nicht als Schaden des Screenings betrachtet werden. Es ist daher davon auszugehen, dass der Effekt des Screenings auf die Rate unerwünschter Ereignisse im Wesentlichen durch den Endpunkt Überdiagnosen abgebildet wird. Schmerzen durch die Kompression der Brust während der Mammografie können ebenfalls als unerwünschtes Ereignis betrachtet werden. Zwar wurden Schmerzen in den vorliegenden Studien nicht erhoben, andere Studien zeigten jedoch, dass viele Frauen die Untersuchung als schmerzhaft empfinden, wobei der Schmerz aber meist nur leicht ist [27]. Diesbezüglich dürfte ein Nachteil des Mammografie-Screenings als belegt gelten.

Auch zur gesundheitsbezogenen Lebensqualität lagen keine Daten vor. Es ist anzunehmen, dass bei Screeningteilnehmerinnen die Mitteilung eines auffälligen Befunds die gesundheitsbezogene Lebensqualität beeinträchtigt. Da dieser Effekt bei falsch-positiven Ergebnissen nur kurzfristig sein dürfte, ist primär bei den Screeningteilnehmerinnen mit richtig-positiven Befunden mit einer relevanten Beeinträchtigung zu rechnen. Eine Beeinträchtigung der gesundheitsbezogenen Lebensqualität durch Diagnose und Behandlung einer Krebserkrankung sind dann als ein durch das Screening verursachten Schaden zu werten, wenn diese auf eine Überdiagnose zurückzuführen sind; ansonsten stellen die (frühe) Diagnose und die nachfolgende Behandlung (auch mit der damit einhergehenden Einschränkung der gesundheitsbezogenen Lebensqualität) ein Ziel des Screenings dar und sollten deshalb nicht als Schaden des Screenings betrachtet werden. Der Effekt des Screenings auf die gesundheitsbezogene Lebensqualität dürfte sich daher im Wesentlichen durch den Endpunkt Überdiagnosen abbilden.

Ob ein mehrjähriges Mammografie-Screening einschließlich Folgediagnostik strahlenschutzrechtlich zulässig ist, wird aktuell durch das Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit (BMU) bewertet. Nach einer älteren Analyse des Bundesamts für Strahlenschutz ist ein Mammografie-Screening vor dem 46. Lebensjahr sehr kritisch zu sehen [28].

In der Gesamtabwägung ergab sich endpunktübergreifend ein Anhaltspunkt für einen Nutzen des Mammografie-Screenings bei 45- bis 49-jährigen Frauen. Angesichts des vergleichsweise kleinen Mortalitätsvorteils in dieser Altersgruppe bleibt eine individuelle Bewertung und Abwägung jedoch unerlässlich.

Fragestellung 2: Altersgruppe ab 70 Jahren

Auf eine tabellarische Gegenüberstellung von Nutzen und Schaden zur Fragestellung 2 wurde verzichtet, weil die Datenlage – vor allem zum möglichen Nutzen – nicht ausreichend war für eine Quantifizierung.

Die vorhandenen Daten aus nur 2 Studien mit insgesamt rund 18 000 ausgewerteten Teilnehmerinnen beziehen sich auf Frauen bis maximal 74 Jahre zu Beginn der Randomisierung. Für eine abschließende Nutzen-Schaden-Abwägung sind diese Daten nicht ausreichend. Insbesondere fehlen hinreichend zuverlässige Daten, um zu beantworten, ob ein Mammografie-Screening über das 69. Lebensjahr hinaus einen zusätzlichen Mortalitätsvorteil bietet, und falls ja, ob dieser die Nachteile eines Mammografie-Screenings in dieser Altersgruppe überwiegt. Hier ist insbesondere das mit dem Alter steigende Überdiagnoserisiko zu beachten, das durch konkurrierende Risiken [3] bedingt wird. Auch wenn zu Überdiagnosen Daten aus 1 RCT vorliegen und sich daraus bereits ein Anhaltspunkt für einen Schaden erkennen lässt, wären weitere Daten nötig, um die Größenordnung an Überdiagnosen in dieser Altersgruppe gut beurteilen zu können.

Zu Konsequenzen aus falsch-positiven Befunden liegen nur Daten zu einer übergreifenden Altersgruppe von Frauen zwischen 40 und 74 Jahren vor. Diese Daten lassen ebenfalls bereits auf einen Anhaltspunkt für einen Schaden in dieser Altersgruppe schließen. Für eine Quantifizierung wären jedoch Daten zur spezifischen Altersgruppe der ab 70-jährigen Frauen nötig.

Zu Mastektomien, unerwünschten Ereignissen und gesundheitsbezogener Lebensqualität lagen keine Daten vor. Wie bereits bei der Altersgruppe der 45- bis 49-jährigen Frauen ist auch in dieser Altersgruppe davon auszugehen, dass der Effekt des Screenings auf unerwünschte Ereignisse und gesundheitsbezogene Lebensqualität im Wesentlichen durch den Endpunkt Überdiagnosen abgebildet werden würde.

Einschätzung des Potenzials

In Bezug auf 50- bis 69-jährige Frauen, für die in Deutschland bereits ein Mammografie-Screening etabliert ist, berichten systematische Reviews übereinstimmend eine Reduktion der brustkrebsspezifischen Mortalität [29]. Basierend darauf, dass die Ergebnisse im Sinne eines Potenzials als übertragbar betrachtet werden, lässt sich für die Altersgruppe der 70- bis 74-jährigen Frauen ein Potenzial erkennen. Die bisher vorliegenden Erkenntnisse zu dieser Altersgruppe, insbesondere zum Schaden durch Konsequenzen aus falsch-positiven Befunden und durch Überdiagnosen, stehen einem Potenzial nicht entgegen. Vor dem Hintergrund der vielversprechenden laufenden Studie AgeX, die Ergebnisse von mindestens etwa 1 Million Frauen im Alter von 71 bis 73 Jahren erwarten lässt, ist die Skizzierung einer Erprobungsstudie zum jetzigen Zeitpunkt jedoch nicht notwendig. Für Frauen ab 75 Jahren konnten im Rahmen dieser Bewertung keine Daten identifiziert werden. Außerdem ist fraglich, ob die Effekte des Mammografie-Screenings für 50- bis 69-jährige Frauen (im Sinne eines Potenzials) auch auf diese noch ältere Population übertragen werden können. Daher ließ sich auf Basis der vorliegenden Erkenntnisse kein Potenzial für die Altersgruppe ab 75 Jahren erkennen.

5 Einordnung des Arbeitsergebnisses

Aussagekraft der Ergebnisse

Bei der Interpretation der Studien sind einige Aspekte zu beachten, die die Aussagekraft der Ergebnisse beeinflussen. Die wesentlichen werden im Folgenden erläutert.

Studiendesigns

Der vorliegende Bericht bewertet den Nutzen und den Schaden einer Ausweitung der Altersgrenzen des Mammografie-Screenings. Im Kern geht es daher um die Frage, ob ein Screening, das früher beginnt bzw. länger durchgeführt wird, einen zusätzlichen Nutzen gegenüber dem bisher etablierten Screeningprogramm aufweist. Im Unterschied dazu zielten die bisher vorliegenden Studien überwiegend darauf ab, den Nutzen eines Screenings gegenüber keinem Screening zu zeigen. Solche Studien wurden als ausreichende Annäherung an die eigentliche Fragestellung des Berichts betrachtet. Dennoch kommen die eingeschlossenen Studien der Berichtsfragestellung unterschiedlich nahe. Studien, in denen nach der Screeningphase ein Abschlussscreening für die Kontrollgruppe angeboten wurde (Göteborg, Stockholm, Swedish Two County) oder bei denen anschließend an die Screeningphase ein allgemeines Screeningprogramm begann (Göteborg bei Frauen ab 50 Jahren, Edinburgh, Malmö, UK Age), bilden die Fragestellung eines früheren gegenüber einem späteren Screening – also Fragestellung 1 – besser ab, als solche Studien, bei denen kein Abschlussscreening und kein Übergang in ein allgemeines Screeningprogramm stattfand (CNBSS; bei der HIP-Studie fehlen Angaben dazu, es kann jedoch vermutet werden, dass sich zunächst kein Screening aller Frauen anschloss). Welchen Einfluss die unterschiedlichen Studiendesigns auf die Interpretation und Anwendbarkeit der Ergebnisse hat, ist bei den betrachteten Endpunkten unterschiedlich. Dies wird im Folgenden ausgeführt.

Brustkrebsspezifische Mortalität

Die brustkrebsspezifische Mortalität wurde in den Studien unterschiedlich ausgewertet. Entweder gingen nur die Brustkrebsfälle in die Analyse ein, die innerhalb der Screeningphase entdeckt wurden (Evaluationsmodell, siehe Abschnitt A1.2). Oder es wurden alle Brustkrebsfälle einbezogen, auch die, die erst nach Ende der Screeningphase auftraten (Follow-up-Modell, siehe Abschnitt A1.2). In Bezug auf Fragestellung 1 bilden dabei Studien, in denen nach der Screeningphase ein allgemeines Screening aller Frauen beginnt und in der die brustkrebsspezifische Mortalität mittels Follow-up-Modell ausgewertet wird, am besten die Fragestellung des Berichts ab. Bei dieser Auswertungsart ist allerdings ein Verdünnungseffekt zu erwarten, da sich durch das spätere Screening in beiden Gruppen die brustkrebsspezifische Mortalität angleicht. Studien, die ein solches Design aufweisen, sind die UK-Age- und die Göteborg-Studie. Beide Studien enthalten zusätzlich Auswertungen nach dem Evaluationsmodell. Vergleicht man die Ergebnisse der verschiedenen Auswertungsarten jeweils innerhalb dieser beiden Studien, sind allerdings keine wesentlichen Unterschiede festzustellen. Dies lässt vermuten, dass die jeweils verwendete Auswertungsart hier keinen wesentlichen Einfluss auf die Ergebnisse hatte.

Überdiagnosen

Überdiagnosen waren nach der für diese Bewertung zugrunde gelegten Methodik nur bei 2 Studien (CNBSS, HIP) sinnvoll berechenbar. Die Ergebnisse zu Überdiagnosen aus diesen beiden Studien sind aber aus anderen Gründen mit Vorsicht zu interpretieren (siehe Abschnitt 4.5.1.5). Zudem wird auch in der Literatur darauf hingewiesen, dass aggregierte Daten die Überdiagnosen oft überschätzen [30]. Auch wenn die Datenlage eine verlässliche Aussage zur Größenordnung der Überdiagnosen erschwert, steht jedoch außer Frage, dass ein Screening zwangsläufig zu Überdiagnosen führt [3].

Überdiagnosen resultieren gemäß Duffy et al. [31] aus 2 unterschiedlichen Krankheitsentitäten. Zum einen sind dies potenziell progrediente Tumore, die aber bei Frauen auftreten, die aus anderen Gründen sterben, bevor der Tumor symptomatisch geworden ist. Den anderen Teil bilden nicht progrediente oder sehr langsam progrediente Tumore. Zu diesen dürfte ein wesentlicher Teil der DCIS-Fälle gehören. Die beiden beschriebenen Entitäten spielen in den zwei betrachteten Altersgruppen eine unterschiedlich große Rolle. So ist beispielsweise erwartbar, dass nur sehr wenige Frauen unter 50 Jahren an anderen Ursachen versterben, bevor der Tumor symptomatisch wird. Diese Form der Überdiagnosen trägt also nur wenig zu den Überdiagnosen insgesamt in dieser Altersgruppe bei, sodass erwartbar ein Großteil der Überdiagnosen aus nicht oder langsam progredienten Tumoren bestehen dürfte. In Bezug auf Fragestellung 1 ist daher erwartbar, dass durch die Vorverlegung des Screeningbeginns um 5 Jahre nicht viele zusätzliche Überdiagnosen resultieren, weil in den meisten Fällen lediglich der Zeitpunkt der Diagnosestellung (und damit auch der Zeitpunkt der Überdiagnose) vorverlegt wird. Die Ergebnisse der Modellierung stützen diese Erwartung: Bei der bisherigen Screeningstrategie (Screening alle 2 Jahre zwischen 50 und 69 Jahren) wird ein Lebenszeitrisiko für eine Überdiagnose von 6,3 % bezogen auf alle teilnehmenden Frauen berechnet, das bei Erweiterung des Screenings auf die Altersgruppe ab 45 Jahren nur moderat auf 7,2 % ansteigt (siehe Tabelle 11 in Abschnitt A4).

Bei den älteren Frauen (Fragestellung 2) hingegen ist zu erwarten, dass das Risiko von Überdiagnosen durch ein Screening deutlich zunimmt [3], vor allem bedingt durch konkurrierende Todesursachen. In den Ergebnissen der Modellierung spiegelt sich dies jedoch nicht wider: Bei einer Verlängerung des Screenings über 69 Jahre hinaus wird je nach Screeningintervall und Screeningende ein Lebenszeitrisiko für Überdiagnosen von 6,3 % bis 6,6 % berechnet (gegenüber 6,3 % bei der Referenzstrategie), also ebenfalls ein allenfalls geringfügiger Unterschied zum bisherigen Screening. Eine mögliche Erklärung dafür, dass in der Modellierung das Überdiagnoserisiko bei älteren Frauen nicht größer ausfällt als bei jüngeren, könnte eine in den Altersgruppen unterschiedliche Häufigkeit von DCIS-Fällen sein. Die Autorinnen und Autoren des Modellierungsberichts weisen aber auch darauf hin, dass wegen der zugrunde liegenden empirischen Daten beim Endpunkt Überdiagnosen große Unsicherheiten bestehen.

Betrachtete Alterssubgruppen

Es liegen für Fragestellung 1 nur wenige Daten vor, die sich auf genau den gesuchten Altersbereich der 45- bis 49-Jährigen beziehen, sodass für die Bewertung auch Daten von Alterssubgruppen herangezogen wurden, die Frauen unter 45 Jahren umfassten (z. B. bei der Alterssubgruppe 39–49 Jahre). Unter der Prämisse, dass der Effekt eines Mammografie-Screenings auf die Mortalität mit jüngerem Alter immer geringer wird, ergibt sich aus dem Einschluss von „zu jungen" Frauen in der betrachteten Alterssubgruppe eine tendenziell konservative Schätzung des Screeningeffekts. Allerdings wächst ein Teil der Frauen, die zu Beginn der Screeningphase im gesuchten Altersbereich (45–49 Jahre) sind, während der Screeningphase in den Altersbereich hinein, in dem in Deutschland das Screeningprogramm beginnt. Das bedeutet, dass der Effekt bei Frauen, die zu Studienbeginn zwischen 40 und 49 Jahren alt sind, zum Teil auf Frauen beruht, deren Brustkrebs erst jenseits des 50. Lebensjahres entdeckt wurde und damit auch mit dem bereits etablierten Screening gefunden worden wäre. Dies führt in Bezug auf die Mortalität tendenziell zu einer Überschätzung des Effekts eines Screenings bei den 45- bis 49-Jährigen. Welcher der beiden gegenläufigen Effekte hier stärker zum Tragen kommt oder ob sie sich gegenseitig aufheben, lässt sich nicht beantworten.

Bei Fragestellung 2 bestand das potenzielle Problem der größeren und überlappenden Altersbereiche nicht, weil sich die vorliegenden Daten auf die Altersgruppe der 70- bis 74-Jährigen bezogen, die vollständig im Bereich der Zielpopulation dieser Fragestellung liegt. Allerdings lagen zu Frauen ab 70 Jahren (zum Zeitpunkt der Randomisierung) in den eingeschlossenen Studien nur sehr wenige Daten vor. Daher stellt sich die Frage, ob der Einbezug von Frauen, die erst im Verlauf der Studie das 70. Lebensjahr erreichen, auch eine Option zur Beantwortung der Fragestellung 2 darstellt, um über eine höhere Grundgesamtheit eine größere Präzision der Schätzung zu erreichen. Separate Daten zur brustkrebsspezifischen Mortalität von Frauen der Altersgruppe 60 bis 69 Jahre liegen aus den 2 Studien Malmö und Swedish Two County vor, Daten von 60- bis 64-Jährigen aus den 3 Studien Edinburgh, HIP und Stockholm. Bei 3 dieser 5 Studien (Malmö, Swedish Two County, Edinburgh) war die Dauer der Screeningphase lang genug, sodass ein Teil der Frauen während der Screeningphase das 70. Lebensjahr erreichte. Anders als bei Fragestellung 1 ist jedoch fraglich, ob die Ergebnisse der zum Zeitpunkt der Randomisierung 60- bis 64-Jährigen bzw. 60- bis 69-Jährigen auf die eigentliche Zielpopulation der ab 70-Jährigen hinreichend übertragbar sind (im Sinne einer Übertragbarkeit des Nutzens). Denn mit zunehmenden Alter spielen konkurrierende Todesursachen eine immer größere Rolle [24], was es erschweren könnte, mithilfe eines Mammografie-Screenings Brustkrebstodesfälle zu vermeiden. Es ist daher keineswegs sicher, dass ein Effekt des Screenings auf die brustkrebsspezifische Mortalität in der Altersgruppe 60 bis 69 Jahre auch für Frauen ab 70 Jahren gilt. Beim Heranziehen der jüngeren Altersgruppe bestünde vielmehr die Gefahr, dass der Effekt des Screenings für die eigentliche Zielpopulation der über 70-Jährigen überschätzt würde.

Alter der Studien und Übertragbarkeit

Der Studienbeginn lag bei den meisten eingeschlossenen Studien in den 70er- und 80er-Jahren des letzten Jahrhunderts. Ausnahmen bilden die HIP-Studie, die 1963 startete und damit sogar

noch 13 Jahre früher als die zweitälteste eingeschlossene Studie – zu dieser Zeit hatte sich noch nicht einmal die adjuvante Chemotherapie bei Brustkrebs etabliert –, sowie die UK-Age-Studie, die 1991 begann. Alle eingeschlossenen Studien fanden also im Rahmen einer anderen medizinischen Versorgungssituation und unter anderen gesellschaftlichen Umständen statt. Es liegt nahe, dass die Übertragbarkeit der Studienergebnisse auf die heutige Situation mit zunehmendem Alter der Studie abnimmt. Zu den wesentlichen Einflussfaktoren dürften dabei zum einen die Weiterentwicklungen in der Brustkrebsdiagnostik und -therapie (siehe z. B. [32]), zum anderen die insgesamt gestiegene Lebenserwartung sowie Verschiebungen in den Häufigkeiten verschiedener Todesursachen gehören. Es ist einerseits möglich, dass der Effekt des Mammografie-Screenings auf die brustkrebsspezifische Mortalität unter heutigen Bedingungen geringer ausfällt, weil sich Diagnostik und Therapie so verbessert haben, dass auch Frauen, deren Brustkrebs erst im fortgeschrittenen Stadium entdeckt wird, gute Überlebenschancen haben (vergleiche [33]). Andererseits könnte der Effekt heute auch größer sein, da die Lebenserwartung insgesamt gestiegen ist und so ein Effekt des Mammografie-Screenings deutlich zutage treten könnte.

Einordnung der Ergebnisse der Modellierung

Die im Rahmen dieser Bewertung durchgeführte Modellierung (siehe Abschnitt A4) führt Daten aus verschiedenen Quellen zusammen und trifft Vorhersagen auch über den Zeitraum hinaus, der durch die bislang vorliegenden Studien abgedeckt wird. Die Modellierung ist damit nicht ausreichend, um für sich allein genommen die Frage nach Nutzen und Schaden zu beantworten. Die Stärke der Modellierung liegt vielmehr darin, dass sie die Abwägung von Nutzen und Schaden unterstützen kann, indem sie zeigt, wie Nutzen und Schaden miteinander in Beziehung stehen und in welche Richtung und in welcher Größenordnung sich Nutzen und Schaden bewegen, wenn bestimmte Screeningparameter verändert werden.

Ein wesentliches Ergebnis der Modellierung ist, dass das Überleben potenziell verlängert wird, je länger und häufiger das Mammografie-Screening erfolgt, dass sich aber gleichzeitig ab einer gewissen Screeningintensität das inkrementelle Schaden-Nutzen-Verhältnis wieder verschlechtert. (Das inkrementelle Schaden-Nutzen-Verhältnis drückt aus, welcher zusätzliche Schaden in Kauf genommen werden muss, um eine zusätzliche Nutzeneinheit zu erzielen.)

Die Ergebnisse der Modellierungen passen insofern zur durchgeführten Metaanalyse, als ein Anheben der oberen Altersgrenze (Screening bis 74 Jahre alle 2 Jahre) im Vergleich zum Absenken der unteren Altersgrenze (Screening ab 45 Jahre alle 2 Jahre) in der Basisfallanalyse einen weniger als halb so großen Vorteil hinsichtlich der im Vergleich zur Referenzstrategie (Screening von 50 bis 69 Jahren alle 2 Jahre) gewonnenen Lebenszeit erbringt (2,4 versus 6,0 Jahre pro 100 Frauen, Tabelle 10 in Abschnitt A4). In der Modellierung der qualitätsadjustierten Lebensjahre (QALYs) wird das bessere Nutzen-Schaden-Verhältnis an der unteren im Vergleich zur oberen Altersgrenze noch deutlicher, weil hier der inkrementelle QALY-Zugewinn sich etwa um den Faktor 5 unterscheidet (3,1 versus 0,6 pro 100 Frauen; Tabelle 10 in Abschnitt A4). Allerdings gibt es auch modellierte Endpunkte, bei denen in dem 5-Jahres-Intervall an der unteren Altersgrenze mehr zusätzlicher Schaden zu erwarten ist als an

der oberen Altersgrenze, z. B. bei der Rate falsch-positiver Screeningbefunde (siehe Tabelle 11 in Abschnitt A4).

Die Modellierung basiert neben empirischen Daten auf einer Reihe von – mehr oder weniger sicheren – Annahmen und stellt damit lediglich ein vereinfachtes Abbild der Realität dar. Unterschiedliche Ergebnisse zwischen Metaanalyse und Modellierung zeigen daher auch nicht unbedingt einen Widerspruch auf. Dass die Modellierung auch bei einer Erweiterung des Screenings über das 70. Lebensjahr hinaus (Fragestellung 2) einen Überlebensvorteil errechnet, während sich dies anhand der vorliegenden Studienergebnisse nicht zeigte, könnte daran liegen, dass der tatsächliche absolute Überlebensvorteil nicht groß genug ist, um mit den bisher vorliegenden Studien eine ausreichende statistische Präzision zu erreichen und damit einen statistisch signifikanten Effekt zeigen zu können. Allerdings ist zu beachten, dass die Konsequenzen der verschiedenen Screeningstrategien in der Modellierung im Wesentlichen auf der Testgüte basieren. Auf deren Grundlage wird unter Einbezug zahlreicher weiterer Inputparameter die Stadienverschiebung modelliert, die aus den Veränderungen der Parameter des Screeningprogramms resultiert. In Verbindung mit dem stadienspezifischen Überleben äußert sich dies in einem Überlebensvorteil der intensiveren Strategien. Anhand der Modellierung allein lässt sich daher auch nicht mit ausreichender Sicherheit sagen, wie gut der errechnete Effekt bei Frauen ab 70 Jahren die Realität abbildet, sodass eine hinreichend sichere Aussage zum Nutzen und Schaden auf Basis der Modellierung nicht möglich ist.

Überlegungen zum Screeningintervall

Fragestellung 1

Bei Erweiterung des Mammografie-Screenings auf 45- bis 49-jährige Frauen wäre eine Festlegung erforderlich, ob das bisher bei den 50- bis 69-Jährigen etablierte Screeningintervall von 2 Jahren beibehalten oder angepasst wird. In den meisten Screeningprogrammen in Ländern der Europäischen Union, die Frauen unter 50 Jahren umfassen, wird ein Mammografie-Screening alle 2 Jahre angeboten [34]. Allerdings wird noch Forschungsbedarf gesehen bei der Frage, welches Screeningintervall insbesondere bei jüngeren Frauen optimal ist [3]. Anhand der eingeschlossenen Studien lässt sich diese Frage nicht zufriedenstellend beantworten. Die Screeningintervalle betrugen in den Studien 1 bis 2 Jahre, jedoch ließ sich – bei Betrachtung von Studien mit einem Screeningintervall von ≤ 1,5 Jahren bzw. > 1,5 Jahre – kein Einfluss auf die untersuchten Endpunkte zeigen.

In der Modellierung (siehe Abschnitt A4) wurde der Einfluss verschiedener Screening-intervalle (jeweils bei Beibehalten des 2-jährigen Screeningintervalls zwischen 50 und 69 Jahren) auf bestimmte Endpunkte jeweils im Vergleich zur bisherigen Screeningstrategie – also dem Screening alle 2 Jahre zwischen 50 und 69 Jahren – untersucht. Die Ergebnisse legen nahe, dass bei Frauen, die das Screening mit einem Alter von 45 Jahren beginnen, der durchschnittliche Gewinn an Lebenszeit je Frau bei einem Screening 1-mal pro Jahr zwischen 45 und 49 Jahren größer ausfallen könnte als bei einem Screening alle 2 Jahre. Der absolute Unterschied zwischen den beiden Strategien ist jedoch gering (38,77 Jahre versus 38,76 Jahre

Restlebenserwartung je Frau; mit der bisherigen Screeningstrategie werden 38,70 Jahre Restlebenserwartung berechnet; siehe Tabelle 10 in Abschnitt A4). Das Lebenszeitrisiko für Überdiagnosen verändert sich gemäß Modellierung durch den 5 Jahre früheren Screeningbeginn nur moderat von 6,3 % je teilnehmender Frau auf 7,2 % (bei beiden untersuchten Screeningintervallen, d. h. für 1 und 2 Jahre). Jedoch werden falsch-positive Screeningbefunde häufiger, nämlich von 86,7 Ereignissen pro 100 teilnehmenden Frauen bei der Referenzstrategie auf 147,4 (bei einem Intervall von 1 Jahr) bzw. 123,1 Ereignisse (bei einem Intervall von 2 Jahren) pro 100 Frauen (siehe Tabelle 11 in Abschnitt A4). Angesichts des eher kleinen Unterschieds bei der Restlebenserwartung zwischen 1- und 2-jährigem Intervall fällt die Veränderung bei den falsch-positiven Befunden am meisten ins Gewicht und spricht dafür, ein 2-jähriges Intervall zu bevorzugen. Auch die inkrementellen Schaden-Nutzen-Verhältnisse zeigen ein günstigeres Ergebnis bei einem 2-jährigem Intervall an.

Fragestellung 2

Die Modellierung betrachtete auch den Einfluss unterschiedlicher Screeningintervalle (2 Jahre und 3 Jahre) jenseits von 69 Jahren auf die untersuchten Endpunkte. Da jedoch noch keine ausreichenden Daten für eine verlässliche Aussage zu Nutzen und Schaden vorliegen, erübrigen sich derzeit Überlegungen zum optimalen Screeningintervall für diese Altersgruppe.

Perspektive der Betroffenen

Im Rahmen des Betroffenengesprächs zum Bericht gaben die teilnehmenden Frauen an, welche Gründe von ihrer Seite einer Teilnahme am Screening entgegenstanden. Dies waren zum einen Schmerzen bei der Untersuchung und eine als sehr unpersönlich empfundene Atmosphäre während des ganzen Termins. Letzteres resultierte nach Angaben der Frauen vor allem daraus, dass vom ausführenden Personal fast keine Kommunikation mit ihnen erfolgte und die Frauen sich allein gelassen fühlten. Auch die Befundübermittlung per Post empfanden die Frauen als belastend. Aus der Literatur sind solche oder ähnliche Probleme bekannt [35,36].

Die von den befragten Frauen angesprochenen Kritikpunkte ließen sich im Rahmen eines Screenings durchaus berücksichtigen. Dies könnte auch dazu beitragen, die Teilnahmerate von derzeit rund 50 % in Deutschland [23] zu erhöhen. Auch weil die individuelle Abwägung zwischen Nutzen und Schaden beim Mammografie-Screening entscheidend ist (insbesondere an der unteren und oberen Altersgrenze), sollten Kommunikation und Beratung optimal organisiert werden. Die deutschlandweit etablierte Entscheidungshilfe kann hier nur Grundlage und Unterstützung sein [37].

Public-Health-Perspektive

Die Weltgesundheitsorganisation (WHO) spricht nur bei ausreichenden Ressourcen eine (bedingte) Empfehlung aus für ein Mammografie-Screening für Frauen von 40 bis 49 Jahren [38]. Bei begrenzten Ressourcen hingegen spricht sie sich in Bezug auf diese Altersgruppe dagegen aus [38]. Bei der Frage, ob im deutschen Versorgungskontext ausreichende Ressourcen für eine Ausweitung der Altersgrenzen vorhanden sind, sollten als mögliche

limitierende Faktoren neben monetären Ressourcen auch ärztliche und nicht ärztliche Personalressourcen bedacht werden.

6 Fazit

Fragestellung 1: Altersgruppe 45 bis 49 Jahre

Für die brustkrebsspezifische Mortalität liegt ein Anhaltspunkt für einen Nutzen des Mammografie-Screenings im Vergleich zu keinem Screening bei Frauen zwischen 45 und 49 Jahren vor. In Hinblick auf die Gesamtmortalität zeigte sich kein statistisch signifikantes Ergebnis. Das Ergebnis zur Gesamtmortalität spricht nicht gegen einen Nutzen bezüglich des Mammografie-Screenings, denn die Gesamtmortalität wird wesentlich durch das Auftreten konkurrierender Todesursachen beeinflusst. Es wird davon ausgegangen, dass der Effekt der verringerten Brustkrebsmortalität durch das Screening insgesamt zu klein war um sich auf die Gesamtmortalität auszuwirken. Das Mammografie-Screening führt jedoch über falsch-positive Screeningbefunde zu negativen Konsequenzen (Hinweis auf einen Schaden). Auch kommt es zu Überdiagnosen (Anhaltspunkt für einen Schaden). In Bezug auf Mastektomien ergab sich kein Anhaltspunkt für einen Nutzen oder Schaden des Mammografie-Screenings. Für unerwünschte Ereignisse und gesundheitsbezogene Lebensqualität lagen keine Daten vor, sodass sich auch hier jeweils kein Anhaltspunkt für einen Nutzen oder Schaden ergab. Der Effekt des Screenings auf die Rate unerwünschter Ereignisse und auf die gesundheitsbezogene Lebensqualität dürfte jedoch im Wesentlichen durch den Endpunkt Überdiagnosen erfasst sein.

Zusammenfassend wird ein Anhaltspunkt für einen Nutzen des Mammografie-Screenings für Frauen zwischen 45 und 49 Jahren gegenüber keinem Screening festgestellt und somit, dass der Nutzen des Mammografie-Screenings den Schaden überwiegt. Eine individuelle Bewertung und Abwägung bleibt aber angesichts des sehr kleinen Mortalitätsvorteils in dieser Altersgruppe unerlässlich. Deshalb müssen alle Voraussetzungen getroffen werden, um den Frauen eine informierte Entscheidung zu ermöglichen.

Fragestellung 2: Altersgruppe ab 70 Jahren

Ein Einfluss des Mammografie-Screenings auf die Mortalität im Vergleich zu keinem Screening ist bei Frauen ab 70 Jahren nicht belegt (kein Anhaltspunkt für einen Nutzen oder Schaden).

Das Mammografie-Screening führt zu Überdiagnosen und zu negativen Konsequenzen aus falsch-positiven Screeningbefunden (jeweils ein Anhaltspunkt für einen Schaden). Für Mastektomien, unerwünschte Ereignisse und die gesundheitsbezogene Lebensqualität lagen keine Daten vor, sodass sich hier jeweils kein Anhaltspunkt für einen Nutzen oder Schaden ergab.

Der endpunktübergreifende Nutzen bzw. Schaden eines Mammografie-Screenings bei Frauen ab 70 Jahren bleibt daher wegen unzureichender Daten insgesamt unklar.

Jedoch lässt sich für die Altersgruppe der 70- bis 74-jährigen Frauen ein Potenzial erkennen. Eine Erprobungsstudie zu dieser Fragestellung erscheint dennoch verzichtbar, weil eine laufende randomisierte kontrollierte Studie (AgeX) in absehbarer Zeit aussagekräftige Daten erwarten lässt.

Details des Berichts

A1 Projektverlauf

A1.1 Zeitlicher Verlauf des Projekts

Der Gemeinsame Bundesausschuss (G-BA) hat am 22.04.2021 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Überprüfung der Altersgrenzen im Mammografie-Screening-Programm beauftragt.

In die Bearbeitung des Projekts wurden externe Sachverständige eingebunden.

Am 07.06.2021 wurden Betroffene, also Frauen im Alter zwischen 45 und 49 Jahren sowie Frauen über 70 Jahre, zu Erfahrungen und Erwartungen zum Mammografie-Screening konsultiert.

Der Berichtsplan in der Version 1.0 vom 21.07.2021 wurde am 28.07.2021 auf der Website des IQWiG veröffentlicht.

Bei dem vorliegenden Vorbericht handelt es sich um eine vorläufige Nutzenbewertung. Er wird zur Anhörung gestellt. Im Anschluss an diese Anhörung wird der Abschlussbericht erstellt. Dieser Bericht wird an den G-BA übermittelt und 4 Wochen später auf der Website des IQWiG veröffentlicht. Der Zeitplan für alle Arbeitsschritte der Berichterstellung ist auf der Website des IQWiG unter „Projekte & Ergebnisse“ dargelegt.

A1.2 Spezifizierungen und Änderungen im Projektverlauf

Vorbericht im Vergleich zum Berichtsplan 1.0

Neben redaktionellen Änderungen ergaben sich folgende Spezifizierungen oder Änderungen im Vorbericht:

- Änderung in Bezug auf einzuschließende Studientypen: Studien, die von den Autorinnen und Autoren als randomisiert bezeichnet, jedoch als quasirandomisiert eingeschätzt wurden, wurden eingeschlossen. Hintergrund ist, dass im vorliegenden Fall der Unterschied in der Ergebnissicherheit zwischen den vorliegenden RCTs und der betreffenden Quasi-RCT als gering eingeschätzt wurde (siehe Abschnitt A5.3.1).
- Spezifizierung der Selektion von Publikationen zu relevanten Studien aus der systematischen Übersichtsarbeit, deren Informationsbeschaffung als Grundlage für die Bewertung verwendet wird (Canelo-Aybar 2021 [39]): Aufgrund der großen Zahl an eingeschlossenen Publikationen in Canelo-Aybar 2021 wurden aus pragmatischen Gründen nur diejenigen Publikationen zu relevanten Studien selektiert, aus denen Daten für die Metaanalysen dieser Übersichtarbeit herangezogen wurden oder die für die Beschreibung der Studie essenziell sind (insbesondere Designpublikationen) (siehe Abschnitt A2.2.3).

- Spezifizierung der ausgewählten Messzeitpunkte bei der brustkrebsspezifischen Mortalität: Es wurden je Studie mehrere Messzeitpunkte zur Darstellung ausgewählt. Denn einerseits muss einige Jahre abgewartet werden, um eine ausreichende Anzahl Ereignisse beobachten und daraus eine verlässliche Aussage zu einem möglichen Effekt ableiten zu können. Andererseits ist zu erwarten, dass sich die brustkrebsspezifische Mortalität von Interventions- und Vergleichsgruppe nach Ende der Screeningphase wieder annähert. Daher wurde angestrebt, den in den Studien untersuchten Beobachtungszeitraum möglichst breit abzudecken.
  Für die Metaanalysen zur brustkrebsspezifischen Mortalität wurden die Daten jeweils den Messzeitpunkten ca. 5, 10, 15 und 20 Jahre nach Randomisierung zugeordnet und entsprechend zusammengefasst. Konnten bei einer Studie 2 verschiedene Messzeitpunkte demselben Metaanalysen-Zeitpunkt zugeordnet werden (Beispiel: Daten zu 9 und 12 Jahren Beobachtungsdauer aus derselben Studie zur Metaanalyse ca. 10 Jahre nach Randomisierung), wurde für die Metaanalyse primär der spätere Zeitpunkt verwendet (in diesem Beispiel: 12 Jahre) und der frühere Zeitpunkt (in diesem Beispiel: 9 Jahre) im Rahmen von Sensitivitätsanalysen berücksichtigt (siehe Abschnitt A2.1.3).
- Spezifizierung der Auswahl von Daten für die Metaanalysen: Lagen in einer Studie zu demselben Zeitpunkt Daten sowohl der entsprechenden 10-Jahres-Altersgruppe (z. B. 40–49 Jahre) als auch der 5-Jahres-Altersgruppe (z. B. 45–49 Jahre) vor, wurden für die Metaanalysen primär die Daten der 5-Jahres-Altersgruppe verwendet, weil diese die Population der Fragestellung besser abdeckt. Die Daten der 10-Jahres-Altersgruppe wurden im Rahmen von Sensitivitätsanalysen berücksichtigt (siehe Abschnitt A2.1.3).
- Spezifizierung zum Einschluss von Studien, die die Einschlusskriterien nicht vollständig erfüllen: Für Fragestellung 1 (Frauen von 45 bis 49 Jahren) wurden Studien auch dann eingeschlossen, wenn weniger als 80 % der Frauen der relevanten Altersgruppe entsprachen bzw. es wurde auch auf Daten von Subgruppen zurückgegriffen, die sich auf die entsprechende 10-Jahres-Altersgruppe (z. B. 40 bis 49 Jahre) bezogen. Hintergrund war, dass hinreichend sicher plausibel ist, dass die in einer 10-Jahres-Altersgruppe beobachteten Effekte auf die Zielpopulation der Frauen von 45 bis 49 Jahren anwendbar sind. Nur wenn in einer Fragestellung zu einem Endpunkt ansonsten gar keine verwertbaren Daten der relevanten Altersgruppe oder der entsprechenden 10-Jahres-Altersgruppe verfügbar waren, wurde auf Daten einer Subgruppe mit einem noch breiteren Altersbereich (z. B. 40 bis 74 Jahre) zurückgegriffen. Dies war der Fall bei Fragestellung 2 in Bezug auf Konsequenzen aus falsch-positiven Befunden. Dabei wurden Fragen der Übertragbarkeit bei der Interpretation berücksichtigt (siehe Abschnitt A2.1.9).
- Spezifizierung der darzustellenden Ergebnisse zur brustkrebsspezifischen Mortalität: Die Auswertung der brustkrebsspezifischen Mortalität kann auf unterschiedliche Weise erfolgen [40,41]. Gehen in die Auswertung der brustkrebsspezifischen Mortalität nur diejenigen Brustkrebsfälle ein, die innerhalb der Screeningphase diagnostiziert wurden, wird dafür im vorliegenden Bericht – angelehnt an Nyström et al. [41] – der Begriff

Evaluationsmodell verwendet. Dabei wird zwischen einem Evaluationsmodell mit und ohne Abschlussscreening in der Kontrollgruppe unterschieden. Gehen in die Auswertung der brustkrebsspezifischen Mortalität alle Brustkrebsfälle ein, die bis zum Ende des Nachbeobachtungszeitraums diagnostiziert wurden, wird dafür im vorliegenden Bericht – ebenfalls angelehnt an Nyström et al. [41] – der Begriff Follow-up-Modell verwendet. Die Daten wurden jeweils dem in der betreffenden Studie angewendeten Auswertungsmodell zugeordnet (Evaluationsmodell mit oder ohne Abschlussscreening oder Follow-up-Modell). Enthielt eine Studie Ergebnisse aus mehreren Modellen, wurden vorzugsweise Daten aus dem Evaluationsmodell mit Abschlussscreening und aus dem Follow-up-Modell dargestellt. Für die Metaanalysen wurden vorzugsweise Daten aus dem Follow-up-Modell herangezogen, da diese Ergebnisse die vorliegende Fragestellung am besten abbilden (siehe Abschnitt A2.3.2).

- Spezifizierung der Effektmodifikatoren: Der potenzielle Effektmodifikator Ausgestaltung des Screeningprogramms wurde anhand der Merkmale Screeningintervall (bis 1,5 Jahre und über 1,5 Jahre) und Screeningphasendauer (bis 5 Jahre und über 5 Jahre) untersucht. Eine zusätzliche Untersuchung des Merkmals Anzahl der Screeningrunden erschien nicht sinnvoll, da dies zum einen stark vom Screeningintervall und von der Screeningphasendauer abhängt und zum anderen in vielen Studien nicht präspezifiziert, sondern abhängig vom Zeitpunkt der Randomisierung der Teilnehmenden war (siehe Abschnitt A2.3.4).

A2 Methodik gemäß Berichtsplan 1.0

Die folgenden Abschnitte geben den Wortlaut der Berichtsmethodik aus dem Berichtsplan wieder. Über diese Methodik hinausgehende Spezifizierungen oder Änderungen der Methoden im Projektverlauf werden in Abschnitt A1.2 erläutert. Im folgenden Text wird an den entsprechenden Stellen auf diesen Abschnitt verwiesen.

Diese Bewertung wird auf Grundlage der Allgemeinen Methoden 6.0 [42] erstellt. Darüber hinaus werden Ergebnisse eigener Modellierungen in die Bewertung einfließen.

A2.1 Kriterien für den Einschluss von Studien in die Untersuchung

A2.1.1 Population

In die Bewertung werden Studien mit folgenden Teilnehmerinnen eingeschlossen:

- Fragestellung 1: Frauen im Alter zwischen 45 und 49 Jahren, bei denen kein Verdacht auf Brustkrebs besteht und kein spezifisch erhöhtes Brustkrebsrisiko vorliegt.
- Fragestellung 2: Frauen im Alter von 70 Jahren und mehr, bei denen kein Verdacht auf Brustkrebs besteht und kein spezifisch erhöhtes Brustkrebsrisiko vorliegt.

A2.1.2 Prüf- und Vergleichsintervention

Die Prüfintervention ist das Mammografie-Screening. Als Vergleichsintervention gilt kein Screening oder die alleinige Palpation der Brust.

A2.1.3 Darstellung der Einzelstudien

Alle für die Nutzenbewertung notwendigen Informationen werden aus den Unterlagen zu den eingeschlossenen Studien in standardisierte Tabellen extrahiert. Die Ergebnisse zu den in den Studien berichteten patientenrelevanten Endpunkten werden im Bericht vergleichend beschrieben.

Je Endpunkt wird vorzugsweise der späteste Messzeitpunkt dargestellt.

Zu diesem Vorgehen gab es Spezifizierungen im Projektverlauf, siehe Abschnitt A1.2.

Die relevanten Ergebnisse werden endpunktspezifisch pro Studie auf ihr jeweiliges Verzerrungspotenzial überprüft. Anschließend werden die Informationen zusammengeführt und analysiert. Wenn möglich werden über die Gegenüberstellung der Ergebnisse der Einzelstudien hinaus die in den Abschnitten A2.3.2 bis A2.3.4 beschriebenen Verfahren eingesetzt.

Ergebnisse fließen in der Regel nicht in die Nutzenbewertung ein, wenn diese auf weniger als 70 % der in die Auswertung einzuschließenden Patientinnen basieren, das heißt, wenn der Anteil der Patientinnen, die nicht in der Auswertung berücksichtigt werden, größer als 30 % ist.

Die Ergebnisse werden auch dann nicht in die Nutzenbewertung einbezogen, wenn der Unterschied der Anteile nicht berücksichtigter Patientinnen zwischen den Gruppen größer als 15 Prozentpunkte ist.

Für den Endpunkt Überdiagnosen werden die Ergebnisse für 2 unterschiedliche Bezugsgrößen dargestellt:

1) Der Anteil der Überdiagnosen unter den während der Screeningphase mit Brustkrebs diagnostizierten Frauen wird folgendermaßen berechnet: Von der Anzahl der Diagnosen der Interventionsgruppe (aus dem gesamten Zeitraum von Screening- und Nachbeobachtungsphase) wird die Anzahl der Diagnosen der Kontrollgruppe (aus demselben genannten Zeitraum) subtrahiert. Für die Nachbeobachtungsphase wird der längste Beobachtungszeitraum zugrunde gelegt. Die Differenz wird durch die Anzahl der Diagnosen in der Interventionsgruppe während des Screeningzeitraums dividiert. Das Ergebnis ist der Anteil der Überdiagnosen (Überdiagnoserisiko) unter den Diagnosen des Screenings. In Fällen, in denen die Behandlungsgruppen nicht gleich groß sind, wird bei der Berechnung die Zahl der Diagnosen in der Kontrollgruppe entsprechend gewichtet.
2) Der Anteil der Überdiagnosen unter den zum Screening eingeladenen Frauen wird folgendermaßen berechnet: Von der Anzahl der Diagnosen der Interventionsgruppe (aus dem gesamten Zeitraum von Screening- und Nachbeobachtungsphase) wird die Anzahl der Diagnosen der Kontrollgruppe (aus demselben genannten Zeitraum) subtrahiert. Für die Nachbeobachtungsphase wird der längste Beobachtungszeitraum zugrunde gelegt. Die Differenz wird durch die Anzahl der Frauen der Interventionsgruppe dividiert. Das Ergebnis ist der Anteil der Überdiagnosen (Überdiagnoserisiko) unter den zum Screening eingeladenen Teilnehmerinnen. In Fällen, in denen die Behandlungsgruppen nicht gleich groß sind, wird bei der Berechnung die Zahl der Diagnosen in der Kontrollgruppe entsprechend gewichtet.

A2.1.4 Patientenrelevante Endpunkte

Für die Untersuchung werden folgende patientenrelevante Endpunkte betrachtet:

- Mortalität (zum Beispiel Gesamtmortalität und brustkrebsspezifische Mortalität)
- Morbidität (zum Beispiel Brustamputationen, Konsequenzen aus falschen Screeningbefunden, Überdiagnosen)
- gesundheitsbezogene Lebensqualität
- unerwünschte Ereignisse

A2.1.5 Studientypen

Randomisierte kontrollierte Studien (RCTs) sind, sofern sie methodisch adäquat und der jeweiligen Fragestellung angemessen durchgeführt wurden, mit der geringsten Ergebnisunsicherheit behaftet. Sie liefern daher die zuverlässigsten Ergebnisse für die Bewertung des Nutzens einer medizinischen Intervention.

Für alle in Abschnitt A2.1.2 genannten Interventionen und alle in A2.1.3 genannten Endpunkte ist eine Evaluation im Rahmen von RCTs möglich und praktisch durchführbar.

Für den zu erstellenden Bericht werden daher RCTs als relevante wissenschaftliche Literatur in die Nutzenbewertung einfließen.

Zu diesem Vorgehen gab es eine Änderung im Projektverlauf, siehe Abschnitt A1.2.

A2.1.6 Studiendauer

Hinsichtlich der Studiendauer besteht keine Einschränkung.

A2.1.7 Publikationssprache

Die Publikation muss in deutscher oder englischer Sprache verfasst sein.

A2.1.8 Tabellarische Darstellung der Kriterien für den Studieneinschluss

In der folgenden Tabelle sind die Kriterien aufgelistet, die Studien erfüllen müssen, um in die Bewertung eingeschlossen zu werden.

Tabelle 4: Übersicht über die Kriterien für den Studieneinschluss

Einschlusskriterien	
E1a	Frauen im Alter von 45–49 Jahren ohne Verdacht auf Brustkrebs und ohne spezifisch erhöhtes Brustkrebsrisiko (siehe auch Abschnitt A2.1.1)
E1b	Frauen im Alter von 70 Jahren und älter ohne Verdacht auf Brustkrebs und ohne spezifisch erhöhtes Brustkrebsrisiko (siehe auch Abschnitt A2.1.1)
E2	Prüfintervention: Mammografie-Screening (siehe auch Abschnitt A2.1.2)
E3	Vergleichsintervention: kein Mammografie-Screening oder die alleinige Palpation der Brust (siehe auch Abschnitt A2.1.2)
E4	patientenrelevante Endpunkte wie in Abschnitt A2.1.3 formuliert
E5	Studientyp: RCTs (siehe auch Abschnitt A2.1.5)
E6	Publikationssprache: Deutsch oder Englisch
E7	Vollpublikation verfügbar[a]

a. Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [43] oder ein Bericht über die Studie, der den Kriterien des CONSORT- [44], TREND- [45] oder STROBE-Statements [46] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zur Studienmethodik und zu den Studienergebnissen nicht vertraulich sind.

CONSORT: Consolidated Standards of Reporting Trials; ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; RCT: randomisierte kontrollierte Studie; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TREND: Transparent Reporting of Evaluations with Nonrandomized Designs

A2.1.9 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen

Für die Einschlusskriterien E1a und E1b (Population), E2 (Prüfintervention, bezogen auf die Interventionsgruppe der Studie) und E3 (Vergleichsintervention, bezogen auf die Vergleichsgruppe der Studie) reicht es aus, wenn bei mindestens 80 % der eingeschlossenen Patientinnen diese Kriterien erfüllt sind. Liegen für solche Studien Subgruppenanalysen für Patientinnen vor, die die Einschlusskriterien erfüllen, wird auf diese Analysen zurückgegriffen. Studien, bei denen die Einschlusskriterien E1a bzw. E1b, E2 und E3 bei weniger als 80 % erfüllt sind, werden nur dann eingeschlossen, wenn Subgruppenanalysen für Patientinnen vorliegen, die die Einschlusskriterien erfüllen.

Zu diesem Vorgehen gab es Spezifizierungen im Projektverlauf, siehe Abschnitt A1.2.

A2.2 Informationsbeschaffung

A2.2.1 Fokussierte Informationsbeschaffung von systematischen Übersichten

Parallel zur Erstellung des Berichtsplans erfolgte eine Recherche nach systematischen Übersichten in der Datenbank MEDLINE (umfasst auch die Cochrane Database of Systematic Reviews), HTA Database sowie auf den Websites des National Institute for Health and Care Excellence (NICE) und der Agency for Healthcare Research and Quality (AHRQ). Die Suche wurde auf das Publikationsjahr ab 2011 eingeschränkt. Die Suche fand am 01.03.2021 statt. Die Suchstrategien für die Suche in bibliografischen Datenbanken finden sich in A9.1.

Die Selektion erfolgte durch 1 Person und wurde anschließend von einer 2. Person überprüft. Diskrepanzen wurden durch Diskussion zwischen beiden aufgelöst.

Es wird geprüft, ob mindestens eine hochwertige und aktuelle systematische Übersicht infrage kommt, deren Informationsbeschaffung als Grundlage für die Bewertung verwendet werden kann. Dafür erfolgt eine Bewertung der Qualität der Informationsbeschaffung dieser systematischen Übersicht(en). Kann mindestens 1 diesbezüglich hochwertige und aktuelle Übersicht identifiziert werden, werden die zugrunde liegenden Studien beziehungsweise Dokumente von 1 Person auf ihre Relevanz für die vorliegende Bewertung geprüft und das Ergebnis von einer 2. Person überprüft. Bewertungen der eingeschlossenen Studien oder die Datenextraktion werden nicht übernommen.

Die finale Entscheidung, welche systematische(n) Übersicht(en) für die Bewertung herangezogen werden, erfolgt nach Fertigstellung des Berichtsplans anhand der darin festgelegten Kriterien.

A2.2.2 Umfassende Informationsbeschaffung von Studien

Für die umfassende Informationsbeschaffung wird eine systematische Recherche nach relevanten Studien beziehungsweise Dokumenten durchgeführt.

Für den Fall, dass mindestens 1 systematische Übersicht als Basis für die Informationsbeschaffung verwendet werden kann (siehe Abschnitt A2.2.1), wird diese für die Informationsbeschaffung von Studien für den von der Übersicht abgedeckten Zeitraum herangezogen. Dieser Teil der Informationsbeschaffung wird ergänzt um eine systematische Recherche nach relevanten Studien beziehungsweise Dokumenten für den nicht von der Übersicht abgedeckten Zeitraum.

Folgende primäre und weitere Informationsquellen sowie Suchtechniken werden dabei berücksichtigt:

Primäre Informationsquellen

- bibliografische Datenbanken
  - MEDLINE
  - Embase
  - Cochrane Central Register of Controlled Trials
- Studienregister
  - U.S. National Institutes of Health. ClinicalTrials.gov
  - World Health Organization. International Clinical Trials Registry Platform Search Portal

Weitere Informationsquellen und Suchtechniken

- durch den G-BA übermittelte Dokumente
- Anwendung weiterer Suchtechniken
  - Sichten von Referenzlisten identifizierter systematischer Übersichten
- Anhörung zum Vorbericht
- Autorenanfragen

A2.2.3 Selektion relevanter Studien aus der umfassenden Informationsbeschaffung

Selektion relevanter Studien bzw. Dokumente aus den Ergebnissen der bibliografischen Datenbanken

Die in bibliografischen Datenbanken identifizierten Treffer werden in einem 1. Schritt anhand ihres Titels und, sofern vorhanden, Abstracts in Bezug auf ihre potenzielle Relevanz bezüglich der Einschlusskriterien (siehe Tabelle 4) bewertet. Als potenziell relevant erachtete Dokumente werden in einem 2. Schritt anhand ihres Volltextes auf Relevanz geprüft. Beide Schritte erfolgen durch 2 Personen unabhängig voneinander. Diskrepanzen werden durch Diskussion zwischen den beiden aufgelöst.

Selektion relevanter Studien bzw. Dokumente aus weiteren Informationsquellen

Die Rechercheergebnisse aus den folgenden Informationsquellen werden von 2 Personen unabhängig voneinander in Bezug auf ihre Relevanz bewertet:

- Studienregister
- durch den G-BA übermittelte Dokumente

Die Rechercheergebnisse aus den darüber hinaus berücksichtigten Informationsquellen werden von 1 Person auf Studien hin gesichtet. Die identifizierten Studien werden dann auf Relevanz geprüft. Der gesamte Prozess wird anschließend von einer 2. Person überprüft. Sofern in einem

der genannten Selektionsschritte Diskrepanzen auftreten, werden diese jeweils durch Diskussion zwischen den beiden aufgelöst.

Zu diesem Vorgehen gab es Spezifizierungen im Projektverlauf, siehe Abschnitt A1.2.

A2.3 Informationsbewertung und -synthese

A2.3.1 Bewertung des Verzerrungspotenzials der Ergebnisse

Das Verzerrungspotenzial der Ergebnisse wird endpunktspezifisch für jede in die Nutzenbewertung eingeschlossene Studie bewertet. Dazu werden insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Kriterien systematisch extrahiert und bewertet:

A: Kriterien für die endpunktübergreifende Bewertung des Verzerrungspotenzials der Ergebnisse

- Erzeugung der Randomisierungssequenz
- Verdeckung der Gruppenzuteilung
- Verblindung der Patientin sowie der behandelnden Personen
- ergebnisunabhängige Berichterstattung

B: Kriterien für die endpunktspezifische Bewertung des Verzerrungspotenzials der Ergebnisse

- Verblindung der Endpunkterheber
- Umsetzung des ITT-Prinzips
- ergebnisunabhängige Berichterstattung

Für die Ergebnisse randomisierter Studien wird das Verzerrungspotenzial zusammenfassend als niedrig oder hoch eingestuft. Wird bereits hinsichtlich der unter (A) aufgeführten Kriterien ein endpunktübergreifend hohes Verzerrungspotenzial festgestellt, gilt dieses damit für alle Ergebnisse aller Endpunkte als hoch, unabhängig von der Bewertung endpunktspezifischer Aspekte. Andernfalls finden anschließend die unter (B) genannten Kriterien pro Endpunkt Berücksichtigung.

A2.3.2 Metaanalysen

Die geschätzten Effekte und Konfidenzintervalle aus den Studien werden mittels Forest Plots zusammenfassend dargestellt. Anschließend wird die Heterogenität des Studienpools anhand des statistischen Tests auf Vorliegen von Heterogenität [47] untersucht. Ergibt der Heterogenitätstest ein statistisch nicht signifikantes Ergebnis ($p \geq 0{,}05$), wird davon ausgegangen, dass die Schätzung eines gemeinsamen (gepoolten) Effekts sinnvoll ist. Im Fall von mindestens 5 Studien erfolgt die Metaanalyse mithilfe des Modells mit zufälligen Effekten nach der Methode von Knapp und Hartung unter Verwendung des Heterogenitätsschätzers nach Paule und Mandel [48]. Als Ergebnis wird der gemeinsame Effekt inklusive Konfidenzintervall

dargestellt. Weil die Heterogenität im Fall weniger Studien nicht verlässlich geschätzt werden kann, werden bei 4 oder weniger Studien gegebenenfalls Modelle mit festem Effekt verwendet. Dazu müssen die Studien ausreichend ähnlich sein und es darf keine Gründe geben, die gegen die Anwendung eines Modells mit festem Effekt sprechen. Ist ein Modell mit festem Effekt nicht vertretbar, so wird wie folgt verfahren: Da das bevorzugte Verfahren von Knapp und Hartung im Fall von sehr wenigen (< 5) Studien extrem breite und nicht informative Konfidenzintervalle liefern kann, wird die Schätzung nach diesem Verfahren im Fall von 2 Studien im Regelfall nicht dargestellt. Im Fall von 3 oder mehr Studien wird zunächst geprüft, ob das Konfidenzintervall des Knapp-Hartung-Verfahrens schmaler ist als das Konfidenzintervall mittels des DerSimonian-Laird-Verfahrens. Ist dies der Fall, wird die Schätzung von Knapp und Hartung mit Ad-hoc-Varianzkorrektur weiter betrachtet. Des Weiteren wird geprüft, ob das 95 %-Konfidenzintervall des Verfahrens von Knapp und Hartung (ggf. mit Ad-hoc-Varianzkorrektur) zu breit ist, z. B. wenn die Vereinigung der 95 %-Konfidenzintervalle der Einzelstudien enthalten ist. Ist dies nicht der Fall, wird das Knapp-Hartung-Konfidenzintervall zur Ableitung einer Nutzenaussage herangezogen, wenn das Ergebnis statistisch signifikant ist. Wird die Schätzung mittels Knapp und Hartung (ggf. mit Ad-hoc-Varianzkorrektur) nicht herangezogen (regelhaft im Fall von 2 Studien bzw. bei zu breitem 95 %-Konfidenzintervall) oder liegt kein statistisch signifikantes Ergebnis vor, so wird geprüft, ob das Verfahren von DerSimonian und Laird ein statistisch signifikantes Ergebnis liefert. Ist dies der Fall, werden die Studienergebnisse qualitativ zusammengefasst. Ist das Ergebnis nach DerSimonian und Laird nicht statistisch signifikant, wird auf die qualitative Zusammenfassung verzichtet, da in diesem Fall auch ein korrektes Verfahren keinen statistisch signifikanten Effekt ergeben kann.

Ergibt der Heterogenitätstest ein statistisch signifikantes Ergebnis ($p < 0,05$), wird im Fall von mindestens 5 Studien nur das Prädiktionsintervall dargestellt. Bei 4 oder weniger Studien erfolgt eine qualitative Zusammenfassung. In beiden Fällen wird außerdem untersucht, welche Faktoren diese Heterogenität möglicherweise verursachen. Dazu zählen methodische Faktoren (siehe Abschnitt A2.3.3) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt A2.3.4).

Abgesehen von den genannten Modellen können in bestimmten Situationen und mit besonderer Begründung Alternativen wie z. B. das Betabinomialmodell bei binären Daten [49] angewendet werden.

Zu diesem Vorgehen gab es Spezifizierungen im Projektverlauf, siehe Abschnitt A1.2.

A2.3.3 Sensitivitätsanalysen

Bestehen Zweifel an der Robustheit von Ergebnissen wegen methodischer Faktoren, die beispielsweise durch die Wahl bestimmter Cut-off-Werte, Ersetzungsstrategien für fehlende Werte, Erhebungszeitpunkte oder Effektmaße begründet sein können, ist geplant, den Einfluss solcher Faktoren in Sensitivitätsanalysen zu untersuchen. Das Ergebnis solcher Sensitivitätsanalysen kann die Sicherheit der aus den beobachteten Effekten abgeleiteten Aussagen beeinflussen. Ein als nicht robust eingestufter Effekt kann zum Beispiel dazu führen, dass nur

ein Hinweis auf anstelle eines Belegs für einen (höheren) Nutzen attestiert wird (zur Ableitung von Aussagen zur Beleglage siehe Abschnitt A2.3.5).

A2.3.4 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse werden hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht. Ziel ist es, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Für einen Nachweis unterschiedlicher Effekte ist die auf einem Homogenitäts- beziehungsweise Interaktionstest basierende statistische Signifikanz Voraussetzung. In die Untersuchung werden die vorliegenden Ergebnisse aus Regressionsanalysen, die Interaktionsterme beinhalten, und aus Subgruppenanalysen einbezogen. Außerdem erfolgen eigene Analysen in Form von Metaregressionen oder Metaanalysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren. Subgruppenanalysen werden nur durchgeführt, falls jede Subgruppe mindestens 10 Personen umfasst und bei binären Daten mindestens 10 Ereignisse in einer der Subgruppen aufgetreten sind. Es ist vorgesehen, folgende Faktoren bezüglich einer möglichen Effektmodifikation in die Analysen einzubeziehen:

- Alter (zum Beispiel 70 bis 74 vs. ≥ 75 Jahre)
- Ausgestaltung des Screeningprogramms (zum Beispiel Screeningintervall, Anzahl der Screeningrunden)
- Brustdichte

Sollten sich aus den verfügbaren Informationen weitere mögliche Effektmodifikatoren ergeben, können diese ebenfalls begründet einbezogen werden.

Bei Identifizierung möglicher Effektmodifikatoren erfolgt gegebenenfalls eine Präzisierung der aus den beobachteten Effekten abgeleiteten Aussagen. Beispielsweise kann der Beleg eines (höheren) Nutzens auf eine spezielle Subgruppe von Patientinnen eingeschränkt werden (zur Ableitung von Aussagen zur Beleglage siehe Abschnitt A2.3.5).

Zu diesem Vorgehen gab es Spezifizierungen im Projektverlauf, siehe Abschnitt A1.2.

A2.3.5 Aussagen zur Beleglage

Für jeden Endpunkt wird eine Aussage zur Beleglage des (höheren) Nutzens oder (höheren) Schadens getroffen. Dabei sind 4 Abstufungen der Aussagesicherheit möglich: Es liegt entweder ein Beleg (höchste Aussagesicherheit), ein Hinweis (mittlere Aussagesicherheit), ein Anhaltspunkt (schwächste Aussagesicherheit) oder keine dieser 3 Situationen vor. Der letzte Fall tritt ein, wenn keine Daten vorliegen oder die vorliegenden Daten keine der 3 übrigen Aussagen zulassen. In diesem Fall wird die Aussage „Es liegt kein Anhaltspunkt für einen (höheren) Nutzen oder (höheren) Schaden vor“ getroffen.

Die regelhaft abzuleitende Aussagesicherheit ist von den in Tabelle 5 dargestellten Kriterien abhängig. Die qualitative Ergebnissicherheit ist abhängig vom Design der Studie. Ergebnisse randomisierter Studien mit niedrigem Verzerrungspotenzial haben eine hohe, Ergebnisse randomisierter Studien mit hohem Verzerrungspotenzial eine mäßige qualitative Ergebnissicherheit. Ergebnisse nicht randomisierter vergleichender Studien haben eine geringe qualitative Ergebnissicherheit.

Tabelle 5: Regelhaft abgeleitete Aussagesicherheiten für verschiedene Evidenzsituationen beim Vorliegen von Studien derselben qualitativen Ergebnissicherheit

		Anzahl Studien				
		1 (mit statistisch signifikantem Effekt)	≥ 2			
			homogen	heterogen		
			Metaanalyse statistisch signifikant	gleichgerichtete Effekte[a]		
				deutlich	mäßig	nein
Qualitative Ergebnis-sicherheit	hoch	Hinweis	Beleg	Beleg	Hinweis	–
	mäßig	Anhaltspunkt	Hinweis	Hinweis	Anhaltspunkt	–
	gering	–	Anhaltspunkt	Anhaltspunkt	–	–

a. Gleichgerichtete Effekte liegen vor, wenn trotz Heterogenität eine deutliche oder mäßige Richtung der Effekte erkennbar ist.

Abschließend erfolgt eine endpunktübergreifende Bewertung des Nutzens. In diese übergreifende Bewertung wird auch die Datenvollständigkeit und die sich daraus möglicherweise ergebende Verzerrung aufgrund von Publikationsbias einbezogen, die zur Einschränkung des Fazits führen kann.

Falls kein Anhaltspunkt für einen (höheren) Nutzen oder (höheren) Schaden ableitbar ist, wird eine Aussage zum Potenzial der Behandlungsmethode getroffen und es werden ggf. Eckpunkte einer Erprobungsstudie formuliert.

A2.4 Modellierung

Neben dem Einschluss von RCTs werden zusätzlich Modellierungsstudien durchgeführt. Diese sollen auf bereits bestehende etablierte Modelle (z. B. ein CISNET-Modell) für Brustkrebs-Screening aufbauen und unter Verwendung aktueller Daten für Deutschland angepasst werden. Mittels Modellierung werden die Ergebnisse ausgewählter patientenrelevanter Endpunkte für (1.) einen früheren Screeningbeginn ab einem Alter von 45 Jahren und (2.) für eine Fortsetzung des Screenings über das 69. Lebensjahr hinaus geschätzt. Dabei wird der nationale Versorgungskontext mitberücksichtigt. Die Modellierungsstudien werden durch eine externe Expertengruppe durchgeführt.

Für die durchgeführte Modellierung erfolgt eine umfassende Dokumentation des Vorgehens mit Darstellung und Begründung der getroffenen Entscheidungen und der Auswahl der im Modell eingehenden Zustände, Parameter und Daten. Es wird einen detaillierten technischen

Bericht mit der Beschreibung aller Modellierungsschritte geben [50]. Ebenfalls dargestellt und begründet werden die verwendeten mathematischen Beziehungen der Modellkomponenten einschließlich aller verwendeten Variablen. Der Einfluss getroffener Annahmen und Modellinputs wird durch Sensitivitätsanalysen überprüft und berichtet [51,52]. Auch die vorhandenen Schätzunsicherheiten der Inputparameter werden im Rahmen der Modellierungsstudie berücksichtigt und die Auswirkungen werden nachvollziehbar dargestellt. Eine Modellvalidierung wird durchgeführt und es wird ein Überblick über die eingesetzten Validierungstechniken und ihre Ergebnisse gegeben [53].

Bei Modellierungen wird aufgrund der fehlenden direkten empirischen Evidenz von einer unzureichenden Ergebnissicherheit ausgegangen, sodass sich hieraus regelhaft kein Anhaltspunkt für einen Nutzen oder Schaden ableiten lässt. Bei Vorliegen gewisser Datensituationen (z. B. ein valides und hinreichend robustes Modell, das die RCT-Ergebnisse zum Screening bei 50- bis 69-Jährigen als Ankerpunkt verwendet) können die Ergebnisse der Modellierungsstudie jedoch die Ableitung der Beleglage unterstützen.

A3 Details der Ergebnisse

A3.1 Informationsbeschaffung

A3.1.1 Fokussierte Informationsbeschaffung nach systematischen Übersichten

Von den 8 identifizierten systematischen Übersichten (siehe Abschnitt A7.1) wurde 1 systematische Übersicht (siehe Tabelle 6) als ausreichend aktuell und hochwertig bewertet und zum Zweck der Identifizierung von Primärstudien berücksichtigt. Die Bewertung der Qualität der Informationsbeschaffung dieser systematischen Übersicht findet sich in Kapitel A8.

In der Basis-SÜ wurden 30 Publikationen identifiziert, für die daraufhin geprüft wurde, inwieweit sie die Einschlusskriterien dieses Berichts erfüllen (siehe Abschnitt A2.1). 4 Publikationen zu relevanten Studien erfüllten nicht die Einschlusskriterien des Berichts. Die Referenzen der als Volltexte geprüften, aber ausgeschlossenen Treffer finden sich mit Angabe des jeweiligen Ausschlussgrundes in Abschnitt A7.2. Somit wurden insgesamt 7 Studien (26 Publikationen) als relevant für die vorliegende Nutzenbewertung identifiziert. Dies entspricht allen Studien, die als RCTs in die Basis-SÜ eingeschlossen wurden.

Tabelle 6: Berücksichtigte systematische Übersichten

Systematische Übersicht	Vollpublikation
Canelo-Aybar 2021	ja [39]

A3.1.2 Umfassende Informationsbeschaffung

In einem nächsten Schritt erfolgte eine ergänzende Suche nach Primärstudien in bibliografischen Datenbanken für den Zeitraum (ab 01.04.2016), der nicht durch die in Abschnitt A3.1.1 identifizierte Basis-SÜ abgedeckt war.

Für alle weiteren Informationsquellen wurde ohne zeitliche Beschränkung gesucht.

A3.1.2.1 Primäre Informationsquellen

A3.1.2.1.1 Bibliografische Datenbanken

Abbildung 1 zeigt das Ergebnis der systematischen Literaturrecherche in den bibliografischen Datenbanken und der Studienselektion gemäß den Kriterien für den Studieneinschluss. Die Suchstrategien für die Suche in bibliografischen Datenbanken finden sich in Abschnitt A9.1. Die letzte Suche fand am 17.05.2021 statt.

Die Referenzen der als Volltexte geprüften, aber ausgeschlossenen Treffer finden sich mit Angabe des jeweiligen Ausschlussgrundes in Abschnitt A7.3.

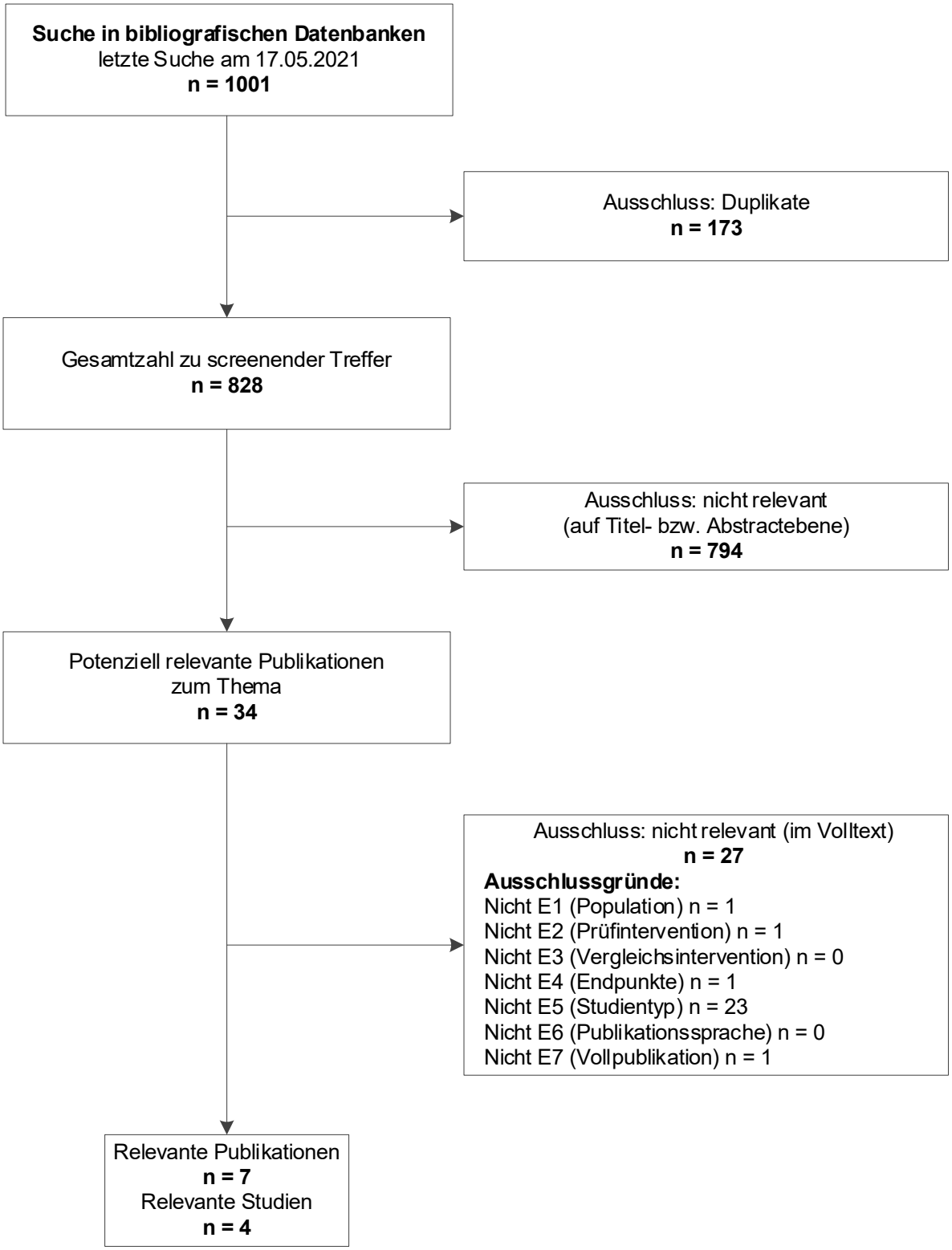


Abbildung 1: Ergebnis der umfassenden Informationsbeschaffung aus den bibliografischen Datenbanken und der Studienselektion

A3.1.2.1.2 Studienregister

Durch die Suche in Studienregistern wurden folgende relevante Studien beziehungsweise Dokumente identifiziert (Tabelle 7):

Tabelle 7: In Studienregistern identifizierte relevante Studien bzw. Dokumente

Studie	Studienregister-ID	Studienregister	Ergebnisbericht in Studienregister vorhanden
Swedish Two County	NCT03217539	ClinicalTrials.gov [54]	nein
UK Age	ISRCTN24647151	ISRCTN [55]	nein

Zur Pilotstudie AgeX [56] wurde zu einem deutlich späteren Zeitpunkt eine Ergebnispublikation [57] identifiziert, die Angaben zu patientenrelevanten Endpunkten (z. B. zu Biopsien mit benignem Befund) enthält. Daher ist diese Studie als relevant einzuschätzen. Diese Studie ist im Rahmen des Vorberichts noch nicht im Studienpool enthalten, wird jedoch im Abschlussbericht berücksichtigt und ggf. ausführlicher dargestellt werden. Eine relevante Änderung des Berichtsfazits ist dadurch jedoch nicht erwartbar, insbesondere da Ergebnisse zur Mortalität nicht enthalten sind und angesichts der kurzen Nachbeobachtungszeit von nur 12 Monaten auch nicht sehr aussagekräftig wären.

In den Studienregistern wurde darüber hinaus eine Studie ohne berichtete Ergebnisse identifiziert (siehe Abschnitt A3.1.4).

Die Suchstrategien für die Suche in Studienregistern finden sich in Abschnitt A9.2. Die letzte Suche in Studienregistern fand am 17.05.2021 statt.

A3.1.2.2 Weitere Informationsquellen und Suchtechniken

Über weitere Informationsquellen und Suchtechniken identifizierte relevante Studien beziehungsweise Dokumente werden nachfolgend nur dargestellt, wenn sie nicht bereits über die primären Informationsquellen gefunden wurden.

A3.1.2.2.1 Durch den G-BA übermittelte Dokumente

Im Rahmen der Auftragsbearbeitung wurden Dokumente vom G-BA an das IQWiG weitergeleitet. Diese wurden auf Duplikate zur bibliografischen Recherche überprüft. Die im Rahmen der Volltextsichtung als nicht relevant ausgeschlossenen Dokumente finden sich mit Angabe des jeweiligen Ausschlussgrundes in Abschnitt A7.4.

Es wurden folgende relevante Studien beziehungsweise Dokumente identifiziert, die nicht über andere Rechercheschritte gefunden werden konnten (Tabelle 8):

Tabelle 8: In vom G-BA übermittelten Dokumenten identifizierte relevante Studien bzw. Dokumente

Studie	Verfügbare Dokumente
Swedish Two County	Publikation [58]

A3.1.2.2.2 Anwendung weiterer Suchtechniken

Im Rahmen der Informationsbeschaffung wurden systematische Übersichten identifiziert – die entsprechenden Referenzen finden sich in Abschnitt A7.1. Die Referenzlisten dieser systematischen Übersichten wurden gesichtet.

Es wurden folgende relevante Studien beziehungsweise Dokumente identifiziert, die nicht über andere Rechercheschritte gefunden werden konnten:

Tabelle 9: Durch Anwendung weiterer Suchtechniken identifizierte relevante Studien bzw. Dokumente

Studie	Verfügbare Dokumente
CNBSS	Publikation [59]
Edinburgh	Publikation [60]
HIP	Publikation [61-63]
Malmö	Publikation [64]
Swedish Two County	Publikation [65]
UK Age	Publikation [66]
CNBSS: Canadian National Breast Screening Study; HIP: Health Insurance Plan of Greater New York	

A3.1.2.2.3 Autorenanfragen

Autorenanfragen bezüglich zusätzlicher Informationen zu relevanten Studien waren nicht erforderlich, da davon auszugehen war, dass solche Informationen keinen relevanten Einfluss auf die Bewertung haben würden.

A3.1.2.2.4 Zusätzliche relevante Studien bzw. Dokumente

Es wurden folgende relevante Studien beziehungsweise Dokumente identifiziert, die nicht über andere Rechercheschritte gefunden werden konnten (Tabelle 10):

Tabelle 10: Zusätzlich identifizierte relevante Studien bzw. Dokumente

Studie	**Verfügbare Dokumente**
CNBSS	Publikation [12]
Edinburgh	Publikation [13,67]
Malmö	Publikation [41]
Stockholm	Publikation [68]
Swedish Two County	Publikation [69-71]
UK Age	Publikation [72]
CNBSS: Canadian National Breast Screening Study	

Es handelt sich um Publikationen zu bereits bekannten Studien, die über Referenzierung in anderen eingeschlossenen Publikationen identifiziert wurden.

A3.1.3 Resultierender Studienpool

Durch die verschiedenen Rechercheschritte konnten insgesamt 8 relevante Studien identifiziert werden (siehe Tabelle 11).

Tabelle 11: Studienpool der Nutzenbewertung

Studie	**Verfügbare Dokumente**	
	Vollpublikation (in Fachzeitschriften)	**Registereintrag / Ergebnisbericht aus Studienregistern**
CNBSS	ja [12,59,73-79]	nein / nein
Edinburgh	ja [13,60,67]	nein / nein
Göteborg	ja [15,80,81]	nein / nein
HIP	ja [14,61-63,82-84]	nein / nein
Malmö	ja [16,41,64,85,86]	nein / nein
Stockholm	ja [17,68,87,88]	nein / nein
Swedish Two County	ja [18,58,65,69-71,89-93]	ja [54] / nein
UK Age	ja [19,25,66,72,94-97]	ja [55] / nein
CNBSS: Canadian National Breast Screening Study; HIP: Health Insurance Plan of Greater New York		

A3.1.4 Studien ohne berichtete Ergebnisse

In Tabelle 12 sind alle durch die Informationsbeschaffung identifizierten Studien ohne bisher berichtete Ergebnisse dargestellt.

Tabelle 12: In der Informationsbeschaffung identifizierte Studien ohne berichtete Ergebnisse

Studie	Dokumentart, Studienregister-ID	Frage-stellung	Studien-typ	Fallzahl	Status
AgeX	Studienregistereintrag, NCT01081288 [98], ISRCTN33292440 [26]	1 und 2	Cluster-RCT	4,4 Mio.	laufend, Rekrutierung abgeschlossen
RCT: randomisierte kontrollierte Studie					

Die noch laufende Studie AgeX aus Großbritannien [98] lässt weitere Erkenntnisse zu beiden Fragestellungen des Berichts erwarten. Eine entsprechende Pilotstudie zur Machbarkeit mit rund 60 000 Teilnehmerinnen wurde bereits abgeschlossen [57] (siehe Abschnitt A3.1.2.1.2). Das Studienprotokoll [99] und weitere Dokumente der englandweit laufenden Studie sind über die Webseite der Studie öffentlich verfügbar. Die Studie soll Erkenntnisse zu der Frage liefern, inwiefern Frauen außerhalb des derzeitigen Altersbereichs des nationalen Screeningprogramms von einem Screening profitieren. Zu diesem Zweck werden Frauen im Alter von 47 bis 49 bzw. von 71 bis 73 Jahren eingeschlossen und per Clusterrandomisierung entweder einem Screening oder einer Vergleichsgruppe ohne Screening zugeteilt. In der jüngeren Altersgruppe erhalten die Frauen der Interventionsgruppe 1 eine zusätzliche Screeningeinladung etwa 3 Jahre vor Start des nationalen Screnningprogramms, in der älteren Altersgruppe Einladungen für bis zu 3 weitere Screenings alle 3 Jahre nach Ende des nationalen Screeningprogramms. Die Ausgestaltung des Screenings in der Studie entspricht dem etablierten Vorgehen im nationalen Screeningprogramm. Frauen im Alter von 50 bis 70 Jahren (zum Zeitpunkt der Randomisierung) sind nicht Teil der Studienpopulation und erhalten das reguläre Screeningangebot.

Primärer Endpunkt der Studie ist die brustkrebsspezifische Mortalität. Die Ergebnisse für jüngere und ältere Frauen sollen jeweils separat analysiert werden. Es sind regelmäßige Interimsanalysen geplant sowie deren Veröffentlichung, sobald damit die Fragestellung der Studie klar beantwortet werden kann. Erste Ergebnisse werden Mitte der 2020-er Jahre erwartet.

Nachdem das nationale Screeningprogramm – und damit einhergehend die Rekrutierung für die Studie – im März 2020 pandemiebedingt unterbrochen wurde, wurde entschieden, die Randomisierung gänzlich zu beenden, um eine Überlastung des wieder anlaufenden Screenings zu vermeiden [100]. Statt der geplanten rund 6 Millionen Frauen wurde die Rekrutierung daher nach 4,4 Millionen Frauen beendet, die Nachbeobachtung wird jedoch fortgeführt. Die Studiengruppe betrachtet diese Anzahl bei entsprechend langer Nachbeobachtungszeit als immer noch ausreichend.

A3.2 Charakteristika der in die Bewertung eingeschlossenen Studien

A3.2.1 Fragestellung 1 – Studiendesign und Studienpopulationen

Tabelle 13: Fragestellung 1 – Charakterisierung der eingeschlossenen Studien (mehrseitige Tabelle)

Studie	**Studien-design**	**Personenzahl N** Gesamt-population / Alterskohorte[a]	**Ort und Zeitraum der Rekrutierung**	**Dauer der Screeningphase / Beobachtungs-dauer (seit Randomisierung)**	**relevante Endpunkte**
CNBSS	RCT, multizen-trisch	40–49 Jahre: 50 430 / 45–49 Jahre: 23 803[b]	Kanada, 1980–1985	5 Jahre / 25 Jahre	Gesamtmortalität, Brustkrebsmortalität, Mastektomien, Brustkrebsinzidenz, Biopsien mit benignem Befund
Edinburgh	Cluster-RCT	45–64 Jahre: 52 253[c] / 45–49 Jahre: 11 723	Edinburgh, Vereinigtes Königreich, 1979–1985	7 Jahre / 14 Jahre	Brustkrebsmortalität
Göteborg	RCT	39–59 Jahre: 52 222[c] / 39–49 Jahre: 26 113[b]	Göteborg, Schweden, 1982–1984	7 Jahre[c] / 24 Jahre	Gesamtmortalität, Brustkrebsmortalität, Brustkrebsinzidenz, Biopsien mit benignem Befund
HIP	RCT	40–64 Jahre: 61 004[c]	New York, USA, 1963–1966	3 Jahre / 18 Jahre	Brustkrebsmortalität, Brustkrebsinzidenz, Biopsien mit benignem Befund
Malmö (MMST 1 und 2)	RCT	44–70 Jahre: 42 283 / 44–54 Jahre: MMST 1: 16 063 MMST 2: 17 786	Malmö, Schweden, MMST 1: 1976–1978 MMST 2: 1978–1990	MMST 1: 10 bzw. 14 Jahre[e] / 23 Jahre MMST 2: 7 / 10 Jahre	Brustkrebsmortalität, Brustkrebsinzidenz
Stockholm	quasi-randomi-siert	40–64 Jahre: 60 318 / 40–49 Jahre: 21 945[b]	Stockholm, Schweden, 1981–1983	4 Jahre / 15 Jahre	Brustkrebsmortalität, Biopsien mit benignem Befund
Swedish Two County	Cluster-RCT, multi-zentrisch	40–74 Jahre: 134 867[b, c] / 40–49 Jahre: 35 448[b]	Dalarna (vormals Kopparberg) und Östergötland, Schweden, 1977–1980	6–8 Jahre / 28 Jahre	Gesamtmortalität, Brustkrebsmortalität, Brustkrebsinzidenz, Biopsien mit benignem Befund
UK Age	RCT, multi-zentrisch	39–41 Jahre: 160 836	Vereinigtes Königreich, 1991–1997	7–9 Jahre / 23 Jahre	Gesamtmortalität, Brustkrebsmortalität, Brustkrebsinzidenz, Biopsien mit benignem Befund

Tabelle 13: Fragestellung 1 – Charakterisierung der eingeschlossenen Studien (mehrseitige Tabelle)

Studie	**Studien-design**	**Personenzahl N** Gesamt-population / Alterskohorte[a]	**Ort und Zeitraum der Rekrutierung**	**Dauer der Screeningphase / Beobachtungs-dauer (seit Randomisierung)**	**relevante Endpunkte**
a. Gesamtpopulation der in der Studie eingeschlossenen Frauen (Altersspanne: n) / für die Fragestellung relevante Alterskohorte (Altersspanne: n) b. eigene Berechnung c. abweichende Angaben zwischen den Publikationen d. Angabe bezieht sich auf die Gruppe der 39- bis 49-jährigen Frauen. e. Dauer der Screeningphase: 10 Jahre für Frauen, geboren zwischen 1908 und 1917; 14 Jahre für Frauen, geboren zwischen 1918 und 1922 bzw. geboren zwischen 1923 und 1932 CNBSS: Canadian National Breast Screening Study; HIP: Health Insurance Plan of Greater New York; N: Anzahl randomisierter bzw. eingeschlossener Frauen; RCT: randomisierte kontrollierte Studie					

Tabelle 14: Fragestellung 1 – Charakterisierung der Interventionen in den eingeschlossenen Studien (mehrseitige Tabelle)

Studie	Intervention	Vergleich
CNBSS	▪ initiale körperliche Untersuchung der Brust sowie Anleitung zur Selbstuntersuchung ▪ Mammografie: jährlich ▪ Anzahl der Screeningrunden: 4–5 ▪ Aufnahmetechnik: 2 Aufnahmen (Obliqueaufnahme und kraniokaudale Ansicht) ▪ Anzahl der befundenden Personen: 1 ▪ zusätzliche klinische Untersuchung: ja, Untersuchungen der Brust jährlich ▪ Selbstuntersuchung / Intervall: ja / k. A.	▪ initiale körperliche Untersuchung der Brust sowie Anleitung zur Selbstuntersuchung ▪ kein Screening; jährliches Follow-up durch einen selbst auszufüllenden Fragebogen ▪ Selbstuntersuchung: ja ▪ Einladung der Kontrollgruppe zum Screening nach Beendigung der Screeningphase: nein
Edinburgh	▪ initiale körperliche Untersuchung der Brust sowie Anleitung zur Selbstuntersuchung ▪ Mammografie: 2-jährlich ▪ Anzahl der Screeningrunden: 2–4 ▪ Aufnahmetechnik: zu Beginn 2 Aufnahmen (Obliqueaufnahme und kraniokaudale Ansicht), bei nachfolgenden Untersuchungen nur einzelne Aufnahmen (Obliqueaufnahmen) ▪ Anzahl der befundenden Personen: 2 (geschulte Ärztinnen und Ärzte, Auffälligkeiten werden durch Radiologinnen oder Radiologen untersucht) ▪ zusätzliche klinische Untersuchung: ja (Untersuchung der Brust), jährlich ▪ Selbstuntersuchung: ja, monatlich	▪ kein Screening; ▪ Selbstuntersuchung: ja ▪ Einladung der Kontrollgruppe zum Screening nach Beendigung der Screeningphase: nein
	Die Frauen der Kontrollgruppe < 65 Jahre wurden ca. 1 Jahr nach Ende der Screeningphase zum organisierten Mammografie-Screening eingeladen, die Frauen der Interventionsgruppe < 65 Jahre 3 Jahre nach ihrem letzten Studienscreening.	
Göteborg	▪ Mammografie: alle 18 Monate ▪ Anzahl der Screeningrunden: 5 ▪ Aufnahmetechnik: Zu Beginn zwei Aufnahmen (Obliqueaufnahme und kraniokaudale Ansicht), bei nachfolgenden Untersuchungen nur einzelne Aufnahme (Obliqueaufnahme), außer bei dichtem Brustgewebe (dann 2 aufnahmen) ▪ Anzahl der befundenden Personen: 1 in den ersten 3. Screeningrunden, danach zweifache Prüfung ▪ Zusätzliche klinische Untersuchung: k. A. ▪ Selbstuntersuchung: k. A.	▪ Grundversorgung ▪ Selbstuntersuchung: k. A. ▪ Einladung der Kontrollgruppe zum Screening nach Beendigung der Screeningphase: ja, ▫ 1 Screening für alle Altersklassen nach Abschluss der Screeningphase der Interventionsgruppe
	Kurz nach Abschluss der Screeningphase wurde ein Routine-Screening für alle Frauen ab 50 Jahren eingeführt.	

Tabelle 14: Fragestellung 1 – Charakterisierung der Interventionen in den eingeschlossenen Studien (mehrseitige Tabelle)

Studie	Intervention	Vergleich
HIP	▪ Mammografie: alle 12 Monate ▪ Anzahl der Screeningrunden: 4 ▪ Aufnahmetechnik: kraniokaudale und laterale Aufnahme ▪ Anzahl der befundenden Personen: 2 ▪ zusätzliche klinische Untersuchung: jährlich ▪ Selbstuntersuchung: k. A.	▪ kein Screening ▪ Selbstuntersuchung: k. A. ▪ Einladung der Kontrollgruppe zum Screening nach Beendigung der Screeningphase: k. A.
Malmö (MMST 1 und 2)	▪ Mammografie: alle 18–24 Monate ▪ Anzahl der Screeningrunden: 9 (MMST 1) bzw. 5 (MMST 2) ▪ Aufnahmetechnik: 2-Ansicht-Mammografie für die erste Aufnahme; in folgenden Screeningrunden 1 oder 2 Ansichten in Abhängigkeit von der Dichte des Brustgewebes ▪ Anzahl der befundenden Personen: 2 ▪ zusätzliche klinische Untersuchung: k. A. ▪ Selbstuntersuchung: k. A.	▪ Selbstuntersuchung: k. A. ▪ Einladung der Kontrollgruppe zum Screening nach Beendigung der Screeningphase: ja ▫ Ab 1990: Screeningintervall 18–24 Monate für Frauen ab 50 Jahren, abhängig von Alter und Parenchymmuster des Brustgewebes
	MMST 1 und 2: Einladung zum nationalen Screening für alle Frauen geboren zwischen 1923 und 1932 nach Studienende 1990, kein Screening für Frauen, die 71 Jahre und älter sind	
Stockholm	▪ Mammografie: ca. alle 2 Jahre ▪ Anzahl der Screeningrunden: 2 ▪ Aufnahmetechnik: einzelne oblique Aufnahme, bei Verdacht auf Malignität erneute Einladung mit 3-Ansicht-Mammografie ▪ Anzahl der befundenden Personen: 1 ▪ zusätzliche klinische Untersuchung: nein, aber selbst auszufüllender Fragebogen ▪ Selbstuntersuchung: k. A.	▪ kein Screening ▪ Selbstuntersuchung: k. A. ▪ Einladung der Kontrollgruppe zum Screening nach Beendigung der Screeningphase: ja ▫ 1 Screening 1986 nach Abschluss der 2 Screeningrunden der Interventionsgruppe
Swedish Two County	▪ Mammografie: alle 2 Jahre ▪ Anzahl der Screeningrunden: 4 ▪ Aufnahmetechnik: Niedrigdosis-Mammografie in einer Ansicht (mediolateral oblique) ▪ Anzahl der befundenden Personen: 1 Radiologin bzw. 1 Radiologe ▪ zusätzliche klinische Untersuchung: k. A. ▪ Selbstuntersuchung: k. A.	▪ kein Screening, Grundversorgung ▪ Selbstuntersuchung: k. A. ▪ Einladung der Kontrollgruppe zum Screening nach Beendigung der Screeningphase: ja ▫ 1 Screening nach Abschluss der Screeningphase der Interventionsgruppe
	Ab 1986 schrittweise Einführung des nationalen Screenings für Frauen im Alter zwischen 40 und 74 Jahren	

Tabelle 14: Fragestellung 1 – Charakterisierung der Interventionen in den eingeschlossenen Studien (mehrseitige Tabelle)

Studie	Intervention	Vergleich
UK Age	▪ Mammografie: jährlich[a] bis zum Erreichen des 48. Lebensjahres ▪ Anzahl der Screeningrunden: 8 ▪ Aufnahmetechnik: 2-Ansicht-Mammografie für die erste Aufnahme, danach Einzelansicht ▪ Anzahl der befundenden Personen: 2 ▪ zusätzliche klinische Untersuchung: k. A. ▪ Selbstuntersuchung: k. A. Alle Studienteilnehmerinnen wurden ab einem Alter zwischen 50 und 52 Jahren zum Mammografie-Screening im Rahmen des nationales Screeningprogramms NHSBSP eingeladen.	▪ kein Screening, Grundversorgung ▪ Selbstuntersuchung: k. A. ▪ Einladung der Kontrollgruppe zum Screening nach Beendigung der Screeningphase: ja
a. Frauen, die ihren Termin innerhalb von 2 Monaten nicht wahrnehmen konnten, wurden erst wieder zum nächsten jährlichen Screening eingeladen. CNBSS: Canadian National Breast Screening Study; HIP: Health Insurance Plan of Greater New York; k. A.: keine Angaben; MMST: Malmö Mammographic Screening Trial; NHSBSP: National Health Service Breast Screening Programme		

Tabelle 15: Fragestellung 1 – Ein- / Ausschlusskriterien für die zu rekrutierenden Frauen

Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
CNBSS	▪ Frauen im Alter von 40 bis 49 Jahren (CNBSS 1)	▪ früheres Mammakarzinom ▪ Mammografie in den letzten 12 Monaten ▪ bestehende Schwangerschaft
Edinburgh	▪ Frauen aus Edinburgh im Alter von 45 bis 64 Jahren	▪ bereits Mammakarzinom diagnostiziert
Göteborg	▪ Frauen aus Göteborg im Alter von 39 bis 59 Jahren	▪ früheres Mammakarzinom[a]
HIP	▪ mind. 1 Jahr Mitgliedschaft bei HIP-Versicherung ▪ Alter: 40–64 Jahre	▪ k. A.
Malmö Malmö (MMST 1 und 2)	▪ 1977–1978 in Malmö lebende Frauen im Alter von 44 bis 70 Jahren (MMST 1) ▪ 1978–1990 in Malmö lebende Frauen im Alter von 44 bis 49 Jahren (MMST 2)	▪ k. A.
Stockholm	▪ Frauen aus dem südöstlichen Teil von Stockholm, die zum Studienbeginn 1981 zwischen 40 und 64 Jahren alt waren	▪ k. A.
Swedish Two County	▪ Frauen aus den Provinzen Dalarna (ehemals Kopparberg) und Östergötland, die zum Zeitpunkt des Studienbeginns zwischen 40 und 74 Jahre alt waren	▪ k. A.
UK Age	▪ Frauen im Alter von 39 bis 41 Jahren	▪ k. A.

a. Der Ausschluss erfolgte erst nach Randomisierung.
b. abweichende Angaben zwischen den Publikationen

CNBSS: Canadian National Breast Screening Study; HIP: Health Insurance Plan; k. A.: keine Angabe; MMST: Malmö Mammographic Screening Trial

Tabelle 16: Fragestellung 1 – Charakterisierung Teilnehmerfluss innerhalb der Studien (mehrseitige Tabelle)

Studie **Alter**	**N**	**N (%)** **Screeningadhärenz**	**N (%)** **Kontamination**	**N (%)** **Wieder-einbestellungs-rate der gescreenten Teilnehmerinnen**	**N (%) invasive Abklärungs-diagnostik der Teilnehmerinnen**
CNBSS 40–49 Jahre	45–49 Jahre: 23 803[a]	Altersgruppe 40–49 Jahre: Screeningrunde 1: I: k. A. (100) weitere Screeningrunden: I: k. A. (89–86)	k. A.	Altersgruppe 40–49 Jahre: k. A. (7)[b]	Altersgruppe 40–49 Jahre: Screeningrunde 1: I: k. A. (7,2)[a, c] V: k. A. (5,2)[a, c, d] Screeningrunde 2: I: k. A. (4,1)[a, c] Screeningrunde 3: I: k. A. (3,5)[a, c] Screeningrunde 4: I: k. A. (3,2)[a, c] Screeningrunde 5: I: k. A. (3,5)[a, c]
Edinburgh 45–64 Jahre	45–49 Jahre: 11 723	Altersgruppe 45–49 Jahre: Screeningrunde 1: k. A. (63,8) Screeningrunde 2: k. A. (56,8) Screeningrunde 3: k. A. (55,9)	k. A.	k. A.	k. A.
Göteborg 39–59 Jahre	39–59 Jahre: 52 222 39–49 Jahre: 26 113	Altersgruppe 39–49 Jahre: Screeningrunde 1: I: 9921 (85) V: 9167 (66) Screeningrunde 2: I: 9157 (78) Screeningrunde 3: I: 9150 (79) Screeningrunde 4: I: 8914 (75) Screeningrunde 5: I: 9167 (66)	k. A.	Altersgruppe 39–49 Jahre: Screeningrunde 1: I: 582 (5,9)[e] V: 228 (2,5)[e] Screeningrunde 2: I: 343 (3,7)[e] Screeningrunde 3: I: 277 (3,0)[e] Screeningrunde 4: I: 157 (1,8)[e] Screeningrunde 5: I: 183 (2,1)[e]	Altersgruppe 39–49 Jahre: Screeningrunde 1: I: 148 (1,5)[f] V: 123 (1,3)[f] Screeningrunde 2: I: 123 (1,3)[f] Screeningrunde 3: I: 138 (1,5)[f] Screeningrunde 4: I: 59 (0,7)[f] Screeningrunde 5: I: 123 (1,3)[f]
HIP 40–64 Jahre	40–64 Jahre: 61 004	Screeningrunde 1: ca. 20 200 (65) Screeningrunde 2–4: k. A. (k. A.)	k. A.	k. A.	Biopsien: I: 624 (k. A.)

Tabelle 16: Fragestellung 1 – Charakterisierung Teilnehmerfluss innerhalb der Studien (mehrseitige Tabelle)

Studie **Alter**	**N**	**N (%)** **Screeningadhärenz**	**N (%)** **Kontamination**	**N (%)** **Wieder-einbestellungs-rate der gescreenten Teilnehmerinnen**	**N (%) invasive Abklärungs-diagnostik der Teilnehmerinnen**
Malmö MMST 1: 44–70 Jahre MMST 2: 44–49 Jahre	Screening-runde 1 MMST 1: I: 3954 V: 4030 MMST 2: I:9574 V: 8212	Screeningrunde 1: I: MMST 1: k. A. (74) I: MMST 2: k. A. (75–80)	k. A.	k. A. (4)	k. A.
Stockholm 40–64 Jahre	40–49 Jahre: 21 945[a] (I: 14 842, V: 7103)	Altersgruppe 45–49 Jahre: Screeningrunde 1: I: 5530 (80,4) V: k. A. (77,0) Screeningrunde 2: I: k. A. (80,0)	k. A.	1655 (5,1)	Screeningrunde 1: 482 (1,5)[g] Screeningrunde 2: 246 (0,8)[g]
Swedish Two County 40–49 Jahre	I: 19 844 V: 15 604	Altersgruppe 40–49 Jahre: Screeningrunde 1: I: k. A. (93,2) V: k. A. (90,4) Screeningrunde 2: I: k. A. (89,2) Screeningrunde 3 I: k. A. (88,4)	Altersgruppe 40–74 Jahre: k. A. (13)	Altersgruppe 40–74 Jahre: Screeningrunde 1: I: k. A. (5,0) V: k. A. (4,6) Screeningrunde 2: I: k. A. (2,5)	Altersgruppe 40–74 Jahre: Screeningrunde 1: I: k. A. (1,2)[h] V: k. A: (0,9)[h] Screeningrunde 2: I: k. A. (0,4)[h]
UK Age 39–41 Jahre	160 836 (I: 53 914, V: 107 007)	Screeningrunde 1: 36 622 (68,1)	über alle Screeningrunden: k. A. (18[i] / 8,4[j])	über alle Screeningrunden: 7893 (14,6)[k]	Screeningrunde 1: 323[l] (k. A.) 237 (0,5) weitere Screeningrunde: 560[m] (k. A.)

Tabelle 16: Fragestellung 1 – Charakterisierung Teilnehmerfluss innerhalb der Studien (mehrseitige Tabelle)

Studie Alter	N	N (%) Screeningadhärenz	N (%) Kontamination	N (%) Wieder-einbestellungs-rate der gescreenten Teilnehmerinnen	N (%) invasive Abklärungs-diagnostik der Teilnehmerinnen
a. eigene Berechnung b. Anteil der Frauen, die eine erneute Mammografie für diagnostische Zwecke erhielten c. Anzahl der tatsächlich durchgeführten invasiven diagnostischen Prozeduren. Die invasive Abklärungsdiagnostik umfasste Flüssig- oder Gewebebiopsien, Nadelbiopsien sowie offene chirurgische Biopsien. d. Teilnehmerinnen in der Kontrollgruppe erhielten nur vor Randomisierung eine initiale körperliche Untersuchung, sodass keine weiteren Daten zur Abklärungsdiagnostik in weiteren Screeningrunden für diese Gruppe berichtet wurden. e. Anzahl der Frauen, die eine erneute Mammografie nach auffälligem Befund erhielten f. Anzahl der Frauen, bei denen eine körperliche Untersuchung und eine Feinnadelbiopsie nach auffälligem Befund durchgeführt wurde g. Die invasive Abklärungsdiagnostik umfasste Feinnadelbiopsien und stereotaktische Punktionen, bei unklarer histopathologischer Diagnose auch offene chirurgische Biopsien. h. Anzahl der durchgeführten chirurgischen Biopsien i. Anzahl der Frauen, die eine Mammografie aufgrund einer auffälligen Symptomatik erhielten j. Anzahl der Frauen, die eine Screeningmammografie erhielten k. Anteil der falsch-positiven Fälle l. eigene Berechnung; Anzahl der Frauen mit einem falsch-positiven Befund, die eine Zytologie (n = 237) oder eine chirurgische oder nicht chirurgische Biopsie erhielten (n = 86) m. eigene Berechnung; Anzahl der Frauen mit einem falsch-positiven Befund, die eine Zytologie (n = 346) oder eine chirurgische oder nicht chirurgische Biopsie erhielten (n = 214) CNBSS: Canadian National Breast Screening Study; HIP: Health Insurance Plan of Greater New York; I: Intervention; k. A.: keine Angabe; MMST: Malmö Mammographic Screening Trial; N: Anzahl randomisierter (bzw. eingeschlossener) Frauen; V: Vergleich					

A3.2.2 Fragestellung 2 – Studiendesign und Studienpopulationen

Tabelle 17: Fragestellung 2 – Charakterisierung der eingeschlossenen Studien

Studie	Studien-design	Personenzahl N Gesamtpopulation / Alterskohorte[a]	Ort und Zeitraum der Rekrutierung	Dauer der Screeningphase / Beobachtungdauer (seit Randomisierung)	Relevante Endpunkte
Malmö (MMST1)	RCT	42 283 / 70 Jahre: 587[b]	Malmö, Schweden, 1976–1978	10 Jahre / 23 Jahre	Brustkrebsmortalität
Swedish Two County	Cluster-RCT, multi-zentrisch	40–74 Jahre: 133 065[b,c] / 70–74 Jahre: 17 646[b]	Dalarna (vormals Kopparberg) und Östergötland, Schweden, 1977–1980	6–7 Jahre / k. A.	Gesamtmortalität Brustkrebsmortalität, Brustkrebsinzidenz, chirurgische Biopsien mit benignem Befund

a. Gesamtpopulation der in der Studie eingeschlossenen Frauen (Altersspanne: n) / für die Fragestellung relevante Alterskohorte (Altersspanne: n)
b. eigene Berechnung
c. abweichende Angaben zwischen den Publikationen
k. A.: keine Angabe; MMST: Malmö Mammographic Screening Trial; N: Anzahl randomisierter bzw. eingeschlossener Frauen; RCT: randomisierte kontrollierte Studie

Tabelle 18: Fragestellung 2 – Charakterisierung der Interventionen in den eingeschlossenen Studien

Studie	Intervention	Vergleich
Malmö (MMST 1)	▪ Mammografie: alle 18–24 Monate ▪ Anzahl der Screeningrunden: 6 ▪ Aufnahmetechnik: 2-Ansicht-Mammografie für die 1. Aufnahme; in folgenden Screeningrunden 1 oder 2 Ansichten in Abhängigkeit von der Dichte des Brustgewebes ▪ Anzahl der befundenden Personen: 2 ▪ zusätzliche klinische Untersuchung: k. A. ▪ Selbstuntersuchung: k. A. Einladung zum nationalen Screening für alle Frauen geboren zwischen 1923 und 1932 nach Studienende 1990; kein Screening für Frauen, die 71 Jahre und älter sind	▪ Selbstuntersuchung: k. A. ▪ Einladung der Kontrollgruppe zum Screening nach Beendigung der Screeningphase: ja ▫ MMST 1 und 2: Einladung zum nationalen Screening in der Kontrollgruppe für Frauen geboren ab 1923 nach Studienende 1990
Swedish Two County	▪ Mammografie: alle 33 Monate (bezogen auf Frauen ab 50 Jahren) ▪ Anzahl der Screeningrunden: 2 ▪ Aufnahmetechnik: Niedrigdosis-Mammografie in einer Ansicht (mediolateral oblique) ▪ Anzahl der befundenden Personen: 1 Radiologin bzw. 1 Radiologe ▪ zusätzliche klinische Untersuchung: k. A. ▪ Selbstuntersuchung: k. A. Ab 1986 schrittweise Einführung des nationalen Screenings für Frauen im Alter zwischen 40 und 74 Jahren	▪ kein Screening, Grundversorgung ▪ Selbstuntersuchung: k. A. ▪ Einladung der Kontrollgruppe zum Screening nach Beendigung der Screeningphase: ja ▫ 1 Screening nach Abschluss der Screeningphase der Interventionsgruppe

k. A.: keine Angaben; MMST: Malmö Mammographic Screening Trial

Tabelle 19: Fragestellung 2 – Ein- / Ausschlusskriterien für die zu rekrutierenden Frauen

Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Malmö (MMST 1)	▪ 1977–1978 in Malmö lebende Frauen, im Alter von 45 bis 70 Jahren[a]	▪ k. A.
Swedish Two County	▪ Frauen aus den Provinzen Kopparberg (1998 Umbenennung in Dalarna) und Östergötland, die zum Zeitpunkt des Studienbeginns zwischen 40 und 74 Jahre alt waren	▪ k. A.
a. abweichende Angaben zwischen den Publikationen k. A.: keine Angabe; MMST: Malmö Mammographic Screening Trial		

Tabelle 20: Fragestellung 2 – Charakterisierung Teilnehmerfluss innerhalb der Studien

Studie Alter	N	N (%) Screening-adhärenz	N (%) Kontamination	N (%) Wiedereinbestellungs-rate der gescreenten Teilnehmerinnen	N (%) invasive Abklärungs-diagnostik der Teilnehmerinnen
Malmö (MMST 1) 45–70 Jahre	Screening-runde 1 I: 3954 V: 4030	Screeningrunde 1: I: k. A. (74)	k. A.	k. A. (4)	k. A.
Swedish Two County 70–74 Jahre	I: 10 339 V: 7307	Altersgruppe 70–74 Jahre: Screeningrunde 1: I: k. A. (78,7) V: k. A. (k. A.) Screeningrunde 2: I: k. A. (66,8) Screeningrunde 3 I: k. A. (k. A.)	Altersgruppe 40–74 Jahre: k. A. (13 %)	Altersgruppe 40–74 Jahre: Screeningrunde 1: I: k. A. (5,0) V: k. A. (4,6) Screeningrunde 2: I: k. A. (2,5)	Altersgruppe 40–74 Jahre: Screeningrunde 1: I: k. A. (1,2)[a] V: k.A: (0,9)[a] Screeningrunde 2: I: k. A. (0,4)[a]
a. Anzahl der durchgeführten chirurgischen Biopsien I: Intervention; k. A.: keine Angabe; MMST: Malmö Mammographic Screening Trial; N: Anzahl randomisierter (bzw. eingeschlossener) Frauen; V: Vergleich					

A3.2.3 Bewertung endpunktübergreifender Kriterien des Verzerrungspotenzials

Die Einschätzung endpunktübergreifender Kriterien des Verzerrungspotenzials ist in der folgenden Tabelle 21 dargestellt.

Tabelle 21: Endpunktübergreifendes Verzerrungspotenzial

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Studienebene[a]
			Personen	Behandelnde Personen			
Fragestellung 1							
CNBSS	ja	unklar	n. b.	n. b.	n. b.	n. b.	hoch
Edinburgh	unklar	unklar	n. b.	n. b.	n. b.	n. b.	hoch
Göteborg	unklar[b]	unklar	n. b.	n. b.	n. b.	n. b.	hoch
HIP	unklar	unklar	n. b.	n. b.	n. b.	n. b.	hoch
Malmö	ja	unklar	n. b.	n. b.	n. b.	n. b.	hoch
Stockholm	nein[c]	n. b.[c]	n. b.	n. b.	n. b.	n. b.	hoch
Swedish Two County	unklar	unklar	n. b.	n. b.	n. b.	n. b.	hoch
UK Age	ja	ja	nein	nein	ja	ja	niedrig
Fragestellung 2							
Malmö	ja	unklar	n. b.	n. b.	n. b.	n. b.	hoch
Swedish Two County	unklar	unklar	n. b.	n. b.	n. b.	n. b.	hoch

a. Sobald sich aus der Bewertung der Kriterien „Adäquate Erzeugung der Randomisierungssequenz“ und / oder „Verdeckung der Gruppenzuteilung“ ein endpunktübergreifend hohes Verzerrungspotenzial der Ergebnisse ergibt, wird auf eine Bewertung der übrigen Kriterien verzichtet.
b. 18 % der Frauen wurden anhand des Geburtstags quasirandomisiert, und es ist unklar, wie die je Geburtsjahrgang unterschiedlichen Randomisierungsverhältnisse erzeugt wurden
c. Die Frauen wurden anhand ihres Geburtstags den Gruppen zugeteilt, und die Studie wurde daher als quasi-randomisiert eingeschätzt.

CNBSS: Canadian National Breast Screening Study; HIP: Health Insurance Plan of Greater New York; n. b.: nicht bewertet

A3.3 Fragestellung 1 – Altersgruppe 45 bis 49 Jahre

A3.3.1 Mortalität

A3.3.1.1 Gesamtmortalität

Verzerrungspotenzial der Ergebnisse zur Gesamtmortalität

In der folgenden Tabelle 22 wird die Bewertung des Verzerrungspotenzials der Ergebnisse zur Gesamtmortalität dargestellt.

Tabelle 22: Fragestellung 1 – endpunktspezifisches Verzerrungspotenzial: Gesamtmortalität

Studie	Endpunktübergreifendes Verzerrungspotenzial	Verblindung Endpunkt-erheber	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Endpunktspezifisches Verzerrungspotenzial[a]
CNBSS	hoch	n. b.	n. b.	n. b.	n. b.	hoch
Göteborg	hoch	n. b.	n. b.	n. b.	n. b.	hoch
Swedish Two County	hoch	n. b.	n. b.	n. b.	n. b.	hoch
UK Age	niedrig	unklar	ja	ja	ja	niedrig

a. Falls ein hohes endpunktübergreifendes Verzerrungspotenzial der Ergebnisse vorliegt, überträgt sich diese Bewertung direkt auf das endpunktspezifische Verzerrungspotenzial der Ergebnisse. In diesen Fällen wird auf die Bewertung endpunktspezifischer Kriterien verzichtet.

CNBSS: Canadian National Breast Screening Study; ITT: Intention to treat; n. b.: nicht bewertet

Ergebnisse zur Gesamtmortalität

Die Ergebnisse zur Gesamtmortalität der Studien mit verwertbaren Daten sind in Tabelle 23 dargestellt. Die metaanalytische Zusammenfassung dieser Ergebnisse ist in Abbildung 2 dargestellt. In den Abbildungen zeigen die Angaben in eckigen Klammern hinter dem Studiennamen jeweils den für die Auswertung ausgewählten Zeitpunkt an, sofern in einer Studie Daten zu mehreren Zeitpunkten vorlagen.

Tabelle 23: Fragestellung 1 – Ergebnisse: Gesamtmortalität (mehrseitige Tabelle)

Studie Alter Zeitpunkt[a]	Intervention					Vergleich					Intervention vs. Vergleich		
	N	Frauen mit Ereignissen		Personen-jahre	Ereignisse pro 10 000 Personen-jahre	N	Frauen mit Ereignissen		Personen-jahre	Ereignisse pro 10 000 Personen-jahre	IDR	95 %-KI	p-Wert
		n	%[b]				n	%[b]					
CNBSS													
40–49 Jahre													
8,5 Jahre (Max.) / ca. 3,5 Jahre													
	25 214	159	0,63	k. A.	k. A.	25 216	156	0,62	k. A.	k. A.	1,02[c, d]	[0,82; 1,27][d]	0,914[d]
13 Jahre (Max.) / 8 Jahre (Max.)													
	25 214	413	1,64	k. A.	k. A.	25 216	413	1,64	k. A.	k. A.	1,00[c, d]	[0,87; 1,14][d]	> 0,999[d]
Göteborg													
39–49 Jahre													
11 Jahre (Max.) / ca. 4 Jahre													
	11 792	409	3,47	138 402	29,55	14 321	506	3,53	168 025	30,11	0,98	[0,86; 1,12]	k. A.
Swedish Two County													
40–49 Jahre													
7,9 Jahre (MW) / 0 Jahre													
	19 844[e]	392[e]	1,98	k. A.	k. A.	15 604[e]	301[e]	1,93	k. A.	k. A.	1,02[d]	[0,88; 1,19][d]	0,777[d]
UK Age													
39–41 Jahre													
10,7 Jahre (MW) / ca. 3 Jahre													
	53 884	960	1,78	578 390	16,60	106 956	1975	1,85	1 149 380	17,18	0,97	[0,89; 1,04]	k. A.
17,7 Jahre (Median) / ca. 10 Jahre													
	53 883	2127	3,95	940 969	22,60	106 953	4320	4,04	1 868 717	23,12	0,98	[0,93; 1,03]	k. A.
22,8 Jahre (Median) / ca. 15 Jahre													
	53 883	3507	6,51	1 201 010	29,20	106 953	6932	6,48	2 385 006	29,06	1,01	[0,96; 1,05]	0,66

Tabelle 23: Fragestellung 1 – Ergebnisse: Gesamtmortalität (mehrseitige Tabelle)

Studie Alter Zeitpunkt[a]	Intervention					Vergleich					Intervention vs. Vergleich		
	N	Frauen mit Ereignissen		Personen-jahre	Ereignisse pro 10 000 Personen-jahre	N	Frauen mit Ereignissen		Personen-jahre	Ereignisse pro 10 000 Personen-jahre	IDR	95 %-KI	p-Wert
		n	%[b]				n	%[b]					

a. Beobachtungszeit seit Randomisierung / seit Ende der Screeningphase
b. Die Prozentangaben in dieser Tabelle wurden selbst berechnet.
c. RR
d. eigene Berechnung von RR, 95 %-KI (asymptotisch) und p-Wert (unbedingter exakter Test, CSZ-Methode nach [101])
e. eigene Berechnung

CNBSS: Canadian National Breast Screening Study; IDR: Inzidenzdichtequotient; k. A.: keine Angabe; KI: Konfidenzintervall; n: Anzahl Frauen mit Ereignis; N: Anzahl ausgewerteter Frauen; Max.: Maximum; MW: Mittelwert; RR: relatives Risiko; vs.: versus

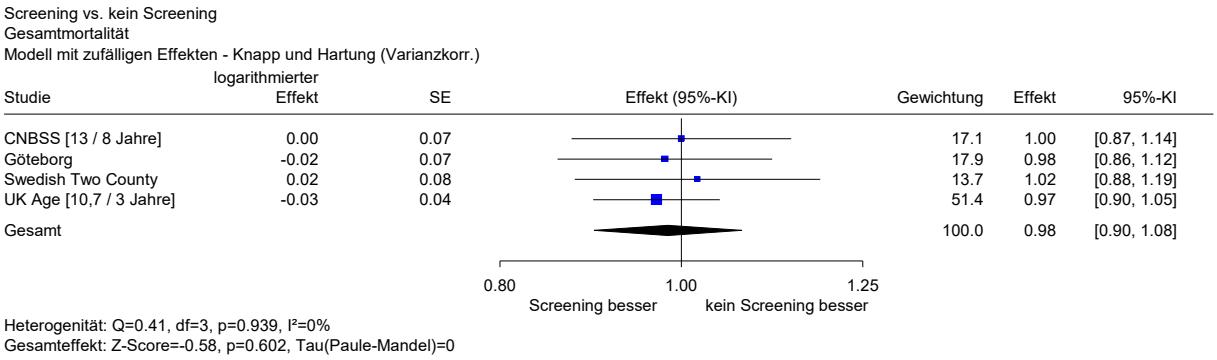


Abbildung 2: Fragestellung 1 – Gesamtmortalität

Sensitivitätsanalysen

Es wurden keine Sensitivitätsanalysen durchgeführt.

Subgruppenmerkmale und andere Effektmodifikatoren

Als potenzielle Effektmodifikatoren wurden das Screeningintervall (≤ 1,5 Jahre versus > 1,5 Jahre) und die Screeningphasendauer (≤ 5 Jahre versus > 5 Jahre) untersucht. Es zeigte sich keine Effektmodifikation (Ergebnisse nicht dargestellt).

A3.3.1.2 Brustkrebsspezifische Mortalität

Verzerrungspotenzial der Ergebnisse zur brustkrebsspezifischen Mortalität

In der folgenden Tabelle 24 wird die Bewertung des Verzerrungspotenzials der Ergebnisse zur brustkrebsspezifischen Mortalität dargestellt.

Tabelle 24: Fragestellung 1 – endpunktspezifisches Verzerrungspotenzial: brustkrebsspezifische Mortalität

Studie	Endpunktübergreifendes Verzerrungspotenzial	Verblindung Endpunkt-erheber	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Endpunktspezifisches Verzerrungspotenzial[a]
CNBSS	hoch	n. b.	n. b.	n. b.	n. b.	hoch
Edinburgh	hoch	n. b.	n. b.	n. b.	n. b.	hoch
Göteborg	hoch	n. b.	n. b.	n. b.	n. b.	hoch
HIP	hoch	n. b.	n. b.	n. b.	n. b.	hoch
Malmö	hoch	n. b.	n. b.	n. b.	n. b.	hoch
Stockholm	hoch	n. b.	n. b.	n. b.	n. b.	hoch
Swedish Two County	hoch	n. b.	n. b.	n. b.	n. b.	hoch
UK Age	niedrig	unklar	ja	ja	ja	niedrig

a. Falls ein hohes endpunktübergreifendes Verzerrungspotenzial der Ergebnisse vorliegt, überträgt sich diese Bewertung direkt auf das endpunktspezifische Verzerrungspotenzial der Ergebnisse. In diesen Fällen wird auf die Bewertung endpunktspezifischer Kriterien verzichtet.

CNBSS: Canadian National Breast Screening Study; HIP: Health Insurance Plan of Greater New York; ITT: Intention to treat; n. b.: nicht bewertet

Ergebnisse zur brustkrebsspezifischen Mortalität

Die Ergebnisse zur brustkrebsspezifischen Mortalität der Studien mit verwertbaren Daten sind in Tabelle 25 dargestellt. Die metaanalytische Zusammenfassung dieser Ergebnisse nach ca. 5 Jahren, 10 Jahren, 15 Jahren und 20 Jahren ist in Abbildung 3 bis Abbildung 6 dargestellt.

Tabelle 25: Fragestellung 1 – Ergebnisse: brustkrebsspezifische Mortalität (mehrseitige Tabelle)

Studie **Alter** **Zeitpunkt**[a] *Auswertungsart*	**Intervention**					**Vergleich**					**Intervention vs. Vergleich**		
	N	**Frauen mit Ereignissen**		**Personen-jahre**	**Ereignisse pro 10 000 Personen-jahre**	**N**	**Frauen mit Ereignissen**		**Personen-jahre**	**Ereignisse pro 10 000 Personen-jahre**	**IDR**	**95 %-KI**	**p-Wert**
		n	**%**[b]				**n**	**%**[b]					
CNBSS **40–49 Jahre** **8,5 Jahre (MW) / ca. 3,5 Jahre**													
Follow-up-Modell	25 214	38	0,15	k. A.	1,47	25 216	28	0,11	k. A.	1,04	1,36[c]	[0,84; 2,21]	k. A.
16 Jahre (Max.) / 11 Jahre (Max.)													
Evaluationsmodell o. A., Brustkrebsfälle bis Jahr 7	25 214	86	0,34	282 606	3,04	25 216	82	0,33	282 575	2,90	1,05	[0,78; 1,42]	k. A.
Evaluationsmodell o. A., Brustkrebsfälle bis Jahr 9	25 214	99	0,39	282 606	3,50	25 216	97	0,38	282 575	3,43	1,02	[0,77; 1,35]	k. A.
Follow-up-Modell	25 214	105	0,42	282 606	3,72	25 216	108	0,43	282 575	3,82	0,97	[0,74; 1,27]	k. A.
13 Jahre (Max.) / 8 Jahre (Max.)													
Follow-up-Modell	25 214	84	0,33	k. A.	k. A.	25 216	75	0,30	k. A.	k. A.	1,12	[0,82; 1,53]	k. A.
Edinburgh **45–49 Jahre (Kohorte 1)** **7 Jahre / 0 Jahre**													
n. a., da Screeningphase gerade abgeschlossen	5913	13	0,22	40 851	3,18	5810	13	0,22	40 009	3,25	0,98[d]	[0,45; 2,11][d]	0,958[d]
14 Jahre / 7 Jahre													
Follow-up-Modell	5777	27	0,47	78 761	3,43	5594	33	0,59	75 726	4,36	0,78[e]	[0,46; 1,32]	k. A.
45–49 Jahre (Kohorten 1, 2 und 3) **10–14 Jahre / 3–7 Jahre**													
Follow-up-Modell	11 479[f]	47	0,41	139 868[f]	3,35	10 267[f]	53	0,52	126 413[f]	4,19	0,83[g]	[0,54; 1,27]	k. A.

Tabelle 25: Fragestellung 1 – Ergebnisse: brustkrebsspezifische Mortalität (mehrseitige Tabelle)

Göteborg **39–49 Jahre** **11 Jahre (Max.) / ca. 4 Jahre**												
Evaluationsmodell[h]	11 792	18	0,15	138 402	1,30[f]	14 321	40	0,28	168 025	2,38[f]	0,55	[0,31; 0,96]	k. A.
13 Jahre (Max.) / ca. 6 Jahre												
Evaluationsmodell	11 792	25	0,21	k. A.	k. A.	14 321	46	0,32	k. A.	k. A.	0,65[c]	[0,40; 1,05]	k. A.
Follow-up-Modell[h]	11 792	34	0,29	k. A.	k. A.	14 321	59	0,41	k. A.	k. A.	0,69[c]	[0,45; 1,05]	k. A.
24 Jahre (Median) / 17 Jahre												
Evaluationsmodell	11 792	33	0,28	k. A.	k. A.	14 321	68	0,47	k. A.	k. A.	0,60[c]	[0,43; 0,85]	0,003
45–49 Jahre **13 Jahre (Max.) / ca. 6 Jahre**												
Evaluationsmodell	5157	10	0,19	k. A.	k. A.	5995	22	0,37	k. A.	k. A.	0,51[c]	[0,24; 1,08]	k. A.
Follow-up-Modell[h]	5157	17	0,33	k. A.	k. A.	5995	29	0,48	k. A.	k. A.	0,67[c]	[0,37; 1,23]	k. A.
HIP **40–49 Jahre**[i] **5 Jahre / 1,5 Jahre**												
Evaluationsmodell o. A.	13 740[f]	19	0,14	k. A.	k. A.	13 740	20	0,15	k. A.	k. A.	0,95[j]	[0,51; 1,78][j]	0,922[j]
9 Jahre / 5,5 Jahre												
Evaluationsmodell o. A., Brustkrebsfälle bis Jahr 5	13 740[f]	39	0,28	k. A.	k. A.	13 740	48	0,35	k. A.	k. A.	0,81[j]	[0,53; 1,24][j]	0,496[j]
14 Jahre / 10,5 Jahre												
Evaluationsmodell o. A., Brustkrebsfälle bis Jahr 5	13 740[f]	46	0,33	k. A.	k. A.	13 740	61	0,44	k. A.	k. A.	0,75[j]	[0,51; 1,10][j]	0,152[j]
18 Jahre / 14,5 Jahre												
Evaluationsmodell o. A., Brustkrebsfälle bis Jahr 5	13 740[f]	50	0,36	k. A.	k. A	13 740	66	0,48	k. A.	k. A.	0,76[j]	[0,53; 1,09][j]	0,148[j]
45–49 Jahre[i] **5 Jahre / 1,5 Jahre**												
Evaluationsmodell o. A.	7240	10	0,14	k. A.	k. A.	7240	9	0,12	k. A.	k. A.	1,11[j]	[0,45; 2,73][j]	0,883[j]

Tabelle 25: Fragestellung 1 – Ergebnisse: brustkrebsspezifische Mortalität (mehrseitige Tabelle)

14 Jahre / 10,5 Jahre													
Evaluationsmodell o. A., Brustkrebsfälle bis Jahr 5	7240	28	0,39	k. A.	k. A.	7240	35	0,48	k. A.	k. A.	0,80[j]	[0,49; 1,31][j]	0,526[j]
Malmö **44–49 Jahre (Kollektiv: MMST 1 und MMST 2)** **10–15,5 Jahre (MW) / ca. 1–6,5 Jahre**													
Evaluationsmodell	13 528	57	0,42[k]	165 596	3,44	12 242	78	0,64[k]	144 036	5,42[f]	0,64	[0,45; 0,89]	0,009
Stockholm **40–49 Jahre** **7,4 Jahre (MW) / ca. 3 Jahre**													
Evaluationsmodell	14 375	16	0,11	99 155	1,61[f]	7103	8	0,11	54 446	1,47[f]	1,10[d]	[0,47; 2,57][d]	0,829[d]
11,4 Jahre (MW) / ca. 7 Jahre													
Evaluationsmodell	14 842	24	0,16	173 866	1,38[f]	7103	12	0,17	87 826	1,37[f]	1,01[d]	[0,51; 2,02][d]	0,977[d]
Swedish Two County **40–49 Jahre**													
7,9 Jahre (MW) / 0 Jahre													
Evaluationsmodell	19 844	28	0,14	k. A.	k. A.	15 604	24	0,15	k. A.	k. A.	0,92[l]	[0,52; 1,60][l]	0,8[l]
10,8 Jahre (MW) / ca. 4 Jahre													
Evaluationsmodell	19 844	42[f]	0,21	225 801[f]	1,86[f]	15 604	32[f]	0,21	177 231[f]	1,81[f]	1,03[d]	[0,65; 1,63][d]	0,899[d]
13 Jahre (MW) / ca. 6 Jahre													
Evaluationsmodell	19 844	45	0,23	k. A.	k. A.	15 604	39	0,25	k. A.	k. A.	0,87[m]	[0,54; 1,41]	k. A.
20 Jahre / ca. 13 Jahre													
Evaluationsmodell	19 844	57	0,29	k. A.	k. A.	15 604	48	0,31	k. A.	k. A.	0,93[j]	[0,64; 1,37][j]	0,743[j]
UK Age **39–41 Jahre** **4 Jahre / n. a.**													
n. a., da Screeningphase noch nicht abgeschlossen	53 883	27	0,05	267 852	1,01[f]	106 953	69	0,06	532 066	1,30[f]	0,78	[0,50; 1,21]	k. A.

Tabelle 25: Fragestellung 1 – Ergebnisse: brustkrebsspezifische Mortalität (mehrseitige Tabelle)

9 Jahre / 0 Jahre													
n. a., da Screeningphase gerade abgeschlossen	53 883	83	0,15	532 729	1,56[f]	106 953	221	0,21	1 058 236	2,09[f]	0,75	[0,58; 0,96]	k. A.
10,7 Jahre (MW) / ca. 3 Jahre													
Evaluationsmodell	53 884	105	0,19	578 390	1,82[f]	106 956	251	0,23	1 149 380	2,18[f]	0,83	[0,66; 1,04]	k. A.
14 Jahre / ca. 6 Jahre													
Follow-up-Modell	53 883	182	0,34	793 852	2,29f	106 953	406	0,38	1 576 346	2,58[f]	0,89	[0,75; 1,06]	k. A.
17,7 Jahre (Median) / ca. 10 Jahre													
Follow-up-Modell	53 883	242	0,45	940 969	2,57[f]	106 953	515	0,48	1 876 717	2,74[f]	0,93	[0,80; 1,09]	k. A.
22,8 Jahre (Median) / ca. 15 Jahre													
Evaluationsmodell	53 883	216	0,40	1 201 010	1,80[f]	106 953	498	0,47	2 385 006	2,09[f]	0,86	[0,73; 1,01]	0,068
Follow-up-Modell	53 883	338	0,63	1 201 010	2,81	106 953	743	0,69	2 385 006	3,12[f]	0,90	[0,79; 1,03]	k. A.

a. Beobachtungszeit seit Randomisierung / seit Ende der Screeningphase
b. Die Prozentangaben in dieser Tabelle wurden selbst berechnet.
c. relatives Risiko
d. eigene Berechnung von IDR, 95 %-KI (asymptotisch) und p-Wert (asymptotisch)
e. Ergebnis ohne weitere Adjustierung bezüglich des sozioökonomischen Status. Ergebnisse unter Berücksichtigung des sozioökonomischen Status sind sehr ähnlich (Modell 1: IDR 0,73, 95 %-KI [0,43; 1,24], Modell 2: IDR 0,70, 95 %-KI [0,41; 1,20]).
f. eigene Berechnung
g. Ergebnis ohne weitere Adjustierung bezüglich des sozioökonomischen Status. Ergebnisse unter Berücksichtigung des sozioökonomischen Status sind sehr ähnlich (Modell 1: IDR 0,78, 95 %-KI [0,50; 1,21], Modell 2: 0,75, 95 %-KI [0,48; 1,18])
h. Klassifikation der Todesursache gemäß schwedischem nationalem Krebsregister
i. Die Werte für N sind ausschließlich in Habbema 1986 [83], Tabelle 1, angegeben und wurden für diese Ergebnistabelle herangezogen. Entsprechend der Angaben der Autorinnen und Autoren sollen die Werte für die Kontroll- und die Interventionsgruppe vergleichbar sein.
j. eigene Berechnung von RR, 95 %-KI (asymptotisch) und p-Wert (unbedingter exakter Test, CSZ-Methode nach [101])
k. Angaben aus Publikation
l. adjustiert für Region und Alter
m. RR, adjustiert für Region

CNBSS: Canadian National Breast Screening Study; HIP: Health Insurance Plan of Greater New York; IDR: Inzidenzdichtequotient; k. A.: keine Angabe; KI: Konfidenzintervall; Max.: Maximum; MW: Mittelwert; n: Anzahl Frauen mit Ereignis; N: Anzahl ausgewerteter Frauen; n. a.: nicht anwendbar; o. A.: ohne Abschlussscreening; RR: relatives Risiko; vs.: versus

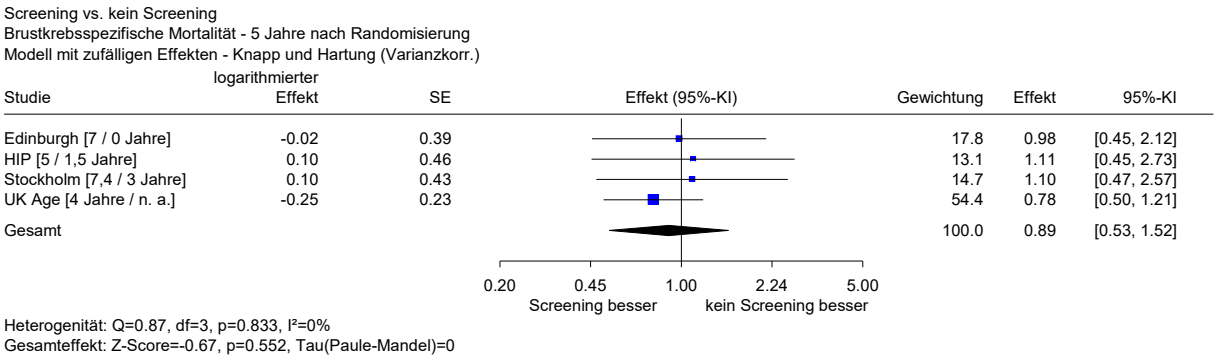


Abbildung 3: Fragestellung 1 – brustkrebsspezifische Mortalität, 5 Jahre nach Randomisierung

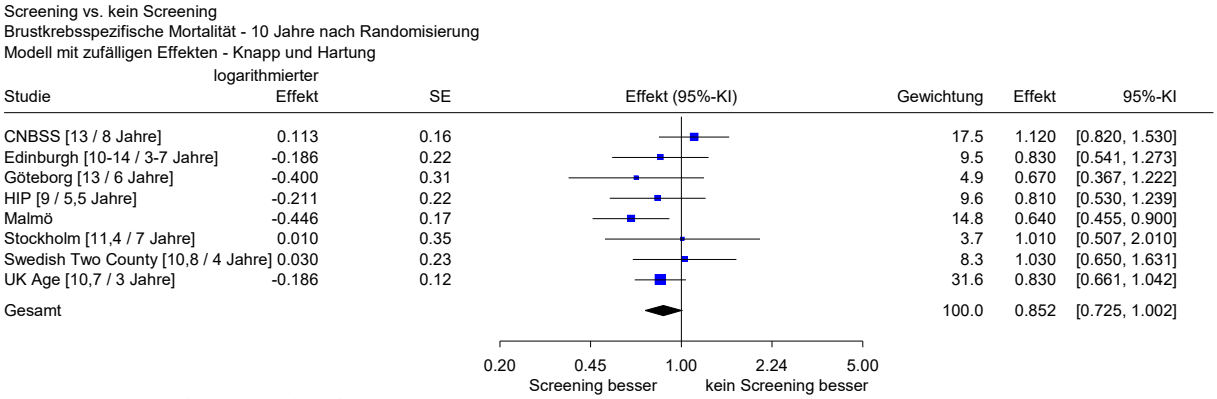


Abbildung 4: Fragestellung 1 – brustkrebsspezifische Mortalität, 10 Jahre nach Randomisierung

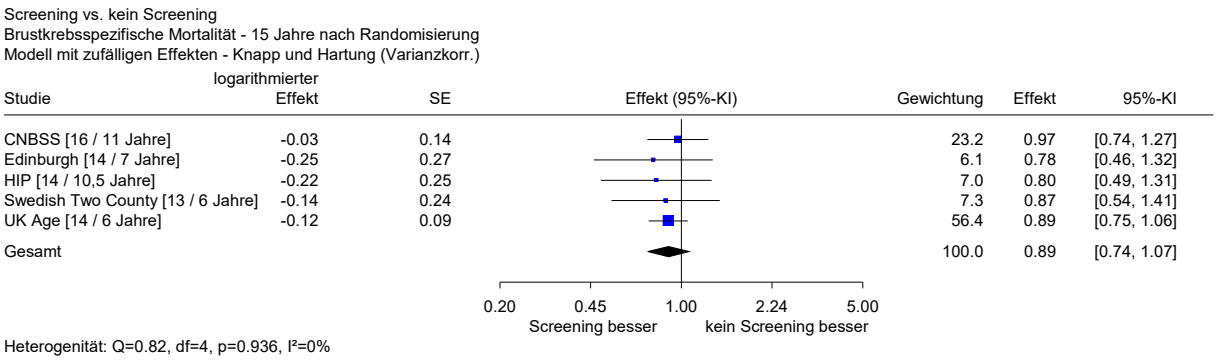


Abbildung 5: Fragestellung 1 – brustkrebsspezifische Mortalität, 15 Jahre nach Randomisierung

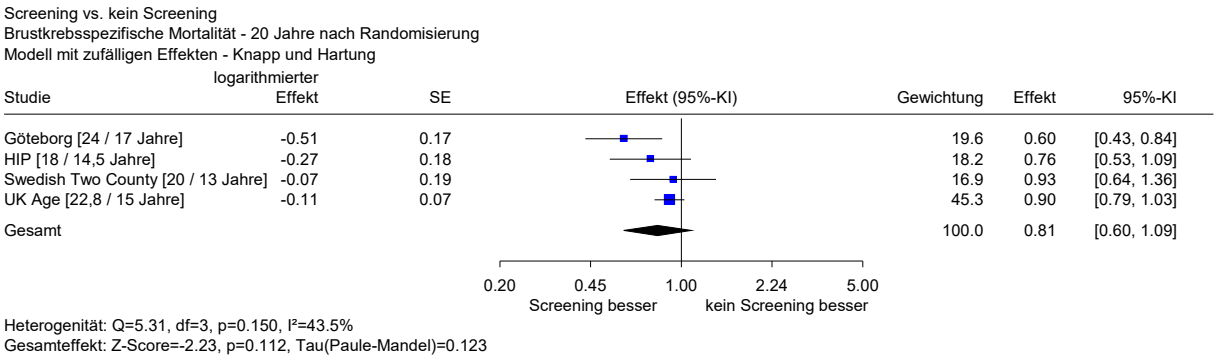


Abbildung 6: Fragestellung 1 – brustkrebsspezifische Mortalität, 20 Jahre nach Randomisierung

Sensitivitätsanalysen

Abbildung 7 zeigt die Sensitivitätsanalyse unter Berücksichtigung der 9-Jahres-Daten der Studie UK Age zum Zeitpunkt ca. 10 Jahre nach Randomisierung.

Abbildung 8 zeigt die Sensitivitätsanalyse ohne Daten der Edinburgh-Studie zum Zeitpunkt ca. 10 Jahre.

Abbildung 9 zeigt die Sensitivitätsanalyse unter Berücksichtigung der Daten der 10-Jahres-Altersgruppe (statt der 5-Jahres-Altersgruppe) der Göteborg-Studie zum Zeitpunkt ca. 10 Jahre.

Entsprechende Sensitivitätsanalysen zu den übrigen Zeitpunkten zeigten – in Übereinstimmung mit den Hauptanalysen – keinen statistisch signifikanten Effekt (nicht dargestellt).

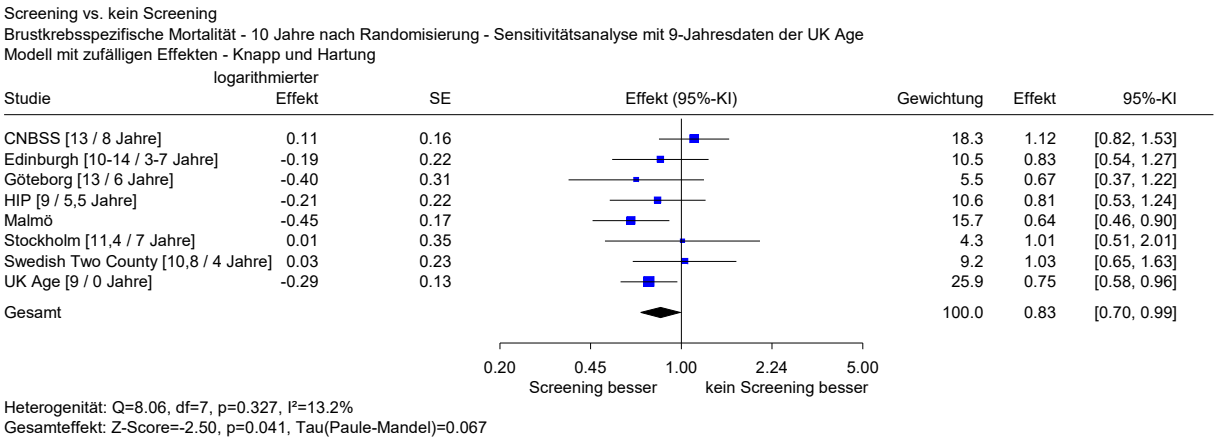


Abbildung 7: Fragestellung 1 – brustkrebsspezifische Mortalität, 10 Jahre nach Randomisierung, Sensitivitätsanalyse mit 9-Jahres-Daten der UK Age

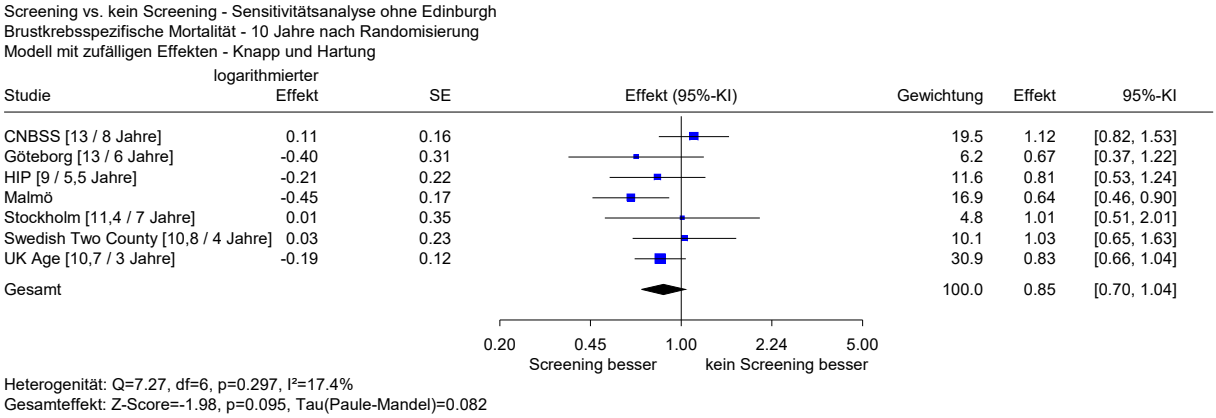


Abbildung 8: Fragestellung 1 – brustkrebsspezifische Mortalität, 10 Jahre nach Randomisierung, Sensitivitätsanalyse ohne Daten der Edinburgh-Studie

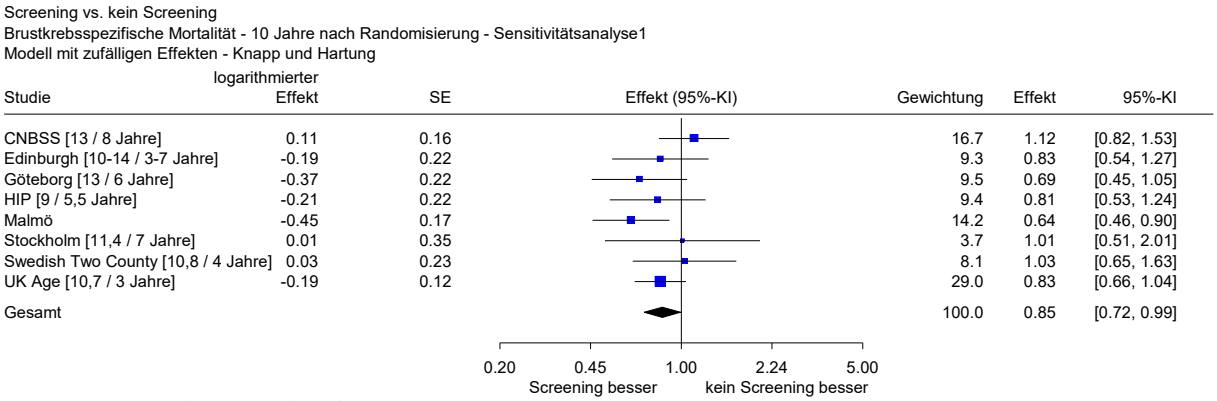


Abbildung 9: Fragestellung 1 – brustkrebsspezifische Mortalität, 10 Jahre nach Randomisierung, Sensitivitätsanalyse mit Daten der 10-Jahres-Altersgruppe der Göteborg-Studie

Subgruppenmerkmale und andere Effektmodifikatoren

Als potenzielle Effektmodifikatoren wurden das Screeningintervall (≤ 1,5 versus > 1,5 Jahre) und die Screeningphasendauer (≤ 5 versus > 5 Jahre) untersucht. Es zeigte sich keine Effektmodifikation (Ergebnisse nicht dargestellt).

A3.3.2 Mastektomien

Verzerrungspotenzial der Ergebnisse zu Mastektomien

In der folgenden Tabelle 26 wird die Bewertung des Verzerrungspotenzials der Ergebnisse zu Mastektomien dargestellt.

Tabelle 26: Fragestellung 1 – endpunktspezifisches Verzerrungspotenzial: Mastektomien

Studie	Endpunktübergreifendes Verzerrungspotenzial	Verblindung Endpunkt-erheber	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Endpunktspezifisches Verzerrungspotenzial[a]
CNBSS	hoch	n. b.	n. b.	n. b.	n. b.	hoch

a. Falls ein hohes endpunktübergreifendes Verzerrungspotenzial der Ergebnisse vorliegt, überträgt sich diese Bewertung direkt auf das endpunktspezifische Verzerrungspotenzial der Ergebnisse. In diesen Fällen wurde auf die Bewertung endpunktspezifischer Kriterien verzichtet.

CNBSS: Canadian National Breast Screening Study; ITT: Intention to treat; n. b.: nicht bewertet

Ergebnisse zu Mastektomien

Die Ergebnisse zu Mastektomien in den Studien mit verwertbaren Daten sind in Tabelle 27 dargestellt.

Tabelle 27: Fragestellung 1 – Ergebnisse: Brustkrebs-Operationen einschließlich Mastektomien

Studie Alter	Zeitpunkt	Intervention			Vergleich			Intervention vs. Vergleich[a]		
		N	Frauen mit Ereignissen		N	Frauen mit Ereignissen		OR	95 %-KI	p-Wert
			n	%[b]		n	%[b]			
CNBSS **40–49 Jahre**										
OPs gesamt[c]	1. Screeningrunde	25 214	98	0,39	25 216	62	0,25	1,58	[1,15; 2,18]	0,004
brusterhaltende OP	1. Screeningrunde	25 214	38	0,15	25 216	26	0,10	1,46	[0,89;2,41]	0,141
Mastektomie	1. Screeningrunde	25 214	60[d]	0,24	25 216	36[e]	0,14	1,67	[1,10;2,52]	0,014
OPs gesamt[c]	7 Jahre	25 214	415	1,65	25 216	313	1,24	1,33	[1,15; 1,54]	< 0,001
brusterhaltende OP	7 Jahre	25 214	232	0,92	25 216	156	0,62	1,49	[1,22;1,83]	< 0,001
Mastektomie	7 Jahre	25 214	183[f]	0,73	25 216	157[g]	0,62	1,17	[0,94;1,45]	0,158

a. eigene Berechnung von OR, 95 %-KI (asymptotisch) und p-Wert (unbedingter exakter Test, CSZ-Methode nach [101])
b. Die Prozentangaben wurden selbst berechnet.
c. brusterhaltende OPs und Mastektomien zusammen
d. Anteil Mastektomien an allen Operationen: 60/98 = 61,2 %
e. Anteil Mastektomien an allen Operationen: 36/62 = 58,1 %
f. Anteil Mastektomien an allen Operationen: 183/415 = 44,1 %
g. Anteil Mastektomien an allen Operationen: 157/313 = 50,2 %

CNBSS: Canadian National Breast Screening Study; KI: Konfidenzintervall; n: Anzahl Frauen mit Ereignis; N: Anzahl ausgewerteter Frauen; OP: Operation; OR: Odds Ratio; vs.: versus

Sensitivitätsanalysen

Es wurden keine Sensitivitätsanalysen durchgeführt.

Subgruppenmerkmale und andere Effektmodifikatoren

Es wurden keine Analysen von Subgruppenmerkmalen und anderen Effektmodifikatoren durchgeführt.

A3.3.3 Konsequenzen aus falsch-positiven Screeningbefunden

Verzerrungspotenzial der Ergebnisse zu Konsequenzen aus falsch-positiven Screeningbefunden

In der folgenden Tabelle 28 wird die Bewertung des Verzerrungspotenzials der Ergebnisse zu Konsequenzen aus falsch-positiven Befunden dargestellt.

Tabelle 28: Fragestellung 1 – endpunktspezifisches Verzerrungspotenzial: Konsequenzen aus falsch-positiven Screeningbefunden

Studie	Endpunktübergreifendes Verzerrungspotenzial	Verblindung Endpunkt-erheber	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Endpunktspezifisches Verzerrungspotenzial[a]
CNBSS	hoch	n. b.	n. b.	n. b.	n. b.	hoch
Göteborg	hoch	n. b.	n. b.	n. b.	n. b.	hoch
HIP	hoch	n. b.	n. b.	n. b.	n. b.	hoch
Stockholm	hoch	n. b.	n. b.	n. b.	n. b.	hoch
Swedish Two County	hoch	n. b.	n. b.	n. b.	n. b.	hoch
UK Age	niedrig	unklar	unklar	ja	ja	hoch

a. Falls ein hohes endpunktübergreifendes Verzerrungspotenzial der Ergebnisse vorliegt, überträgt sich diese Bewertung direkt auf das endpunktspezifische Verzerrungspotenzial der Ergebnisse. In diesen Fällen wird auf die Bewertung endpunktspezifischer Kriterien verzichtet.

CNBSS: Canadian National Breast Screening Study; HIP: Health Insurance Plan of Greater New York; ITT: Intention to treat; n. b.: nicht bewertet

Ergebnisse zu Konsequenzen aus falsch-positiven Screeningbefunden

Die Ergebnisse zu Konsequenzen aus falsch-positiven Screeningbefunden aus den Studien mit verwertbaren Daten sind in Tabelle 29 und Tabelle 30 dargestellt.

Tabelle 29: Fragestellung 1 – Ergebnisse: Konsequenzen aus falsch-positiven Screeningbefunden: Biopsien mit benignem Befund (mehrseitige Tabelle)

Studie Alter	Zeitpunkt	N	Frauen mit Ereignissen	
			n	%
CNBSS **40–49 Jahre**				
Biopsien mit benignem Befund	Jahr 1	25 214	847	3,4
Biopsien mit benignem Befund	Jahr 2	22 424	392[a]	1,8
Biopsien mit benignem Befund	Jahr 3	22 066	289[a]	1,3
Biopsien mit benignem Befund	Jahr 4	21 839	262[a]	1,2
Biopsien mit benignem Befund	Jahr 5	14 146	187[a]	1,3
HIP **40–49 Jahre**				
Biopsien mit benignem Befund	k. A.	k. A.	k. A.	k. A.[d]
Malmö **45–49 Jahre (Kollektiv: MMST 1)**				
Biopsien mit benignem Befund	1. Screeningrunde	3795	12	0,32[a]
44–49 Jahre (Kollektiv: MMST 1 und MMST 2)				
Biopsien mit benignem Befund	10–15,5 Jahre / 1–6,5 Jahre	13 528	93[a]	0,69[a]
Stockholm **45–49 Jahre**				
Biopsien mit benignem Befund	1. Screeningrunde	6875	66	0,96
Biopsien mit benignem Befund	2. Screeningrunde	6875	44	0,64
UK Age **39–41 Jahre**				
Biopsien mit benignem Befund	1. Einladung	35 846[b]	56[c]	0,16[a]
Biopsien mit benignem Befund	1. Screeningrunde bei vorherigen Nichtteilnehmerinnen	7679[b]	31[c]	0,40[a]
Biopsien mit benignem Befund	2. Screeningrunde	39 912[b]	29[c]	0,07[a]
Biopsien mit benignem Befund	3. Screeningrunde	37 128[b]	35[c]	0,09[a]
Biopsien mit benignem Befund	4. Screeningrunde	34 082[b]	38[c]	0,11[a]

Tabelle 29: Fragestellung 1 – Ergebnisse: Konsequenzen aus falsch-positiven Screeningbefunden: Biopsien mit benignem Befund (mehrseitige Tabelle)

Studie Alter	Zeitpunkt	N	Frauen mit Ereignissen	
			n	%
Biopsien mit benignem Befund	5. Screeningrunde	29 818[b]	24[c]	0,08[a]

a. eigene Berechnung
b. Anzahl durchgeführter Screeninguntersuchungen
c. Anzahl Biopsien; es ist davon auszugehen, dass dies annähernd der Anzahl der Frauen mit Biopsie entspricht
d. Anteil benigner Biopsien an allen durchgeführten Biopsien: 90 % (eigene Berechnung)

CNBSS; Canadian National Breast Screening Study; HIP; Health Insurance Plan of Greater New York; k. A.: keine Angabe; MMST: Malmö Mammographic Screening Trial; n: Anzahl Frauen mit Ereignis; N: Anzahl ausgewerteter Frauen in der Interventionsgruppe

Tabelle 30: Fragestellung 1 – Ergebnisse: Konsequenzen aus falsch-positiven Screeningbefunden: Operationen mit benignem Befund

Studie Alter	Zeitpunkt	N	Frauen mit Ereignissen	
			n	%[a]
Göteborg **39–49 Jahre**				
OP mit benignem Befund	1. Screeningrunde	11 720[b]	32[c]	0,27
OP mit benignem Befund	2. Screeningrunde	11 679[b]	15[d]	0,13
OP mit benignem Befund	3. Screeningrunde	11 624[b]	12[e]	0,10
OP mit benignem Befund	4. Screeningrunde	11 571[b]	4[f]	0,03
OP mit benignem Befund	5. Screeningrunde	11 519[b]	9[g]	0,08

a. Die Prozentangaben in dieser Tabelle wurden selbst berechnet.
b. Anzahl durchgeführter Screeninguntersuchungen
c. von insgesamt 49 Operationen; Anteil mit benignem Befund an allen Operationen: 65 %
d. von insgesamt 25 Operationen; Anteil mit benignem Befund an allen Operationen: 60 %
e. von insgesamt 27 Operationen; Anteil mit benignem Befund an allen Operationen: 44 %
f. von insgesamt 25 Operationen; Anteil mit benignem Befund an allen Operationen: 16 %
g. von insgesamt 29 Operationen; Anteil mit benignem Befund an allen Operationen: 31 %

k. A.: keine Angabe; n: Anzahl Frauen mit Ereignis; N: Anzahl ausgewerteter Frauen in der Interventionsgruppe

Sensitivitätsanalysen

Es wurden keine Sensitivitätsanalysen durchgeführt.

Subgruppenmerkmale und andere Effektmodifikatoren

Es wurden keine Analysen von Subgruppenmerkmalen und anderen Effektmodifikatoren durchgeführt.

A3.3.4 Unerwünschte Ereignisse

Zu diesem Endpunkt lagen keine Daten vor.

A3.3.5 Überdiagnosen

Verzerrungspotenzial der Ergebnisse zu Überdiagnosen

In der folgenden Tabelle 31 wird die Bewertung des Verzerrungspotenzials der Ergebnisse zu Überdiagnosen dargestellt.

Tabelle 31: Fragestellung 1 – endpunktspezifisches Verzerrungspotenzial: Überdiagnosen

Studie	Endpunktübergreifendes Verzerrungspotenzial	Verblindung Endpunkt-erheber	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Endpunktspezifisches Verzerrungspotenzial[a]
CNBSS	hoch	n. b.	n. b.	n. b.	n. b.	hoch
Göteborg	hoch	n. b.	n. b.	n. b.	n. b.	hoch
HIP	hoch	n. b.	n. b.	n. b.	n. b.	hoch
Malmö	hoch	n. b.	n. b.	n. b.	n. b.	hoch
Swedish Two County	hoch	n. b.	n. b.	n. b.	n. b.	hoch
UK Age	niedrig	unklar	ja	ja	ja	niedrig

a. Falls ein hohes endpunktübergreifendes Verzerrungspotenzial der Ergebnisse vorliegt, überträgt sich diese Bewertung direkt auf das endpunktspezifische Verzerrungspotenzial der Ergebnisse. In diesen Fällen wird auf die Bewertung endpunktspezifischer Kriterien verzichtet.

CNBSS: Canadian National Breast Screening Study; HIP: Health Insurance Plan of Greater New York; ITT: Intention to treat; n. b.: nicht bewertet

Ergebnisse zu Überdiagnosen

Die Ergebnisse zu Überdiagnosen in den Studien mit verwertbaren Daten sind in Tabelle 32 dargestellt.

Tabelle 32: Fragestellung 1 – Ergebnisse: Überdiagnosen (mehrseitige Tabelle)

Studie Alter bei Einschluss Zeitpunkt	Intervention					Vergleich			Überdiagnoserisiko	
	N_S	Frauen mit Ereignissen[a]				N_C	Frauen mit Ereignissen[a]		Bezogen auf zum Screening eingeladene Frauen $(S_2 - C_2) / N_S$[b]	Bezogen auf Frauen mit Brustkrebsdiagnose während der Screeningphase $(S_2 - C_2) / S_1$[b]
		nach der Screeningphase		zum Ende der Beobachtungszeit			zum Ende der Beobachtungszeit			
		alle Brustkrebsdiagnosen S_1 (%)	– davon invasive Karzinome n (%)	alle Brustkrebs-diagnosen S_2 (%)	– davon invasive Karzinome n (%)		alle Brustkrebsdiagnosen C_2 (%)	– davon invasive Karzinome n (%)	% 95 %-KI	% 95 %-KI
CNBSS										
40–49 Jahre										
S_1: 5 Jahre nach Studienbeginn, entspricht etwa dem Ende der Screeningphase; S_2/C_2: 5 Jahre nach Ende der Screeningphase										
	25 214	326 (1,29)	284 (87)	590 (2,33)	544 (92)	25 216	487 (1,93)	476 (98)	0,41 [0,16; 0,66]	31,61 [12,10; 51,19]
Göteborg										
39–49 Jahre										
$S_1/S_2/C_2$: nach 1. Screening der Kontrollen, entspricht etwa 7 Jahren nach Studienbeginn										
	11 792	149[c] (1,26)	125[d] (84)	149[c] (1,26)	125[d] (84)	14 321	196[c] (1,37)	183[d] (93)	–[e]	–[e]
45–49 Jahre										
$S_1/S_2/C_2$: nach 1. Screening der Kontrollen, entspricht etwa 7 Jahren nach Studienbeginn										
	5157	67[c] (1,30)	57 (85)	67[c] (1,30)	57 (85)	5995	84 (1,40)	78 (93)	–[e]	–[e]
HIP										
40–49 Jahre[f]										
S_1: 3 Jahre nach Studienbeginn, entspricht etwa dem Ende der Screeningphase; S_2/C_2: etwa 3 Jahre nach Ende der Screeningphase										
	13 740[c]	62 (0,45)	k. A.	149[c] (1,08)	k. A.	13 740[c]	144[c] (1,05)	k. A.	0,04 [0; 0,28]	8,06 [0; 62,25]
45–49 Jahre[f]										
S_1: 3 Jahre nach Studienbeginn, entspricht etwa dem Ende der Screeningphase, S_2/C_2: etwa 3 Jahre nach Ende der Screeningphase										
	7240	38 (0,52)	k. A.	82 (1,13)	k. A.	7240	81 (1,12)	k. A.	0,01 [0; 0,36]	2,63 [0; 68,86]

Tabelle 32: Fragestellung 1 – Ergebnisse: Überdiagnosen (mehrseitige Tabelle)

Studie Alter bei Einschluss Zeitpunkt	Intervention					Vergleich			Überdiagnoserisiko	
	N_S	Frauen mit Ereignissen[a]				N_C	Frauen mit Ereignissen[a]		Bezogen auf zum Screening eingeladene Frauen $(S_2 - C_2) / N_S$[b]	Bezogen auf Frauen mit Brustkrebsdiagnose während der Screeningphase $(S_2 - C_2) / S_1$[b]
		nach der Screeningphase		zum Ende der Beobachtungszeit			zum Ende der Beobachtungszeit			
		alle Brustkrebsdiagnosen S_1 (%)	– davon invasive Karzinome n (%)	alle Brustkrebs-diagnosen S_2 (%)	– davon invasive Karzinome n (%)		alle Brustkrebsdiagnosen C_2 (%)	– davon invasive Karzinome n (%)	% 95 %-KI	% 95 %-KI
Malmö **45–54 Jahre** **S_1: vor Screening der Kontrollen, maximal 14 Jahren nach Studienbeginn; S_2/C_2: maximal 11 Jahre nach dem Screening der Kontrollen**										
	7981[c]	303 (3,80)	k. A.	540 (6,77)	481 (89)	8082[c]	507 (6,27)	454 (90)	–[e]	–[e]
Swedish Two County **40–49 Jahre** **S_1: vor dem Screening der Kontrollen, etwa 7 Jahre nach Studienbeginn; S_2/C_2: nach dem Screening der Kontrollen, etwa 7–8 Jahre nach Studienbeginn**										
	19 844	238 (1,20)	k. A.	256 (1,29)	228 (89)[g]	15 604	162 (1,04)	152 (94)[g]	–[e]	–[e]

Tabelle 32: Fragestellung 1 – Ergebnisse: Überdiagnosen (mehrseitige Tabelle)

Studie Alter bei Einschluss Zeitpunkt	Intervention					Vergleich			Überdiagnoserisiko	
	N_S	Frauen mit Ereignissen[a]				N_C	Frauen mit Ereignissen[a]		Bezogen auf zum Screening eingeladene Frauen $(S_2 - C_2) / N_S$[b]	Bezogen auf Frauen mit Brustkrebsdiagnose während der Screeningphase $(S_2 - C_2) / S_1$[b]
		nach der Screeningphase		zum Ende der Beobachtungszeit			zum Ende der Beobachtungszeit			
		alle Brustkrebsdiagnosen S_1 (%)	– davon invasive Karzinome n (%)	alle Brustkrebs-diagnosen S_2 (%)	– davon invasive Karzinome n (%)		alle Brustkrebsdiagnosen C_2 (%)	– davon invasive Karzinome n (%)	% 95 %-KI	% 95 %-KI
UK Age										
39–41 Jahre										
S_1: vor 1. Routinescreening, etwa 8 Jahre nach Studienbeginn; S_2/C_2: zum 1. Routinescreening, etwa 10,7 Jahre nach Studienbeginn (MW)										
	53 883	953 (1,77)	835 (88)	1132 (2,10)	970 (86)	106 953	2278 (2,13)	2021 (89)	–[e]	–[e]
S_1: vor 1. Routinescreening, etwa 8 Jahre nach Studienbeginn; S_2/C_2: 17,7 Jahre nach Studienbeginn (Median)										
	53 883	953 (1,77)	835 (88)	1906 (3,54)	1654 (87)	106 953	3855 (3,60)	3382 (88)	–[e]	–[e]
S_1: vor 1. Routinescreening, etwa 8 Jahre nach Studienbeginn; S_2/C_2: 22,8 Jahre nach Studienbeginn (Median)										
	53 883	953 (1,77)	835 (88)	2617 (4,86)	2288 (87)	106 953	5260 (4,92)	4640 (88)	–[e]	–[e]

a. Die Prozentangaben in dieser Tabelle wurden selbst berechnet.
b. Bei ungleichen Anzahlen von Frauen in Interventions- und Vergleichsgruppe wurde C_2 mit dem Faktor N_S/N_C multipliziert, um die Vergleichbarkeit von S_2 und C_2 zu gewährleisten.
c. eigene Berechnung
d. eigene Berechnung aus den Angaben zu den Nodalstadien negativ, positiv und unbekannt
e. Aufgrund des nach Abschluss der Screeningphase stattfindenden Screenings im Kontrollarm ließ sich der Anteil an Überdiagosen nicht genau berechnen, da auch im Kontrollarm Überdiagnosen auftreten können.
f. Die Werte für N sind ausschließlich in Habbema 1986 [83], Tabelle 1, angegeben und wurden für diese Ergebnistabelle herangezogen. Entsprechend der Angaben der Autorinnen und Autoren sollen die Werte für die Kontroll- und die Interventionsgruppe vergleichbar sein.
g. eigene Berechnung aus der Anzahl der DCIS-Fälle

C_2: Anzahl der Frauen mit Brustkrebsdiagnose in der Vergleichsgruppe zum Ende der Beobachtungszeit; CNBSS; Canadian National Breast Screening Study; DCIS: duktales Carcinoma in situ; HIP: Health Insurance Plan of Greater New York; KI: Konfidenzintervall; n: Anzahl der Frauen mit Ereignis; N_C: Anzahl der Frauen in der Vergleichsgruppe; N_S: Anzahl der Frauen in der Interventionsgruppe; S_1: Anzahl der Frauen mit Brustkrebsdiagnose in der Screeninggruppe nach der Screeningphase; S_2: Anzahl der Frauen mit Brustkrebsdiagnose in der Screeninggruppe zum Ende der Beobachtungszeit

Sensitivitätsanalysen

Es wurden keine Sensitivitätsanalysen durchgeführt.

Subgruppenmerkmale und andere Effektmodifikatoren

Es wurden keine Analysen von Subgruppenmerkmalen und anderen Effektmodifikatoren durchgeführt.

A3.3.6 Gesundheitsbezogene Lebensqualität

Zu diesem Endpunkt lagen keine Daten vor.

A3.4 Fragestellung 2 – Altersgruppe ab 70 Jahren

A3.4.1 Mortalität

A3.4.1.1 Gesamtmortalität

Verzerrungspotenzial der Ergebnisse zur Gesamtmortalität

In der folgenden Tabelle 33 wird die Bewertung des Verzerrungspotenzials der Ergebnisse zur Gesamtmortalität dargestellt.

Tabelle 33: Fragestellung 2 – endpunktspezifisches Verzerrungspotenzial: Gesamtmortalität

Studie	Endpunktübergreifendes Verzerrungspotenzial	Verblindung Endpunkterheber	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Endpunktspezifisches Verzerrungspotenzial[a]
Swedish Two County	hoch	n. b.	n. b.	n. b.	n. b.	hoch

a. Falls ein hohes endpunktübergreifendes Verzerrungspotenzial der Ergebnisse vorliegt, überträgt sich diese Bewertung direkt auf das endpunktspezifische Verzerrungspotenzial der Ergebnisse. In diesen Fällen wurde auf die Bewertung endpunktspezifischer Kriterien verzichtet.

ITT: Intention to treat; n. b.: nicht bewertet

Ergebnisse zur Gesamtmortalität

Die Ergebnisse zur Gesamtmortalität in der Studie mit verwertbaren Daten sind in Tabelle 34 dargestellt.

Tabelle 34: Fragestellung 2 – Ergebnisse: Gesamtmortalität

Studie Alter Zeitpunkt[a]	Intervention					Vergleich					Intervention vs. Vergleich		
	N	Frauen mit Ereignissen		Personen-jahre	Ereignisse pro 10 000 Personen-jahre	N	Frauen mit Ereignissen		Personen-jahre	Ereignisse pro 10 000 Personen-jahre	RR	95 %-KI	p-Wert
		n	%[b]				n	%[b]					
Swedish Two County **70–74 Jahre** **7,9 Jahre (MW) / 0 Jahre**													
	10 339	2869	27,75	k. A.	k. A.	7307	2073	28,37	k. A.	k. A.	0,98[c]	[0,93; 1,03][c]	0,384[c]

a. Beobachtungszeit seit Randomisierung / seit Ende der Interventionsphase
b. Die Prozentangaben in dieser Tabelle wurden selbst berechnet.
c. eigene Berechnung von RR, 95 %-KI (asymptotisch) und p-Wert (unbedingter exakter Test, CSZ-Methode nach [101])

k. A.: keine Angabe; KI: Konfidenzintervall; MW: Mittelwert; n: Anzahl Frauen mit Ereignis; N: Anzahl ausgewerteter Frauen; RR: relatives Risiko

Sensitivitätsanalysen

Es wurden keine Sensitivitätsanalysen durchgeführt.

Subgruppenmerkmale und andere Effektmodifikatoren

Es wurden keine Analysen von Subgruppenmerkmalen und anderen Effektmodifikatoren durchgeführt.

A3.4.1.2 Brustkrebsspezifische Mortalität

Verzerrungspotenzial der Ergebnisse zur brustkrebsspezifischen Mortalität

In der folgenden Tabelle 35 wird die Bewertung des Verzerrungspotenzials der Ergebnisse zur brustkrebsspezifischen Mortalität dargestellt.

Tabelle 35: Fragestellung 2 – endpunktspezifisches Verzerrungspotenzial: brustkrebsspezifische Mortalität

Studie	Endpunktübergreifendes Verzerrungspotenzial	Verblindung Endpunkterheber	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Endpunktspezifisches Verzerrungspotenzial[a]
Malmö	hoch	n. b.	n. b.	n. b.	n. b.	hoch
Swedish Two County	hoch	n. b.	n. b.	n. b.	n. b.	hoch

a. Falls ein hohes endpunktübergreifendes Verzerrungspotenzial der Ergebnisse vorliegt, überträgt sich diese Bewertung direkt auf das endpunktspezifische Verzerrungspotenzial der Ergebnisse. In diesen Fällen wurde auf die Bewertung endpunktspezifischer Kriterien verzichtet.

ITT: Intention to treat; n. b.: nicht bewertet

Ergebnisse zur brustkrebsspezifischen Mortalität

Die Ergebnisse zur Gesamtmortalität der Studien mit verwertbaren Daten sind in Tabelle 36 dargestellt. Die metaanalytische Zusammenfassung dieser Ergebnisse ist in Abbildung 10 dargestellt.

Tabelle 36: Fragestellung 2 – Ergebnisse: brustkrebsspezifische Mortalität

Studie **Alter** **Zeitpunkt**[a] ***Auswertungsart***	**Intervention**					**Vergleich**					**Intervention vs. Vergleich**		
	N	**Frauen mit Ereignissen**		**Personen-jahre**	**Ereignisse pro 10 000 Personen-jahre**	**N**	**Frauen mit Ereignissen**		**Personen-jahre**	**Ereignisse pro 10 000 Personen-jahre**	**IDR**	**95 %-KI**	**p-Wert**
		n	**%**[b]				**n**	**%**[b]					
Malmö **> 69 Jahre (Kollektiv MMST 1)** **13,6 Jahre / ca. 5 Jahre**[c]													
Evaluationsmodell o. A.	296	3	1,01	4000	k. A.	291	3	1,03	4000	k. A.	0,98	[0,15; 6,60]	k. A.
Swedish Two County **70–74 Jahre** **7,9 Jahre (MW) / 0 Jahre**													
Evaluationsmodell	10 339	35	0,34	k. A.	k. A.	7307	31	0,42	k. A.	k. A.	0,77[d]	[0,47; 1,27][d]	0,3[d]
10,8 Jahre (MW) / ca. 4 Jahre													
Evaluationsmodell	10 339	52[e]	0,50	96 783[e]	5,37[e]	7307	40[e]	0,68	68 086[e]	5,87[e]	0,91	[0,61; 1,38]	0,671
20 Jahre / ca. 13 Jahre													
Evaluationsmodell	10 339	61	0,59	k. A.	k. A.	7307	57	0,78	k. A.	k. A.	0,76[f]	[0,53; 1,08][f]	0,113[f]

a. Beobachtungszeit seit Randomisierung / seit Ende der Screeningphase
b. Die Prozentangaben in dieser Tabelle wurden selbst berechnet.
c. Die mittlere Dauer der Screeningphase von 8,8 Jahren bezieht sich auf das Gesamtkollektiv. Die genaue Dauer für die Altersgruppe > 69 Jahre ist unbekannt. Bekannt ist nur, dass diesen Frauen bis zu 6 Screenings im Abstand 18–24 Monate angeboten wurden. Hieraus lässt sich aber keine sichere mittlere Interventionsdauer für diese Altersgruppe ableiten.
d. adjustiert für Region und Alter
e. eigene Berechnung
f. eigene Berechnung von RR, 95 %-KI (asymptotisch) und p-Wert (unbedingter exakter Test, CSZ-Methode nach [101])

IDR: Inzidenzdichtequotient; k. A.: keine Angabe; KI: Konfidenzintervall; MMST: Malmö Mammographic Screening Trial; MW: Mittelwert; n: Anzahl Frauen mit Ereignis; N: Anzahl ausgewerteter Frauen; o. A.: ohne Abschlussscreening; RR: relatives Risiko

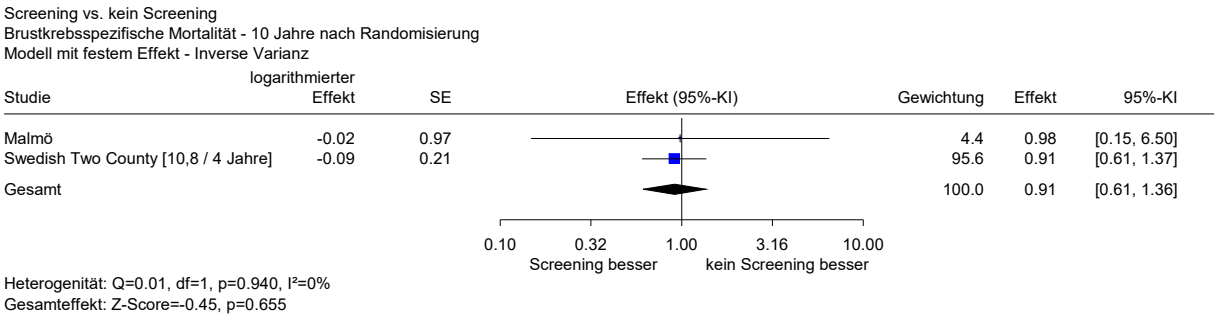


Abbildung 10: Fragestellung 2 – brustkrebsspezifische Mortalität, 10 Jahre nach Randomisierung

Sensitivitätsanalysen

Es wurden keine Sensitivitätsanalysen durchgeführt.

Subgruppenmerkmale und andere Effektmodifikatoren

Es wurden keine Analysen von Subgruppenmerkmalen und anderen Effektmodifikatoren durchgeführt.

A3.4.2 Mastektomien

Zu diesem Endpunkt lagen keine verwertbaren Daten vor.

A3.4.3 Konsequenzen aus falsch-positiven Screeningbefunden

Verzerrungspotenzial der Ergebnisse zu Konsequenzen aus falsch-positiven Screeningbefunden

In der folgenden Tabelle 37 wird die Bewertung des Verzerrungspotenzials der Ergebnisse zu Konsequenzen aus falsch-positiven Befunden dargestellt.

Tabelle 37: Fragestellung 2 – endpunktspezifisches Verzerrungspotenzial: Konsequenzen aus falsch-positiven Screeningbefunden

Studie	Endpunktübergreifendes Verzerrungspotenzial	Verblindung Endpunkt-erheber	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Endpunktspezifisches Verzerrungspotenzial[a]
Swedish Two County	hoch	n. b.	n. b.	n. b.	n. b.	hoch

a. Falls ein hohes endpunktübergreifendes Verzerrungspotenzial der Ergebnisse vorliegt, überträgt sich diese Bewertung direkt auf das endpunktspezifische Verzerrungspotenzial der Ergebnisse. In diesen Fällen wurde auf die Bewertung endpunktspezifischer Kriterien verzichtet.

ITT: Intention to treat; n. b.: nicht bewertet

Ergebnisse zu Konsequenzen aus falsch-positiven Screeningbefunden

Die Ergebnisse zu Konsequenzen aus falsch-positiven Screeningbefunden aus der Studie mit verwertbaren Daten sind in Tabelle 38 dargestellt.

Tabelle 38: Fragestellung 2 – Ergebnisse: Konsequenzen aus falsch-positiven Screeningbefunden: chirurgische Biopsien mit benignem Befund

Studie Alter	Zeitpunkt	N	Frauen mit Ereignissen	
			n	%
Swedish Two County **40–74 Jahre**				
chirurgische Biopsie mit benignem Befund	1. Screeningrunde	k. A.	k. A.	0,6
chirurgische Biopsie mit benignem Befund	spätere Screeningrunden	k. A.	k. A.	0,1
k. A.: keine Angabe; n: Anzahl Frauen mit Ereignis; N: Anzahl ausgewerteter Frauen in der Interventionsgruppe				

Sensitivitätsanalysen

Es wurden keine Sensitivitätsanalysen durchgeführt.

Subgruppenmerkmale und andere Effektmodifikatoren

Es wurden keine Analysen von Subgruppenmerkmalen und anderen Effektmodifikatoren durchgeführt.

A3.4.4 Unerwünschte Ereignisse

Zu diesem Endpunkt lagen keine Daten vor.

A3.4.5 Überdiagnosen

Verzerrungspotenzial der Ergebnisse zu Überdiagnosen

In der folgenden Tabelle 39 wird die Bewertung des Verzerrungspotenzials der Ergebnisse zu Überdiagnosen dargestellt.

Tabelle 39: Fragestellung 2 – endpunktspezifisches Verzerrungspotenzial: Überdiagnosen)

Studie	Endpunktübergreifendes Verzerrungspotenzial	Verblindung Endpunkt-erheber	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Endpunktspezifisches Verzerrungspotenzial[a]
Swedish Two County	hoch	n. b.	n. b.	n. b.	n. b.	hoch

a. Falls ein hohes endpunktübergreifendes Verzerrungspotenzial der Ergebnisse vorliegt, überträgt sich diese Bewertung direkt auf das endpunktspezifische Verzerrungspotenzial der Ergebnisse. In diesen Fällen wurde auf die Bewertung endpunktspezifischer Kriterien verzichtet.

ITT: Intention to treat; n. b.: nicht bewertet

Ergebnisse zu Überdiagnosen

Die Ergebnisse zu Überdiagnosen aus der Studie mit verwertbaren Daten sind in Tabelle 40 dargestellt.

Tabelle 40: Fragestellung 2 – Ergebnisse: Überdiagnosen

Studie Alter bei Einschluss Zeitpunkt	Intervention					Vergleich			Überdiagnoserisiko	
	N_S	Frauen mit Ereignissen[a]				N_C	Frauen mit Ereignissen[a]		Bezogen auf zum Screening eingeladene Frauen	Bezogen auf Frauen mit Brustkrebsdiagnose während der Interventionsphase
		nach der Screeningphase		zum Ende der Beobachtungszeit			zum Ende der Beobachtungszeit			
		alle Brustkrebsdiagnosen S_1 (%)[b]	– davon invasive Karzinome n (%)	alle Brustkrebsdiagnosen S_2 (%)[b]	– davon invasive Karzinome n (%)		alle Brustkrebsdiagnosen C_2 (%)[b]	– davon invasive Karzinome n (%)	$(S_2 - C_2) / N_S$[b]	$(S_2 - C_2) / S_1$[b]
									%[b]	%[b]
Swedish Two County **70–74 Jahre** **S_1: vor dem Screening der Kontrollen, etwa 7 Jahre nach Studienbeginn; S_2/C_2: nach dem Screening der Kontrollen, etwa 7–8 Jahre nach Studienbeginn**										
	10 339	255 (2,47)	k. A.	259 (2,51)	k. A.	7307	148 (2,03)	k. A.	–[c]	–[c]

a. Die Prozentangaben in dieser Tabelle wurden selbst berechnet.
b. Bei ungleichen Anzahlen von Frauen in Interventions- und Vergleichsgruppe wurde C_2 mit dem Faktor N_S/N_C multipliziert, um die Vergleichbarkeit von S_2 und C_2 zu gewährleisten.
c. Aufgrund des nach Abschluss der Interventionsphase stattfindenden Screenings im Kontrollarm ließ sich der Anteil an Überdiagosen nicht genau berechnen, da auch im Kontrollarm Überdiagnosen auftreten können.

C_2: Anzahl der Frauen mit Brustkrebsdiagnose in der Vergleichsgruppe zum Ende der Beobachtungszeit; DCIS: duktales Carcinoma in situ; n: Anzahl der Frauen mit Ereignis; N_C: Anzahl der Frauen in der Vergleichsgruppe; N_S: Anzahl der Frauen in der Interventionsgruppe; S_1: Anzahl der Frauen mit Brustkrebsdiagnose in der Screeninggruppe nach der Screeningphase; S_2: Anzahl der Frauen mit Brustkrebsdiagnose in der Screeninggruppe zum Ende der Beobachtungszeit

Sensitivitätsanalysen

Es wurden keine Sensitivitätsanalysen durchgeführt.

Subgruppenmerkmale und andere Effektmodifikatoren

Es wurden keine Analysen von Subgruppenmerkmalen und anderen Effektmodifikatoren durchgeführt.

A3.4.6 Gesundheitsbezogene Lebensqualität

Zu diesem Endpunkt lagen keine Daten vor.

A4 Bericht der externen Expertengruppe zur Modellierung

Überprüfung der Altersgrenzen im Mammographie-Screening-Programm (S21-01) – Modellierungsstudie

31. Dezember 2021

Autoren: Gaby Sroczynski, Lára R. Hallsson, Nikolai Mühlberger, Felicitas Kühne, Beate Jahn, Uwe Siebert

Institution: UMIT – Private Universität für Gesundheitswissenschaften, Medizinische Informatik und Technologie

Projekttitel: „Überprüfung der Altersgrenzen im Mammographie-Screening-Programm“ (S21-01) – Modellierungsstudie

Erstellt für das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)

Danksagung

Wir bedanken uns bei Frau Dr. med. Silke Siebert für die interne Begutachtung und Qualitätsprüfung des Berichts.

Inhalt

1. Zusammenfassung 10
2. Hintergrund 13
3. Fragestellungen 13
4. Methoden 14
   4.1. Studiendesign 14
      4.1.1. Entscheidungsanalytisches Modell 14
      4.1.2. Zielpopulation 18
      4.1.3. Evaluierte Strategien 18
      4.1.4. Perspektive 18
      4.1.5. Analytischer Zeithorizont und Zykluslänge 19
      4.1.6. Zielparameter (Endpunkte) 19
   4.2. Daten für die entscheidungsanalytische Modellierung 19
      4.2.1. Datenerhebung 20
      4.2.2. Daten zur Krankheitsentwicklung 20
      4.2.3. Klinische Daten 23
      4.2.4. Lebensqualitätsdaten 26
      4.2.5. Kalibrierung 29
      4.2.6. Validierung 30
   4.3. Analysen 32
      4.3.1. Basisfallanalyse 32
      4.3.2. Sensitivitätsanalysen 34
      4.3.3. Software 35
5. Ergebnisse 35
   5.1. Kalibrierung 35
   5.2. Validierung 38
   5.3. Basisfallanalysen 40
      5.3.1. Restlebenserwartung und Lebensqualitätsadjustierte Restlebenserwartung 40
      5.3.2. Nutzen und Schaden 44

5.3.3. Schaden-Nutzen-Verhältnisse 47
5.4. Sensitivitätsanalysen 55
5.4.1. Lebensqualität 55
5.4.1. Testgüte 59
6. Diskussion 74
7. Schlussfolgerungen 78
8. Literaturverzeichnis 80
9. Appendix 83
9.1. Externe Validierung: zusätzliche Ergebnisse 83
9.2. Sensitivitätsanalysen: zusätzliche Ergebnisse 84

Abbildungen

Abbildung 1 Natürlicher Krankheitsverlauf von Brustkrebs im Markov Zustand-Übergangs-Modell. 16

Abbildung 2 Kalibrierungsergebnis: Altersspezifische jährliche Inzidenzrate von symptomatisch entdecktem DCIS pro 100.000 Frauen. 36

Abbildung 3 Kalibrierungsergebnis: Altersspezifische jährliche Inzidenzrate von symptomatisch entdecktem Brustkrebs pro 100.000 Frauen. 37

Abbildung 4 Kalibrierungsergebnis: UICC-Stadienverteilung entdeckter Brustkrebsfälle. 38

Abbildung 5 Ergebnisse der externen Validierung: UICC-Stadienverteilung entdeckter Brustkrebsfälle. 40

Abbildung 6 Ergebnisse der Basisfallanalyse: Gewonnene Lebensjahre und lebensqualitätsadjustierte Lebensjahre verschiedener Mammographie-Screeningstrategien im Vergleich zu keinem Screening. 43

Abbildung 7 Ergebnisse der Basisfallanalyse: Gesamtanzahl an zusätzlichen Mammographien versus gewonnene Lebensjahre. 48

Abbildung 8 Ergebnisse der Basisfallanalyse: Gesamtanzahl an zusätzlichen positiven Mammographien versus gewonnene Lebensjahre. 49

Abbildung 9 Ergebnisse der Basisfallanalyse: Gesamtanzahl an zusätzlichen falsch-positiven Screeningbefunden versus gewonnene Lebensjahre. 50

Abbildung 10 Ergebnisse der Basisfallanalyse: Gesamtanzahl an zusätzlichen Mammographien versus verhinderte Krebstodesfälle. 52

Abbildung 11 Ergebnisse der Basisfallanalyse: Gesamtanzahl an zusätzlichen positiven Mammographien versus verhinderte Krebstodesfälle. 53

Abbildung 12 Ergebnisse der Basisfallanalyse: Gesamtanzahl an zusätzlichen falsch-positiven Mammographien versus verhinderte Krebstodesfälle. 54

Abbildung 13 Ergebnisse der Sensitivitätsanalyse zur Testgüte: Gesamtanzahl an zusätzlichen positiven Mammographien versus gewonnene Lebensjahre bei einer um relative 20% erhöhte falsch-negativen Rate der Mammographie. 62

Abbildung 14 Ergebnisse der Sensitivitätsanalyse zur Testgüte: Gesamtanzahl an zusätzlichen positiven Mammographien versus gewonnene Lebensjahre bei einer um relative 20% reduzierte falsch-negativen Rate der Mammographie. 63

Abbildung 15 Ergebnisse der Sensitivitätsanalyse zur Testgüte: Gesamtanzahl an zusätzlichen positiven Mammographien versus gewonnene Lebensjahre bei einer um relative 20% erhöhte falsch-positiven Rate der Mammographie. 64

Abbildung 16 Ergebnisse der Sensitivitätsanalyse zur Testgüte: Gesamtanzahl an zusätzlichen positiven Mammographien versus gewonnene Lebensjahre bei einer um relative 20% reduzierte falsch-positiven Rate der Mammographie..... 65

Abbildung 17 Ergebnisse der Sensitivitätsanalyse Szenario 1: Gesamtanzahl an zusätzlichen positiven Mammographien versus gewonnene Lebensjahre bei einer um absolute 5% erhöhter Sensitivität und gleichzeitig um absolute 5% reduzierter Spezifität der Mammographie. 70

Abbildung 18 Ergebnisse der Sensitivitätsanalyse Szenario 2: Gesamtanzahl an zusätzlichen positiven Mammographien versus gewonnene Lebensjahre bei einer um absolute 5% reduzierte Sensitivität und gleichzeitig um absolute 5% erhöhte Spezifität der Mammographie. 71

Abbildung 19 Ergebnisse der Sensitivitätsanalyse Szenario 1: Gesamtanzahl an zusätzlichen falsch-positiven Mammographien versus gewonnene Lebensjahre bei einer um absolute 5% erhöhter Sensitivität und gleichzeitig um absolute 5% reduzierter Spezifität der Mammographie. 72

Abbildung 20 Ergebnisse der Sensitivitätsanalyse Szenario 2: Gesamtanzahl an zusätzlichen falsch-positiven Mammographien versus gewonnene Lebensjahre bei einer um absolute 5% reduzierte Sensitivität und gleichzeitig um absolute 5% erhöhte Spezifität der Mammographie. 73

Tabellen

Tabelle 1 Modellparameter des natürlichen Krankheitsverlaufs. 22

Tabelle 2 Relative Überlebenswahrscheinlichkeit für Patientinnen mit entdecktem Brustkrebs. ... 23

Tabelle 3 Altersspezifischer Anteil (in %) der BI-RADS Brustdichte Kategorien in der weiblichen Bevölkerung in Deutschland (23-25). 24

Tabelle 4 Brustdichte- und Alters- spezifische Testgenauigkeit (Sensitivität und Spezifität) von Mammographie (26). 25

Tabelle 5 Relatives Risiko für Brustkrebs in Abhängigkeit von BI-RADS Brustdichte Kategorien (23). 26

Tabelle 6 Nutzwerte für die gesundheitsbezogene Lebensqualität. 27

Tabelle 7 Zielparameter der Modellkalibrierung. 30

Tabelle 8 Validierungsparameter der externen Modellvalidierung. 32

Tabelle 9 Ergebnisse der externen Validierung unter Annahme einer Teilnahme am Screening von 50%. 39

Tabelle 10 Basisfallanalyse: Restlebenserwartung und für die lebensqualitätsadjustierte Restlebenserwartung pro 100 Frauen. 42

Tabelle 11 Ergebnisse der Basisfallanalyse: Ereignishäufigkeiten verschiedener Nutzen- und Schaden-Parameter 46

Tabelle 12 Sensitivitätsanalyse: Auswirkung der Lebensqualität bei Alter, Screening, positivem Screeningbefund und Behandlung von entdeckten DCIS- und Brustkrebsfällen auf den Nutzen (pro 100 Frauen). 57

Tabelle 13 Sensitivitätsanalyse: Auswirkung der gleichzeitigen Erhöhung bzw. Reduktion der Lebensqualität um relative 5 % auf den Nutzen gemessen in lebensqualitätsadjustierte Lebenserwartung (pro 100 Frauen). 58

Tabelle 14 Sensitivitätsanalyse: Auswirkung der Variation der falsch-negativen Rate der Mammographie auf selektierte Nutzen- und Schadenparameter (pro 100 Frauen). 60

Tabelle 15 Sensitivitätsanalyse: Auswirkung der Variation der falsch-positiven Rate der Mammographie auf Nutzen- und Schadenparameter (pro 100 Frauen). 61

Tabelle 16 Sensitivitätsanalyse: Auswirkung der gleichzeitigen Erhöhung (bzw. Reduktion) der Sensitivität und Reduktion (bzw. Erhöhung) der Spezifität der Mammographie auf die Restlebenserwartung und lebensqualitätsadjustierte Restlebenserwartung (pro 100 Frauen). 68

Tabelle 17 Sensitivitätsanalyse: Auswirkung der gleichzeitigen Erhöhung (bzw. Reduktion) der Sensitivität und Reduktion (bzw. Erhöhung) der Spezifität der Mammographie auf die positiven und falsch-positiven Befunde (pro 100 Frauen). 69

Appendix Tabellen

Tabelle A. 1 Appendix: Zusätzliche Ergebnisse der externen Validierung 83

Tabelle A. 2 Appendix: Sensitivitätsanalyse zur Variation der Lebensqualitätsdaten für alle Alterskategorien pro 100 Frauen. 85

Tabelle A. 3 Appendix: Sensitivitätsanalyse zur Variation der Lebensqualitätsdekremente aufgrund von Screening pro 100 Frauen. 86

Tabelle A. 4 Appendix: Sensitivitätsanalyse zur Variation der Lebensqualitätsdekremente aufgrund von positiven Befunden pro 100 Frauen. 87

Tabelle A. 5 Appendix: Sensitivitätsanalyse zur Variation der relativen Lebensqualitätsdekremente aufgrund von entdeckter und behandelter DCIS und Brustkrebs pro 100 Frauen. 88

Tabelle A. 6 Appendix: Sensitivitätsanalyse Szenario 1: Lebenszeitrisiko für diagnostizierten Brustkrebs, DCIS und Tod durch diagnostizierten Brustkrebs bei einer gleichzeitigen Erhöhung der Sensitivität und Reduktion der Spezifität um absolut 5%. 89

Tabelle A. 7 Appendix: Sensitivitätsanalyse Szenario 2: Lebenszeitrisiko für diagnostizierten Brustkrebs, DCIS und Tod durch diagnostizierten Brustkrebs bei einer gleichzeitigen Reduktion der Sensitivität und Erhöhung der Spezifität um absolut 5%. 90

Tabelle A. 8 Appendix: Sensitivitätsanalyse: Verhinderte Brustkrebstodesfälle und positive Screeningbefunde bei einer Erhöhung oder Reduktion der falsch-negativen Rate (1-Sensitivität) um relativ 5% des Ausgangswertes in der Basisfallanalyse. 91

Tabelle A. 9 Appendix: Sensitivitätsanalyse: Verhinderte Brustkrebstodesfälle und positive Screeningbefunde bei einer Erhöhung oder Reduktion der falsch-positiven Rate (1-Sensitivität) um relativ 5% des Ausgangswertes in der Basisfallanalyse. 92

Appendix Abbildungen

Abbildung A. 1. Appendix: Zusätzliche Ergebnisse der externen Validierung: Stadienverteilung bei unterschiedlichen Teilnahmeraten. 84

1. Zusammenfassung

Hintergrund/Ziel

In Deutschland haben Frauen im Alter von 50 bis 69 Jahren einen Anspruch, alle zwei Jahre an dem Programm zur Früherkennung von Brustkrebs (Mammographie-Screening) teilzunehmen. Im März 2021 wurde die europäische Brustkrebsleitlinie der EU-Kommission aktualisiert, mit einer Empfehlung auch Frauen im Alter von 45 bis 49 sowie zwischen 70 und 74 Jahren in das Früherkennungsprogramm einzubeziehen. Ziel dieser Studie war die Entwicklung und der Einsatz eines entscheidungsanalytischen Modells zur evidenzbasierten Evaluation des Langzeitnutzens, -schadens sowie des Nutzen-Schaden-Verhältnisses eines Programms zur Früherkennung von Brustkrebs bei Frauen im Alter von 45 bis 49 sowie zwischen 70 und 74 Jahren oder darüber hinaus in Deutschland im Vergleich zu einem Vorgehen ohne Früherkennungsuntersuchung in der jeweiligen Altersgruppe.

Methoden

Ein entscheidungsanalytisches Markov-Zustands-Übergangsmodell, das die Entwicklung von Brustkrebs inklusive ductales Carcinoma in situ (DCIS) simuliert, wurde entwickelt und für den deutschen klinischen und epidemiologischen Kontext angepasst, kalibriert, validiert und angewandt, um verschiedene Screening-Strategien zu evaluieren, die sich durch das Alter zu Beginn und Ende des Mammographie-Screenings sowie durch das Screeningintervall unterscheiden. Es wurden epidemiologische und klinische Daten aus dem Kontext Deutschlands sowie Daten aus internationalen Studien zur Lebensqualität, Brustdichte und Testgüte (Sensitivität und Spezifität) der Mammographie verwendet. Die Endpunkte waren das Lebenszeitrisiko für entdeckte invasive Brustkrebskarzinome, entdeckte DCIS und an entdecktem invasivem Brustkrebskarzinom zu versterben, verbleibende Restlebenserwartung und lebensqualitätsadjustierte Restlebenserwartung, Anzahl von Mammographien, Anzahl positiver und falsch-positiver Mammographien, Lebenszeitrisiko für eine Überdiagnose (DCIS und Brustkrebs), sowie die Number-Needed-To-Screen (NNS) und das inkrementelle Schaden-Nutzen-Verhältnis (ISNV). Umfassende Einweg- und Mehrweg-Sensitivitätsanalysen, insbesondere für die Testgüte und für Lebensqualitätsdaten, wurden durchgeführt.

Ergebnisse

Im Vergleich zu einem Vorgehen ohne Screening im Alter unter 50 Jahren oder über 69 Jahren wird der höchste Gewinn an Restlebenserwartung mit einem Mammographie Screening im Alter von 45 bis 79 Jahren (jährlich für Frauen unter 50 Jahren und zweijährlich für Frauen ab 50 Jahren) erzielt. Diese Strategie führt zu 10 gewonnenen Lebensjahren pro 100 Frauen im Vergleich zum

derzeitig etablierten Screeningprogramm mit einem Screening im Alter von 50 bis 69 Jahren. Die höchste für die Lebensqualität adjustierte Restlebenserwartung (quality-adjusted life years, QALYs) erzielt ein zweijährliches Mammographie-Screening im Alter von 45 bis 74 Jahren. Diese Strategie führt zu 3,5 gewonnenen QALYs pro 100 Frauen im Vergleich zum etablierten Screeningprogramm.

Mit dem Absenken des Alters für den Beginn des Mammographie-Screenings kann ein größerer inkrementeller Nutzen hinsichtlich der gewonnenen Restlebenserwartung und der für die Lebensqualität adjustierten Restlebenserwartung erzielt werden als mit einem Heraufsetzen des Alters für das Ende des Mammographie-Screenings.

Wird die Gesamtanzahl an Screeningbefunden (Mammographien) als potenzielle Belastung (Burden) bzw. Schaden herangezogen, so ergibt sich für das etablierte Mammographie-Screening (Alter: 50-69 Jahre, Intervall:2 Jahre) bei einer vollständigen und regelmäßigen Teilnahme aller screeningberechtigter Frauen ab 50 Jahren im Durchschnitt eine Anzahl von 903 Mammographien pro 100 Frauen. Wird die Altersgrenze für den Beginn des Mammographie-Screenings auf 45 Jahre (Alter: 45-69 Jahre, Intervall:2 Jahre) herabgesetzt, ist die Anzahl an Mammographien 1186 pro 100 Frauen. Bleibt die Altersgrenze für den Beginn erhalten und wird die Altersgrenze für das Ende des etablierten Mammographie-Screening angehoben auf 74 Jahre (Alter: 50-74 Jahre, Intervall:2 Jahre), ist die Gesamtanzahl an Mammographien 1135 pro 100 Frauen. Bleibt die Altersgrenze für den Beginn erhalten und wird die Altersgrenze für das Ende des etablierten Mammographie-Screening angehoben auf 79 Jahre (Alter: 50-79 Jahre, Intervall:2 Jahre), ist die Gesamtanzahl an Mammographien 1273 pro 100 Frauen. Unter den untersuchten Screeningstrategien nimmt also die Anzahl der Mammographien mit der Intensität und Länge des Mammographie-Screenings zu. Die höchste Anzahl an Mammographien erzielt ein jährliches Mammographie-Screening im Alter von 45 bis 49 Jahren gefolgt von einem zweijährlichen Mammographie-Screening im Alter von 50 bis 79 Jahren mit 1741 Mammographien pro 100 Frauen.

Hieraus resultieren für die inkrementellen Schaden-Nutzen-Verhältnisse folgende Ergebnisse: Im Vergleich zum etablierten Mammographie-Screening müssen für das Herabsetzen der Altersgrenze für den Beginn des Mammographie-Screening auf 45 Jahre (Screening im Alter 45 - 69 Jahren, zweijährlich) im Durchschnitt 47 zusätzliche Mammographien pro gewonnenes Lebensjahr akzeptiert werden. Im Vergleich zu diesem Screening müssen für ein zusätzliches Heraufsetzen der Altersgrenze für das Ende des Mammographie-Screening auf 74 Jahre (Screening im Alter 45 - 74 Jahren, zweijährlich) im Durchschnitt 96 weitere zusätzliche Mammographien pro weiteres gewonnenes Lebensjahr akzeptiert werden. Durch ein kürzeres Screeningintervall im Alter von 45 bis 49 Jahren und/oder eine Verlängerung des Screenings bis zu einem Alter von 79 Jahren kann trotz deutlich zunehmender Anzahl an Mammographien und der weiteren Schadensendpunkte nur noch ein vergleichsweise geringfügiger Nutzen hinsichtlich der Restlebenserwartung erzielt

werden. Die inkrementellen Schaden-Nutzen-Verhältnisse dieser Strategien sind daher deutlich höher.

Schlussfolgerungen

Basierend auf internationalen Daten und Daten aus dem Gesundheitssystem Deutschland konnte ein entscheidungsanalytisches Modell zur Abschätzung des Nutzens und des möglichen Schadens mit einem lebenslangen analytischen Zeithorizont entwickelt und anhand von Daten des deutschen Gesundheitssystems kalibriert und validiert werden.

Basierend auf den Ergebnissen der entscheidungsanalytischen Modellierung kann ein Mammographie-Screening im Alter von 45 bis 49 Jahren und im Alter von 70 bis 79 Jahren zusätzlich zum derzeitigen Mammographie-Screening von 50 bis 69 Jahre alle zwei Jahre weitere Todesfälle durch Brustkrebs verhindern und die Restlebenserwartung weiter vergrößern. Unter Berücksichtigung der Lebensqualität kann bei Beibehaltung des derzeitigen Screeningintervalls von zwei Jahren eine Ausweitung des Screening-Programms auf ein Alter von 45 bis 74 Jahre empfohlen werden. Dabei kommt dem früheren Screeningstart ab 45 Jahre ein größerer Nutzen zu als dem späteren Screeningende mit 74 Jahren. Bis zu welchem Alter und mit welchem Screeningintervall gescreent werden sollte, hängt entscheidend von der Akzeptanz der Frauen in Deutschland bezüglich des Nutzen- Schaden-Verhältnisses ab. Die individuellen Präferenzen und Risikoeinstellungen zum akzeptablen inkrementellen Nutzen-Schaden-Verhältnis der Frauen sind im Rahmen einer partizipativen Entscheidungsfindung zu berücksichtigen.

2. Hintergrund

Bisher haben in Deutschland Frauen im Alter von 50 bis 69 Jahren einen Anspruch, alle 2 Jahre an dem Programm zur Früherkennung von Brustkrebs (Mammographie-Screening) teilzunehmen (1). Im März 2021 wurde die europäische Brustkrebsleitlinie der EU-Kommission aktualisiert, mit einer Empfehlung auch Frauen im Alter von 45 bis 49 sowie zwischen 70 und 74 Jahren in das Früherkennungsprogramm einzubeziehen (2). Der Gemeinsame Bundesausschuss (G-BA) hat die Beratungen über eine Anpassung der Altersgrenzen beim Mammographie-Screening eingeleitet und das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) beauftragt, Studien zu dieser Fragestellung zu recherchieren und auszuwerten (3, 4). Ziel soll die Nutzenbewertung eines Screenings auf Brustkrebs mittels Mammographie für Frauen im Alter zwischen 45 und 49, 70 und 74 sowie 75 und 79 im Vergleich zu keinem Screening in diesen Altersbereichen sein.

Da die Evidenzlage basierend auf randomisierten klinischen Studien und Kohortenstudien begrenzt ist, soll basierend auf einer evidenzbasierten entscheidungsanalytischen Modellierungsstudie für den deutschen Kontext der Einfluss einer Ausweitung der Altersgrenzen im etablierten Brustkrebs-Mammographie-Screening in Deutschland hinsichtlich des potenziellen Nutzens und Schadens geprüft werden.

3. Fragestellungen

Ziel ist die Entwicklung und der Einsatz eines entscheidungsanalytischen Modells zur evidenzbasierten Evaluation der Langzeiteffektivität eines Programms zur Früherkennung von Brustkrebs bei Frauen im Alter von 45 bis 49 sowie zwischen 70 und 74 Jahren oder darüber hinaus in Deutschland.

Dabei sollen die folgenden Fragestellungen untersucht werden:

- Welchen zusätzlichen Nutzen und Schaden (bezogen auf Lebenserwartung, Lebensqualität, Diagnosen von frühen Stadien des Brustkrebses und Anzahl von Mammographien) hat der Einsatz einer Früherkennungsuntersuchung auf Brustkrebs mittels Mammographie für Frauen im Alter zwischen 45 und 49 Jahren (zusätzlich zu dem aktuell etablierten Mammographie-Screening im Alter von 50 bis 69 Jahren) in Deutschland im Vergleich zu keiner Früherkennungsuntersuchung in dieser Altersgruppe?

- Welchen zusätzlichen Nutzen und Schaden (bezogen auf Lebenserwartung, Lebensqualität, Diagnosen von frühen Stadien des Brustkrebses und Anzahl von Mammographien) hat der

Einsatz einer Früherkennungsuntersuchung auf Brustkrebs mittels Mammographie für Frauen im Alter zwischen 70 und 74 bzw. zwischen 70 und 79 Jahren (zusätzlich zu dem etablierten Mammographie-Screening im Alter von 50 bis 69 Jahren) in Deutschland im Vergleich zu keiner Früherkennungsuntersuchung in dieser Altersgruppe?

4. Methoden

4.1. Studiendesign

4.1.1. Entscheidungsanalytisches Modell

Für die Evaluation klinischer Langzeitkonsequenzen einer Brustkrebs-Früherkennungsuntersuchung für Frauen zwischen 45 und 49 Jahren, 70 und 74 sowie zwischen 75 und 79 Jahren wurde ein entscheidungsanalytisches Markov-Zustands-Übergangs-Modell (5-10) entwickelt und für den deutschen klinischen und epidemiologischen Kontext angepasst, kalibriert und validiert. Ein entscheidungsanalytisches Modell dient der systematischen, expliziten und quantitativen Evaluation von möglichen Handlungsstrategien. Durch eine Synthese von Evidenz aus verschiedenen Quellen und eine Berücksichtigung multipler Zielparametern zu Schaden und Nutzen kann ein entscheidungsanalytisches Modell basierend auf der verfügbaren Evidenz die „optimale“ Handlungsstrategie aus einer bestimmten Perspektive (Individuum, Gesellschaft, etc.) für einen vorgegebenen Zeithorizont identifizieren (7, 11, 12).

Das vorliegende entscheidungsanalytische Modell stellt die natürliche Entwicklung und Progression von Brustkrebs sowie die Auswirkungen von Screening und Therapie auf verschiedene gesundheitliche Endpunkte dar. Im Modell durchläuft eine hypothetische Kohorte von Frauen in Deutschland im Laufe ihrer Leben unterschiedliche Gesundheitszustände (Markov-Zustände), wie beispielsweise einen krebsfreien Zustand und verschiedene Stadien der Krebsentwicklung sowie Tod aufgrund von Brustkrebs oder aus anderer Ursache. Der natürliche Verlauf der Brustkrebsentwicklung wird im Modell durch Wahrscheinlichkeiten für den Übergang von einem Gesundheitszustand zu einem anderen Gesundheitszustand abgebildet. Neben der tumorbedingten stadienspezifischen Mortalitätsrate wird im Modell auch das Versterben aus anderen Gründen auf der Basis altersspezifischer Mortalitätsraten für Frauen in Deutschland berücksichtigt.

Das Markov-Zustands-Übergangs-Modell kann den Krankheitsverlauf des Brustkrebses widerspiegeln, da die Krankheit verschiedenen wohldefinierten, histologischen und klinischen Zuständen mit spezifischen Übergangs- und Ereigniswahrscheinlichkeiten folgt (8, 9, 13, 14). Außerdem kann ein Markov-Modell die wiederkehrenden Ereignisse des Screenings bis zum

finalen Endpunkt abbilden, welcher für die Evaluation von Screening-Programmen erforderlich ist. Dieser Modelltyp wurde daher für die hier beschriebene Analyse gewählt. Da die Anzahl der Gesundheitszustände überschaubar ist, wurde für die Analyse des Modells das Verfahren der Kohortensimulation gewählt. (14).

Die Modellstruktur spiegelt den Krankheitsverlauf sowie die Auswirkungen von Screening und Überwachung wider und ist in Abbildung 1 dargestellt. Der Krankheitsverlauf wird ausgehend von gesunden Frauen über die mögliche Entwicklung von ductalem Carcinoma in situ (DCIS) und einer Progression zu invasivem Brustkrebs oder direkter Entwicklung von invasivem Brustkrebs modelliert. DCIS kann sich zu invasivem Brustkrebs entwickeln oder regredieren. Präklinische (d. h. nicht entdeckte) Karzinome können von Stadium I bis Stadium IV progredieren. Die modellierten Stadien basieren auf der Klassifikation der Union for International Cancer Control (UICC).

Sowohl DCIS als auch invasive Karzinome können sowohl durch Früherkennungsuntersuchungen als auch durch das Auftreten von Symptomen bzw. Auffälligkeiten entdeckt werden. Im Modell beinhaltet eine symptomatische Entdeckung die Annahme, dass eine Mammographie zur Diagnose durchgeführt wird. Im Falle des entdeckten DCIS und entdeckter invasiver Karzinome wird im Modell die Durchführung einer den deutschen Leitlinien (1, 2) getreuen Abklärung und Behandlung angenommen.

Frauen mit einer initialen Krebsdiagnose können aufgrund des invasiven Brustkrebses basierend auf stadienspezifischen Mortalitätsraten oder aus anderen Ursachen sterben oder verbleiben für die Restlebenszeit in dem entsprechenden Krankheitszustand. Entsprechend der Kodierung in der Gesundheitsberichterstattung und der Dokumentation des Krankheitsverlaufs werden Behandlungen, Folgeuntersuchungen und mögliche Krankheitsprogression basierend auf dem initial entdeckten Krebsstadium berücksichtigt. Zu jedem Zeitpunkt können Frauen aus anderen Gründen als Brustkrebs basierend auf der alters- und geschlechtsspezifischen Mortalitätswahrscheinlichkeit in Deutschland versterben.

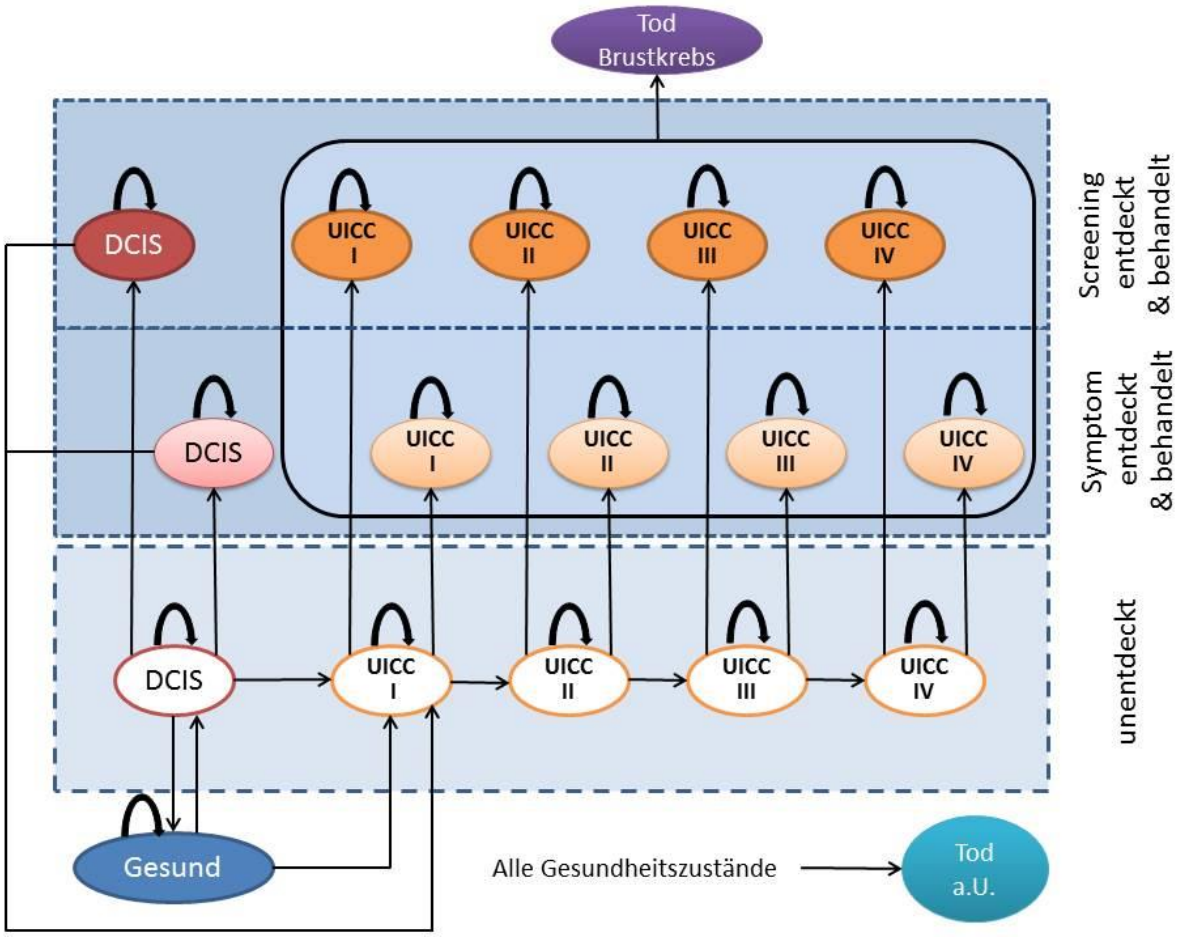


Abbildung 1 Natürlicher Krankheitsverlauf von Brustkrebs im Markov Zustand-Übergangs-Modell.

DCIS - Ductales Carcinoma in situ, UICC - Klassifikation der Union for International Cancer Control, a.U. andere Ursachen.
Jede Blase steht für einen Gesundheitszustand. Jeder Pfeil steht für mögliche Übergänge zwischen Gesundheitszuständen, die jedes Jahr mit bestimmten Übergangswahrscheinlichkeiten auftreten können. Alle Frauen beginnen im gesunden Zustand. Im Laufe der Zeit können die Frauen unentdecktes DCIS oder unentdeckten Brustkrebs im Stadium UICC I entwickeln. Unentdecktes DCIS kann zum Zustand Gesund regredieren. Unentdeckter Brustkrebs im Stadium UICC I kann zu fortgeschrittenen Stadien UICC II-IV fortschreiten. Unentdecktes DCIS und unentdeckter Krebs kann in jedem Stadium durch Symptome oder Screening diagnostiziert werden. Frauen mit diagnostiziertem DCIS oder diagnostiziertem Krebs (durch Symptome oder Screening) gehen in den jeweiligen diagnostizierten Gesundheitszustand und erhalten eine der Diagnose entsprechende Behandlung. Frauen mit diagnostiziertem DCIS können fortschreiten zu einem Zustand mit unentdecktem Brustkrebs im Stadium UICC I. Frauen mit diagnostiziertem Brustkrebs können an Brustkrebs versterben, wobei die Sterbewahrscheinlichkeit und somit die Zeit bis zum Tod vom initial entdecken Tumorstadium bestimmt wird. Alle Frauen können aus einem beliebigen Zustand heraus an anderen Ursachen sterben entsprechend der alters- und geschlechtsspezifischen Sterblichkeit in Deutschland.

Die Gesamteffektivität des Früherkennungsprogramms wird über die Brustdichte- und altersspezifische Testgenauigkeit (Sensitivität und Spezifität) der Mammographie, der Testgenauigkeit der weiteren Abklärungsverfahren sowie der stadienspezifischen Effektivität der Behandlung (berücksichtigt über das stadienspezifische Überleben) abgebildet. Dabei wird berücksichtigt, dass eine frühzeitige Entdeckung eines invasiven Brustkrebses im Modell durch ein höheres Überleben bei geringgradigen Krebsstadien zu einem Überlebensvorteil (durch die günstigere Prognose geringgradiger Krebsstadien) führen kann. Die Verschiebung der initial diagnostizierten Krebsstadien zu geringgradigen Stadien (hier durch eine Früherkennungs-maßnahme) wird Stadienverschiebung (Stage-Shift) genannt.

In dem vorliegenden Modell wurden mehrere Annahmen über den natürlichen Verlauf der Krebserkrankung getroffen:

- Frauen können DCIS entwickeln und dieses kann ohne Behandlung wieder regredieren.
- Die Krankheitsentwicklung ist ausschließlich abhängig vom aktuellen Gesundheits-zustand und dem Alter der Frau sowie basierend auf der Gewebedichte der Brust, aber nicht von vorherigen Gesundheitszuständen.
- Frauen mit entdecktem invasivem Krebs verbleiben entsprechend der Diagnosepraxis und -dokumentation im Modell bis zum Tod in dem initial diagnostizierten Krankheitsstadium und können auf Basis der stadienspezifischen Mortalitätsraten des initialen Krebsstadiums an Krebs oder auf Basis der altersspezifischen Mortalitätsraten für Frauen in Deutschland aus anderer Ursache als Krebs versterben.
- Innerhalb des gleichen initialen UICC-Stadiums unterscheiden sich die stadien-spezifischen Mortalitätsraten von Frauen mit durch Screening entdecktem invasivem Krebs gegenüber Frauen mit symptomatisch entdecktem invasivem Krebs (15).
- Frauen mit entdecktem DCIS oder entdecktem invasivem Brustkrebs werden gemäß der deutschen Leitlinie behandelt. Der Behandlungsverlauf wird nicht explizit modelliert. Stattdessen werden durchschnittliche stadienspezifische Überlebens-wahrscheinlichkeiten nach Erstdiagnose angewandt.
- Frauen mit symptomatisch entdecktem DCIS oder invasivem Krebs erhalten zur Abklärung der Diagnose eine Mammographie, die in die evaluierte Gesamtzahl der Mammographien einer Strategie einfließen.
- Der Einfluss eines positiven Screeningbefunds auf die Lebensqualität wird durch ein 3-wöchiges Lebensqualitätsdekrement modelliert. Frauen mit einem falsch-positiven Befund wird nach 3 Wochen wieder die Lebensqualität einer gesunden Frau zugeordnet.

Die Modellentwicklung und –analyse folgt internationalen Standards und Empfehlungen zur Durchführung und dem Reporting von entscheidungsanalytischen Studien (8, 9, 13, 14, 16, 17).

4.1.2. Zielpopulation

Zielpopulation für eine Brustkrebs-Früherkennungsuntersuchung in dieser Entscheidungsanalyse sind Frauen in Deutschland in verschiedenen Altersgruppen.

4.1.3. Evaluierte Strategien

Folgende Strategien werden in der vorliegenden entscheidungsanalytischen Studie verglichen, wobei der derzeitige Screening-Standard in Deutschland als Referenzstrategie dient.

1) Kein Screening
2) Referenzstrategie: Mammographie, alle 2 Jahre im Alter 50 - 69 Jahre (Standard in Deutschland)
3) Mammographie, jährlich im Alter 45 - 49 Jahre und alle zwei Jahre im Alter 50 - 69 Jahre
4) Mammographie alle zwei Jahre im Alter 45 - 69 Jahre
5) Mammographie, jährlich im Alter 45 - 49 Jahre und alle zwei Jahre im Alter 50 - 74 Jahre
6) Mammographie alle zwei Jahre im Alter 45 - 74 Jahre
7) Mammographie, jährlich im Alter 45 - 49 Jahre und alle zwei Jahre im Alter 50 - 79 Jahre
8) Mammographie alle zwei Jahre im Alter 45 - 79 Jahre
9) Mammographie, jährlich im Alter 45 - 49 Jahre, alle zwei Jahre im Alter 50 - 69 Jahre und alle drei Jahre im Alter 70 - 74 Jahre
10) Mammographie, jährlich im Alter 45 - 49 Jahre, alle zwei Jahre im Alter 50 - 69 Jahre und alle drei Jahre im Alter 70 - 79 Jahre
11) Mammographie alle zwei Jahre im Alter 50 - 74 Jahre
12) Mammographie alle zwei Jahre im Alter 50 - 79 Jahre
13) Mammographie alle zwei Jahre im Alter 50 - 69 und alle drei Jahre im Alter 70 – 74 Jahre
14) Mammographie alle zwei Jahre im Alter 50 - 69 und alle drei Jahre im Alter 70 – 79 Jahre

4.1.4. Perspektive

In den Analysen wird die Perspektive der Zielpopulation, d. h. der Frauen ab einem Alter von 45 Jahren in Deutschland eingenommen und es werden die Ergebnisse für eine regelmäßige und vollständige Teilnahme an der Früherkennungsuntersuchung der jeweils untersuchten Screeningstrategien und den Folgebehandlungen bestimmt.

4.1.5. Analytischer Zeithorizont und Zykluslänge

Da ein erwarteter Nutzen und ein möglicher erwarteter Schaden eines Früherkennungsprogramms lebenslang relevant sind und die damit verbundenen gesundheitsbezogenen Ereignisse im zeitlichen Verlauf zu unterschiedlichen Zeitpunkten und wiederholt auftreten können, wird internationalen Empfehlungen folgend ein lebenslanger Zeithorizont für die entscheidungsanalytische Modellierung gewählt (8, 9). Für das Markov-Modell und die entsprechenden Übergangswahrscheinlichkeiten wurde eine Zykluslänge von einem Jahr gewählt.

4.1.6. Zielparameter (Endpunkte)

Es werden verschiedene Zielparameter, die den langzeitigen Nutzen oder möglichen Schaden darstellen, in den Analysen berechnet:

1. Lebenszeitrisiko für entdeckte invasive Brustkrebskarzinome (BK) (in %)
2. Lebenszeitrisiko für entdeckte ductale Carcinoma in situ (DCIS) (in %)
3. Lebenszeitrisiko, an entdecktem invasivem Brustkrebskarzinom zu versterben (in %)
4. Verbleibende Restlebenserwartung (in Jahre) pro 100 Frauen
5. Verbleibende lebensqualitätsadjustierte Restlebenserwartung (in qualitätsadjustierte Lebensjahre (QALY)) pro 100 Frauen
6. Absolute Anzahl von Screeningbefunden (Mammographien) pro 100 Frauen
7. Absolute Anzahl positiver und falsch-positiver Screeningbefunde (Mammographien) pro 100 Frauen
8. Lebenszeitrisiko für eine Überdiagnose (DCIS und Brustkrebs) (in %)
9. Number-Needed-To-Screen (NNS), um einen Brustkrebs-bedingten Todesfall zu vermeiden

4.2. Daten für die entscheidungsanalytische Modellierung

In diesem Kapitel werden die Daten und der Umgang mit diesen Daten für die Modellierung beschrieben. Für jeden Übergang von einem zu einem anderen Zustand im Modell werden jährliche Übergangswahrscheinlichen zugeordnet. Vorab ist zu erwähnen, dass Übergangs-wahrscheinlichkeiten im Modell nicht immer basierend auf Beobachtungsdaten bezogen werden können. Dies gilt insbesondere für die frühen Übergänge der nicht symptomatischen Erkrankung, da es hierzu naturgemäß wenige bzw. keine systematischen Studien bzw. Evidenz gibt. Für diese Übergangswahrscheinlichkeiten wurde wie in Screeningmodellen üblich ein Kalibrierungsprozess angewandt. Da im Modell für symptomatisch entdeckte und durch Screening diagnostizierte Krebsfälle unterschieden wird, wird die Kalibrierung der Übergangswahrscheinlichkeiten zunächst

in einem Setting ohne Screening durchgeführt. Hierfür werden alters- und stadienspezifische Beobachtungsdaten aus Deutschland vor Einführung des Screenings (Jahre 1999-2004) herangezogen und angenommen, dass sich die momentane Brustkrebsinzidenz ohne Screening von der damaligen Inzidenz ohne Screening nicht maßgeblich unterscheidet. Die Effektivität des Mammographie-Screenings wird anhand von klinischen Parametern modelliert. Selektierte modellberechnete Daten für die Referenzstrategie werden anhand von aktuellen alters- und stadienspezifischen Beobachtungsdaten in der weiblichen Bevölkerung bzw. in der am Mammographie-Screening teilnehmenden weiblichen Bevölkerung in Deutschland validiert.

Die folgenden Unterkapitel beschreiben die Datenerhebung und die Input-, Kalibrierungs- und Validierungsparameter, sowie den Kalibrierungs- und Validierungsprozess und für das Modell erforderliche Parametertransformationen.

4.2.1. Datenerhebung

Für die Modellierung sind unterschiedliche Daten aus verschiedenen Evidenzquellen erforderlich (z.B. Krebsregisterdaten, Literaturdaten). Alle Daten, die für die Modellierung und die Evaluation benötigt wurden, wurden mittels systematischer Recherchen in zwei Datenbanken zusammengestellt. In diesen Datenbanken wurden Parameterwerte entsprechend der derzeitigen Evidenzlage zu Früherkennung- und Behandlungsrichtlinien, sowie der Krankheitsprogression und Epidemiologie und den Mammographie-Testeigenschaften zusammengetragen.

Bei der ersten Datenbank handelt es sich um die epidemiologische Datenbank, welche Prävalenzen, Krankheitsprogressionsparameter, Inzidenzen, Stadienverteilungen, brustkrebsstadienabhängige relative Überlebenswahrscheinlichkeiten, Hintergrundsterblichkeit etc. beinhaltet. Die zweite Datenbank ist die klinische Datenbank, welche Parameter wie Testgütewerte, Lebensqualitätsdaten etc. enthält. Selektierte Daten aus diesen Datenbanken wurden in das mathematische Modell wie in den nachfolgenden Abschnitten beschrieben eingepflegt.

4.2.2. Daten zur Krankheitsentwicklung

Die Krankheitsprogressionswahrscheinlichkeiten (Tabelle 1) wurden aus der Literatur entnommen und im Sinne der Evidenzbasierung im Kontext des deutschen Gesundheitssystems anhand empirisch beobachteter epidemiologischer Daten aus Deutschland in den Jahren vor der Einführung des Mammographie-Screenings angepasst. Die Übergangswahrscheinlichkeiten von „Gesund“ zu „DCIS“ und zu „unentdeckter Brustkrebs UICC-Stadium I“ wurden kalibriert, d. h. die Übergangswahrscheinlichkeiten wurden systematisch und iterativ solange variiert, bis das Modell beobachtete Daten zu den Zielparametern altersspezifische Häufigkeit von entdecktem DCIS (18)

(siehe Kapitel 4.2.5) und altersspezifische Brustkrebsinzidenz (19) vor der Einführung des Mammographie-Screenings in Deutschland gut wiedergab. Ebenso wurden die Übergangswahrscheinlichkeiten für die Progression von unentdecktem invasivem Brustkrebs von Stadium UICC I bis UICC III nach Stadium UICC II bis UICC IV sowie die symptomatische Entdeckungsrate so kalibriert, dass die beobachtete Verteilung der UICC-Stadien entdeckter Brustkrebsfälle in Deutschland vor Screening-Einführung abgebildet wurde (siehe Kapitel 4.2.5). Als Startkonfiguration der Parameter im Kalibrierungsprozess dienten hier jeweils die jährlichen Übergangswahrscheinlichkeiten aus einer österreichischen Modellierungsstudie (20). Die Übergangswahrscheinlichkeiten vom Markov Zustand „DCIS“ zu „Gesund“ (Regression), entdecktem DCIS und unentdecktem invasivem Brustkrebs (Stadium UICC I) basieren ebenfalls auf einem österreichischen Sub-Modell (20).

Die aus der Literatur entnommenen und die in der Kalibrierung ermittelten Übergangswahrscheinlichkeiten des natürlichen Krankheitsverlaufs sind in Tabelle 1 aufgeführt. Die Zielparameter der Kalibrierung sind in Absatz 4.2.5 in Tabelle 7 zusammengefasst.

Für die Hintergrundsterblichkeit wurden die altersspezifischen Mortalitätswahrscheinlichkeiten basierend auf der Sterbetafel für die deutsche weibliche Bevölkerung herangezogen (Destatis. Sterbetafel nach Alter und Geschlecht 2018/2020 (21)). In den Markov-Zuständen „entdeckter invasiver Brustkrebs Stadium UICC I bis IV“ wurde diese altersspezifische Hintergrundmortalität adjustiert für die altersspezifische Brustkrebsmortalität (19). Für Modellierung der erkrankungsspezifischen Mortalität wurden stadienspezifische relative Überlebensraten basierend auf Krebsregisterdaten (22) herangezogen (Tabelle 2). Diese Daten mussten für die Modellierung von der TNM-Klassifikation (Tumor, Nodes, Metastasen) in die UICC-Klassifikation transformiert werden. Hierbei wurde aufgrund fehlender weiterführender Information die TNM-Stadien T1N+M0 und T2N+M0 jeweils zur Hälfte auf das UICC-Stadium II und das UICC-Stadium III verteilt.

Da sich symptomatisch entdeckte von durch Screening entdeckten Krebsfälle bezüglich ihrer Charakteristika und damit verbundenen Prognose systematisch unterscheiden können, muss der Überlebensvorteil durch einen Stage-Shift im Modell für den unterschiedlichen Entdeckungsmodus adjustiert werden. Hierzu wurden die stadienspezifischen Überlebenswahrscheinlichkeiten spezifisch für den jeweiligen Entdeckungsmodus (Screening versus Symptome) adjustiert und ins Modell eingesetzt. Für diese Adjustierung wurde basierend auf Daten von drei großen Brustkrebs-Screening Studien (15) für symptomatisch entdeckte Brustkrebsfälle ein relatives Risiko an Brustkrebs zu versterben von 1,36 zugrunde gelegt.

Tabelle 1 Modellparameter des natürlichen Krankheitsverlaufs.

Übergang von	nach	Alter (Jahre)	Jährliche Übergangs-wahrscheinlichkeit	Quelle
Gesund	DCIS unentdeckt	0-12	0	kalibriert*[1]
		13-29	0,000068	
		30-39	0,002047	
		40-45	0,004536	
		45-49	0,004353	
		50-54	0,004288	
		55-59	0,003184	
		60-64	0,002534	
		65+	0,000122	
Gesund	UICC I unentdeckt	0-14	0	kalibriert*[2]
		15-19	0,000001	
		20-24	0,000018	
		25-29	0,000184	
		30-34	0,000449	
		35-39	0,000663	
		40-44	0,001402	
		45-49	0,001611	
		50-54	0,002361	
		55-59	0,002802	
		60-64	0,002910	
		65+	0,002993-0,003311	
DCIS unentdeckt	Gesund		0,068184	Schiller-Fruehwirth 2017 (20), kalibriert*[2]
	DCIS symptomatisch entdeckt*[2]		0,009446	
	UICC I unentdeckt		0,123799	
UICC I unentdeckt	UICC I symptomatisch entdeckt*[2]		0,163963	
	UICC II unentdeckt*[2]		0,339696	
UICC II unentdeckt	UICC II symptomatisch entdeckt*[2]		0,394884	
	UICC III unentdeckt*[2]		0,391166	
UICC III unentdeckt	UICC III symptomatisch entdeckt*[2]		0,528699	
	UICC IV unentdeckt*[2]		0,619020	
UICC IV unentdeckt	UICC IV symptomatisch entdeckt*[2]		0,829060	

UICC - Union for International Cancer Control classification, DCIS - Ductales Carcinoma in situ

*[1] Kalibriert auf altersspezifische DCIS Inzidenz für Frauen in Deutschland vor Einführung des Screenings (Jahr 2003) (18)

*[2] Kalibriert auf altersspezifische Inzidenz und UICC-Stadienverteilung entdeckter Brustkrebsfälle vor Einführung des Screenings (19)

Tabelle 2 Relative Überlebenswahrscheinlichkeit für Patientinnen mit entdecktem Brustkrebs.

Jahr nach Erstdiagnose	Relative Überlebenswahrscheinlichkeit für Patienten nach Erstdiagnose von entdecktem Brustkrebs (22)			
	UICC I*	UICC II*	UICC III*	UICC IV*
	1	1	1	1
1. Jahr	1,000	0,998	0,983	0,739
2. Jahr	1,000	0,990	0,960	0,783
3. Jahr	1,000	0,979	0,954	0,763
4. Jahr	1,000	0,979	0,958	0,790
5. Jahr	1,000	0,977	0,961	0,811
6. Jahr	1,000	0,979	0,960	0,834
7. Jahr	1,000	0,978	0,962	0,852
8. Jahr	0,997	0,980	0,969	0,856
9. Jahr	0,997	0,980	0,967	0,884
10. Jahr	0,993	0,978	0,966	0,882
11. Jahr	0,994	0,982	0,970	0,896
12. Jahr	0,994	0,977	0,967	0,967
13. Jahr	0,990	0,980	0,977	0,905
14. Jahr	0,991	0,979	0,974	0,914
15.Jahr	0,991	0,981	0,978	0,948

UICC - Union for International Cancer Control classification
* Daten wurden auf UICC transferiert

4.2.3. Klinische Daten

Frauen unterschiedlichen Alters haben eine unterschiedliche Dichte des Brustgewebes. Die altersspezifischen Häufigkeiten von Frauen in Deutschland mit Brustdichtekategorien basierend auf der Klassifikation des Breast Imaging Reporting and Data System (BI-RADS) wurden einer Modellierungsstudie (23-25) entnommen (Tabelle 3). Die Dichte des Brustgewebes hat Einfluss auf 1) das Risiko DCIS und invasiven Brustkrebs zu entwickeln und 2) die Testgüte der Mammographie.

Tabelle 3 Altersspezifischer Anteil (in %) der BI-RADS Brustdichte Kategorien in der weiblichen Bevölkerung in Deutschland (23-25).

Alter (Jahre)	BI-RADS 1 (%)	BI-RADS 2 (%)	BI-RADS 3 (%)	BI-RADS 4 (%)
45-49§	4,8	31,9	55,3	8,0
50–54	6,1	38,7	48,3	6,9
55–59	7,4	45,5	41,3	5,8
60–64	8,6	48,5	37,9	5,0
65–69	9,45	50,0	36,1	4,45
70+	10,0	50,85	35,0	4,15

BI-RADS - Breast Imaging Reporting and Data System
BI-RADS 1: fast fettig (< 25% drüsenförmig)
BI-RADS 2: verstreute fibroglanduläre Dichten (25-50% drüsenförmig)
BI-RADS 3: ungleichmäßig dicht (51-75% drüsenförmig)
BI-RADS 4: extrem dicht (> 75% drüsenförmig)
§ berechnet aus den absoluten Differenzen der Alterskategorie 50-54 vs. 55-59 in der jeweiligen BI-RADS Kategorie

Die Evidenz zur Testgüte der Mammographie (26) und für das relative Risiko für die Entwicklung von DCIS und invasivem Brustkrebs in Abhängigkeit von der BI-RADS Brustdichte-Kategorie (23) wurde aus der Literatur entnommen. In Tabelle 4 sind die für das Modell herangezogenen Daten zur Sensitivität und Spezifität der Mammographie in Abhängigkeit von Brustdichte und Alter der Frau aufgeführt. Tabelle 5 zeigt das jeweilige relative Risiko für die Entwicklung von Brustkrebs in Abhängigkeit von der BI-RADS Kategorie.

Tabelle 4 Brustdichte- und Alters- spezifische Testgenauigkeit (Sensitivität und Spezifität) von Mammographie (26).

Brustdichte	Altersgruppe (Jahre)	Intervall [1]	Sensitivität (invasiver Krebs)	Sensitivität (DCIS)	Spezifität
BI-RADS 1	40-49	Erste	0,921	0,954	0,872
		Jährlich	0,806	0,919	0,930
		Zweijährlich	0,881	0,942	0,925
	50-64	Erste	0,948	0,955	0,903
		Jährlich	0,868	0,921	0,948
		Zweijährlich	0,921	0,943	0,944
	≥65	Erste	0,963	0,955	0,916
		Jährlich	0,903	0,922	0,955
		Zweijährlich	0,943	0,944	0,952
BI-RADS 1	40-49	Erste	0,894	0,948	0,797
		Jährlich	0,751	0,911	0,884
		Zweijährlich	0,844	0,935	0,876
	50-64	Erste	0,930	0,949	0,843
		Jährlich	0,826	0,912	0,912
		Zweijährlich	0,895	0,937	0,906
	≥65	Erste	0,950	0,950	0,863
		Jährlich	0,871	0,913	0,924
		Zweijährlich	0,924	0,937	0,919
BI-RADS 3	40-49	Erste	0,817	0,964	0,760
		Jährlich	0,615	0,937	0,860
		Zweijährlich	0,740	0,955	0,851
	50-64	Erste	0,876	0,965	0,812
		Jährlich	0,716	0,938	0,894
		Zweijährlich	0,818	0,956	0,886
	≥65	Erste	0,909	0,965	0,836
		Jährlich	0,782	0,938	0,908
		Zweijährlich	0,865	0,956	0,901
BI-RADS 4	40-49	Erste	0,746	0,943	0,815
		Jährlich	0,512	0,902	0,895
		Zweijährlich	0,652	0,929	0,888
	50-64	Erste	0,822	0,944	0,857
		Jährlich	0,623	0,904	0,921
		Zweijährlich	0,747	0,930	0,915
	≥65	Erste	0,868	0,944	0,876
		Jährlich	0,702	0,904	0,932
		Zweijährlich	0,808	0,931	0,927

BI-RADS - Breast Imaging Reporting and Data System, DCIS - Ductales Carcinoma in situ

BI-RADS 1: fast fettig (< 25% drüsenförmig)

BI-RADS 2: verstreute fibroglanduläre Dichten (25-50% drüsenförmig)

BI-RADS 3: ungleichmäßig dicht (51-75% drüsenförmig)

BI-RADS 4: extrem dicht (> 75% drüsenförmig)

[1] durch Screening entdeckter Krebs in einem Jahr diagnostiziert; Cut-off beim nächsten Screening

Tabelle 5 Relatives Risiko für Brustkrebs in Abhängigkeit von BI-RADS Brustdichte Kategorien (23).

Altersgruppe (Jahre)	Brustdichte			
	BI-RADS 1	BI-RADS 2	BI-RADS 3	BI-RADS 4
40-49§	0,376	0,783	1,212	1,573
50–59	0,388	0,807	1,251	1,623
60–64	0,400	0,832	1,291	1,675
65–69	0,581	0,885	1,228	1,283
70+	0,600	0,914	1,268	1,325

BI-RADS - Breast Imaging Reporting and Data System
BI-RADS 1: fast fettig (< 25% drüsenförmig)
BI-RADS 2: verstreute fibroglanduläre Dichten (25-50% drüsenförmig)
BI-RADS 3: ungleichmäßig dicht (51-75% drüsenförmig)
BI-RADS 4: extrem dicht (> 75% drüsenförmig)
§berechnet mit relativen Gewichten der Alterskategorien (50-54J) zu (55-59J): Faktor 0,376 – 1,573

Informationen und Daten zur klinischen Praxis wie beispielsweise die Abfolge und der Prozess der Kontrolluntersuchungen nach entdecktem DCIS sowie zur leitlinienkonformen stadienabhängigen Therapie des Brustkrebs wurden den aktuellen Leitlinien (1) in Deutschland entnommen.

4.2.4. Lebensqualitätsdaten

Im Modell wurden für alle Gesundheitszustände multiattributive Nutzwerte für die gesundheitsbezogene Lebensqualität hinterlegt. Altersspezifische Nutzwerte basierend auf dem Lebensqualitätsinstrument EQ-5D für die deutsche weibliche Population wurden mit den krankheitsspezifischen relativen EQ-5D Nutzwerten in Abhängigkeit vom Alter gewichtet. Die deutsche altersspezifische Lebensqualität basierte auf einer Publikation von Janssen et al. (27) und wurde mit altersspezifischen relativen Nutzwerten für DCIS und Brustkrebs aus einer schwedischen Studie (23, 28, 29) multipliziert. In der Publikation von Janssen et al. (27) wurden drei Studien, in denen deutsche altersabhängige Lebensqualitätsdaten berichtet wurden, gepoolt. Da in diesen Studien die EQ-5D Nutzwerte nicht getrennt für das weibliche Geschlecht berichtet wurden, wurden die Nutzwerte für die Gesamtpopulation angewandt. Janssen et al. (27) berichteten jedoch ähnliche Visuelle Analog Skala Lebensqualitätswerte für Frauen und Männer, so dass die Annahme ähnlicher EQ-5D-Werte akzeptabel erscheint. Zusätzlich wurden absolute Nutzwert-Dekremente für die Mammographie (23) und für einen positiven Testbefund (30) basierend auf der Methode der zeitlichen Abwägung (Time-trade-off, TTO) berücksichtigt. Es wurde angenommen, dass unentdeckter Brustkrebs keine relevanten Symptome zeigt und somit die altersspezifische Lebensqualität nicht beeinflusst. Die Nutzwerte für die gesundheitsbezogene Lebensqualität sind in Tabelle 6 dargestellt.

Tabelle 6 Nutzwerte für die gesundheitsbezogene Lebensqualität.

Lebensqualität deutsche weibliche Bevölkerung (EQ-5D)				
	Alter (Jahre)		**Nutzwert**	**Quelle**
	0-17		1	**(27)**
	18-24		0,950	
	25-34		0,949	
	35-44		0,943	
	45-54		0,908	
	55-64		0,881	
	65-74		0,838	
	75+		0,771	
Nutzwert-Dekrement für Screening (TTO)				
	Nutzwert Dekrement	**Dauer**	**Dekrement über 1 Jahr**	
Mammographie	0,196	2 Wochen	0,0075	**(23)**
Positiver Befund	0,105	5 Wochen	0,0100	**(30)**
Alters- und stadienspezifische relative Nutzwerte für DCIS und Brustkrebs				
	Alter (Jahre)	**Dauer**	**Relativer Nutzwert**	
DCIS	0-39	Erstes Jahr	1	
		Weitere Jahre	1	
	40-49	Erstes Jahr	0,905	**Berechnet anhand von schwedischen Daten (23, 28, 29)**
		Weitere Jahre	1	
	50-59	Erstes Jahr	0,889	
		Weitere Jahre	1	
	60-69	Erstes Jahr	0,904	
		Weitere Jahre	1	
	70-79	Erstes Jahr	0,904	
		Weitere Jahre	1	
	80+	Erstes Jahr	0,904	
		Weitere Jahre	1	
UICC I	0-39	Erstes Jahr	1	
		Weitere Jahre	1	
	40-49	Erstes Jahr	0,846	
		Weitere Jahre	0,980	
	50-59	Erstes Jahr	0,846	
		Weitere Jahre	0,980	
	60-69	Erstes Jahr	0,846	
		Weitere Jahre	0,980	
	70-79	Erstes Jahr	0,847	
		Weitere Jahre	0,980	
	80+	Erstes Jahr	0,847	
		Weitere Jahre	0,980	

Alters- und stadienspezifische relative Nutzwerte für DCIS und Brustkrebs

	Alter (Jahre)	**Dauer**	**Relativer Nutzwert**
UICC II	0-39	Erstes Jahr	1
		Weitere Jahre	1
	40-49	Erstes Jahr	0,753
		Weitere Jahre	0,905
	50-59	Erstes Jahr	0,753
		Weitere Jahre	0,905
	60-69	Erstes Jahr	0,753
		Weitere Jahre	0,905
	70-79	Erstes Jahr	0,753
		Weitere Jahre	0,905
	80+	Erstes Jahr	0,753
		Weitere Jahre	0,906
UICC III	0-39	Erstes Jahr	1
		Weitere Jahre	1
	40-49	Erstes Jahr	0,753
		Weitere Jahre	0,905
	50-59	Erstes Jahr	0,753
		Weitere Jahre	0,905
	60-69	Erstes Jahr	0,753
		Weitere Jahre	0,905
	70-79	Erstes Jahr	0,753
		Weitere Jahre	0,905
	80+	Erstes Jahr	0,753
		Weitere Jahre	0,906
UICC IV	0-39	Erstes Jahr	1
		Weitere Jahre	1
	40-49	Erstes Jahr	0,753
		Weitere Jahre	0,832
	50-59	Erstes Jahr	0,753
		Weitere Jahre	0,832
	60-69	Erstes Jahr	0,753
		Weitere Jahre	0,833
	70-79	Erstes Jahr	0,753
		Weitere Jahre	0,833
	80+	Erstes Jahr	0,753
		Weitere Jahre	0,832

EQ-5D - European Quality of Life 5 Dimensions; UICC - Union for International Cancer Control classification; TTO - time trade-off

4.2.5. Kalibrierung

Die Übergangswahrscheinlichkeiten von „Gesund“ zu „DCIS“ oder zu „unentdecktem invasivem Brustkrebs UICC I“ sowie die Entdeckungsrate von DCIS und invasivem Brustkrebs und Progressionswahrscheinlichkeiten von unentdecktem invasivem Brustkrebs wurden kalibriert. Hierzu wurden folgende Kalibrierungszielparameter benutzt (siehe auch Tabelle 7):

- Altersspezifische Inzidenz von entdecktem DCIS vor Screening-Einführung (Jahr 2003)
- Altersspezifische Inzidenz entdeckter Brustkrebsfälle in Deutschland vor Screening-Einführung (Durchschnitt über die Jahre 1999 bis 2004)
- Stadienverteilung von entdeckten Brustkrebsfällen (UICC I-IV) in Deutschland vor Screening-Einführung (Durchschnitt über die Jahre 2002 bis 2004)

Der Übergang von Gesund zu DCIS (unentdeckt) wurde anhand von altersspezifischer Inzidenz von DCIS vor Screening-Einführung (Daten aus dem Jahr 2003) kalibriert. Als Startparameter dienten berechnete Werte aus altersspezifischen DCIS Inzidenzen des Jahres 2019, also nach Einführung des Brustkrebsscreenings. Diese altersspezifischen Übergangswahrscheinlichkeiten wurden in der Kalibrierung solange systematisch verändert bis in der Strategie „Kein Screening“ des entscheidungsanalytischen Modells die beobachtete altersspezifische DCIS Inzidenz für Frauen in Deutschland vor Einführung des Mammographie-Screenings basierend auf Daten von DESTATIS (aus dem Jahr 2003) (18) hinreichend abgebildet wurde.

Ferner wurden im entscheidungsanalytischen Modell die Übergangswahrscheinlichkeiten von „Gesund“ sowie von „DCIS“ zu „unentdecktem invasivem Brustkrebs UICC I“ anhand der beobachteten altersspezifischen Inzidenz entdeckter Brustkrebsfälle und beobachteter Stadienverteilung entdeckter Brustkrebsfälle in Deutschland vor Screening-Einführung basierend auf Krebsregisterdaten (19) kalibriert. Die Wahrscheinlichkeit der Progression von nicht-entdeckten Brustkrebsfällen von UICC I nach UICC II, von UICC II nach UICC III und von UICC III nach UICC IV wurde so kalibriert, dass das Modell die beobachteten altersspezifischen Brustkrebsinzidenzen und die stadienspezifische Verteilung (31) in Deutschland prädiziert. Als Startkonfigurationsparameterwerte für die Kalibrierung wurden die Wahrscheinlichkeiten aus der österreichischen Modellierungsstudie (20) eingesetzt.

Die Übergangswahrscheinlichkeiten von unentdecktem zu entdecktem Brustkrebs in der Strategie „Kein Screening“ des Modells basiert auf den stadienspezifischen Wahrscheinlichkeiten aus einer österreichischen Modellierungsstudie (20). Diese Startparameterwerte wurden so angepasst, dass die vom Modell berechneten entdeckten Krebsfälle die beobachteten altersspezifischen Brustkrebsinzidenzen (19) und die Stadienverteilung entdeckter Krebsfälle (31) in Deutschland vor Screening-Einführung widerspiegeln.

Tabelle 7 Zielparameter der Modellkalibrierung.

Altersspezifische Inzidenz DCIS vor Screening-Einführung (Jahr 2003)		
Alter (Jahre)	**Inzidenz**	**Quelle**
40-44	0,000148	(18)
45-49	0,000257	
50-54	0,000318	
55-59	0,000345	
60-64	0,000351	
65-69	0,000282	
70-74	0,000191	
75-79	0,000119	
80-84	0,000066	
85-89	0,000148	
Altersspezifische Inzidenz von Brustkrebs in Deutschland (Durchschnitt der Jahre 1999-2004)		
Alter (Jahre)	**Inzidenz**	
15-19	0,000092	(19)
20-24	0,000266	
25-29	0,000531	
30-34	0,001001	
35-39	0,001696	
40-44	0,002146	
45-49	0,002695	
50-54	0,003068	
55-59	0,003033	
60-64	0,003062	
65-69	0,003307	
70-74	0,003311	
75-79	0,000092	
80-84	0,000266	
85 +	0,000531	
Stadienverteilung von entdecktem Brustkrebs (Durchschnitt über die Jahre 2002 bis 2004)		
UICC I	0,3805	(31)
UICC II	0,3953	
UICC III	0,1465	
UICC IV	0,0775	

UICC - Union for International Cancer Control classification
DCIS - Ductales Carcinoma in situ

4.2.6. Validierung

Da das Ziel einer jeden Modellierung ist, die Realität zwar vereinfacht, aber realistisch darzustellen, kommt der Bewertung der Validität des Modells eine wichtige Rolle zu. Die internationale ISPOR-SMDM Modeling Good Research Practices Task Force hat hierzu Empfehlungen erarbeitet und publiziert (32). Hier werden fünf Arten der Validierung empfohlen,

wobei nicht immer alle fünf Schritte erfolgen können bzw. müssen. Diese fünf Schritte beinhalten Augenscheinvalidität (Face validity), interne Validität (internal validity), Verifikation (verification), externe Validität (external validity) und voraussagende Validität (32). Zusätzlich wurden selektierte Endpunkte mit denselben Endpunkten aus anderen Modellierungsstudien (23, 26) verglichen (Cross-Model Validation).

Bei dem vorliegenden Screening-Modell wurde die Augenscheinvalidität geprüft, indem die Modellstruktur und Modellparameter auf validierten und publizierten Studien basieren und in einem Team von Mitarbeiterinnen und Mitarbeitern der UMIT TIROL und des IQWIG diskutiert wurden. Die interne Validität wurde geprüft, indem für das Modell verwendete Inputparameter reproduziert durchgeführt wurden. Für die technische Verifikation wurden im Sinne eines Debuggings verschiedene systematische Verfahren eingesetzt. Unter anderem wurden die Modell-Outputs (sogenannte Markovspur) durch systematische Inspektion analysiert und auf Plausibilität geprüft. Ferner wurden Modellparameter ausgeschaltet bzw. Extremwertanalysen durchgeführt und das Verhalten des Modells und seiner Ergebnisse auf Plausibilität überprüft. Für die externe Validität wurden mittels des Modells errechnete epidemiologische Parameter systematisch mit externen epidemiologischen Beobachtungsdaten verglichen, die nicht im Modell als Input- oder Kalibrierungsparameter verwendet wurden (Tabelle 8). Als externe Validierungszielparameter dienten die beobachteten Daten in Deutschland zum Lebenszeitrisiko für die Entwicklung eines invasiven Brustkrebses und daran zu versterben, als auch die Brustkrebsstadienverteilung basierend auf aktuellen Daten des Robert-Koch Instituts (33). Da die Beobachtungsdaten auf einer Population von Frauen basiert, die nur zum Teil, also nicht vollständig, am Mammographie-Screening teilnahmen, und die Teilnahme am Screening auch altersabhängig variierte, wurden für den Validierungsprozess entsprechende Annahmen für eine durchschnittliche Teilnahmerate gemacht.

Tabelle 8 Validierungsparameter der externen Modellvalidierung.

	Wert	Quelle
Inzidenz entdeckter Brustkrebs pro 100.000 (Jahr 2016)	112,2	(33)
Stadienverteilung entdeckter Brustkrebsfälle in Deutschland (Jahr 2016)		
UICC I	41%	(33)
UICC II	39%	
UICC III	13%	
UICC IV	7%	
Brustkrebs Lebenszeitrisiko (Jahr 2016)	12,2%	(33)
Lebenszeitrisiko an Brustkrebs zu versterben (Jahr 2016)	3,6%	(33)

UICC - Union for International Cancer Control Classification

4.3. Analysen

4.3.1. Basisfallanalyse

Mit Hilfe des entscheidungsanalytischen Modells werden deterministische Kohortensimulationen durchgeführt und für jede der alternativen Strategien die Erwartungswerte zu den folgenden gesundheitsbezogenen Endpunkten berechnet:

1. Nutzenendpunkte
   - Lebenszeitrisiko (in %) für entdeckte invasive Brustkrebskarzinome (BK)
   - Lebenszeitrisiko (in %) für entdeckte ductales Carcinoma in situ (DCIS)
   - Lebenszeitrisiko (in %), an entdecktem invasivem Brustkrebskarzinom zu versterben
   - Verbleibende Restlebenserwartung (in Jahre) pro 100 Frauen
   - Verbleibende lebensqualitätsadjustierte Restlebenserwartung (in QALY) pro 100 Frauen
2. Schadenendpunkte
   - Absolute Anzahl von Screeningbefunde (Mammographien) pro 100 Frauen
   - Absolute Anzahl positiver und falsch-positiver Screeningbefunde (Mammographien) pro 100 Frauen
   - Lebenszeitrisiko (in %) für eine Überdiagnose (DCIS und Brustkrebs)
3. Abwägungsmaße
   - Number-Needed-To-Screen (NNS) um einen Todesfall zu vermeiden
   - Inkrementelles Schaden-Nutzen-Verhältnis (ISNV)

Die lebensqualitätsadjustierte Restlebenserwartung (in QALY) wurde berechnet, indem die Zeit, die Frauen in einem bestimmten Gesundheitszustand verbracht haben, mit den jeweiligen Nutzwerten für die gesundheitsbezogene Lebensqualität in diesem Zustand gewichtet wurden. Ein

Gewicht von 1 bedeutet vollständige Lebensqualität und ein Gewicht von 0 ist gleichbedeutend mit Tod. Die kumulierten gewichteten Lebensjahre über den lebenslangen Zeithorizont spiegeln die lebensqualitätsadjustierte Restlebenserwartung wider.

Als Überdiagnose wird die Diagnose einer Erkrankung bezeichnet, die ohne Untersuchung nie bemerkt worden wäre bzw. Symptome verursacht hätte. Eine Überdiagnose verbessert die Gesundheit nicht, bewirkt aber eventuelle Anspannung, unnötige Sorge und Lebensqualitätseinbuße. In dieser entscheidungsanalytischen Studie wurde eine Diagnose von DCIS oder invasivem Krebs durch die Mammographie als Überdiagnose eingestuft, wenn sie in der Strategie „Kein Screening“ zu keinem Zeitpunkt während des Lebens einer Frau diagnostiziert wurde.

Als Primärendpunkt einer Schaden-Nutzen-Evaluation wird als „Abwägungsmaß“ bzw. Effizienzmaß („efficiency ratio“) üblicherweise das so genannte inkrementelle Schaden-Nutzen-Verhältnis (ISNV) herangezogen. Das ISNV drückt aus, welcher zusätzlicher Schaden im Durchschnitt in Kauf genommen werden muss, um eine zusätzliche Nutzeneinheit zu erzielen, beispielsweise zusätzliche (falsch-)positive Screeningbefunde pro verhinderten Krebsfall oder pro gewonnenes Lebensjahr. Zur Berechnung wird die Differenz des gesundheitlichen Schadens (inkrementeller Schaden) durch die Differenz des gewählten Gesundheitsnutzens (inkrementeller Gesundheitsnutzen) zweier Handlungsalternativen dividiert. Das ISNV gibt somit das Verhältnis von zusätzlichem Schaden für eine zusätzliche gesundheitliche Nutzeneinheit (hier quantifiziert als gewonnenes Lebensjahr) wieder.

Von einer so genannten ‘dominanten’ Behandlungsoption ist die Rede, wenn diese im Vergleich zur Alternativoption sowohl einen geringeren erwarteten Schaden als auch einen höheren erwarteten Gesundheitsnutzen aufweist. Im Gegensatz dazu wird eine Handlungsalternative als ‘dominiert’ bezeichnet, wenn diese verglichen mit der Alternativoption sowohl schädlicher als auch weniger Nutzen bringt ist. Eine erweiterte Dominanz liegt dann vor, wenn die (Linear-) Kombination zweier Strategien (d. h. ein Teil der Population erhält eine Strategie und der Rest eine andere) eine einzelne andere Strategie dominiert, d. h. das ISNV einer bestimmten Behandlungsoption höher ist als das ISNV der gewichteten Kombination der beiden anderen Strategien.

Um die Unterschiede von inkrementellen Schaden-Nutzen-Verhältnissen der evaluierten Screeningstrategien grafisch in zwei Dimensionen zu visualisieren, wurde eine Effizienzkurve (bzw. Effizienzgrenze) auf einer Schaden-Nutzen-Ebene herangezogen. In der Schaden-Nutzen-Ebene können mehr als zwei Handlungsalternativen dargestellt werden. Die Linie, die aufeinanderfolgende nicht dominierte Alternativen verbindet, wird Schaden-Nutzen-Grenze oder Effizienzgrenze für die Schaden-Nutzen-Verhältnisse genannt. Es ist die Linie, auf der nach

Elimination aller dominieren Strategien alle verbleibenden in Frage kommenden Strategien liegen. Aus dem Steigungsdreieck eines Segments der Grenzlinie kann das ISNV des Strategievergleichs zwischen zwei Alternativen errechnet werden. Handlungsalternativen, die nicht auf der Schaden-Nutzen-Grenze liegen, haben ein schlechteres Verhältnis von Nutzen zu Schaden und stellen somit dominierte (nicht zu präferierende) Handlungsalternativen dar. Welche auf der Effizienzgrenze liegende Strategie als optimal zu bewerten ist, hängt von dem ISNV ab, welches man bereit ist zu akzeptieren. Im vorliegenden Bericht wurde kein Schwellenwert für ein ISNV verwendet, sondern es werden die resultierenden ISNV berichtet. Die Entscheidung für oder gegen eine bestimmte Strategie liegt damit beim Entscheidungsträger (bzw. den zu screenenden Frauen).

4.3.2. Sensitivitätsanalysen

In der vorliegenden Entscheidungsanalyse wurden zur Darstellung der Unsicherheit relevante Modellparameter in deterministischen Einweg- und Mehrweg-Sensitivitätsanalysen variiert, um die Robustheit der Ergebnisse aus der Basisfallanalyse systematisch zu untersuchen und sensitive Parameter zu identifizieren. Hierzu wurden insbesondere die Parameter der Testgüte (Sensitivität und Spezifität) der Mammographie in Mehrweg-Sensitivitätsanalysen und die Lebensqualitätsdaten in Einweg- und Mehrweg-Sensitivitätsanalysen variiert. Für die Nutzwerte und Nutzwertdekremente (Lebensqualität) lagen keine 95%-Konfidenzintervalle vor. In mehreren Mehrweg-Sensitivitätsanalysen wurden die Nutzwerte und Nutzwertdekremente zunächst schrittweise kombiniert eingesetzt, d. h. zunächst wurden ausschließlich die Nutzwerte in Abhängigkeit vom Alter berücksichtigt und in weiteren Analysen schrittweise die Nutzwertdekremente für den Gesundheitszustand DCIS oder Brustkrebs, für die Teilnahme am Screening sowie für einen positiven Screeningbefund einbezogen, um den Anteil der einzelnen Lebensqualitätsparameter am Ergebnis gemessen in QALY im Vergleich zur Lebenserwartung (ohne Adjustierung für die gesundheitsbezogene Lebensqualität) zu ermitteln. In weiteren Mehrweg-Sensitivitätsanalysen zur Lebensqualität wurden alle Nutzwerte und Nutzwertdekremente gleichzeitig im Bereich relative ± 5% des Ausgangswertes der Basisfallanalyse variiert. In dieser Mehrweg-Sensitivitätsanalyse wurden gleichzeitig die Nutzwerte für die Lebensqualität in Abhängigkeit vom Alter (altersspezifische Lebensqualität) und die relativen Nutzwertdekremente für die Gesundheitszustände entdeckter DCIS- und Krebsfälle erhöht (bzw. reduziert), während die Nutzwertdekremente für die Teilnahme am Screening und für einen positiven Screeningbefund reduziert (bzw. erhöht) wurden. Für die Parameter der Sensitivität und Spezifität der Mammographie für die Entdeckung von DCIS und Brustkrebs in Abhängigkeit vom Alter und BI-RADS-Kategorie lagen keine 95%-Konfidenzintervalle vor. In Mehrweg-Sensitivitätsanalysen wurde die falsch-negative Rate (1-Sensitivität) und in separaten Analysen in weiteren Sensitivitätsanalysen die falsch-positive Rate (1-Spezifität) um relative ± 5% und ± 20% des Ausgangswerts aus der Basisfallanalyse variiert, um die Auswirkung auf den Nutzen

gemessen an verhinderten Brustkrebstodesfällen sowie auf selektierte Schadensparameter (positive Screeningbefunde) zu untersuchen. Um einen Schwellenwerteffekt (verschiedene Punkte auf der ROC Kurve der Mammographie) zu simulieren, wurde in zusätzlichen Mehrweg-Sensitivitätsanalysen die Auswirkung einer gleichzeitigen Variation von Sensitivität und Spezifität der Mammographie jeweils in entgegengesetzte Richtungen um jeweils absolute ± 5% auf den Nutzen, gemessen an der Restlebenserwartung (in Lebensjahre pro 100 Frauen) und lebensqualitätsadjustierte Restlebenserwartung (in QALY pro 100 Frauen), untersucht. Im Szenario 1 wurden die Sensitivität und Spezifität gleichzeitig um jeweils absolute 5% erhöht bzw. reduziert, während im Szenario 2 die Sensitivität und Spezifität gleichzeitig um jeweils absolute 5% reduziert bzw. erhöht wurden.

4.3.3. Software

Für die Programmierung und Kalibrierung des entscheidungsanalytischen Modells sowie für die Basisfall- und Sensitivitätsanalysen wurde die Software TreeAge Pro Healthcare (Version 2021, R2.0) (34) verwendet. Für die Berechnung und Transformation von Modellparametern wurde die Software Excel (35) verwendet.

5. Ergebnisse

5.1. Kalibrierung

In Abbildung 2 sind die Ergebnisse der Modellkalibrierung für die Inzidenz von symptomatisch entdeckten DCIS in der Strategie „Kein Screening“ im Vergleich zur beobachteten altersspezifischen Inzidenz vor Einführung des Brustkrebsfrüherkennungsprogramms (Jahr 2003) in Deutschland dargestellt. Es ergab sich eine gute Übereinstimmung zwischen den modellierten und den tatsächlich beobachteten Daten (Tabelle 7).

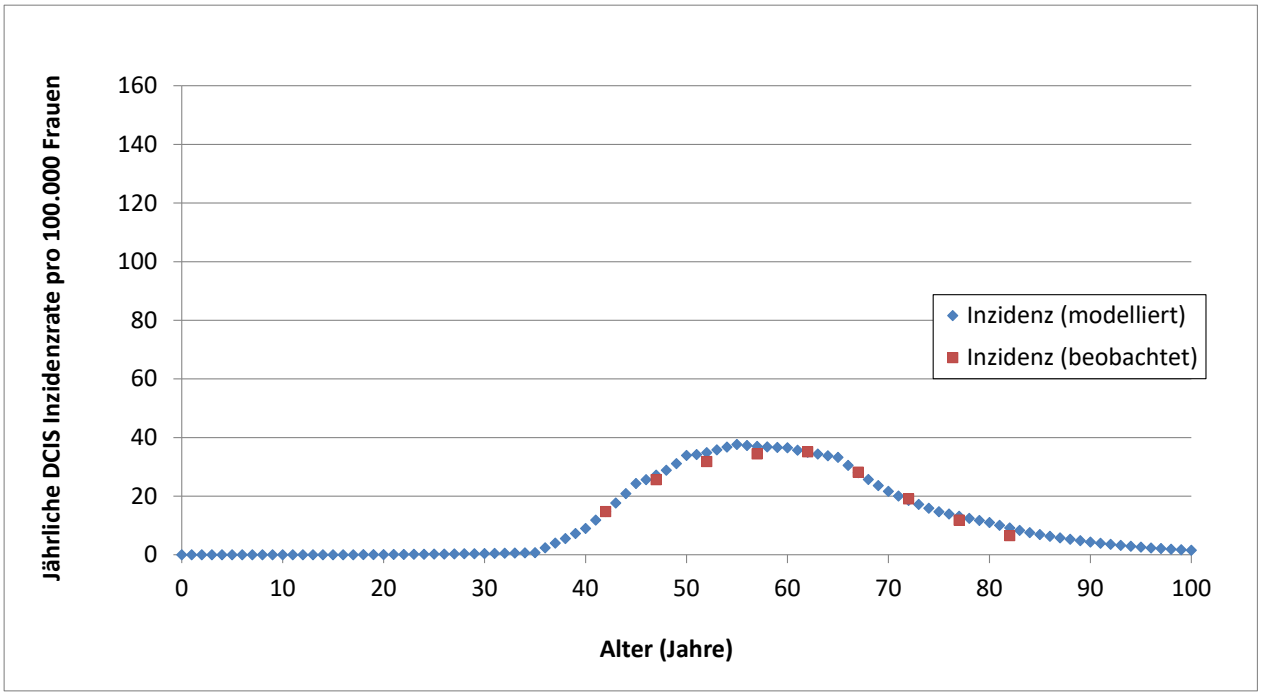


Abbildung 2 Kalibrierungsergebnis: Altersspezifische jährliche Inzidenzrate von symptomatisch entdecktem DCIS pro 100.000 Frauen.

Abgebildet wird die vom entscheidungsanalytischen Modell berechneten jährlichen DCIS-Neuerkrankungen in der Strategie „Kein Screening“ im Vergleich zu beobachteten jährlichen DCIS-Neuerkrankungsfällen vor der Einführung des Brustkrebsfrüherkennungsprogramms in Deutschland aus den Jahren 1999-2004 (19) jeweils pro 100.000 Frauen in Abhängigkeit vom Alter.
DCIS - Ductales Carcinoma in situ

In Abbildung 3 sind die Ergebnisse der Modellkalibrierung für die Inzidenz von symptomatisch entdecktem Brustkrebs in der Strategie „Kein Screening“ im Vergleich zur altersspezifischen beobachteten Inzidenz vor Einführung des Brustkrebsfrüherkennungsprogramms (Durchschnittsdaten aus den Jahren 1999-2004) in Deutschland dargestellt. Es ergab sich eine gute Übereinstimmung zwischen den modellierten und den tatsächlich beobachteten Daten (Tabelle 7).

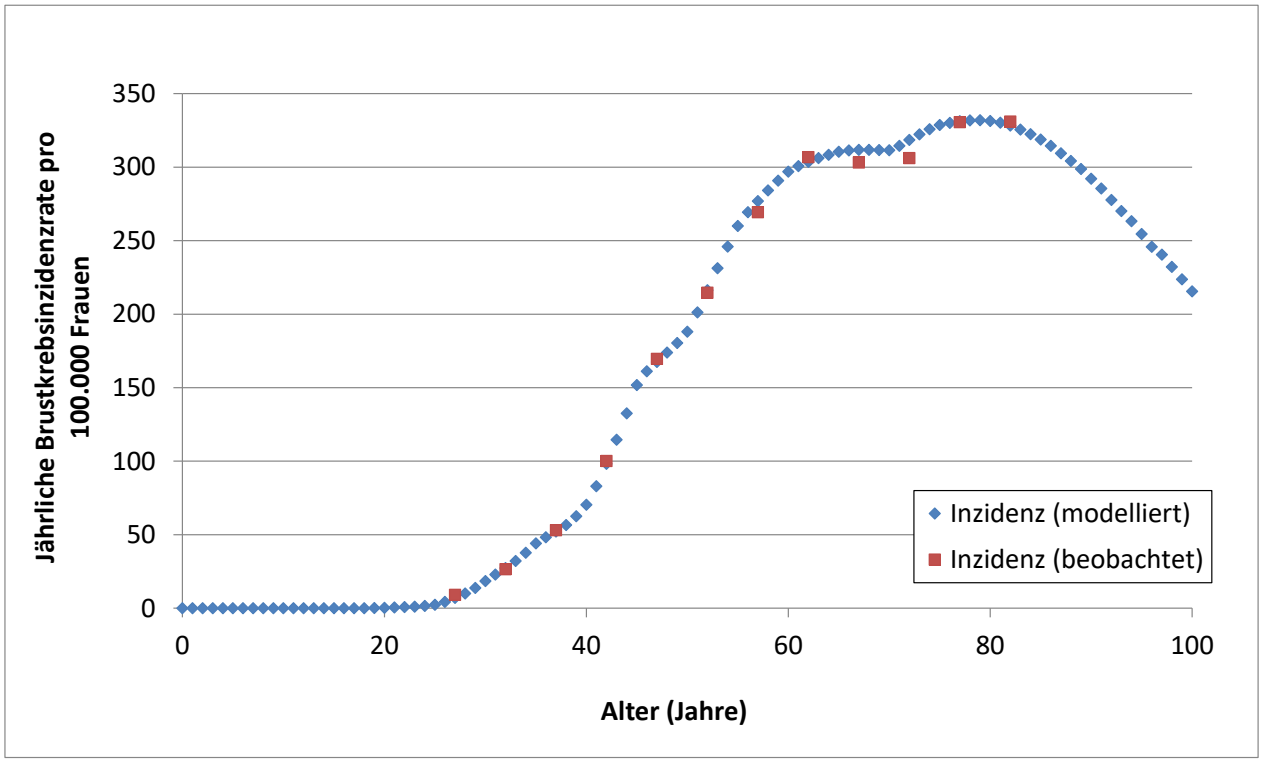


Abbildung 3 Kalibrierungsergebnis: Altersspezifische jährliche Inzidenzrate von symptomatisch entdecktem Brustkrebs pro 100.000 Frauen.

Abgebildet werden die vom entscheidungsanalytischen Modell berechneten Neuerkrankungen in der Strategie „Kein Screening“ im Vergleich zu beobachteten Neuerkrankungsfällen pro Jahr und 100.000 Frauen vor der Einführung des Brustkrebsfrüherkennungsprogramms in Deutschland aus den Jahren 1999-2004 (19) jeweils pro 100.000 Frauen in Abhängigkeit vom Alter.

In Abbildung 4 sind die Ergebnisse der Modellkalibrierung für die Stadienverteilung des initial diagnostizierten Brustkrebs in der Strategie „Kein Screening“ im Vergleich zur beobachteten Stadienverteilung vor Einführung des Brustkrebsfrüherkennungsprogramms (Daten aus den Jahren 2002-2004) in Deutschland dargestellt. Es ergab sich eine gute Übereinstimmung zwischen den modellierten und den tatsächlich beobachteten Daten (Tabelle 7).

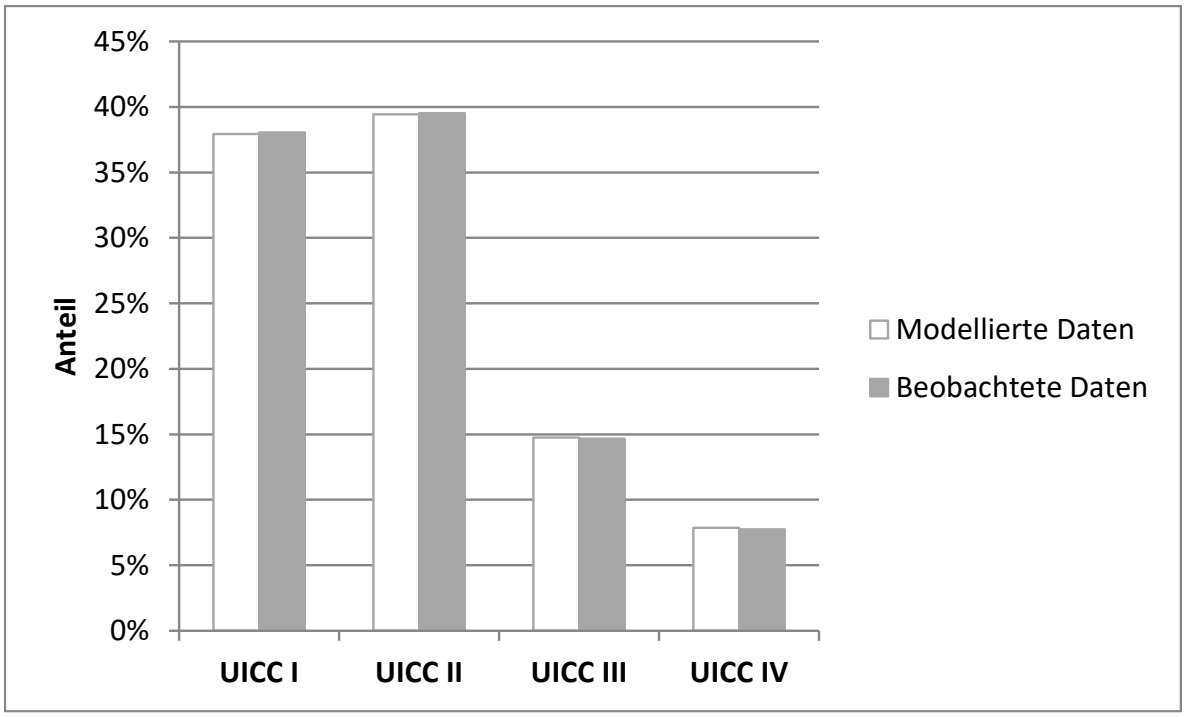


Abbildung 4 Kalibrierungsergebnis: UICC-Stadienverteilung entdeckter Brustkrebsfälle.

Abgebildet wird die vom entscheidungsanalytischen Modell berechnete Anteile von Frauen mit diagnostiziertem Brustkrebs im Stadium UICC I bis IV in der Strategie „Kein Screening“ im Vergleich zu beobachteten Daten vor der Einführung des Darmkrebsfrüherkennungsprogramms in Deutschland (31).
UICC - Union for International Cancer Control classification

5.2. Validierung

In Tabelle 9 werden die vom Modell errechneten Werte für das Lebenszeitrisiko an invasivem Brustkrebs zu erkranken und zu versterben sowie die Stadienverteilung der entdeckten Karzinome in der Referenzstrategie Screening unter der Annahme von 50% Teilnahmerate in einer externen Validierung den beobachteten Daten in Deutschland nach Einführung der Brustkrebsfrüherkennung (Daten aus dem Jahr 2016) gegenübergestellt.

In Abbildung 5 sind die Ergebnisse der externen Validierung der UICC-Stadienverteilung entdeckter Brustkrebsfälle dargestellt. Verglichen werden die vom entscheidungsanalytischen Modell berechnete Anteile von Frauen mit diagnostiziertem Brustkrebs im Stadium UICC I bis IV in der Referenzstrategie „Mammographie, Alter 50-69 Jahre, Intervall: 2 Jahre“ mit beobachteten Daten in Deutschland nach Einführung der Brustkrebsfrüherkennung (Daten aus dem Jahr 2016; RKI 2016 (33)).

Die Teilnahmerate von 50% basiert auf der Beobachtung, dass im Jahr 2018 insgesamt 2.859.139 Frauen von 5.708.761 eingeladenen am Mammographie-Screening-Programm teilnahmen. Das sind rund 50 % der eingeladenen Frauen (36).

Da jedoch in der Realität nicht alle der teilgenommenen Frauen im Alter von 50 Jahren das Mammographie-Screening begonnen haben und regelmäßig im 2-Jahresintervall durchgeführt haben, ist von einer altersabhängigen Teilnahme auszugehen. In weiteren Validierungsanalysen wurden unterschiedliche Annahmen zu durchschnittlichen Teilnahmeraten in Abhängigkeit vom Alter bei Beginn des Screenings durchgeführt. Die Ergebnisse sind im Appendix dargestellt.

Tabelle 9 Ergebnisse der externen Validierung unter Annahme einer Teilnahme am Screening von 50%.

	Modellergebnisse (Teilnahmerate 50%)	Beobachtete Daten	Referenz
Stadienverteilung[1]			RKI 2016 (33)
UICC I	0,49	0,41	
UICC II	0,35	0,39	
UICC III	0,11	0,13	
UICC IV	0,05	0,07	
UICC II+	0,51	0,59	
Stadienverteilung in Frauen Alter 50-69J			RKI 2016 (33)
UICC I	0,60	0,51	
UICC II	0,29	0,34	
UICC III	0,08	0,09	
UICC IV	0,03	0,05	
UICC II+	0,40	0,48	
Stadienverteilung in Frauen Alter 50-69J			Jahresbericht 2018 (36)
DCIS+UICC1	0,824	0,786	
UICC II-IV	0,176	0,210	
Lebenszeitrisiko[1]			RKI 2016 (33)
Invasiver Brustkrebs	0,108	0,122	
Tod durch invasiven Brustkrebs	0,035	0,036	
Inzidenz[1]			RKI 2016 (33)
Invasiver Brustkrebs	0,001294	0,001122	

DCIS - ductales Carcinoma in situ; UICC - UICC - Union for International Cancer Control classification; RKI - Robert Koch Institut

[1] Die beobachteten Werte sind auch in Tabelle 8 im Kapitel Methoden dargestellt.

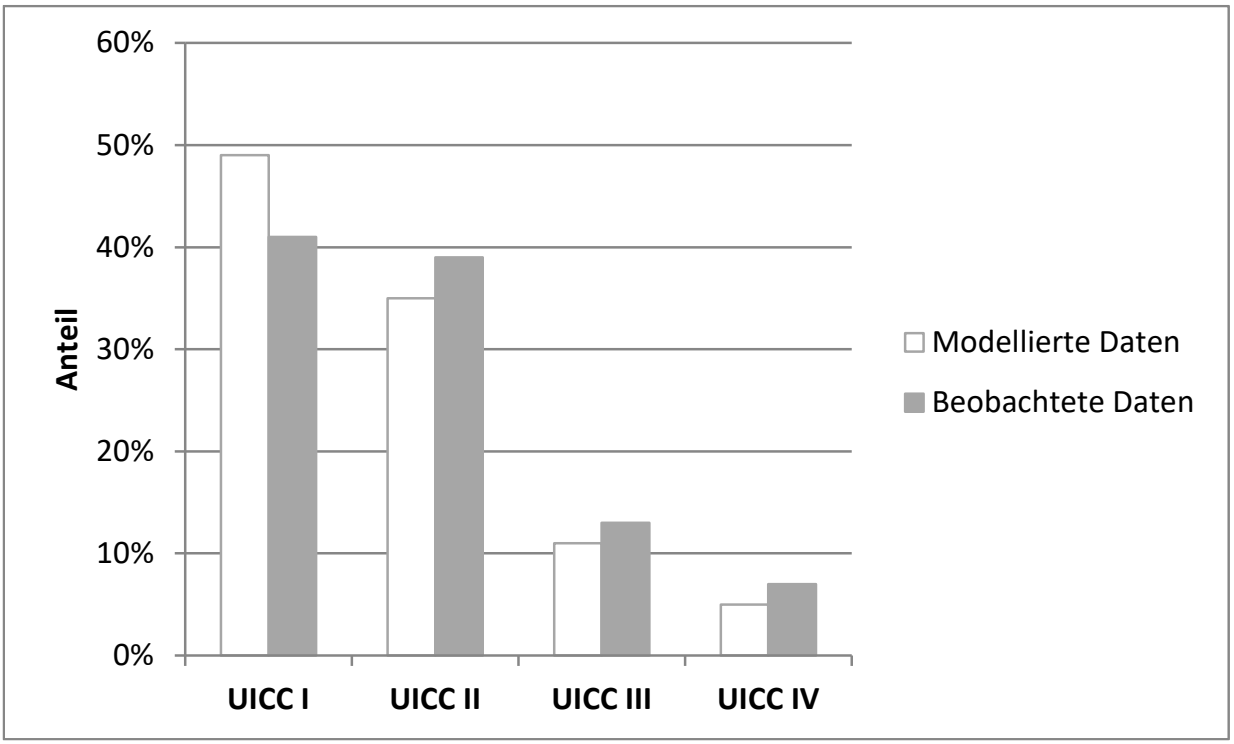


Abbildung 5 Ergebnisse der externen Validierung: UICC-Stadienverteilung entdeckter Brustkrebsfälle.

Abgebildet werden die vom entscheidungsanalytischen Modell berechnete Anteile von Frauen mit diagnostiziertem Brustkrebs im Stadium UICC I bis IV in der Referenzstrategie „Mammographie, Alter 50-69 Jahre, Intervall: 2 Jahre“ unter der Annahme einer durchschnittlichen Teilnahmerate von 50% im Vergleich zu beobachteten Daten in Deutschland nach Einführung der Brustkrebsfrüherkennung (Daten aus dem Jahr 2016; RKI 2016 (33))

5.3. Basisfallanalysen

Die Basisfallanalysen wurden unter der Verwendung einer regelmäßigen und vollständigen Screeningteilnahme und Compliance bei Abklärung von positiven Befunden und nachfolgender Therapie durchgeführt.

5.3.1. Restlebenserwartung und Lebensqualitätsadjustierte Restlebenserwartung

In Tabelle 10 werden für jede der untersuchten Screeningstrategien die durchschnittliche Restlebenserwartung (in Jahre) sowie die lebensqualitätsadjustierte Restlebenserwartung (in QALY) für eine Kohorte von 45-jährigen Frauen dargestellt, die regelmäßig am Mammographie-Screening teilnehmen (100% Teilnahmerate). Zusätzlich wird die Differenz (Inkrement) zur aktuell etablierten Screeningstrategie „Mammographie alle 2 Jahre im Alter 50 bis 69 Jahre“ angegeben. Die Strategien sind nach aufsteigender Restlebenserwartung sortiert.

Ohne Mammographie-Screening beträgt die Restlebenserwartung einer Kohorte von hundert 45-jährigen Frauen im Durchschnitt 3848 Jahre und die lebensqualitätsadjustierte Restlebenserwartung 3310 QALY. Ein Mammographie-Screening erhöht diese durchschnittliche Restlebenserwartung in Abhängigkeit vom Screeningintervall und den Altersgrenzen um durchschnittlich 22 bis 32 Lebensjahre pro 100 Frauen. Die lebensqualitätsadjustierte Restlebenserwartung wird um durchschnittlich 12,5 bis 16 QALYs pro 100 Frauen erhöht.

Abbildung zeigt die gewonnenen Lebensjahre und gewonnenen QALY für jede der verschiedenen Mammographie-Screeningstrategien im Vergleich zu keinem Screening. Der höchste Gewinn an Lebenszeit wird mit dem intensivsten Mammographie-Screening erzielt, d. h. Mammographie im Alter von 45 bis 79 Jahren und dabei jährlich für Frauen unter 50 und zweijährlich für Frauen ab 50 Jahren. Bezieht man die Lebensqualität mit ein, erzielt ein zweijährliches Mammographie-Screening im Alter von 45 bis 74 Jahren die höchste für die Lebensqualität adjustierte Lebenszeit.

Aus Tabelle 10 ist ersichtlich, dass mit dem Absenken des Alters für den Beginn des Mammographie-Screenings ein größerer inkrementeller Nutzen hinsichtlich der gewonnenen Lebenszeit und der für die Lebensqualität adjustierte Lebenszeit erzielt werden kann als mit einem Heraufsetzen des Alters für das Ende des Mammographie-Screenings. So ist einem gegenüber der Referenzstrategie um fünf Jahre früheren Screeningstart (mit 45 Jahren statt 50 Jahren, alle zwei Jahre) ein Zusatznutzen von 3,1 gewonnenen QALY pro 100 Frauen zuzuschreiben. Einer gegenüber der Referenzstrategie um fünf Jahre verlängerten Screeningdauer (bis 74 Jahre statt bis 69 Jahre, alle zwei Jahre) ist ein Zusatznutzen von 0,6 gewonnenen QALY pro 100 Frauen zuzuschreiben. Ein zweijährliches Screening, das sowohl mit 45 Jahren beginnt als auch mit 74 Jahren endet, hat gegenüber der derzeitigen Referenzstrategie einen Gesamtzusatznutzen von 3,5 gewonnenen QALY pro 100 Frauen. Dies entspricht einer NNS von 29 Frauen pro gewonnenem QALY. Aus Tabelle 10 lässt sich ebenfalls ablesen, dass Strategien mit einem zweijährlichen Screening im Alter von 45 bis 49 Jahren eine bessere Nutzen-Schaden-Balance aufweisen als Strategien mit einem jährlichen Screening in dieser Altersperiode. In ähnlicher Weise lässt sich ablesen, dass Strategien mit einem dreijährlichen Screening im Alter ab 70 Jahren eine gleich gute oder bessere Nutzen-Schaden-Balance aufweisen als Strategien mit einem zweijährlichen Screening in dieser Altersperiode.

Tabelle 10 Basisfallanalyse: Restlebenserwartung und für die lebensqualitätsadjustierte Restlebenserwartung pro 100 Frauen.

Strategie	Restlebenserwartung pro 100 Frauen (nach dem 45. Lebensjahr)			Lebensqualitätsadjustierte Restlebenserwartung pro 100 Frauen (nach dem 45. Lebensjahr)		
	Lebensjahre	**Inkrementelle Lebensjahre** vs. Kein Screening	**Inkrementelle Lebensjahre** vs. Referenz	**QALY**	**Inkrementelle QALY** vs. Kein Screening	**Inkrementelle QALY** vs. Referenz
Kein Screening	3847,7	---	---	3310,0	---	---
Screening: Alter 50-69 J, 2 J (Ref)	3870,0	22,4	---	3322,5	12,5	---
Screening: Alter 50-69 J, 2 J; 70-74 J, 3 J	3871,2	23,5	1,1	3322,9	13,0	0,5
Screening: Alter 50-74 J, 2 J	3872,5	24,8	2,4	3323,1	13,1	0,6
Screening: Alter 50-69 J, 2 J; 70-79 J, 3 J	3872,6	25,0	2,6	3323,2	13,3	0,8
Screening: Alter 50-79 J, 2 J	3873,4	25,7	3,3	3322,9	13,0	0,4
Screening: Alter 45-69 J, 2 J	3876,0	28,4	6,0	3325,6	15,6	3,1
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J	3877,1	29,4	7,0	3325,0	15,0	2,5
Screening: Alter 45-74 J, 2 J	3877,7	30,0	7,6	**3326,0**	**16,0**	**3,5**
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-74 J, 3 J	3878,1	30,4	8,1	3325,4	15,4	2,9
Screening: Alter 45-49 J, 1 J; 50-74 J, 2 J	3878,7	31,0	8,6	3325,4	15,4	2,9
Screening: Alter 45-79 J, 2 J	3879,0	31,4	9,0	3325,8	15,8	3,3
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-79 J, 3 J	3879,4	31,7	9,3	3325,6	15,6	3,1
Screening: Alter 45-49 J, 1 J; 50-79 J, 2 J	**3880,0**	**32,4**	**10,0**	3325,2	15,2	2,7

J - Jahre, Ref - Referenzstrategie

Fett gedruckt: die Strategien mit höchstem Nutzenzuwachs hinsichtlich Restlebenserwartung und lebensqualitätsadjustierte Restlebenserwartung.

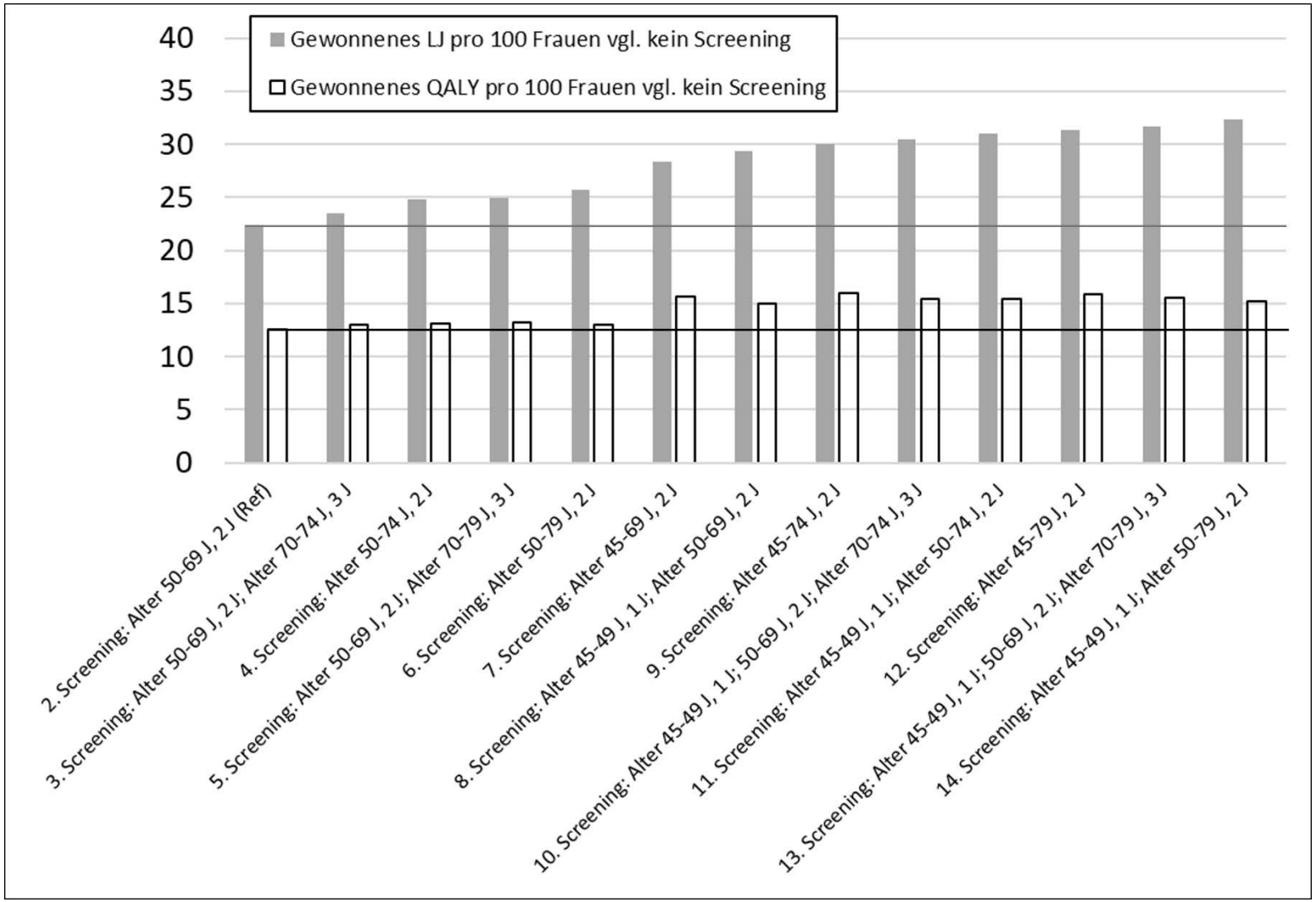


Abbildung 6 Ergebnisse der Basisfallanalyse: Gewonnene Lebensjahre und lebensqualitätsadjustierte Lebensjahre verschiedener Mammographie-Screeningstrategien im Vergleich zu keinem Screening.

LJ – Lebensjahr; QALY – quality-adjusted life year; vgl. – verglichen; J – Jahre; Ref - Referenz

5.3.2. Nutzen und Schaden

Tabelle 11 zeigt die Ergebnisse der Basisfallanalyse im Hinblick auf die durchschnittlichen gesundheitsbezogenen Ereignishäufigkeiten hinsichtlich verschiedener Nutzen- und Schadensendpunkte, die für die verschiedenen evaluierten Screeningstrategien zu erwarten sind.

Im Vergleich zu keinem Screening kann das Lebenszeitrisiko an einem entdeckten Brustkrebs zu versterben mit dem etablierten Mammographie-Screening (Alter: 50-69 Jahre, Intervall:2 Jahre) bei einer vollständigen und regelmäßigen Teilnahme aller screeningberechtigten Frauen ab 50 Jahren im Durchschnitt von 4,27% auf 3,05% (absolute Reduktion um 1,22%) gesenkt werden. Wird die Altersgrenze für den Beginn des Mammographie-Screenings auf 45 Jahre (Alter: 45-69 Jahre, Intervall:2 Jahre) herabgesetzt, kann das Lebenszeitrisiko an einem entdeckten Brustkrebs zu versterben auf 2,80% (absolute Reduktion um 1,47% gegenüber keinem Screening) gesenkt werden. Das bedeutet einen Zusatznutzen im Vergleich zur derzeitigen Referenz von absoluten 0,25%, also die Vermeidung eines zusätzlichen brustkrebsbedingten Todesfalles pro 400 gescreenten Frauen. Bleibt die Altersgrenze für den Beginn erhalten und wird die Altersgrenze für das Ende des etablierten Mammographie-Screening angehoben auf 74 Jahre (Alter: 50-74 Jahre, Intervall:2 Jahre) kann das Lebenszeitrisiko an einem entdeckten Brustkrebs zu versterben ebenfalls auf 2,80% (absolute Reduktion gegenüber keinem Screening um 1,47%) gesenkt werden. Das bedeutet auch für diese Strategie einen Zusatznutzen im Vergleich zur derzeitigen Referenz von absoluten 0,25%, also die Vermeidung eines zusätzlichen brustkrebsbedingten Todesfalles pro 400 gescreenten Frauen. Bleibt die Altersgrenze für den Beginn erhalten und wird die Altersgrenze für das Ende des etablierten Mammographie-Screening angehoben auf 79 Jahre (Alter: 50-79 Jahre, Intervall:2 Jahre) kann das Lebenszeitrisiko an einem entdeckten Brustkrebs zu versterben auf 2,68% (absolute Reduktion gegenüber keinem Screening um 1,59%) gesenkt werden. Das bedeutet auch für diese Strategie einen Zusatznutzen im Vergleich zur derzeitigen Referenz von absoluten 0,37%, also die Vermeidung eines zusätzlichen brustkrebsbedingten Todesfalles pro 270 gescreenten Frauen.

Den größten Nutzen hinsichtlich einer Reduktion im Lebenszeitrisiko, an einem entdeckten Brustkrebs zu versterben, erzielt ein jährliches Mammographie-Screening im Alter von 45 bis 49 Jahren gefolgt von einem zweijährlichen Mammographie-Screening im Alter von 50 bis 79 Jahren. Unter den untersuchten Screeningstrategien nimmt also der Nutzen gemessen an der Reduktion der Todesfälle durch Brustkrebs zu mit Intensität und Länge des Mammographie-Screenings (Tabelle 11).

Wird die Gesamtanzahl an Screeningbefunde (Mammographien) als potenzieller Schaden (Burden) herangezogen, so ergibt sich die Darstellung in Tabelle 11. Die Anzahl an Mammographien im etablierten Mammographie-Screening (Alter: 50-69 Jahre, Intervall:2 Jahre)

bei einer vollständigen und regelmäßigen Teilnahme aller screeningberechtigter Frauen ab 50 Jahren liegt im Durchschnitt bei 903 pro 100 Frauen. Wird die Altersgrenze für den Beginn des Mammographie-Screenings auf 45 Jahre (Alter: 45-69 Jahre, Intervall:2 Jahre) herabgesetzt, ist die Gesamtanzahl an Mammographien 1186 pro 100 Frauen. Bleibt die Altersgrenze für den Beginn erhalten und wird die Altersgrenze für das Ende des etablierten Mammographie-Screening angehoben auf 74 Jahre (Alter: 50-74 Jahre, Intervall:2 Jahre), ist die Gesamtanzahl an Mammographien 1135 pro 100 Frauen. Bleibt die Altersgrenze für den Beginn erhalten und wird die Altersgrenze für das Ende des etablierten Mammographie-Screening angehoben auf 79 Jahre (Alter: 50-79 Jahre, Intervall:2 Jahre), ist die Gesamtanzahl an Mammographien 1273 pro 100 Frauen. Unter den untersuchten Screeningstrategien nimmt also die Anzahl der Mammographien mit der Intensität und Länge des Mammographie-Screenings zu. Die höchste Anzahl an Mammographien erzielt ein jährliches Mammographie-Screening im Alter von 45 bis 49 Jahren gefolgt von einem zweijährlichen Mammographie-Screening im Alter von 50 bis 79 Jahren mit 1741 Mammographien pro 100 Frauen.

Ein ähnliches Bild stellt sich für andere Schadensendpunkte (Gesamtanzahl positiver Screeningbefunde, Gesamtanzahl falsch-positiver Screeningbefunde und Lebenszeitrisiko für eine Überdiagnose von DCIS und Brustkrebs) dar.

Tabelle 11 Ergebnisse der Basisfallanalyse: Ereignishäufigkeiten verschiedener Nutzen- und Schaden-Parameter

Strategie	Lebenszeit-risiko: diagn. BK (%)	Lebenszeit-risiko: diagn. DCIS (%)	Lebenszeit-risiko: Tod durch BK (%)	Screening-befunde pro 100 Frauen (n)	Positive Screening-befunde pro 100 Frauen (n)	Falsch-positive Screening-befunde pro 100 Frauen (n)	Lebenszeit-risiko: Über-diagnose (DCIS+BK) (%)	NNS um einen BK-bedingten Todesfall zu vermeiden (n)
Kein Screening	11,12	1,01	4,27	0	0	0	0	---
Screening: Alter 50-69 J, 2 J (Ref)	10,78	7,65	3,05	903	98,4	86,7	6,30	81
Screening: Alter 50-74 J, 2 J	10,83	7,71	2,80	1135	117,7	104,8	6,41	68
Screening: Alter 50-79 J, 2 J	10,89	8,67	2,68	1273	129,2	115,4	6,55	63
Screening: Alter 45-69 J, 2 J	10,70	7,76	2,80	1186	136,5	123,1	7,18	68
Screening: Alter 45-74 J, 2 J	10,74	8,70	2,64	1341	149,4	135,1	7,26	61
Screening: Alter 45-79 J, 2 J	10,83	8,80	2,45	1548	166,7	151,1	7,46	55
Screening: Alter 50-69 J, 2 J; 70-74 J, 3 J	10,79	7,67	2,94	982	105,0	92,9	6,34	75
Screening: Alter 50-69 J, 2 J; 70-79 J, 3 J	10,86	8,60	2,76	1126	117,3	104,0	6,49	66
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-74 J, 3 J	10,72	8,64	2,66	1457	167,6	153,4	7,29	62
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-79 J, 3 J	10,79	8,76	2,49	1597	179,5	164,2	7,46	56
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J	10,70	7,79	2,77	1379	161,0	147,4	7,24	66
Screening: Alter 45-49 J, 1 J; 50-74 J, 2 J	10,74	8,71	2,60	1534	173,9	159,4	7,32	60
Screening: Alter 45-49 J, 1 J; 50-79 J, 2 J	10,82	8,83	2,41	1741	191,2	175,4	7,52	54

Diagn. - diagnostiziert, J - Jahre, DCIS - ductales Carcinoma in situ; BK - Brustkrebs, NNS - Number needed to screen, Ref. - Referenzstrategie.

5.3.3. Schaden-Nutzen-Verhältnisse

Für die systematische Darstellung der Verhältnisse zwischen potenziellem Schaden und dem Nutzen von verschiedenen Mammographie Screening-Strategien wurden Effizienzgrenzendiagramme verwendet. In den Abbildungen (Abbildung 7-Abbildung 9) werden Nutzen und Schaden gegenübergestellt und die jeweiligen schrittweise inkrementelle Schaden-Nutzen-Verhältnisse (ISNV) berichtet.

In Abbildung 7 werden für die verschiedenen Mammographie-Screeningstrategien die Anzahl an zusätzlichen Mammographien (Tabelle 11) dem Nutzen hinsichtlich der gewonnenen Lebenserwartung (Tabelle 10) gegenübergestellt.

Auf der Effizienz-Grenzlinie liegen die Mammographie-Screeningstrategien, die im Verhältnis von Schaden zu Nutzen gemessen an der Anzahl an Mammographien pro gewonnenes Lebensjahr den Strategien unterhalb der Grenzlinie überlegen sind. Auf der Effizienz-Grenzlinie liegen neben dem etablierten Screening (Mammographie Alter 50-69 J, Screeningintervall: 2 Jahre) die folgenden Screeningstrategien: zweijährliche Mammographie im Alter 45-69 Jahre, zweijährliche Mammographie im Alter 45-74 Jahre, Mammographie im Alter 45-49 jährlich, im Alter 50-69 zweijährlich und im Alter 70-79 dreijährlich, Mammographie im Alter 45-49 jährlich, im Alter 50-79 zweijährlich. Die Strategie zweijährliche Mammographie im Alter 45-79 Jahre wurde knapp erweitert dominiert.

Im Vergleich zum etablierten Mammographie-Screening müssten für den aus dem Herabsetzen der Altersgrenze für den Beginn des Mammographie-Screening auf 45 Jahre (Screening im Alter 45 - 69 Jahren, zweijährlich) resultierenden Nutzen im Durchschnitt 47 zusätzliche Mammographien pro gewonnenes Lebensjahr akzeptiert werden. Im Vergleich zu diesem Screening müsste für ein gleichzeitiges Heraufsetzen der Altersgrenze für das Ende des Mammographie-Screening auf 74 Jahre (Screening im Alter 45 - 74 Jahren, zweijährlich) im Durchschnitt 96 zusätzliche Mammographien pro gewonnenes Lebensjahr in Kauf genommen werden. Durch ein kürzeres Screeningintervall im Alter von 45 bis 49 Jahren und/oder eine Verlängerung des Screenings bis zum Alter von 79 Jahren kann trotz deutlich zunehmender Anzahl an Mammographien nur noch geringfügig der Nutzen hinsichtlich der Lebenserwartung gesteigert werden. Die inkrementellen Schaden-Nutzen-Verhältnisse dieser Strategien sind daher deutlich höher.

Dieses Muster in der Reihenfolge wurde im Allgemeinen auch bei anderen Endpunkten für einen potenziellen Schaden gefunden (siehe Abbildung 8-Abbildung 9).

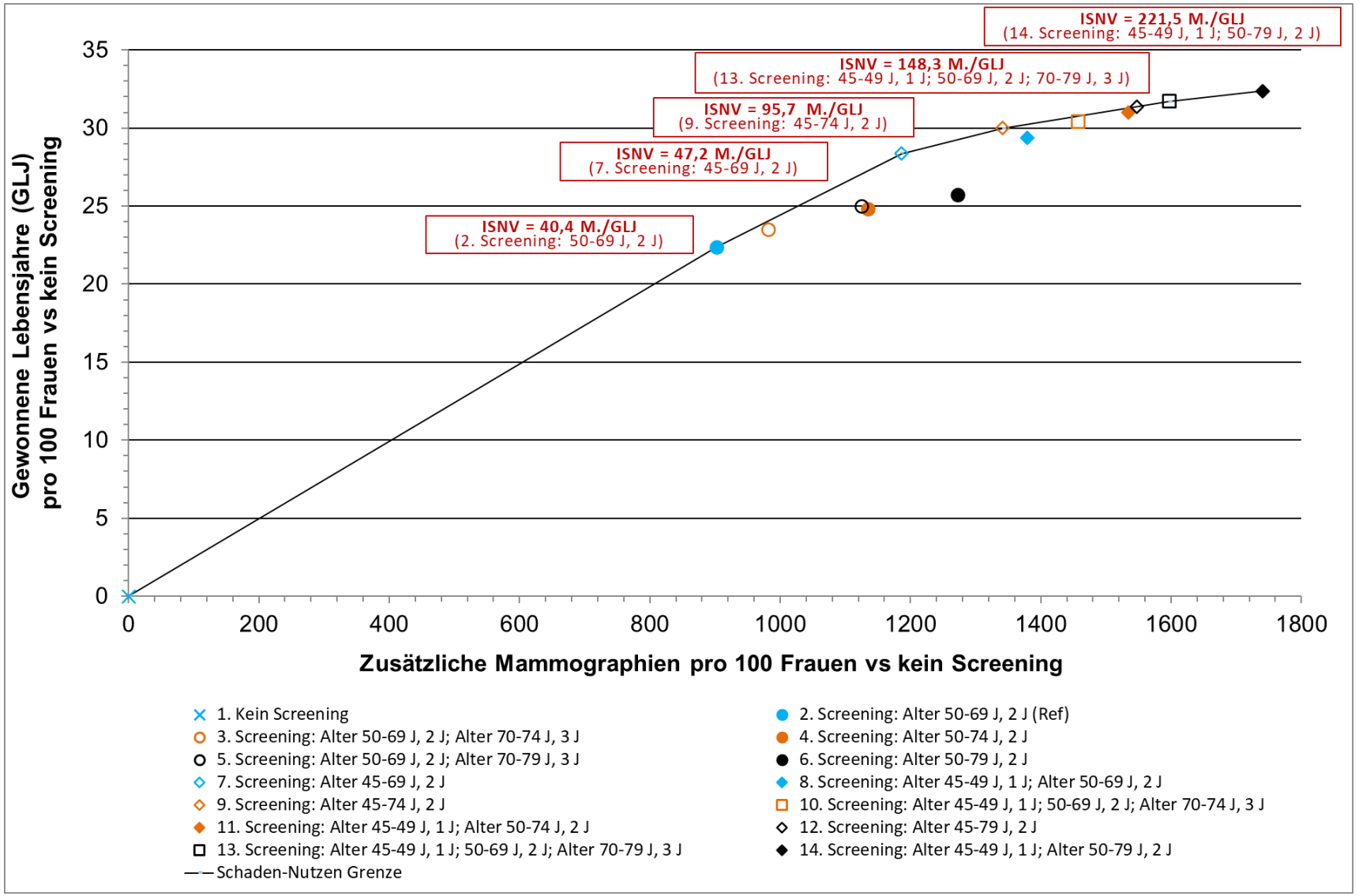


Abbildung 7 Ergebnisse der Basisfallanalyse: Gesamtanzahl an zusätzlichen Mammographien versus gewonnene Lebensjahre.

ISNV - inkrementelle Schaden-Nutzen-Verhältnis; M/GLJ – Mammographie pro gewonnenes Lebensjahr; J - Jahr

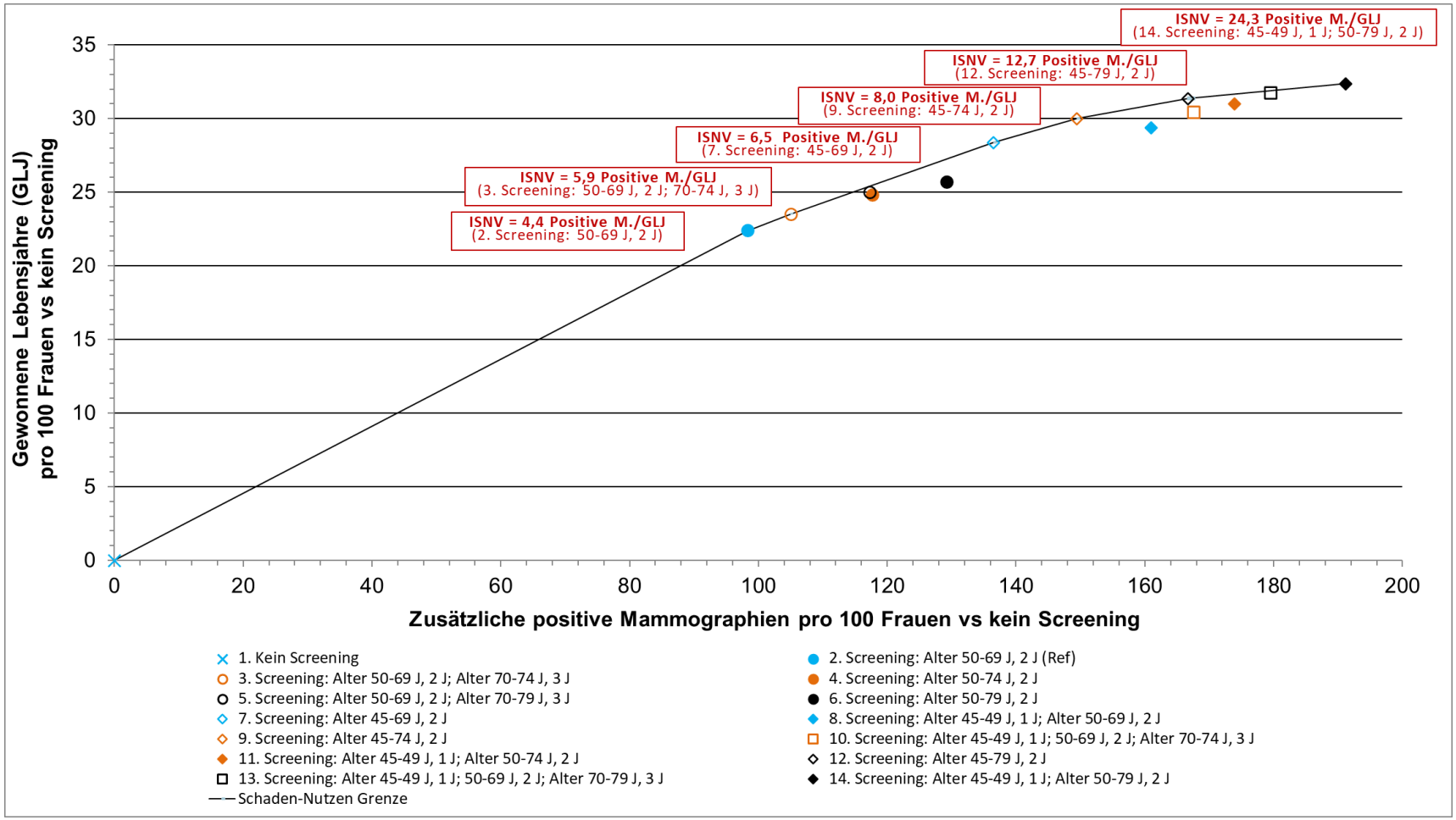


Abbildung 8 Ergebnisse der Basisfallanalyse: Gesamtanzahl an zusätzlichen positiven Mammographien versus gewonnene Lebensjahre.

ISNV - inkrementelle Schaden-Nutzen-Verhältnis; M/GLJ – Mammographie pro gewonnenes Lebensjahr; J - Jahr

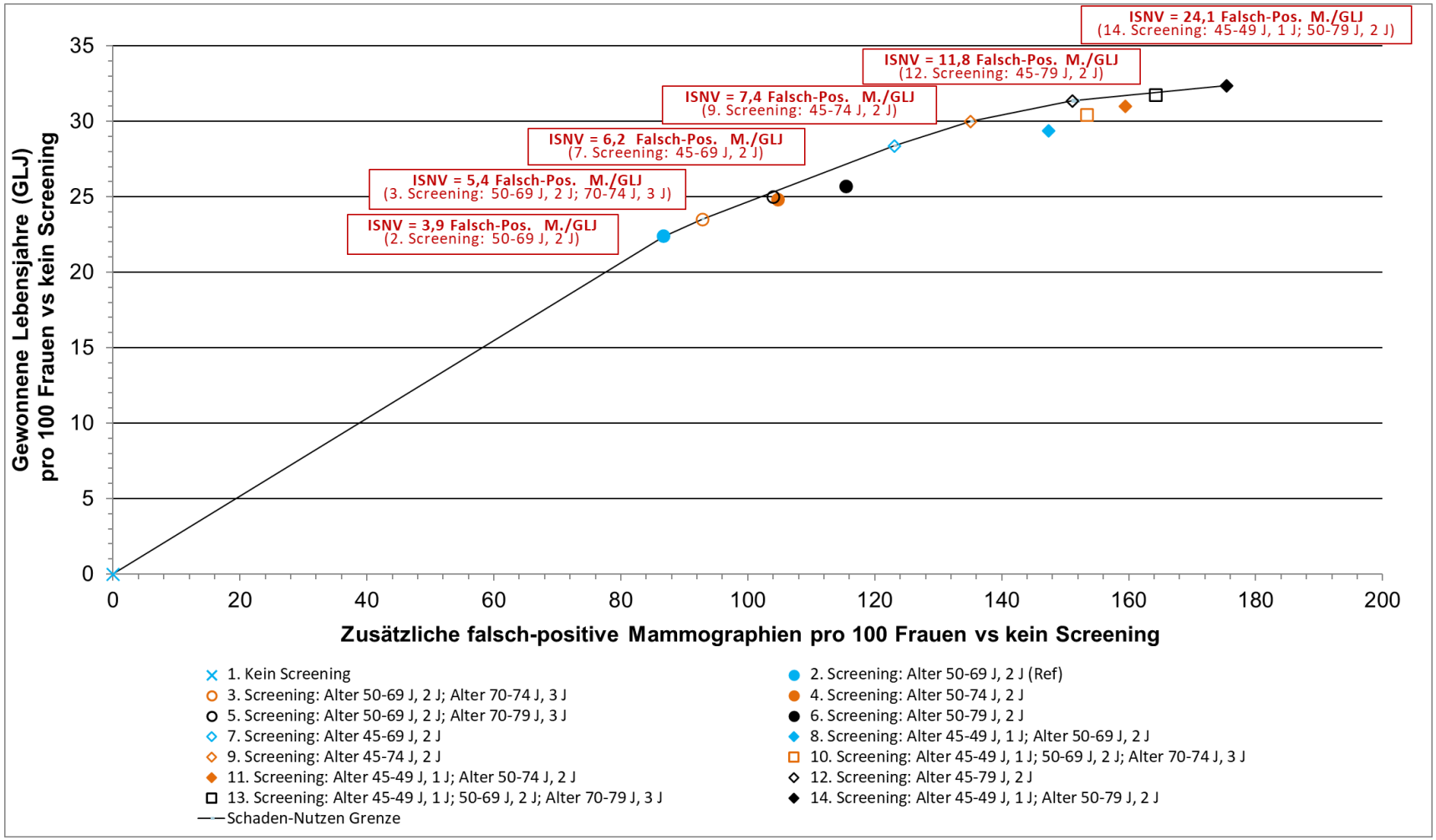


Abbildung 9 Ergebnisse der Basisfallanalyse: Gesamtanzahl an zusätzlichen falsch-positiven Screeningbefunden versus gewonnene Lebensjahre.

ISNV - inkrementelle Schaden-Nutzen-Verhältnis; M/GLJ – Mammographie pro gewonnenes Lebensjahr; J - Jahr

In Abbildung 10 werden für die Screeningstrategien die Gesamtanzahl an zusätzlichen Mammographien dem gesundheitsbezogenen Nutzen hinsichtlich der verhinderten Krebstodesfälle gegenübergestellt. Auf der Effizienzgrenze liegen die Strategien mit einer Mammographie a) zweijährlich im Alter 50-69 Jahren und dreijährlich im Alter 70-74 Jahre, b) zweijährlich im Alter 50-69 Jahren und dreijährlich im Alter 70-79 Jahre, c) zweijährlich im Alter 45-79 Jahre sowie d) jährlich im Alter 45-49 Jahre und zweijährlich im Alter 50-79 Jahre. Das etablierte Mammographie-Screening ist hier rechnerisch gerade eben erweitert dominiert, liegt aber praktisch (fast) auf der Effizienzgrenze.

Auch hier zeigt sich ein ähnliches Muster. Das Herabsetzen der Altersgrenze für den Beginn des Screenings mit 45 Jahren mit gleichzeitigem Heraufsetzen der Altersgrenze für das Ende des Mammographie-Screenings mit 79 Jahren unter Beibehalten des zweijährlichen Screeningintervalls erzielt einen deutlich höheren Nutzen bei akzeptablem Verhältnis zwischen Schaden und Nutzen. Ein intensiveres Screening durch kürzere Intervalle im Alter von 45 bis 49 Jahren erzielt nur geringfügig höheren Nutzen und das inkrementelle Schaden-Nutzen-Verhältnis steigt deutlich an. Dieses Muster in der Reihenfolge wurde im Allgemeinen auch bei anderen Endpunkten für einen potenziellen Schaden gefunden (siehe Abbildung 11- Abbildung 12).

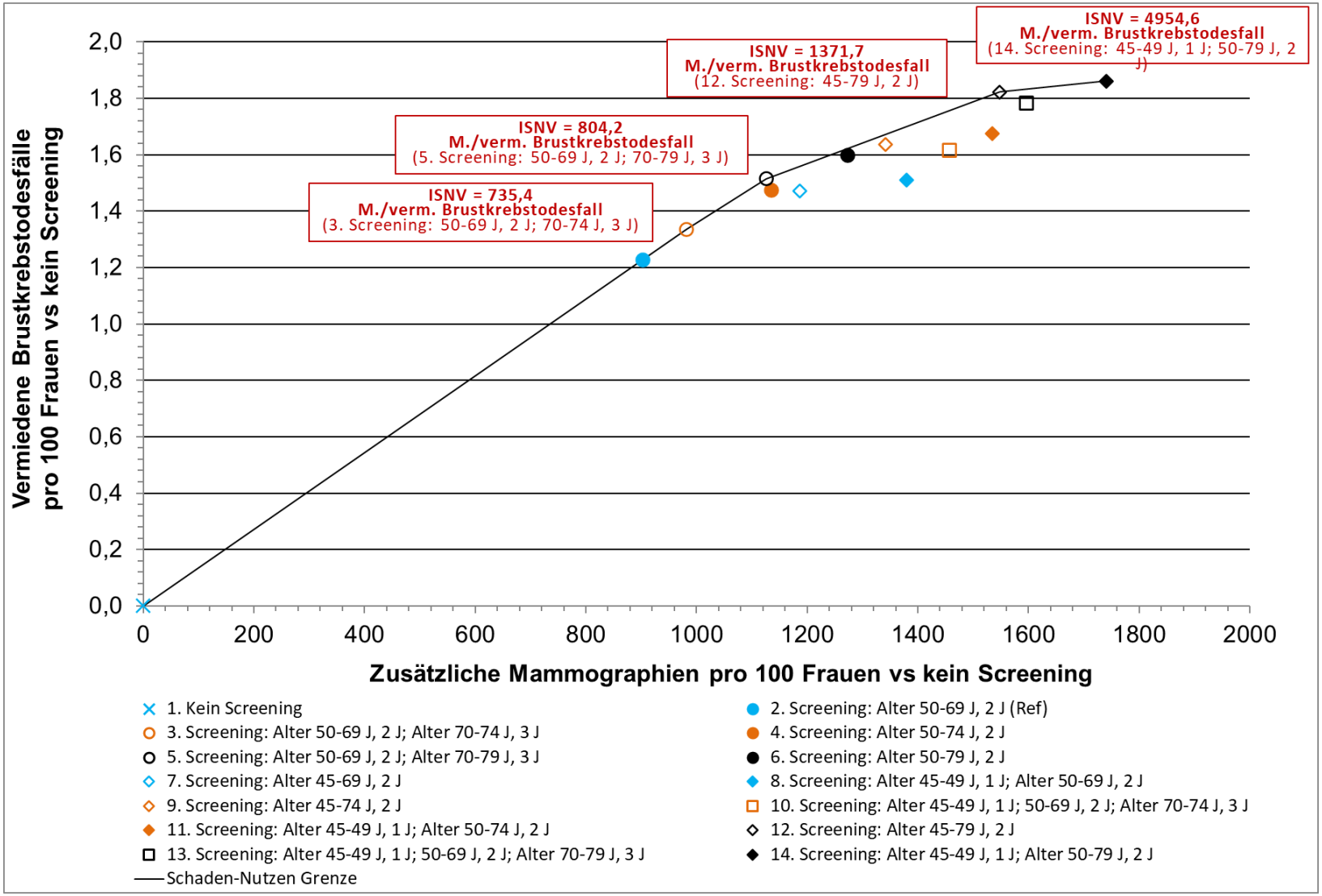


Abbildung 10 Ergebnisse der Basisfallanalyse: Gesamtanzahl an zusätzlichen Mammographien versus verhinderte Krebstodesfälle.

ISNV - inkrementelle Schaden-Nutzen-Verhältnis; M/verm. – Mammographie pro vermiedenem; J - Jahr

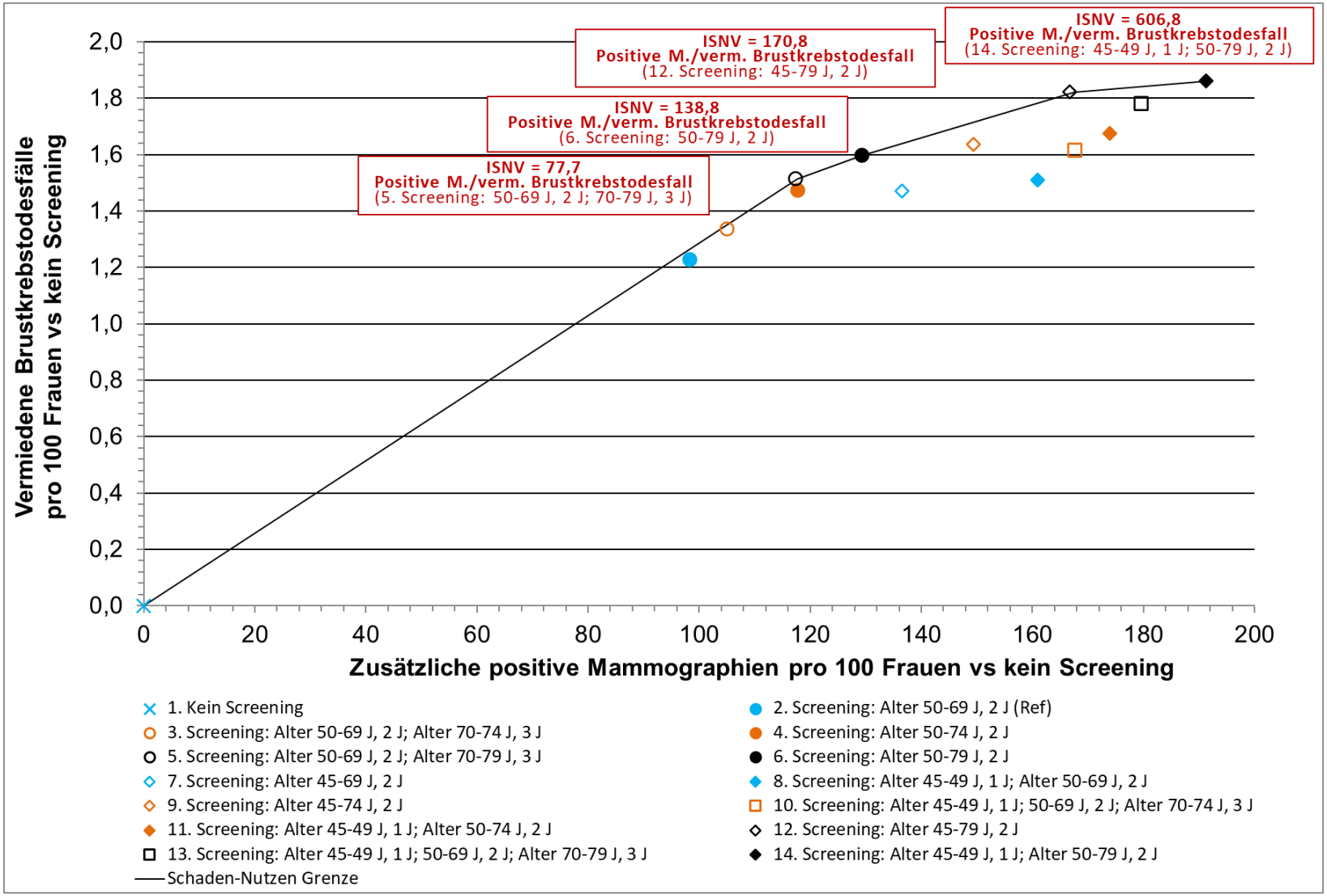


Abbildung 11 Ergebnisse der Basisfallanalyse: Gesamtanzahl an zusätzlichen positiven Mammographien versus verhinderte Krebstodesfälle.

ISNV - inkrementelle Schaden-Nutzen-Verhältnis; M/verm. – Mammographie pro vermiedenem; J - Jahr

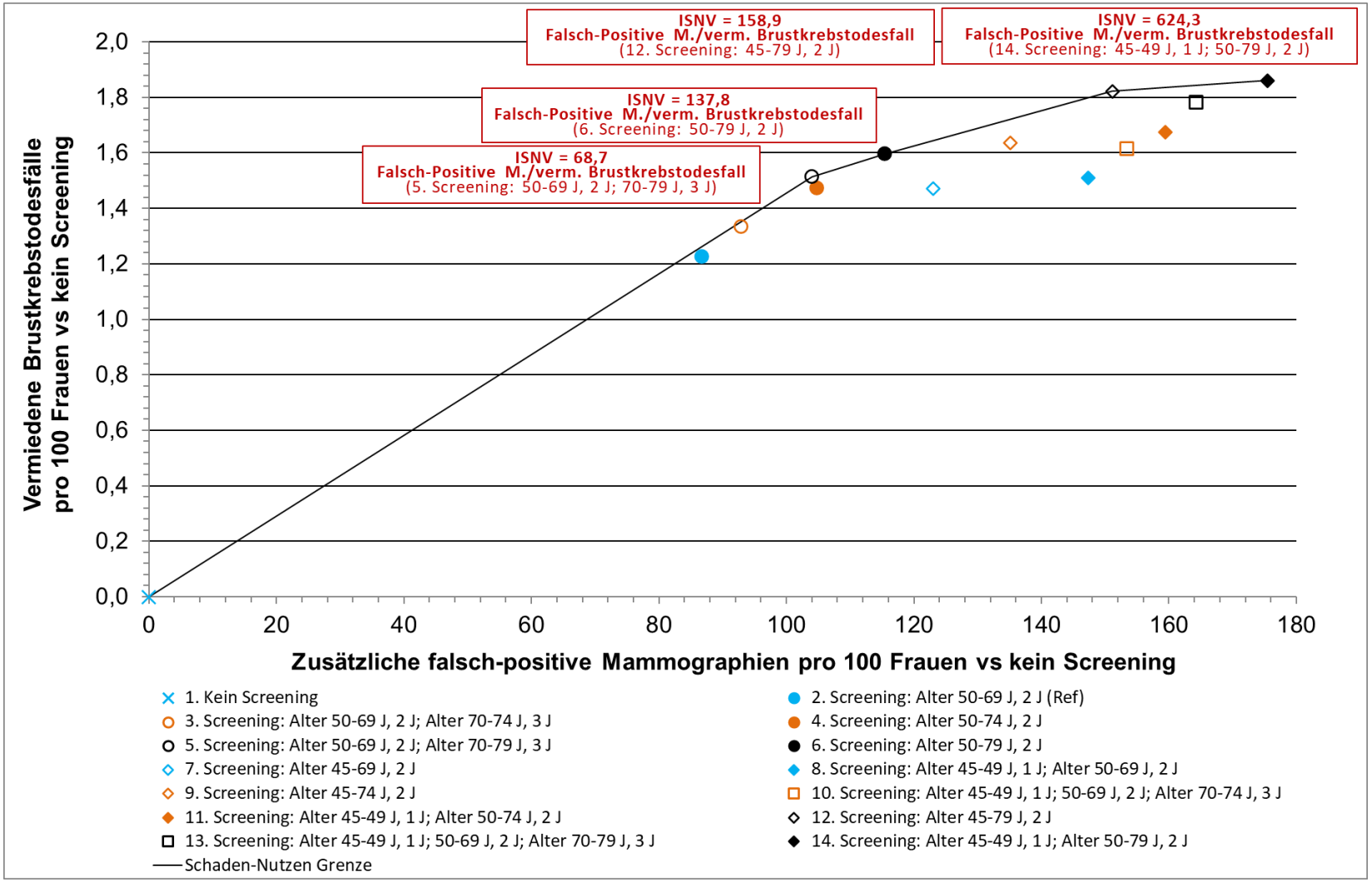


Abbildung 12 Ergebnisse der Basisfallanalyse: Gesamtanzahl an zusätzlichen falsch-positiven Mammographien versus verhinderte Krebstodesfälle.

ISNV - inkrementelle Schaden-Nutzen-Verhältnis; M/verm. – Mammographie pro vermiedenem; J - Jahr

5.4. Sensitivitätsanalysen

In der vorliegenden Entscheidungsanalyse wurden zur Darstellung der Unsicherheit relevante Modellparameter in deterministischen Einweg- und Mehrweg-Sensitivitätsanalysen variiert, um die Robustheit der Ergebnisse aus der Basisfallanalyse zu untersuchen und sensitive Parameter zu identifizieren. In den folgenden Unterkapiteln werden die wichtigsten Ergebnisse zur Auswirkung der Variation der Modellparameter zur Lebensqualität und zur Testgüte der Mammographie dargestellt.

5.4.1. Lebensqualität

Tabelle 12 zeigt die Ergebnisse der Sensitivitätsanalysen zur Auswirkung der Berücksichtigung einzelner und kombinierter Nutzwerte und Nutzwertdekremente für die Lebensqualität in Abhängigkeit vom Alter, Screening, positiven Screeningbefunden sowie der Behandlung von entdeckten DCIS und Brustkrebsfällen auf den Nutzen gemessen in QALY.

Die Berücksichtigung der Lebensqualität in Abhängigkeit vom Alter (aufgrund der altersbedingten Komorbidität und damit verbundenen Verschlechterung des allgemeinen Gesundheitszustands der Frauen) verringerte den zu erwartenden Nutzen gemessen in QALY im Vergleich zur Lebenserwartung deutlich, hatte aber keinen Einfluss auf die Rangfolge der Strategien im Vergleich zum Endpunkt Restlebenserwartung (ohne Berücksichtigung der Lebensqualität).

Die Berücksichtigung der Lebensqualitätsreduktion, die mit einer Teilnahme am Screening, einem positiven Screeningbefund oder einer Behandlung entdeckter DCIS- und Krebsfälle verbunden ist, verringerte den zu erwartenden Nutzen gemessen in QALY im Vergleich zur Lebenserwartung jeweils nur geringfügig, hatte aber Einfluss auf die Rangfolge der Strategien im Vergleich zum Endpunkt Restlebenserwartung (ohne Berücksichtigung der Lebensqualität). Bei Berücksichtigung ausschließlich der Lebensqualitätsreduktion, die mit einer Teilnahme am Screening einhergeht, stellt sich dasselbe Ranking der Strategien ein wie in der Basisfallanalyse.

Die Reduktion in der Lebensqualität mit zunehmendem Alter der Frau hatte den größten Einfluss auf den Endpunkt QALY, d. h. bewirkte die größte relative Differenz zwischen den beiden Nutzenmaßen Restlebenserwartung und lebensqualitätsadjustierte Lebenserwartung (durchschnittliche mittlere relative Reduktion um 14%). Die Auswirkung des Screenings selbst, eines positiven Befundes beim Screening oder auch die negativen Auswirkungen aufgrund der Behandlung eines entdeckten DCIS oder Brustkrebsfalles waren im Vergleich dazu sehr gering (durchschnittliche mittlere relative Reduktion um weniger als 1%). Die QALY-Verluste waren bei den Screeningstrategien mit einem zweijährlichen Screeningintervall, gegenüber denen mit einem jährlichen Screeningintervall, bei Frauen unter 50 Jahren deutlich geringer, da bei Letzteren die Zahl der falsch-positiven Ergebnisse im Verhältnis zum Nutzen höher war als beim

zweijährlichen Screening. Ähnlich verhielt es sich bei einem dreijährlichen gegenüber einem zweijährlichen Screeningintervall bei Frauen ab einem Alter von 70 Jahren. In allen Screeningstrategien wurde im Vergleich zu keinem Screening konsistent ein Gewinn an QALYs erzielt. Dieser Gewinn fiel jedoch deutlich geringer aus, wenn gleichzeitig die Nutzwerte für die Lebensqualität in Abhängigkeit vom Alter sehr niedrig, sowie die Nutzwertdekremente sehr hoch waren wie in Tabelle 13 dargestellt.

Tabelle 13 zeigt die Auswirkung der Variation der Lebensqualität insgesamt auf den Nutzen gemessen in QALY. In dieser Mehrweg-Sensitivitätsanalyse wurden gleichzeitig die Nutzwerte für die Lebensqualität in Abhängigkeit vom Alter (altersspezifische Lebensqualität) und die relativen Nutzwertdekremente für die Gesundheitszustände entdeckter DCIS- und Krebsfälle erhöht (bzw. reduziert), während die Nutzwertdekremente für die Teilnahme am Screening und für einen positiven Screeningbefund reduziert (bzw. erhöht) wurden.

Die Ergebnisse weiterer Sensitivitätsanalysen zur Lebensqualität sind im Appendix tabellarisch dargestellt (Appendix 9.2.)

Tabelle 12 Sensitivitätsanalyse: Auswirkung der Lebensqualität bei Alter, Screening, positivem Screeningbefund und Behandlung von entdeckten DCIS- und Brustkrebsfällen auf den Nutzen (pro 100 Frauen).

		Lebensqualität (Nutzwerte) adjustiert für									
		Alter		Screening		Screening, positive Befunde		Screening, positive Befunde, DCIS-/Brustkrebs		Screening, positive Befunde, DCIS-/Brustkrebs, Alter (Basisfallanalyse)	
Strategie	**Lebens-jahre**	**QALY**	**Rel. Diff. (%)**	**QALY**	**Rel. Diff. (%)**	**QALY**	**Rel. Diff. (%)**	**QALY**	**Rel. Diff. (%)**	**QALY**	**Rel. Diff. (%)**
Kein Screening	3847,7	3319,6	13,72	3847,7	0	3847,7	0	3836,3	0,30	3310,0	13,98
Screening: Alter 50-69 J, 2 J (Ref)	3870,0	3338,2	13,74	3863,3	0,18	3862,4	0,20	3851,9	0,47	3322,5	14,15
Screening: Alter 50-69 J, 2 J; 70-74 J, 3 J	3871,2	3339,2	13,74	3863,8	0,19	3862,8	0,22	3852,5	0,48	3322,9	14,16
Screening: Alter 50-74 J, 2 J	3872,5	3340,2	13,74	3863,9	0,22	3862,9	0,25	3852,7	0,51	3323,1	14,19
Screening: Alter 50-69 J, 2 J; 70-79 J, 3 J	3872,6	3340,4	13,75	3864,2	0,22	3863,1	0,25	3852,9	0,51	3323,2	14,19
Screening: Alter 50-79 J, 2 J	3873,4	3341,0	13,75	3863,8	0,25	3862,6	0,28	3852,6	0,54	3322,9	14,21
Screening: Alter 45-69 J, 2 J	3876,0	3343,3	13,75	3867,1	0,23	3865,9	0,26	3855,8	0,52	3325,6	14,20
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J	3877,1	3344,2	13,75	3866,7	0,27	3865,2	0,31	3855,2	0,56	3325,0	14,24
Screening: Alter 45-74 J, 2 J	3877,7	3344,6	13,75	3867,6	0,26	3866,2	0,30	3856,3	0,55	3326,0	14,23
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-74 J, 3 J	3878,1	3345,0	13,75	3867,2	0,28	3865,6	0,32	3855,7	0,58	3325,4	14,25
Screening: Alter 45-49 J, 1 J; 50-74 J, 2 J	3878,7	3345,5	13,75	3867,1	0,30	3865,5	0,34	3855,8	0,59	3325,4	14,27
Screening: Alter 45-79 J, 2 J	3879,0	3345,7	13,75	3867,4	0,30	3865,9	0,34	3856,2	0,59	3325,8	14,26
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-79 J, 3 J	3879,4	3346,1	13,75	3867,4	0,31	3865,7	0,35	3856,0	0,60	3325,6	14,28
Screening: Alter 45-49 J, 1 J; 50-79 J, 2 J	3880,0	3346,6	13,75	3867,0	0,34	3865,2	0,38	3855,6	0,63	3325,2	14,30

J: Jahr, QALY: quality-adjusted life year, Ref.: Referenzstrategie, Rel. Diff.: relative Differenz im Vergleich zu Lebensjahren

Tabelle 13 Sensitivitätsanalyse: Auswirkung der gleichzeitigen Erhöhung bzw. Reduktion der Lebensqualität um relative 5 % auf den Nutzen gemessen in lebensqualitätsadjustierte Lebenserwartung (pro 100 Frauen).

	Basisfallanalyse			Lebensqualitätsdaten um relative 5% erhöht		reduziert	
Strategie	**Lebensjahre**	**QALY**	**Rel. Diff. (%)**	**QALY**	**Rel. Diff. (%)**	**QALY**	**Rel. Diff. (%)**
Kein Screening	3847,7	3310,0	13,98	3480,4	9,55	3137,2	18,47
Screening: Alter 50-69 J, 2 J (Ref)	3870,0	3322,5	14,15	3494,1	9,71	3139,8	18,87
Screening: Alter 50-69 J, 2 J; 70-74 J, 3 J	3871,2	3322,9	14,16	3494,6	9,73	3140,1	18,89
Screening: Alter 50-74 J, 2 J	3872,5	3323,1	14,19	3494,8	9,75	3140,0	18,91
Screening: Alter 50-69 J, 2 J; 70-79 J, 3 J	3872,6	3323,2	14,19	3495,0	9,75	3140,1	18,92
Screening: Alter 50-79 J, 2 J	3873,4	3322,9	14,21	3494,7	9,78	3139,7	18,94
Screening: Alter 45-69 J, 2 J	3876,0	3325,6	14,20	3497,5	9,77	3140,6	18,97
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J	3877,1	3325,0	14,24	3496,9	9,80	3139,8	19,02
Screening: Alter 45-74 J, 2 J	3877,7	3326,0	14,23	3498,0	9,79	3140,8	19,00
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-74 J, 3 J	3878,1	3325,4	14,25	3497,4	9,82	3140,0	19,03
Screening: Alter 45-49 J, 1 J; 50-74 J, 2 J	3878,7	3325,4	14,27	3497,4	9,83	3139,9	19,05
Screening: Alter 45-79 J, 2 J	3879,0	3325,8	14,26	3497,9	9,83	3140,3	19,04
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-79 J, 3 J	3879,4	3325,6	14,28	3497,7	9,84	3139,9	19,06
Screening: Alter 45-49 J, 1 J; 50-79 J, 2 J	3880,0	3325,2	14,30	3497,3	9,87	3139,5	19,09

J: Jahr, QALY: quality-adjusted life year, Ref.: Referenzstrategie, Rel. Diff.: relative Differenz im Vergleich zu Lebensjahren

5.4.1. Testgüte

Tabelle 14 zeigt die Ergebnisse der Sensitivitätsanalysen zur Auswirkung einer Variation der falsch-negativen Rate (1-Sensitivität) der Mammographie um relative 20% auf selektierte Nutzen- und Schadensparameter.

Die Erhöhung oder Reduktion der falsch-negativen Rate der Mammographie hatte nur eine geringe Auswirkung auf den Nutzen gemessen an verhinderten Brustkrebstodesfällen pro 100 Frauen und den potenziellen Schaden gemessen an positiven Befunden. Die relative Differenz für die selektierten Nutzen- und Schadenparameter im Vergleich zu den Ergebnissen in der Basisfallanalyse lag unter 3,2% für die verhinderten Todesfälle und unter 0,18% für positive Befunde.

Tabelle 15 zeigt die Ergebnisse der Sensitivitätsanalysen zur Auswirkung einer Variation der falsch-positiven Rate (1-Spezifität) der Mammographie um relative 20% auf selektierte Nutzen- und Schadensparameter. Die relative Differenz für die selektierten Nutzen- und Schadenparameter im Vergleich zu den Ergebnissen in der Basisfallanalyse lag bei 17,6-18,3% für die verhinderten Todesfälle und bei 20% für positive Befunde.

Die Erhöhung oder Reduktion der falsch-positiven Rate der Mammographie hatte eine deutliche Auswirkung auf den Nutzen gemessen an verhinderte Brustkrebstodesfälle pro 100 Frauen und auf den potenziellen Schaden gemessen an positiven (und falsch-positiven) Befunden.

Die Variation der falsch-negativen Rate zeigte relativ geringe Auswirkungen auf die Schaden-Nutzen Verhältnisse (Abbildung 13 bis Abbildung 16). Das Ranking der Strategien blieb unverändert im Vergleich zur Basisfallanalyse und die ISNV der einzelnen nicht-dominierten Strategien auf der Effizienzgrenzenlinie veränderten sich nur in geringem Ausmaß.

Die Variation der falsch-positiven Rate zeigte nur geringe Auswirkungen auf die Schaden-Nutzen Verhältnisse (Abbildung 15 bis Abbildung 16). Das Ranking der Strategien blieb unverändert im Vergleich zur Basisfallanalyse. In der Sensitivitätsanalyse zur Testgüte mit erhöhter falsch-positiven Rate (Abbildung 15Abbildung 13) waren die inkrementellen Schaden-Nutzen-Verhältnisse im Vergleich zu den ISNV in der Basisfallanalyse etwas höher, während bei reduzierter falsch-positiven Rate (Abbildung 16) die ISNV im Vergleich zur Basisfallanalyse etwas niedriger ausfielen.

Tabelle 14 Sensitivitätsanalyse: Auswirkung der Variation der falsch-negativen Rate der Mammographie auf selektierte Nutzen- und Schadenparameter (pro 100 Frauen).

	Basisfall		Falsch-negative Rate um relativ 20% erhöht				Falsch-negative Rate um relativ 20% reduziert			
Strategie	**Verhinderte Todesfälle pro 100 Frauen (n)**	**Positive Befunde pro 100 Frauen (n)**	**Verhinderte Todesfälle pro 100 Frauen (n)**	**Positive Befunde pro 100 Frauen (n)**	**Rel. Diff. Verhinderte Todesfälle (%)**	**Rel. Diff. Positive Befunde (%)**	**Verhinderte Todesfälle pro 100 Frauen (n)**	**Positive Befunde pro 100 Frauen (n)**	**Rel. Diff. Verhinderte Todesfälle (%)**	**Rel. Diff. Positive Befunde (%)**
Kein Screening	0	0	0	0	0	0	0	0	0	0
Screening: Alter 50-69 J, 2 J (Ref)	1,23	98	1,19	98	-2,80	-0,15	1,26	98	2,67	0,14
Screening: Alter 50-69 J, 2 J; 70-74 J, 3 J	1,34	105	1,30	105	-2,82	-0,16	1,37	105	2,69	0,15
Screening: Alter 50-74 J, 2 J	1,47	118	1,43	118	-2,68	-0,15	1,51	118	2,56	0,14
Screening: Alter 50-69 J, 2 J; 70-79 J, 3 J	1,51	117	1,47	117	-2,81	-0,17	1,55	117	2,69	0,16
Screening: Alter 50-79 J, 2 J	1,60	129	1,56	129	-2,62	-0,15	1,64	129	2,50	0,14
Screening: Alter 45-69 J, 2 J	1,47	137	1,42	136	-3,16	-0,14	1,51	137	3,00	0,13
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J	1,51	161	1,46	161	-2,99	-0,11	1,55	161	2,80	0,11
Screening: Alter 45-74 J, 2 J	1,64	149	1,59	149	-3,05	-0,14	1,68	150	2,90	0,13
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-74 J, 3 J	1,62	168	1,57	167	-2,99	-0,12	1,66	168	2,81	0,11
Screening: Alter 45-49 J, 1 J; 50-74 J, 2 J	1,68	174	1,63	174	-2,90	-0,12	1,72	174	2,73	0,11
Screening: Alter 45-79 J, 2 J	1,82	167	1,77	167	-2,93	-0,14	1,87	167	2,79	0,13
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-79 J, 3 J	1,78	180	1,73	179	-2,97	-0,13	1,83	180	2,80	0,12
Screening: Alter 45-49 J, 1 J; 50-79 J, 2 J	1,86	191	1,81	191	-2,80	-0,12	1,91	191	2,64	0,11

J: Jahr, Ref.: Referenzstrategie, Rel. Diff.: relative Differenz im Vergleich zu den Ergebnissen in der Basisfallanalyse

Tabelle 15 Sensitivitätsanalyse: Auswirkung der Variation der falsch-positiven Rate der Mammographie auf Nutzen- und Schadenparameter (pro 100 Frauen).

	Basisfall		Falsch-positive Rate um relativ 20% erhöht				Falsch-positive Rate um relativ 20% reduziert			
Strategie	Verhinderte Todesfälle pro 100 Frauen (n)	Positive Befunde pro 100 Frauen (n)	Verhinderte Todesfälle pro 100 Frauen (n)	Positive Befunde pro 100 Frauen (n)	Rel. Diff. Verhinderte Todesfälle (%)	Rel. Diff. Positive Befunde (%)	Verhinderte Todesfälle pro 100 Frauen (n)	Positive Befunde pro 100 Frauen (n)	Rel. Diff. Verhinderte Todesfälle (%)	Rel. Diff. Positive Befunde (%)
Kein Screening	0	0	0	0			0	0		
Screening: Alter 50-69 J, 2 J (Ref)	98	87	116	104	17,6	20,0	81	69	-17,6	-20,0
Screening: Alter 50-69 J, 2 J; 70-74 J, 3 J	105	93	124	111	17,7	20,0	86	74	-17,7	-20,0
Screening: Alter 50-74 J, 2 J	118	105	139	126	17,8	20,0	97	84	-17,8	-20,0
Screening: Alter 50-69 J, 2 J; 70-79 J, 3 J	117	104	138	125	17,7	20,0	96	83	-17,7	-20,0
Screening: Alter 50-79 J, 2 J	129	115	152	139	17,9	20,0	106	92	-17,9	-20,0
Screening: Alter 45-69 J, 2 J	137	123	161	148	18,0	20,0	112	98	-18,0	-20,0
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J	161	147	190	177	18,3	20,0	132	118	-18,3	-20,0
Screening: Alter 45-74 J, 2 J	149	135	176	162	18,1	20,0	122	108	-18,1	-20,0
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-74 J, 3 J	168	153	198	184	18,3	20,0	137	123	-18,3	-20,0
Screening: Alter 45-49 J, 1 J; 50-74 J, 2 J	174	159	206	191	18,3	20,0	142	128	-18,3	-20,0
Screening: Alter 45-79 J, 2 J	167	151	197	181	18,1	20,0	137	121	-18,1	-20,0
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-79 J, 3 J	180	164	212	197	18,3	20,0	147	131	-18,3	-20,0
Screening: Alter 45-49 J, 1 J; 50-79 J, 2 J	191	175	226	211	18,3	20,0	156	140	-18,3	-20,0

J: Jahr, Ref.: Referenzstrategie, Rel. Diff.: relative Differenz im Vergleich zu den Ergebnissen in der Basisfallanalyse

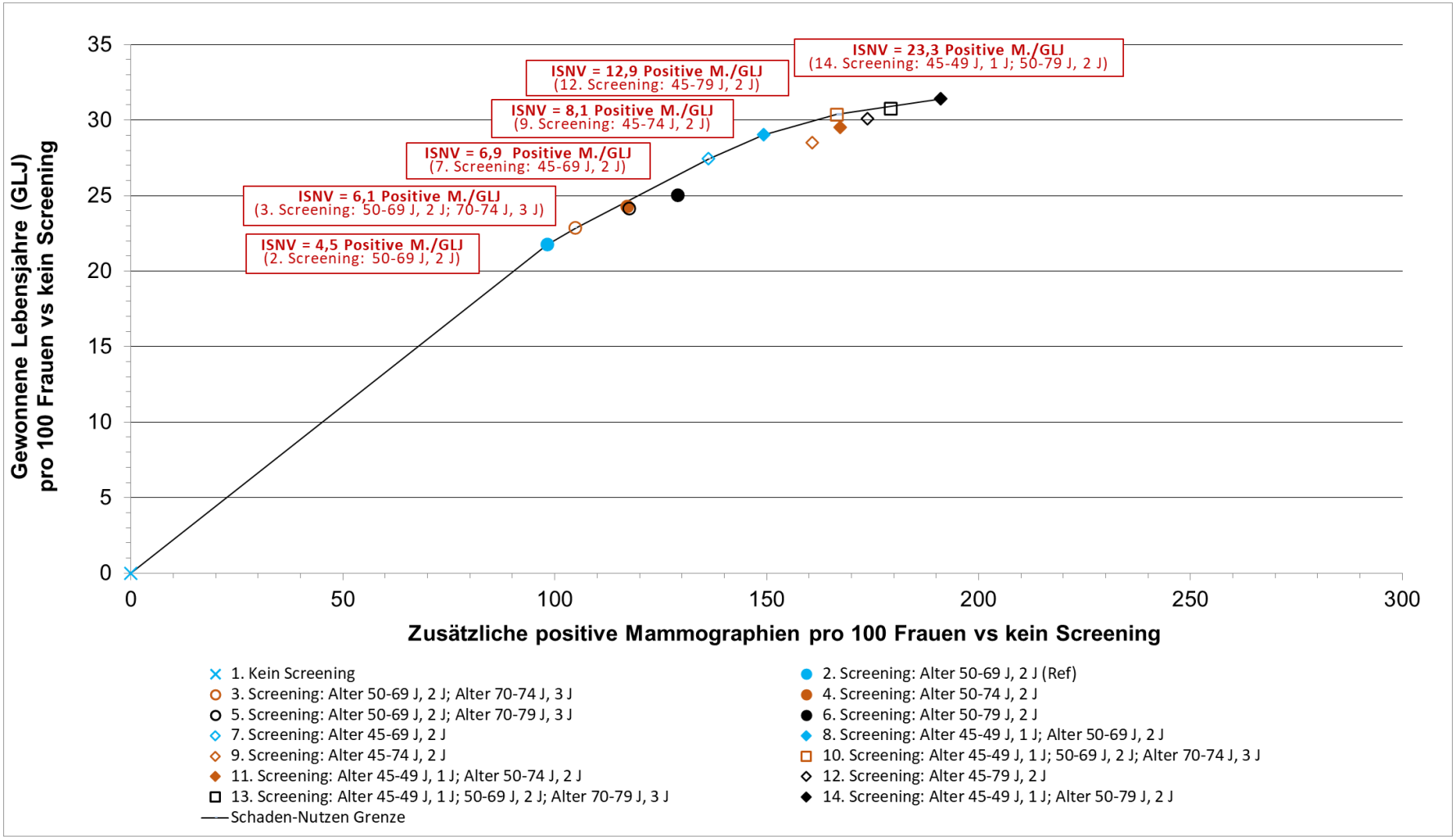


Abbildung 13 Ergebnisse der Sensitivitätsanalyse zur Testgüte: Gesamtanzahl an zusätzlichen positiven Mammographien versus gewonnene Lebensjahre bei einer um relative 20% erhöhte falsch-negativen Rate der Mammographie.

ISNV - inkrementelle Schaden-Nutzen-Verhältnis; M/GLJ – Mammographie pro gewonnenes Lebensjahr; J - Jahr

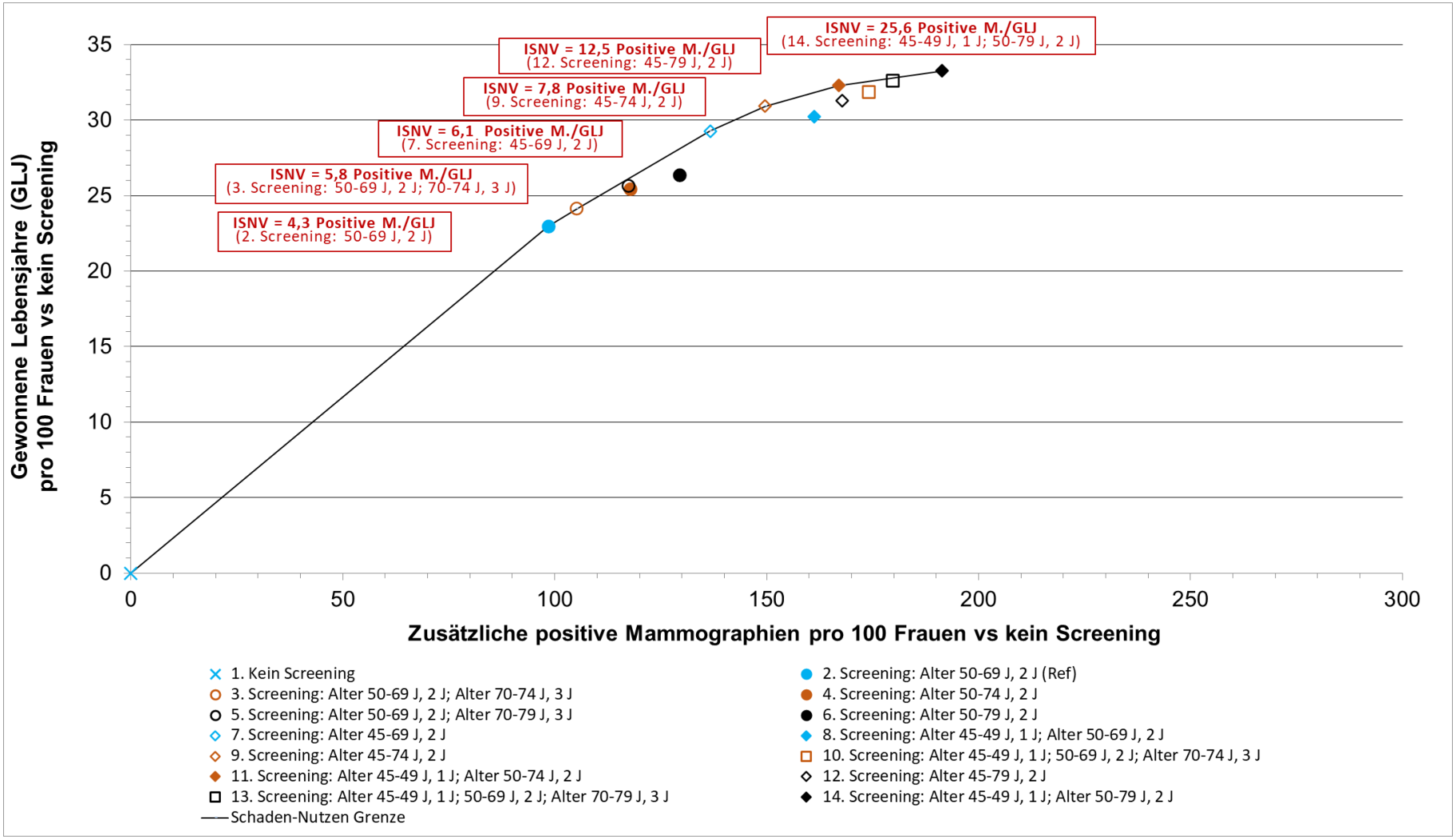


Abbildung 14 Ergebnisse der Sensitivitätsanalyse zur Testgüte: Gesamtanzahl an zusätzlichen positiven Mammographien versus gewonnene Lebensjahre bei einer um relative 20% reduzierte falsch-negativen Rate der Mammographie.

ISNV - inkrementelle Schaden-Nutzen-Verhältnis; M/GLJ – Mammographie pro gewonnenes Lebensjahr; J - Jahr

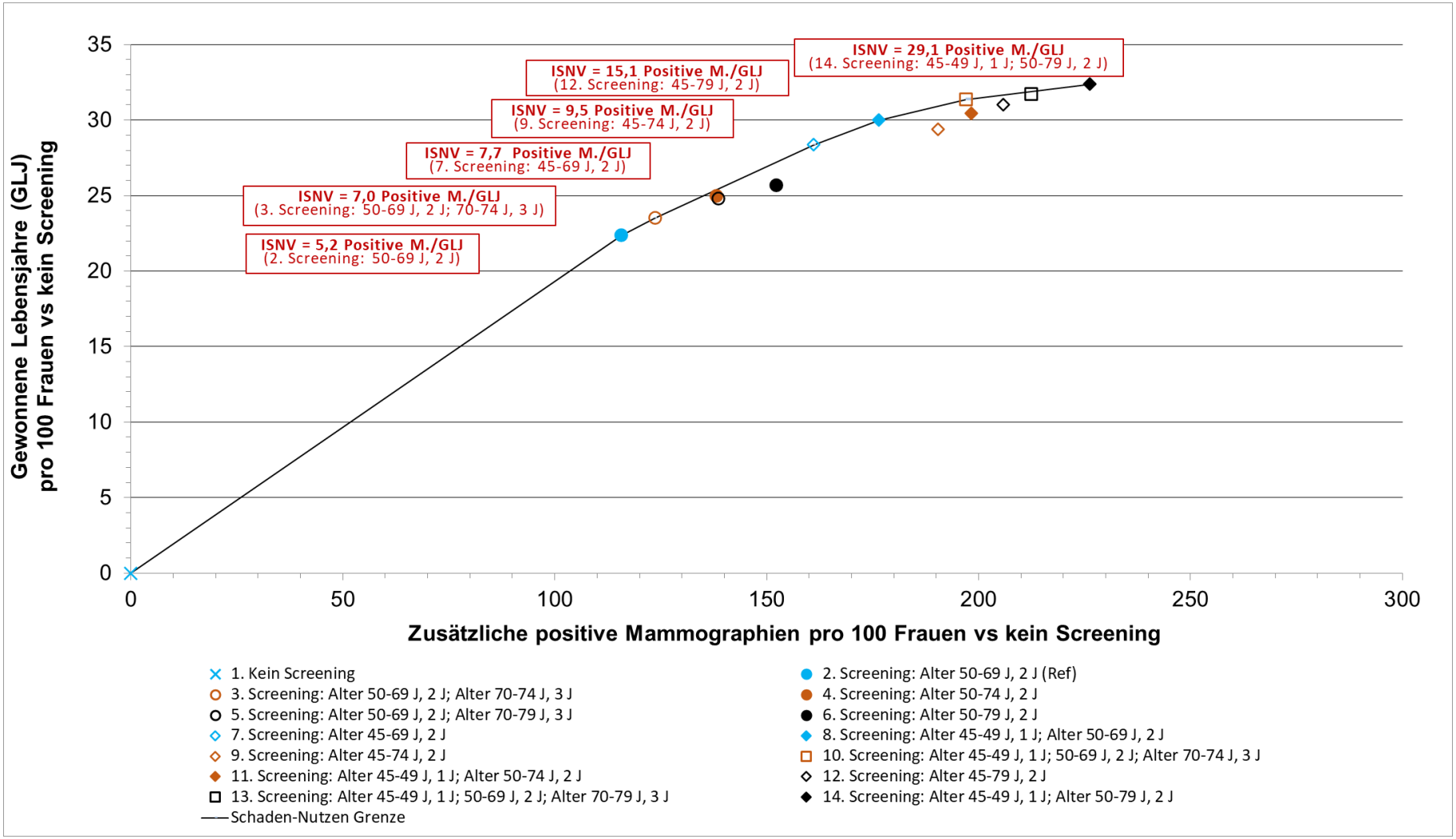


Abbildung 15 Ergebnisse der Sensitivitätsanalyse zur Testgüte: Gesamtanzahl an zusätzlichen positiven Mammographien versus gewonnene Lebensjahre bei einer um relative 20% erhöhte falsch-positiven Rate der Mammographie.

ISNV - inkrementelle Schaden-Nutzen-Verhältnis; M/GLJ – Mammographie pro gewonnenes Lebensjahr; J - Jahr

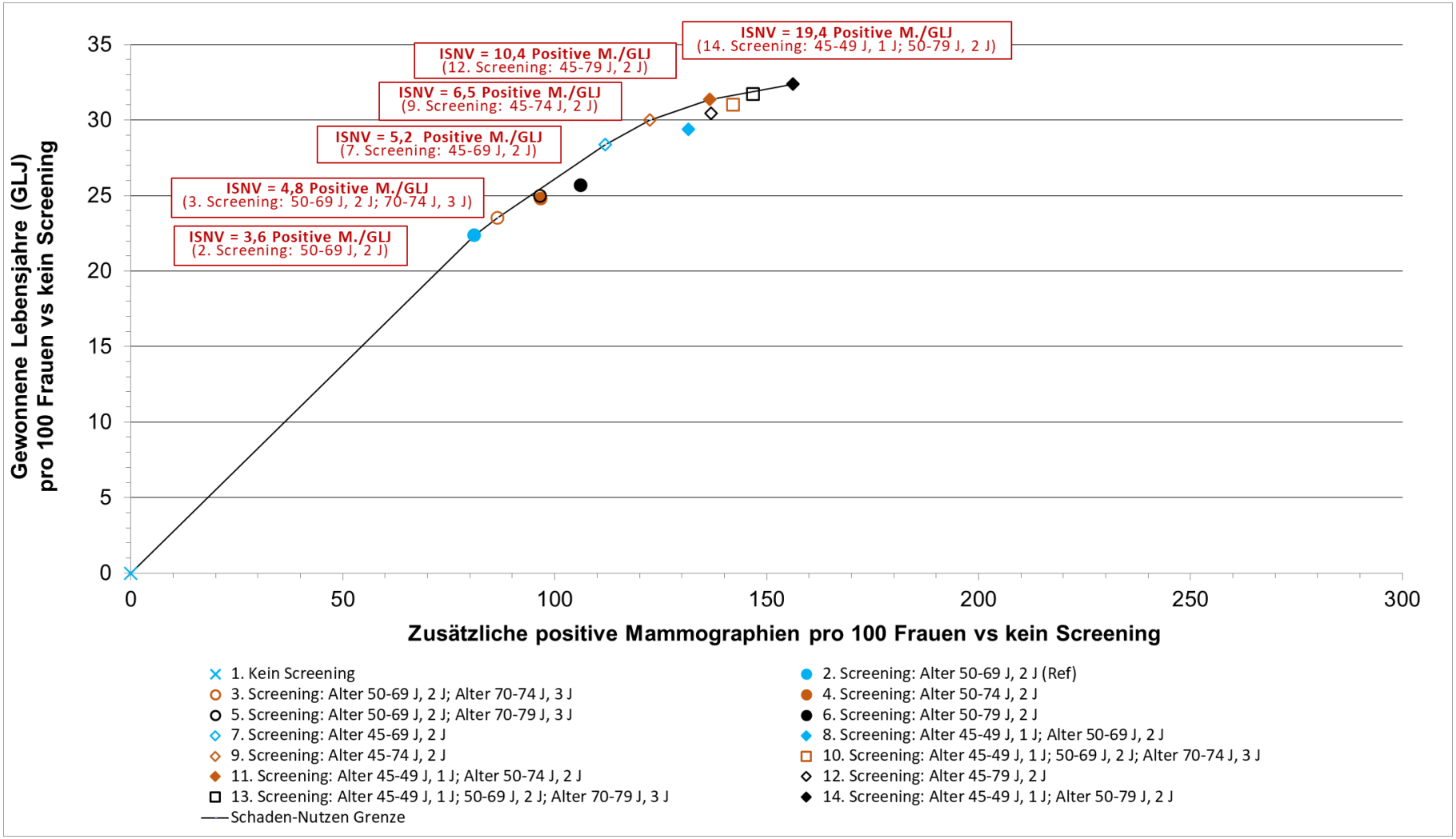


Abbildung 16 Ergebnisse der Sensitivitätsanalyse zur Testgüte: Gesamtanzahl an zusätzlichen positiven Mammographien versus gewonnene Lebensjahre bei einer um relative 20% reduzierte falsch-positiven Rate der Mammographie.

ISNV - inkrementelle Schaden-Nutzen-Verhältnis; M/GLJ – Mammographie pro gewonnenes Lebensjahr; J - Jahr

Tabelle 16 zeigt die Ergebnisse der Sensitivitätsanalysen zur Auswirkung einer gleichzeitigen Variation von Sensitivität und Spezifität der Mammographie in entgegengesetzte Richtungen um jeweils absolute 5% auf den Nutzen gemessen an der Restlebenserwartung (in Lebensjahre pro 100 Frauen) und lebensqualitätsadjustierte Restlebenserwartung (in QALY pro 100 Frauen).

Die gleichzeitige Erhöhung der Sensitivität und Reduktion der Spezifität um absolute 5% (Szenario 1) hatte nur geringe Auswirkung auf die Restlebenserwartung und lebensqualitätsadjustierte Lebenserwartung. Hier liegt die relative Differenz der Ergebnisse in der Restlebenserwartung und lebensqualitätsadjustierte Lebenserwartung im Vergleich zu den Ergebnissen in der Basisfallanalyse bei 0,020-0,028 % (für Lebensjahre) und 0,004-0,011 % (für QALY). Ebenso hat die gleichzeitige Reduktion der Sensitivität und Erhöhung der Spezifität um absolute 5% (Szenario 2) nur geringe Auswirkung auf die Lebenserwartung und lebensqualitätsadjustierte Restlebenserwartung. Hier liegt die relative Differenz der Ergebnisse in der Restlebenserwartung und lebensqualitätsadjustierten Restlebenserwartung im Vergleich zu den Ergebnissen in der Basisfallanalyse bei 0,021-0,029 % (Lebensjahre) und 0,006-0,013 % (QALY).

Tabelle 17 zeigt die Ergebnisse der Sensitivitätsanalysen zur Auswirkung einer gleichzeitigen Variation von Sensitivität und Spezifität der Mammographie in entgegengesetzte Richtungen um jeweils absolute 5% auf die positiven und falsch-positiven Befunde (pro 100 Frauen). Die gleichzeitige Erhöhung der Sensitivität und Reduktion der Spezifität um absolute 5% (Szenario 1) hat eine deutliche Auswirkung auf die positiven und falsch-positiven Befunde. Hier liegt die relative Differenz der Ergebnisse in den positiven und falsch-positiven Befunden im Vergleich zu den Ergebnissen in der Basisfallanalyse bei 42,5-48,9 % (für positive Befunde) und 46,3-54,5 % (für falsch-positive Befunde). Ebenso hat die gleichzeitige Reduktion der Sensitivität und Erhöhung der Spezifität um absolute 5% (Szenario 2) eine deutliche Auswirkung auf die positiven und falsch-positiven Befunden. Hier liegt die relative Differenz der Ergebnisse in den positiven und falsch-positiven Befunden im Vergleich zu den Ergebnissen in der Basisfallanalyse bei 42,4-48,7 % (für positive Befunde) und 46,2-54,2 % (für falsch-positive Befunde).

Die gleichzeitige Variation von Sensitivität und Spezifität in entgegengesetzter Richtung zeigte eine Auswirkung auf die Schaden-Nutzen Verhältnisse (Abbildung 17-Abbildung 20). Im Szenario 1 mit reduzierter Spezifität und erhöhter Sensitivität waren die inkrementellen Schaden-Nutzen-Verhältnisse im Vergleich zu den ISNV in der Basisfallanalyse deutlich höher, während im Szenario 2 mit erhöhter Spezifität und reduzierter Sensitivität die ISNV im Vergleich zur Basisfallanalyse niedriger ausfielen. Im Szenario 1 wurde außerdem im Vergleich zur Basisfallanalyse die Strategie 3 „Screening: Alter 50-69 J, 2 J; Alter 70-74 J, 3 J“ erweitert dominiert. Im Szenario 2 wurde die Strategie 7 „Screening: Alter 45-69 J, 2 J“ dominiert durch die Strategie 5 „Screening: Alter 50-69 J, 2 J; Alter 70-79 J, 3 J“.

Die Ergebnisse weiterer Sensitivitätsanalysen zur Testgüte der Mammographie sind im Appendix tabellarisch dargestellt (Appendix 9.2.)

Tabelle 16 Sensitivitätsanalyse: Auswirkung der gleichzeitigen Erhöhung (bzw. Reduktion) der Sensitivität und Reduktion (bzw. Erhöhung) der Spezifität der Mammographie auf die Restlebenserwartung und lebensqualitätsadjustierte Restlebenserwartung (pro 100 Frauen).

	Basisfall		Szenario 1 (Sens. +5%, Spez.-5%)				Szenario 2 (Sens. -5%, Spez. +5%)			
Strategie	**Lebens-jahre (LJ)**	**QALY**	**Lebens-jahre (LJ)**	**QALY**	**Rel. Diff. LJ (%)**	**Rel. Diff. QALY (%)**	**Lebens-jahre (LJ)**	**QALY**	**Rel. Diff. LJ (%)**	**Rel. Diff. QALY (%)**
Kein Screening	3848	3310	3848	3310	0,000	0,000	3848	3310	0,000	0,000
Screening: Alter 50-69 J, 2 J (Ref)	3870	3322	3871	3323	0,020	0,008	3869	3322	-0,021	-0,010
Screening: Alter 50-69 J, 2 J; 70-74 J, 3 J	3871	3323	3872	3323	0,021	0,009	3870	3323	-0,022	-0,010
Screening: Alter 50-74 J, 2 J	3872	3323	3873	3323	0,021	0,007	3872	3323	-0,023	-0,009
Screening: Alter 50-69 J, 2 J; 70-79 J, 3 J	3873	3323	3874	3324	0,022	0,008	3872	3323	-0,024	-0,010
Screening: Alter 50-79 J, 2 J	3873	3323	3874	3323	0,022	0,006	3872	3323	-0,023	-0,008
Screening: Alter 45-69 J, 2 J	3876	3326	3877	3326	0,026	0,011	3875	3325	-0,027	-0,013
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J	3877	3325	3878	3325	0,024	0,006	3876	3325	-0,025	-0,008
Screening: Alter 45-74 J, 2 J	3878	3326	3879	3326	0,027	0,010	3877	3326	-0,029	-0,012
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-74 J, 3 J	3878	3325	3879	3326	0,025	0,007	3877	3325	-0,027	-0,008
Screening: Alter 45-49 J, 1 J; 50-74 J, 2 J	3879	3325	3880	3326	0,025	0,006	3878	3325	-0,027	-0,007
Screening: Alter 45-79 J, 2 J	3879	3326	3880	3326	0,028	0,009	3878	3325	-0,029	-0,010
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-79 J, 3 J	3879	3326	3880	3326	0,026	0,006	3878	3325	-0,028	-0,008
Screening: Alter 45-49 J, 1 J; 50-79 J, 2 J	3880	3325	3881	3325	0,026	0,004	3879	3325	-0,028	-0,006

Rel. Diff.: relative Differenz, Sens.: Sensitivität, Spez.: Spezifität. Anmerkung: im Szenario 1 wurde die Sensitivität um absolut 5% erhöht und die Spezifität gleichzeitig um 5% reduziert, während im Szenario 2 die Sensitivität um absolut 5% reduziert und die Spezifität gleichzeitig um 5% erhöht wurde.

Tabelle 17 Sensitivitätsanalyse: Auswirkung der gleichzeitigen Erhöhung (bzw. Reduktion) der Sensitivität und Reduktion (bzw. Erhöhung) der Spezifität der Mammographie auf die positiven und falsch-positiven Befunde (pro 100 Frauen).

	Basisfall		Szenario 1 (Sens. + 5%, Spez. - 5%)				Szenario 2 (Sens. - 5%, Spez.+ 5%)			
Strategie	Positive Befunde pro 100 Frauen (n)	Falsch-positive Befunde pro 100 Frauen (n)	Positive Befunde pro 100 Frauen (n)	Falsch-positive Befunde pro 100 Frauen (n)	Rel. Diff. Positive Befunde (%)	Rel. Diff. Falsch-positive Befunde (%)	Positive Befunde pro 100 Frauen (n)	Falsch-positive Befunde pro 100 Frauen (n)	Rel. Diff. Positive Befunde (%)	Rel. Diff. Falsch-positive Befunde (%)
Kein Screening	0	0	0	0			0	0		
Screening: Alter 50-69 J, 2 J (Ref)	98	87	143	131	45,5	51,3	54	42	-45,4	-51,2
Screening: Alter 50-69 J, 2 J; 70-74 J, 3 J	105	93	154	141	46,4	52,2	56	45	-46,2	-52,0
Screening: Alter 50-74 J, 2 J	118	105	174	161	47,8	53,5	62	49	-47,7	-53,3
Screening: Alter 50-69 J, 2 J; 70-79 J, 3 J	117	104	173	160	47,6	53,4	62	49	-47,4	-53,2
Screening: Alter 50-79 J, 2 J	129	115	192	178	48,9	54,5	66	53	-48,7	-54,2
Screening: Alter 45-69 J, 2 J	137	123	195	182	43,1	47,6	78	65	-43,0	-47,5
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J	161	147	229	216	42,5	46,3	93	79	-42,4	-46,2
Screening: Alter 45-74 J, 2 J	149	135	216	201	44,6	49,1	83	69	-44,4	-48,9
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-74 J, 3 J	168	153	240	226	43,2	47,0	95	82	-43,1	-46,9
Screening: Alter 45-49 J, 1 J; 50-74 J, 2 J	174	159	250	235	43,8	47,6	98	84	-43,7	-47,5
Screening: Alter 45-79 J, 2 J	167	151	244	228	46,1	50,7	90	75	-45,9	-50,4
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-79 J, 3 J	180	164	259	243	44,2	48,1	100	85	-44,0	-48,0
Screening: Alter 45-49 J, 1 J; 50-79 J, 2 J	191	175	278	262	45,2	49,1	105	90	-45,1	-48,9

Rel. Diff.: relative Differenz, Sens.: Sensitivität, Spez.: Spezifität. Anmerkung: im Szenario 1 wurde die Sensitivität um absolut 5% erhöht und die Spezifität gleichzeitig um 5% reduziert, während im Szenario 2 die Sensitivität um absolut 5% reduziert und die Spezifität gleichzeitig um 5% erhöht wurde.

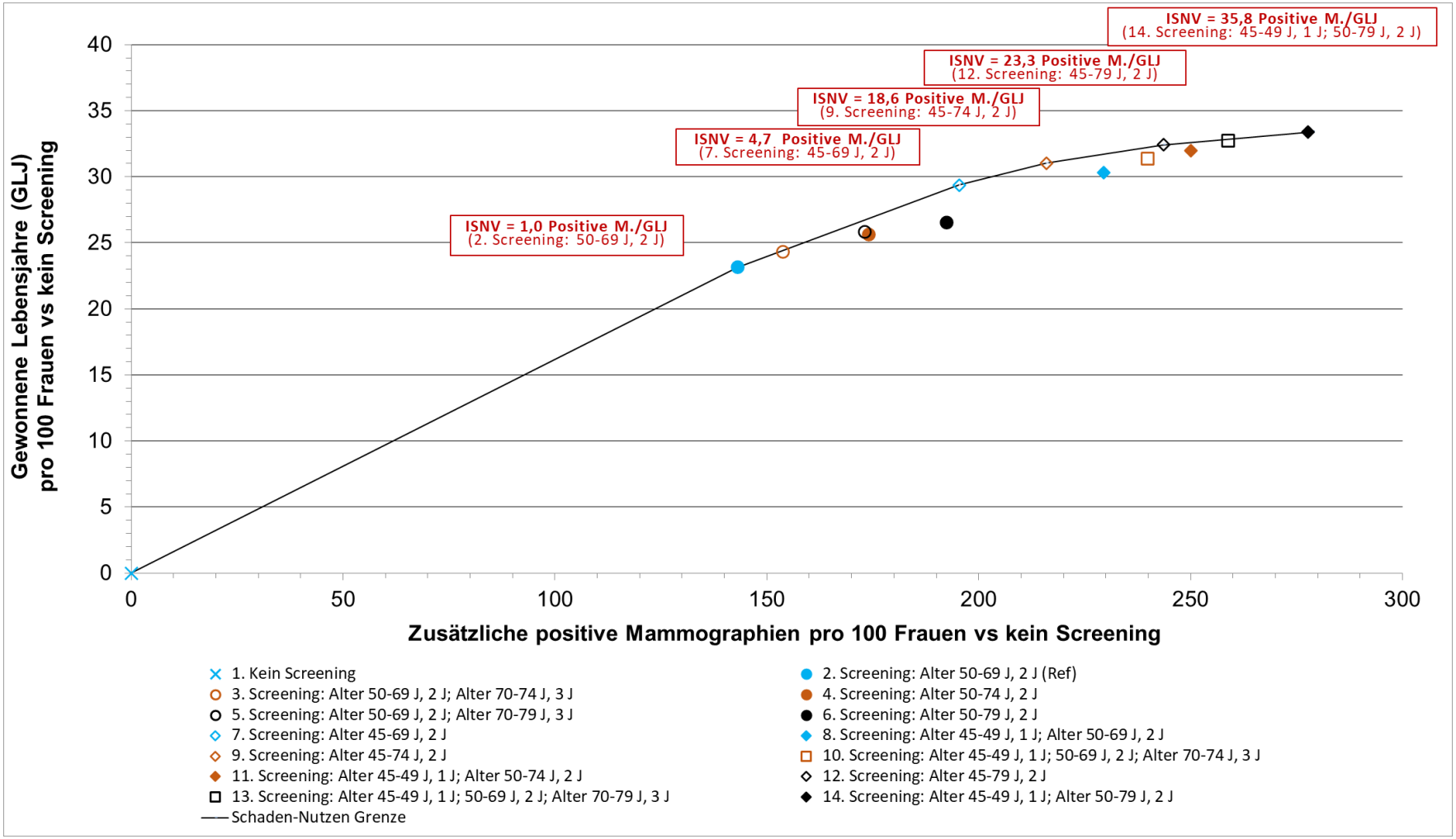


Abbildung 17 Ergebnisse der Sensitivitätsanalyse Szenario 1: Gesamtanzahl an zusätzlichen positiven Mammographien versus gewonnene Lebensjahre bei einer um absolute 5% erhöhter Sensitivität und gleichzeitig um absolute 5% reduzierter Spezifität der Mammographie.

ISNV - inkrementelle Schaden-Nutzen-Verhältnis; M/GLJ – Mammographie pro gewonnenes Lebensjahr; J - Jahr

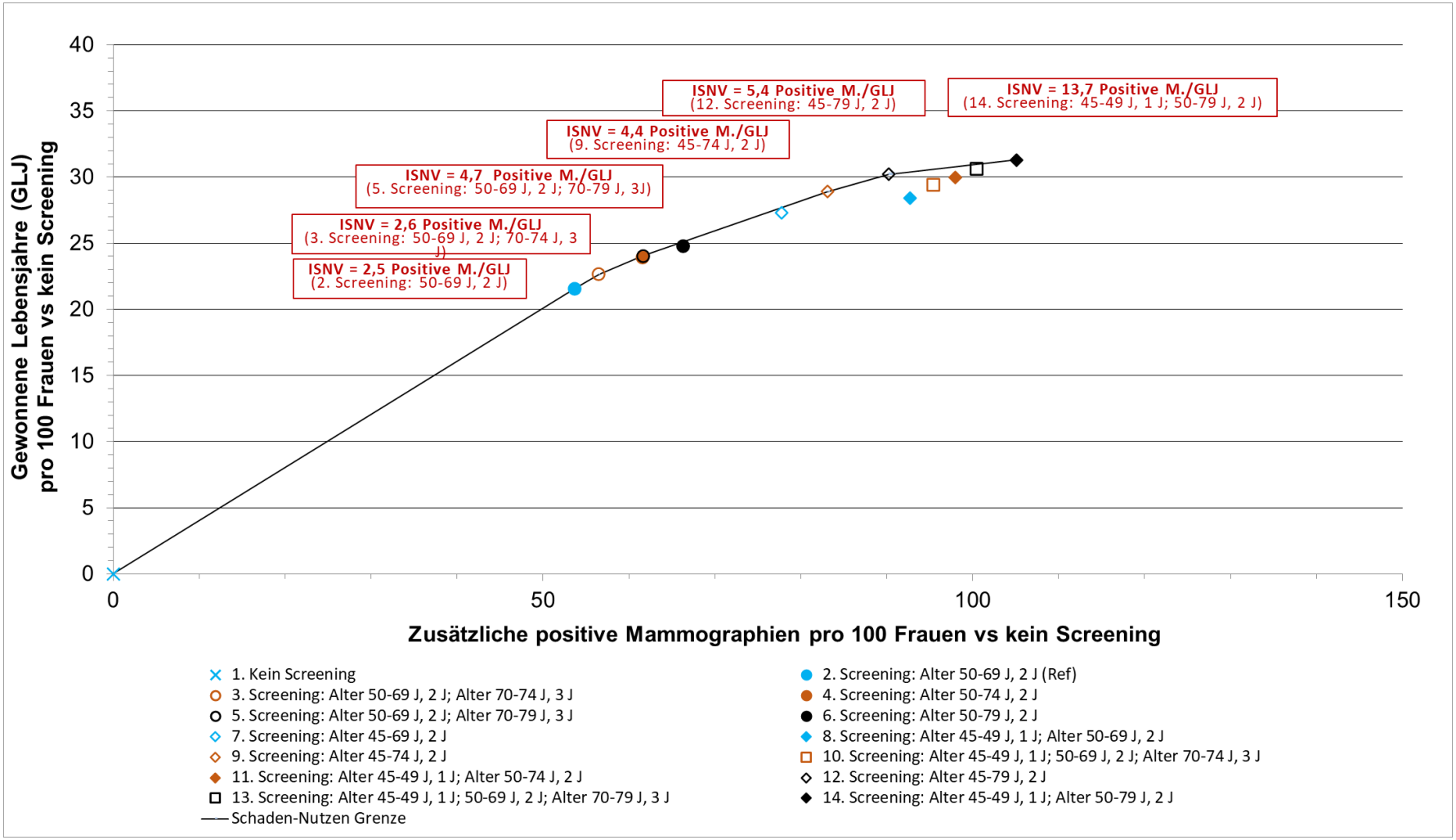


Abbildung 18 Ergebnisse der Sensitivitätsanalyse Szenario 2: Gesamtanzahl an zusätzlichen positiven Mammographien versus gewonnene Lebensjahre bei einer um absolute 5% reduzierte Sensitivität und gleichzeitig um absolute 5% erhöhte Spezifität der Mammographie.

ISNV - inkrementelle Schaden-Nutzen-Verhältnis; M/GLJ – Mammographie pro gewonnenes Lebensjahr; J - Jahr

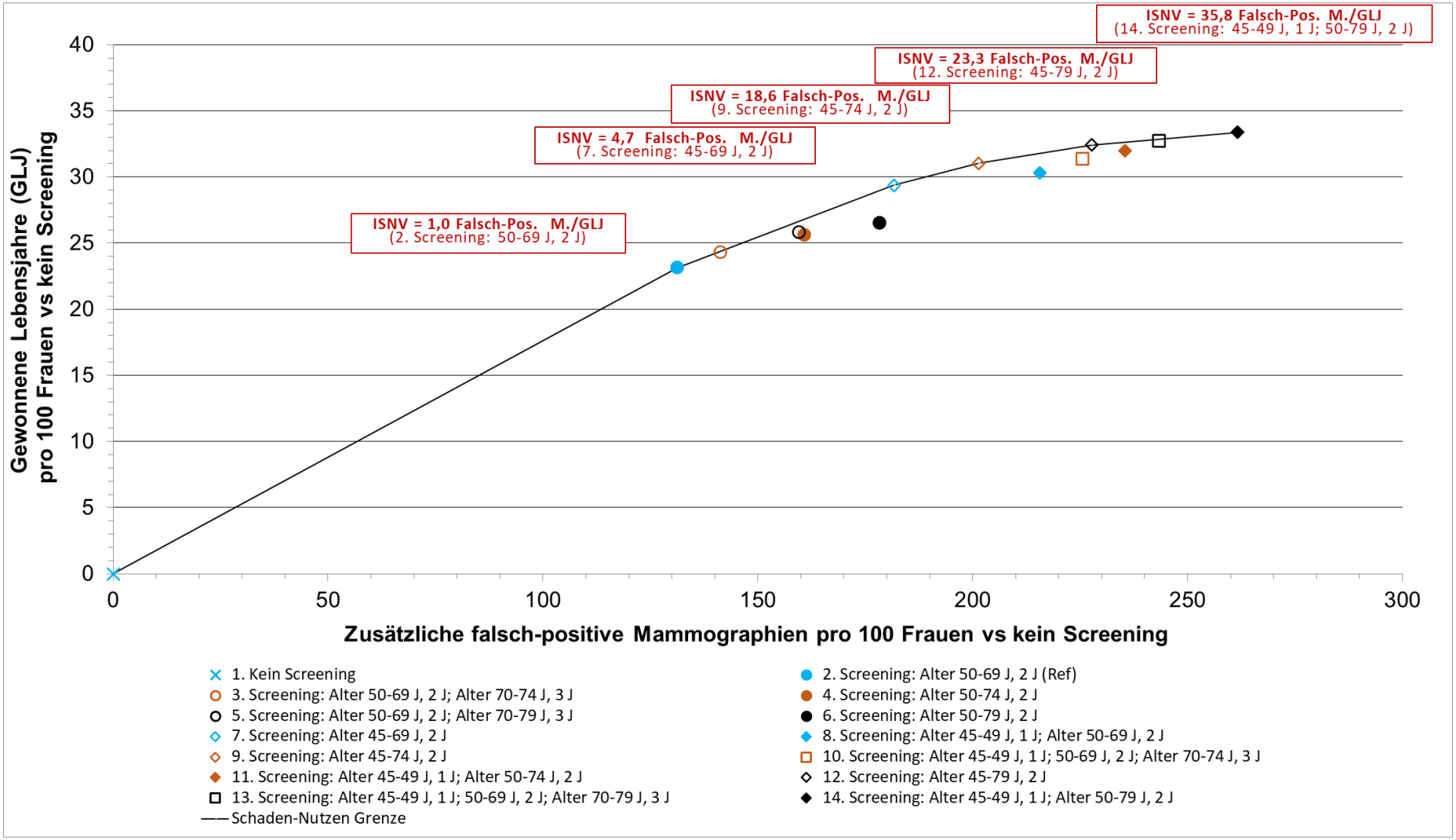


Abbildung 19 Ergebnisse der Sensitivitätsanalyse Szenario 1: Gesamtanzahl an zusätzlichen falsch-positiven Mammographien versus gewonnene Lebensjahre bei einer um absolute 5% erhöhter Sensitivität und gleichzeitig um absolute 5% reduzierter Spezifität der Mammographie.

ISNV - inkrementelle Schaden-Nutzen-Verhältnis; M/GLJ – Mammographie pro gewonnenes Lebensjahr; J – Jahr

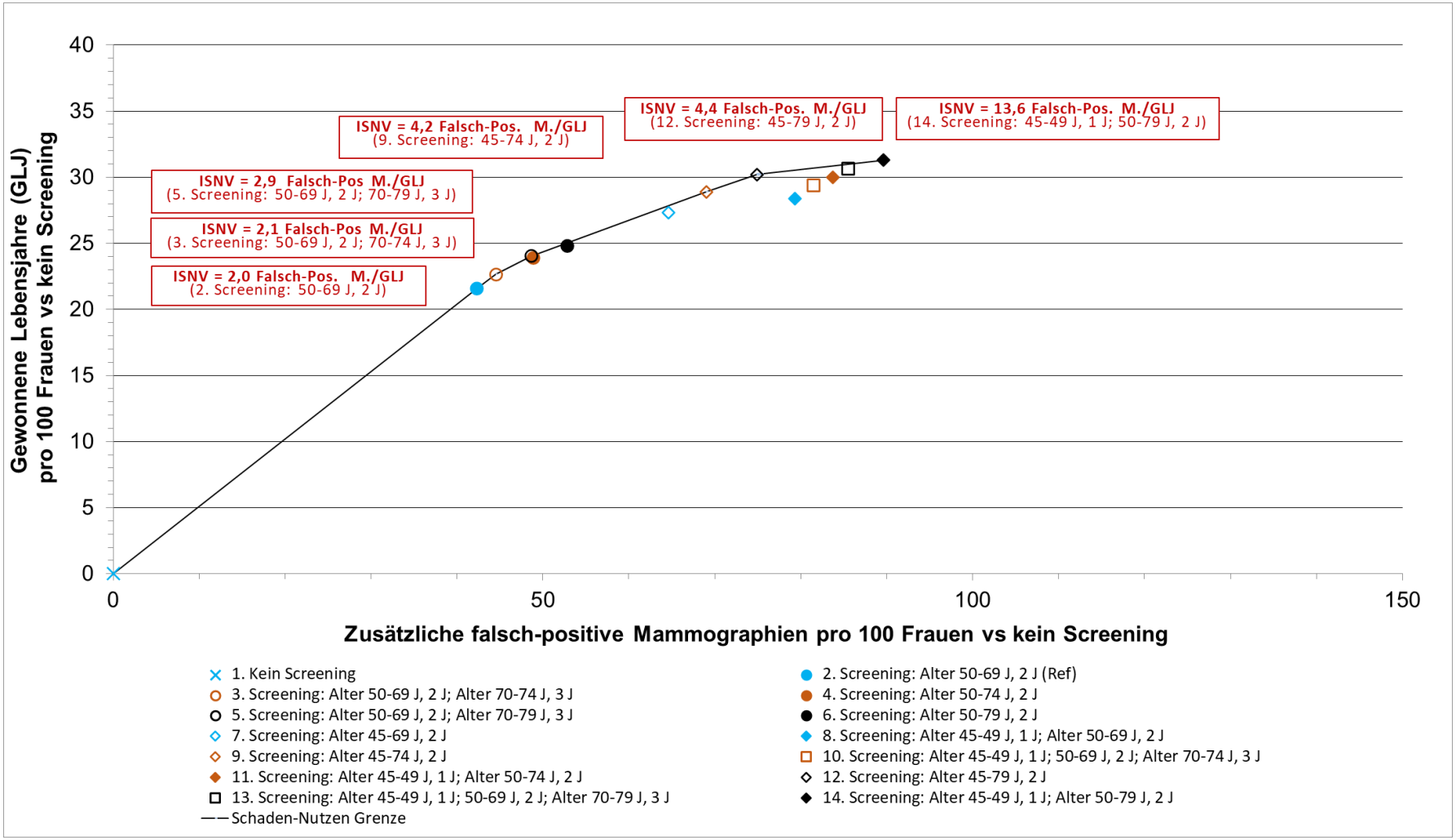


Abbildung 20 Ergebnisse der Sensitivitätsanalyse Szenario 2: Gesamtanzahl an zusätzlichen falsch-positiven Mammographien versus gewonnene Lebensjahre bei einer um absolute 5% reduzierte Sensitivität und gleichzeitig um absolute 5% erhöhte Spezifität der Mammographie.

ISNV - inkrementelle Schaden-Nutzen-Verhältnis; M/GLJ – Mammographie pro gewonnenes Lebensjahr; J - Jahr

6. Diskussion

Ziel dieser Studie war basierend auf einer evidenzbasierten entscheidungsanalytischen Modellierungsstudie für den deutschen Kontext den Einfluss einer Ausweitung der Altersgrenzen im etablierten Brustkrebs-Mammographie-Screening in Deutschland hinsichtlich des potenziellen Nutzens und Schadens zu untersuchen.

Hierzu wurde ein entscheidungsanalytisches Markov-Modell entwickelt, welches auf der besten verfügbaren Evidenz aus verschiedenen Quellen beruht und die Konsequenzen verschiedener untersuchter Brustkrebs-Screeningstrategien im Hinblick auf verschiedene Nutzen- und Schadenendpunkte darstellt.

Auf Basis der Ergebnisse in den Basisfallanalysen reduziert ein Mammographie-Screening in allen untersuchten Altersgruppen und Screeningintervallen das Lebenszeitrisiko, an entdecktem Brustkrebs zu versterben. Je größer die „Intensität“ eines Mammographie-Screening ist, d. h. je früher mit dem Screening begonnen wird, je später das Screening beendet wird und je kürzer die Screeningintervalle gewählt werden, desto größer ist der Nutzen bzgl. verhinderter brustkrebsbezogener Todesfälle.

Ein Screening kann auch negative Aspekte mit sich bringen, wie zum Beispiel die Sorge und Angstgefühle um die Mammographie und der psychologische Schaden durch positive bzw. falsch-positive Mammographien und Überdiagnosen, die wiederum mit einer Übertherapie verbunden sind, die ebenfalls zumindest für die Dauer der Therapie die Lebensqualität maßgeblich einschränken kann. Auch dieser Schaden nimmt in Summe im Erwartungswert mit zunehmender Intensität des Mammographie-Screening zu, was sich in einer Zunahme der im Rahmen dieser Studie modellierten Schadensendpunkte Anzahl der Mammographien, positiven Screeningbefunden und falsch-positiven Screeningbefunden ausdrückt.

Aus diesem Grund wurde eine den Nutzen und Schaden integrierende Analyse mit dem Endpunkt lebensqualitätsadjustierter Lebensjahre durchgeführt, da hierbei die Auswirkung von potenziellen Schadensaspekten des Screenings (z. B. durch die Mammographie, Screeningbefunde) und der darauffolgenden Behandlung in der Reduktion der Lebensqualität berücksichtigt. In dieser Analyse wird zudem der Nutzen des Screenings bezüglich vermiedener Morbidität durch Brustkrebs sowie der Einfluss von Alter auf die allgemeine Lebensqualität berücksichtigt. Unter Einbeziehung der Lebensqualität generierte die Strategie “Screening im Alter von 45-74 Jahren in zwei Jahresintervallen“ den höchsten Erwartungswert für lebensqualitätsadjustierten Lebensjahre. Die Detailergebnisse der Entscheidungsanalyse zeigen, dass dabei gemessen an lebensqualitätsadjustierten Lebensjahren dem früheren Screeningstart

ab 45 Jahren ein etwa fünfmal so großer Zusatznutzen zukommt wie dem späteren Screeningende mit 74 Jahren.

Zusätzlich wurden verschiedene Nutzen- und Schadenendpunkte im Rahmen von expliziten Nuten-Schaden-Gegenüberstellungen mit der Berechnung schrittweiser inkrementeller Schaden-Nutzen-Verhältnis durchgeführt und präsentiert. Hierbei wurde der Nutzen in Form von gewonnenen Lebensjahren dem Schaden gemessen an der Anzahl der Mammographien und positiven Screeningbefunden sowie falsch-positiven Screeningbefunden gegenübergestellt. Mammographie-Screening im Alter von 45 bis 74 Jahren erzielte einen deutlich höheren Nutzen als das etablierte Screening im Alter von 50 bis 69 Jahre oder auch eine Verlängerung des etablierten Screenings bis zum Alter 74 Jahre bei akzeptablen Schaden-Nutzen-Verhältnis. Durch ein kürzeres Screeningintervall im Alter von 45 bis 49 Jahren bzw. eine Verlängerung des Screenings bis zum Alter von 79 Jahren konnte der Nutzen hinsichtlich der Lebenserwartung nur noch geringfügig gesteigert werden und das inkrementelle Schaden-Nutzen-Verhältnis war deutlich höher.

Eine besondere Stärke dieser entscheidungsanalytischen Studie ist, dass langfristige Tradeoffs zwischen gesundheitlichem Nutzen und Belastungen im Zusammenhang mit verschiedenen Screeningstrategien, die sich nicht nur im Alter für den Beginn, sondern auch im Alter für das Ende und in den Screeningintervallen für bestimmte Altersgruppen unterscheiden, und den nachfolgenden Behandlungen quantifiziert und transparent dargestellt werden. Ferner stellte sich heraus, dass die Ergebnisse aus den unterschiedlichen Ansätzen, Schaden und Nutzen zu integrieren, wie der Ansatz mit lebensqualitätsadjustierten Lebensjahren oder der Ansatz mit inkrementellen Nutzen-Schaden-Verhältnissen, zu weitgehend ähnlichen Ergebnisse geführt hat, was die Identifikation optimaler Screeningstrategien angeht. Somit bietet diese entscheidungsanalytische Modellierungsstudie eine transparente und robuste Grundlage für Entscheidungsträger im Gesundheitswesen und ergänzt die Ergebnisse aus randomisierten klinischen Studien, die zum Teil heterogene Studiendesigns und Bedingungen aufweisen können sowie meist nicht über einen längeren Zeitraum die Effekte eines Screenings nachverfolgen.

Wie jedes entscheidungsanalytische Modell basiert auch dieses Modell auf der aktuellen Evidenz zu verschiedenen epidemiologischen und klinischen Parameterwerten sowie auf getroffenen Annahmen, die die Ergebnisse beeinflussen und daher als Limitationen zu nennen sind. Erstens wurde für die Bewertung des Nutzens des Mammographie-Screenings eine vollständige und regelmäßige Teilnahme am Screening, sowie eine vollständige Compliance mit assoziierten Kontrolluntersuchungen und Behandlungen verwendet. Dies spielt zwar eine untergeordnete Rolle bei niedrigeren Teilnahmeraten in der Population, da sich dabei gleichermaßen Nutzen und Schaden proportional zueinander verringern, allerdings kann sich eine nicht eingehaltene Regelmäßigkeit des Screenings innerhalb einer Person negativ auf den Nutzen oder das Schaden-Nutzen-Verhältnis auswirken. Zweitens wurden zwar selektierte

alternative Mammographie-Screeningstrategien mit unterschiedlichen Screeningintervallen in Abhängigkeit vom Alter der Frau analysiert, jedoch nicht alle möglichen Handlungsalternativen vollständig untersucht. Alternative Screeningstrategien könnten jedoch das inkrementelle Schaden-Nutzen Verhältnis verändern und sollten in zukünftigen Studien berücksichtigt werden (37). Beispielsweise wurde keine Strategie untersucht, die durchgehend einem jährlichen oder dreijährlichen Screeningintervall folgt. Auch wurden nicht alle möglichen dynamischen Screeningstrategien untersucht, bei denen sich das Screeningintervall mit dem Alter ändert. So weisen die vergleichenden Ergebnisse der Entscheidungsanalyse beispielsweise darauf hin, dass sich bei einer Verlängerung der Screeningdauer auf 74 oder 79 Jahre mit einem Screeningintervall von drei statt zwei Jahren in dieser Altersperiode ein gleichgroßer oder höherer Zusatznutzen erzielen lässt, wenn man als Zielgröße die qualitätsadjustierten Lebensjahre verwendet. Auch wurden im Rahmen der vorliegenden entscheidungsanalytischen Modellierung kein Vergleich von Screeningstrategien für Frauen durchgeführt, die im Alter von 50 bis 69 Jahren nicht an einem Mammographiescreening teilgenommen haben und in einem Alter von 70 Jahren beginnen möchten. Drittens wurde keine strahleninduzierte Brustkrebsentwicklung modelliert und somit kann über diesen möglichen Schaden und sein Verhältnis zum Nutzen keine Aussage getroffen werden. Der Schaden durch Strahlenbelastung dürfte sich aber ähnlich verhalten wie die Nutzen-Schaden-Analyse in Abhängigkeit von den durchgeführten Mammographien, die als Proxy-Maß für die Strahlenbelastung verstanden werden können. Viertens wurde kein familiäres oder genetisches Risiko für Brustkrebs berücksichtigt und auch der Menopausenstatus wurde nicht explizit modelliert. In dem verwendeten Modell variiert das Risiko für Brustkrebs ausschließlich basierend auf Alter und Brustdichte. Dies bedeutet, dass die Analysen nicht auf eine Population mit überdurchschnittlichem Risiko übertragen werden können; hierzu wären separate Modellierungen erforderlich. Fünftens wurden in dieser Entscheidungsanalyse keine anderen bildgebenden Verfahren berücksichtigt wie beispielsweise Ultraschalluntersuchungen oder computergestützte Auswertungen, Tomosynthese oder Magnetresonanztomographie. Daher basieren die gezogenen Schlussfolgerungen allein auf den untersuchten Mammographie-Screeningstrategien. Sechstens wurde angenommen, dass der natürliche Krankheitsverlauf des Brustkrebses dem Krankheitsverlauf vor Screening-Einführung im Jahr 2005 ähnelt. Diese Annahme wurde getroffen, da im Modell ein natürlicher Verlauf und symptomatische Entdeckungen modelliert wird, während die aktuellen Beobachtungsdaten aber immer auch die Effektivität des derzeitigen Screeningprogramms unter der jeweiligen Teilnahmerate widerspiegeln. Das Modell wurde anhand von aktuellen Beobachtungsdaten validiert und zeigt dabei eine gute Übereinstimmung. Siebtens, das entwickelte Markov-Modell unterliegt der Annahme, dass die Krankheitsentwicklung nur vom aktuellen Gesundheitszustand und Alter abhängig ist, aber nicht von vorherigen Gesundheitszuständen. Andere publizierte und internationale entscheidungsanalytische Modelle machen dieselbe Annahme und haben bei Expertinnen und Experten aus Medizin und Epidemie eine hohe Akzeptanz (26). Unser Modell

zeigte zudem im Validierungsprozess eine gute Übereinstimmung mit aktuellen Beobachtungsdaten. Achtens wurde im vorliegenden Modell die vereinfachende Annahme gemacht, dass Personen mit entdecktem DCIS und invasivem Krebs gemäß den deutschen Leitlinien vollständig abgeklärt und behandelt werden. Der Krankheitsverlauf wird nicht explizit modelliert, sondern es werden die durchschnittlichen alters- und stadienspezifischen deutschen Überlebenswahrscheinlichkeiten angewandt. Da die deutschen Leitlinien zur Brustkrebsbehandlung anerkannt sind und sowohl die Prognose als auch die Behandlung auf UICC- Stadien auf dem Zeitpunkt der Erstdiagnose beruhen, ist diese Annahme als akzeptabel zu bewerten. Es wurde außerdem basierend auf drei großen Brustkrebs-Screening Studien angenommen (15), dass zusätzlich zum positiven Stage-Shift-Effekt des Mammographie-Screenings zu berücksichtigen ist, dass die Überlebenswahrscheinlichkeiten nicht allein vom UICC-Stadium abhängen, sondern zusätzlich davon, ob es sich um einen symptomatisch oder durch Screening entdeckten Brustkrebs handelt. Diese Tatsache wurde im Modell quantitativ berücksichtigt. Neuntens, die größte Unsicherheit ist in den Ergebnissen für die Überdiagnose zu erwarten, da es sich hierbei um ein Maß handelt, für das es derzeit keine validen empirischen fundierten Daten gibt (38, 39). Für die Berechnung der Überdiagnose gibt es unterschiedliche Methoden, aber bisher keinen Goldstandard. Diese Modellierung leistet dennoch einen nützlichen Beitrag zur Abschätzung der Größenordnung von zu erwartenden Überdiagnosen, da ausdrücklich die Lead-Time und die konkurrierende Mortalität berücksichtigt, sowie außerdem eine Lebenszeitperspektive eingenommen wurde. Die Ergebnisse sollten jedoch mit Vorsicht betrachtet und interpretiert werden. Andere entscheidungsanalytische Modelle für den US-amerikanischen Kontext schätzten, dass je nach Screening-Strategie 2 % bis 12 % der Brustkrebsfälle und 30 % bis 50 % der DCIS-Fälle eine Überdiagnose darstellen könnten (26). Die Ergebnisse des vorliegenden Modells für Deutschland schätzten den Anteil der überdiagnostizierten Fälle insgesamt (Brustkrebs und DCIS) auf 49 % bis 57 %. Obwohl die absoluten Schätzungen für Überdiagnosen aus unterschiedlichen Modellen stark variierten, stimmten jedoch alle Ergebnisse (auch die des vorliegenden Modells) darin überein, dass die meisten überdiagnostizierten Fälle DCIS sind. Die aktive Überwachung von DCIS mit geringem Risiko für eine Progression ist daher ein möglicher zukünftiger Ansatz zur Verbesserung des Schaden-Nutzen-Verhältnisses. Die Modellergebnisse für die Überdiagnose sind nicht direkt vergleichbar mit anderen berichteten Schätzungen aus empirischen Studien, da im entscheidungsanalytischen Modell die Frauen ein Leben lang verfolgt werden.

Eine der wichtigsten offenen Fragen bleibt die Fragen nach einem optimalen bzw. akzeptablen Schaden-Nutzen Verhältnis für die jeweiligen Endpunkte in der Entscheidungsfindung. Darüber hinaus gibt es keine empirischen Daten zur Akzeptanzschwelle der screeningberechtigten Frauen in Deutschland, d. h. wie viele Fälle (bzw. Risiko) von Überdiagnose und Überbehandlung Frauen in Deutschland für akzeptabel halten, um einen Todesfall durch Brustkrebs zu verhindern. Diese Akzeptanzschwelle kann selbstverständlich erheblich variieren je nach Alter

der Frauen, ihren individuellen Präferenzen und ihrer individuellen Wahrnehmung der Schwere der Folgen einer Behandlung und ihrer Auswirkungen auf die individuelle Lebensqualität. In der Literatur werden Akzeptanzschwellen von 150 Überdiagnosen für eine 10%ige Reduktion in der brustkrebsspezifischen Mortalität oder 14 Überdiagnosen bzw. 47,8 falsch-positive Screeningbefunde für einen verhinderten brustkrebsbedingten Todesfall genannt (40, 41). Deshalb ist weitere Forschung erforderlich, um den Bereich der Akzeptanzschwelle der Frauen in Deutschland für eine Überbehandlung beim Mammographie-Screening zu ermitteln. Da die Abwägung zur Teilnahme an einem Screening bzw. zwischen verschiedenen Screeningoptionen letztlich von der einzelnen Frau zu treffen ist, sind die im Rahmen dieser Studie ermittelten Ergebnisse in Entscheidungshilfen wie beispielsweise Faktenboxen (42-47) so aufzubereiten, dass sie in einem partizipatorischen Entscheidungsprozess zwischen Frau und betreuender Ärztin bzw. betreuendem Arzt zur transparenten Entscheidungsunterstützung eingesetzt werden können. Dabei ist darauf zu achten, dass nicht alle Abwägungen gegenüber keinem Screening angestellt werden, sondern dass eine inkrementelle Betrachtung der Schaden-Nutzen-Verhältnisse von Screeningstrategien mit geringerer Intensität schrittweise zu Screeningstrategien mit höherer Intensität vorgenommen wird, da nur auf diese Weise die Erwartungswerte der Entscheidungen optimiert werden.

Die Ergebnisse dieser entscheidungsanalytischen Modellierung sollen eine Unterstützung in der Entscheidungsfindung für Entscheidungsträger im öffentlichen Gesundheitswesen darstellen, aber können keinesfalls die individuelle Screening-Entscheidung der Frauen bzw. das ärztliche Gespräch ersetzen.

In zukünftigen Forschungsstudien sollte untersucht werden, wie hoch die Bereitschaft von Frauen ist für die Überdiagnose und damit assoziierte Überbehandlung bzw. wie viele Screeninguntersuchungen (Mammographien) und damit verbundene Belastung Frauen in Deutschland für einen vermiedenen Krebstodesfall akzeptieren. Zusätzlich sollten in entscheidungsanalytischen Studien weitere zusätzliche Screeningstrategien untersucht werden darunter auch personalisierte Strategien, die die Screeninghistorie einer Frau und weitere individuelle Risikofaktoren berücksichtigen.

7. Schlussfolgerungen

Basierend auf internationalen Daten und Daten aus dem Gesundheitssystem Deutschland konnte ein entscheidungsanalytisches Modell zur Abschätzung des Nutzens und des möglichen Schadens mit einem lebenslangen analytischen Zeithorizont entwickelt und anhand von Daten des deutschen Gesundheitssystems kalibriert und validiert werden.

Basierend auf den Ergebnissen der entscheidungsanalytischen Modellierung kann ein Mammographie-Screening im Alter von 45 bis 49 Jahren und im Alter von 70 bis 79 Jahren zusätzlich zum derzeitigen Mammographie-Screening von 50 bis 69 Jahre alle zwei Jahre weitere Todesfälle durch Brustkrebs verhindern und die Restlebenserwartung weiter vergrößern.

Ohne Berücksichtigung der gesundheitsbezogenen Lebensqualität nimmt der Nutzen gemessen in gewonnenen Lebensjahren mit abnehmenden Screeningintervallen und zunehmender Screeningdauer innerhalb der in der vorliegenden entscheidungsanalytischen Studie evaluierten Strategien zu. Bei zusätzlicher Berücksichtigung der Lebensqualität ergibt sich ein differenzierteres Bild. Bei Beibehaltung des derzeitigen Screeningintervalls von zwei Jahren kann eine Ausweitung des Screening-Programms auf ein Alter von 45 bis 74 Jahre empfohlen werden. Dabei kommt dem früheren Screeningstart ab 45 Jahre ein größerer Nutzen zu als dem späteren Screeningende mit 74 Jahren.

Ob eine Verlängerung des Screenings über das Alter von 75 Jahren hinaus angeboten werden sollte und mit welchem Screeningintervall, hängt entscheidend von der Akzeptanz der Frauen in Deutschland ab zum Schaden-Nutzen-Verhältnis ab, d. h. wie viele potenzielle Schadenseinheiten (z. B. gemessen an Mammographien oder erwarteten positiven Testergebnissen) pro zusätzlich gewonnene Nutzeneinheit (z. B. verhinderten Todesfall durch Brustkrebs oder gewonnenem Lebensjahr) sie akzeptieren würden. Ebenso ist die Entscheidung für ein kürzeres Screeningintervall bei Frauen im Alter von 45 bis 49 Jahren von dieser Akzeptanzschwelle abhängig. Die individuellen Präferenzen und Risikoeinstellungen zum akzeptablen inkrementellen Nutzen-Schaden-Verhältnis der Frauen sind im Rahmen einer partizipativen Entscheidungsfindung zu berücksichtigen.

Die Ergebnisse der Entscheidungsanalyse der im Rahmen dieser Modellierung untersuchten und verglichenen Screeningstrategien weisen darauf hin, dass bei einer Verlängerung der Screeningperiode auf 74 oder 79 Jahre ein Screeningintervall von zwei Jahren keinen maßgeblichen Zusatznutzen bzw. eher einen Verlust an Lebensqualität zur Folge hat. Aus diesem Grunde drei Jahren einen höheren Nutzen hinsichtlich der qualitätsadjustierten Lebensjahre erbringt.

Zukünftiger Forschungsbedarf liegt einerseits in der Erhebung der gesellschaftlichen und individuellen Akzeptanz für Überdiagnosen bzw. potenziellem Schaden pro verhinderten Todesfall durch Brustkrebs. Zum anderen sollten weitere dynamische Screeningstrategien in die Entscheidungsanalysen einbezogen werden, bei welchen das Screeningintervall aufgrund des Alters oder anderer individueller Faktoren auf mehr als zwei Jahre verlängert werden könnte, um Überdiagnosen zu minimieren und das Nutzen-Schaden-Verhältnisses weiter zu optimieren.

8. Literaturverzeichnis

1. Deutsche Krebsgesellschaft e.V. (DKG) Deutsche Gesellschaft für Gynäkologie und Geburtshilfe (DGGG). Interdisziplinäre S3-Leitlinie für die Diagnostik, Therapie und Nachsorge des Mammakarzinoms 2012. Leitlinienprogramm Onkologie der AWMF, Deutschen Krebsgesellschaft, e.V. und Deutschen Krebshilfe e.V. Krebs in Deutschland für 2013/2014. 11 ed. Berlin: Robert Koch-Institut und die Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (Hrsg); 2017. Available from: https://www.awmf.org/leitlinien/detail/ll/032-045OL.html
2. Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft, Deutsche Krebshilfe, AWMF). S3-Leitlinie Früherkennung, Diagnose, Therapie und Nachsorge des Mammakarzinoms, Version 4.4. Available from: http://www.leitlinienprogramm-onkologie.de/leitlinien/mammakarzinom/ 2021 [
3. Gemeinsame Bundesausschuss (G-BA). Beschluss des Gemeinsamen Bundesausschusses über die KFE-RL: Überprüfung der Altersgrenzen im Mammographie-Screening gemäß § 135 Absatz 1 Satz 1 i.V. mit § 25 Absatz 2 des Fünften Buches Sozialgesetzbuch. Available from: https://www.g-ba.de/downloads/39-261-4760/2021-03-18_KFE-RL_Einleitung-Beratung_Mammo-Altersgrenzen.pdf: Gemeinsame Bundesausschuss (G-BA); 2021 [
4. Gemeinsame Bundesausschuss (G-BA). Zeitplan: Krebsfrüherkennungs-Richtlinie: MammographieScreening Erweiterung der Altersgrenzen; Stand: 11.03.2021. Available from: https://www.g-ba.de/downloads/40-268-7392/2021-03-18_KFE-RL_Einleitung-Beratung_Mammo-Altersgrenzen_Zeitplan.pdf2021
5. Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-Effectiveness in Health and Medicine: Oxford University Press; 1996.
6. Hunink MGM, Weinstein, M. C., Wittenberg, E., Drummond, M. F., Pliskin, J. S., Wong, J. B., Wong, J. B., & Glasziou, P. P. Decision making in health and medicine: Integrating evidence and values. 2nd ed: Cambridge University Press; 2014.
7. Siebert U. When should decision-analytic modeling be used in the economic evaluation of health care? The European Journal of Health Economics, formerly: HEPAC. 2003;4(3):143-50.
8. Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, et al. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--3. Value Health. 2012;15(6):812-20.
9. Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, et al. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--3. Medical Decision Making. 2012;32:690-77.
10. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Med Decis Making. 1993;13(4):322-38.
11. Siebert U. Transparente Entscheidungen in Public Health mittels systematischer Entscheidungsanalyse. In: Schwartz FW, Ulla W, Siegrist J, Kolip P, Leidl R, Dierks ML, et al., editors. Das Public Health Buch. Gesundheit und Gesundheitswesen. 3rd ed. Munich: Urban & Fischer; 2012
12. Siebert U. Prognosen sind unsicher – besonders wenn sie die Zukunft betreffen. Der Stellenwert von entscheidungsanalytischen Modellen. Das Herbst-Symposium: Übertragung von Evidenz – Spiel ohne Grenzen?; 2017; Köln: IQWIG.
13. Caro JJ, Briggs AH, Siebert U, Kuntz KM. Modeling Good Research Practices—Overview:A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force–1. Value in Health 2012;15:796-803.
14. Caro JJ, Briggs AH, Siebert U, Kuntz KM. Modeling Good Research Practices—Overview:A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force–1. Medical Decision Making. 2012;32(5):667-77.

15. Shen Y, Yang Y, Inoue LY, Munsell MF, Miller AB, Berry DA. Role of detection method in predicting breast cancer survival: analysis of randomized screening trials. J Natl Cancer Inst. 2005;97(16):1195-203.
16. EUnetHTA. Methods for health economic evaluations – A guideline based on current practices in Europe. 2015. Available from: https://www.eunethta.eu/wp-content/uploads/2018/03/Methods_for_health_economic_evaluations.pdf
17. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)--explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. Value Health. 2013;16(2):231-50.
18. GENESIS-Online. Wiesbaden: Statistisches Bundesamt (Destatis). Available from: https://www-genesis.destatis.de. [updated 05.07.2021].
19. Robert Koch-Institut. Zentrum für Krebsregisterdaten im Robert Koch-Institut: Datenbankabfrage mit Schätzung der Inzidenz, Prävalenz und des Überlebens von Krebs in Deutschland auf Basis der epidemiologischen Landeskrebsregisterdaten. Mortalitätsdaten bereitgestellt vom Statistischen Bundesamt. Available from: https://www.krebsdaten.de/Krebs/SiteGlobals/Forms/Datenbankabfrage/datenbankabfrage_stufe2_form.html: www.krebsdaten.de/abfrage;[updated 16.03.2021, accessed: 24.6.2021].
20. Schiller-Fruehwirth I, Jahn B, Einzinger P, Zauner G, Urach C, Siebert U. The Long-Term Effectiveness and Cost Effectiveness of Organized versus Opportunistic Screening for Breast Cancer in Austria. Value Health. 2017;20(8):1048-57.
21. DESTATIS. Sterbetafel (Periodensterbetafel): Deutschland, Jahre, Geschlecht, Vollendetes Alter 2018/20. Available from: https://www-genesis.destatis.de/genesis//online?operation=table&code=12621-0001&bypass=true&levelindex=0&levelid=1606127446797#abreadcrumb
22. Tumorregister München. Überleben ICD-10 C50: Mammakarzinom (Frauen) [Internet]. 2021 Available from: https://www.tumorregister-muenchen.de/facts/surv/sC50f_G-ICD-10-C50-Mammakarzinom-Frauen-Survival.pdf]. [updated 25.08.2021, accessed 17.09.2021.]
23. Arnold M, Pfeifer K, Quante AS. Is risk-stratified breast cancer screening economically efficient in Germany? PLOS ONE. 2019;14(5):e0217213.
24. Weigel S, Heindel W, Heidrich J, Hense HW, Heidinger O. Digital mammography screening: sensitivity of the programme dependent on breast density. Eur Radiol. 2017;27(7):2744-51.
25. Kerlikowske K, Ichikawa L, Miglioretti DL, Buist DS, Vacek PM, Smith-Bindman R, et al. Longitudinal measurement of clinical mammographic breast density to improve estimation of breast cancer risk. J Natl Cancer Inst. 2007;99(5):386-95.
26. Mandelblatt JS, Stout NK, Schechter CB, van den Broek JJ, Miglioretti DL, Krapcho M, et al. Collaborative Modeling of the Benefits and Harms Associated With Different U.S. Breast Cancer Screening Strategies. Ann Intern Med. 2016;164(4):215-25.
27. Janssen MF, Szende A, Cabases J, Ramos-Goñi JM, Vilagut G, König HH. Population norms for the EQ-5D-3L: a cross-country analysis of population surveys for 20 countries. Eur J Health Econ. 2019;20(2):205-16.
28. Schousboe JT, Kerlikowske K, Loh A, Cummings SR. Personalizing Mammography by Breast Density and Other Risk Factors for Breast Cancer: Analysis of Health Benefits and Cost-Effectiveness. Annals of Internal Medicine. 2011;155(1):10-20.
29. Lidgren M, Wilking N, Jönsson B, Rehnberg C. Health related quality of life in different states of breast cancer. Quality of Life Research. 2007;16(6):1073-81.
30. Sankatsing VD, Heijnsdijk EA, van Luijt PA, van Ravesteyn NT, Fracheboud J, de Koning HJ. Cost-effectiveness of digital mammography screening before the age of 50 in The Netherlands. Int J Cancer. 2015;137(8):1990-9.

31. Epidemiologisches Krebsregister NRW. Brustkrebs - Ergebnisse aus dem Reg.-Bez. Münster 1992-2004. Available from: https://www.landeskrebsregister.nrw/fileadmin/user_upload/dokumente/veroeffentlichungen/ekr_nrw_Brustkrebs_web.pdf.
32. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Won JB. Model transparency and validation: A report of the ISPOR-SMDM modeling good research practices task force-4. Value Health. 2012;15:843-50.
33. Krebs in Deutschland für 2015/2016. 12. Ausgabe. Berlin: Robert Koch-Institut und die Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (Hrsg); 2019. Available from: https://www.krebsdaten.de/Krebs/DE/Content/Publikationen/Krebs_in_Deutschland/kid_2019/krebs_in_deutschland_2019.pdf?__blob=publicationFile.
34. TreeAge Software LLC. TreeAge Pro Healthcare (Version 2021). 2021. Available from: https://www.treeage.com
35. Microsoft Corporation. Microsoft Excel für Microsoft 365 MSO. 2021. Available from: https://www.microsoft.com/
36. Deutsches Mammographie-Screening-Programm, Kooperationsgemeinschaft Mammographie Berlin. Jahresbericht Evaluation 2018. Available from: https://www.mammo-programm.de/download/downloads/berichte/Jahresbericht-Evaluation-2018.pdf2020 [updated November 2020.]
37. O'Mahony JF, Naber SK, Normand C, Sharp L, O'Leary JJ, de Kok IM. Beware of Kinked Frontiers: A Systematic Review of the Choice of Comparator Strategies in Cost-Effectiveness Analyses of Human Papillomavirus Testing in Cervical Screening. Value Health. 2015;18(8):1138-51.
38. Etzioni R, Gulati R, Mallinger L, Mandelblatt J. Influence of Study Features and Methods on Overdiagnosis Estimates in Breast and Prostate Cancer Screening. Annals of Internal Medicine. 2013;158(11):831-8.
39. Lee CI, Etzioni R. Missteps in Current Estimates of Cancer Overdiagnosis. Acad Radiol. 2017;24(2):226-9.
40. Van den Bruel A, Jones C, Yang Y, Oke J, Hewitson P. People's willingness to accept overdetection in cancer screening: population survey. BMJ : British Medical Journal. 2015;350:h980.
41. Sicsic J, Pelletier-Fleury N, Moumjid N. Women's Benefits and Harms Trade-Offs in Breast Cancer Screening: Results from a Discrete-Choice Experiment. Value Health. 2018;21(1):78-88.
42. Brick C, McDowell M, Freeman ALJ. Risk communication in tables versus text: a registered report randomized trial on 'fact boxes'. R Soc Open Sci. 2020;7(3):190876.
43. Gigerenzer G, Kolpatzik K. How new fact boxes are explaining medical risk to millions. Bmj. 2017;357:j2460.
44. McDowell M, Gigerenzer G, Wegwarth O, Rebitschek FG. Effect of Tabular and Icon Fact Box Formats on Comprehension of Benefits and Harms of Prostate Cancer Screening: A Randomized Trial. Med Decis Making. 2019;39(1):41-56.
45. Hinneburg J WC, Ellermann C. Methodenpapier für die Entwicklung von Faktenboxen. Stand: 01.06.20212020 01.12.2021. Available from: https://www.hardingcenter.de/sites/default/files/2021-06/Methodenpapier%20des%20Harding-Zentrums_DE_20210616_final_0.pdf.
46. Hillemanns P, Friese K, Dannecker C, Klug S, Seifert U, Iftner T, et al. Prevention of Cervical Cancer: Guideline of the DGGG and the DKG (S3 Level, AWMF Register Number 015/027OL, December 2017) - Part 2 on Triage, Treatment and Follow-up. Geburtshilfe und Frauenheilkunde. 2019;79(2):160-76.
47. Hillemanns P, Friese K, Dannecker C, Klug S, Seifert U, Iftner T, et al. Prevention of Cervical Cancer: Guideline of the DGGG and the DKG (S3 Level, AWMF Register Number 015/027OL, December 2017) - Part 1 with Introduction, Screening and the Pathology of Cervical Dysplasia. Geburtshilfe und Frauenheilkunde. 2019;79(2):148-59.

9. Appendix

9.1. Externe Validierung: zusätzliche Ergebnisse

Tabelle A. 1 Appendix: Zusätzliche Ergebnisse der externen Validierung

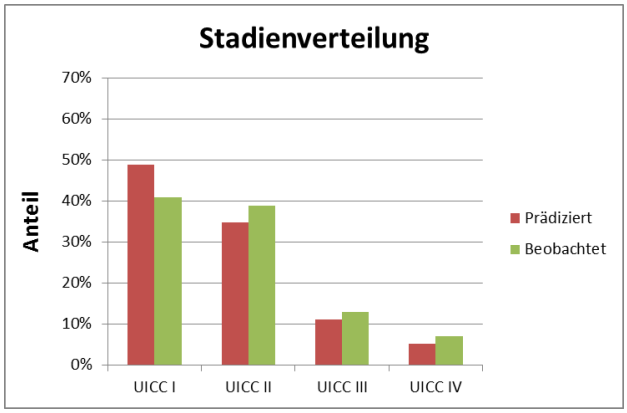
	Modellergebnisse			Beobachtete	Referenz
	30%[1]/50%[2]	42%	50%	Daten	
Stadienverteilung					RKI 2016 (33)
UICC I	0,48	0,47	0,49	0,41	
UICC II	0,35	0,35	0,35	0,39	
UICC III	0,11	0,12	0,11	0,13	
UICC IV	0,05	0,06	0,05	0,07	
UICC II+	0,52	0,53	0,51	0,59	
Stadienverteilung in Frauen Alter 50-69J					
UICC I	0,58	0,57	0,60	0,51	RKI 2016 (33)
UICC II	0,31	0,31	0,29	0,34	
UICC III	0,08	0,09	0,08	0,09	
UICC IV	0,03	0,03	0,03	0,05	
UICC II+	0,42	0,43	0,40	0,48	
Stadienverteilung in Frauen Alter 50-69J					Jahresbericht
DCIS+UICC I		0,824		0,786	2018 (36)
UICC II-IV		0,176		0,21	

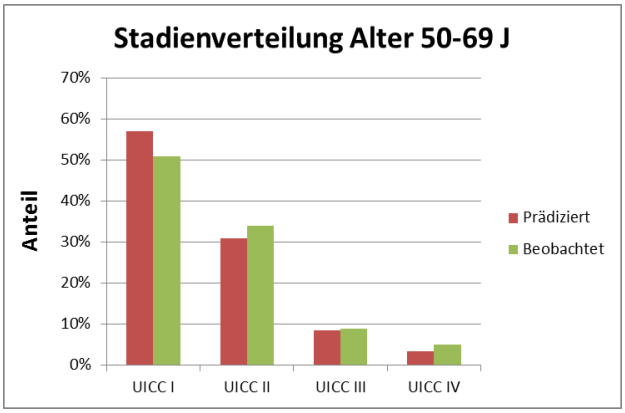
DCIS - ductales Carcinoma in situ; UICC - UICC - Union for International Cancer Control classification; RKI - Robert

[1] im Alter 50-59 Jahre

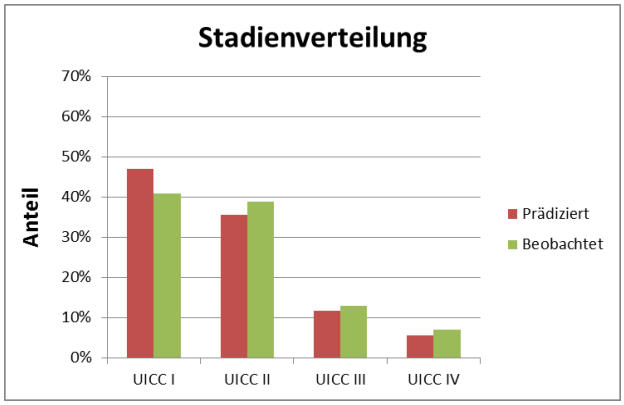
[2] im Alter 60-69 Jahre

Teilnahmerate 42 %





Teilnahmerate 30% im Alter 50-59 Jahre, 50% im Alter 60-69 Jahre



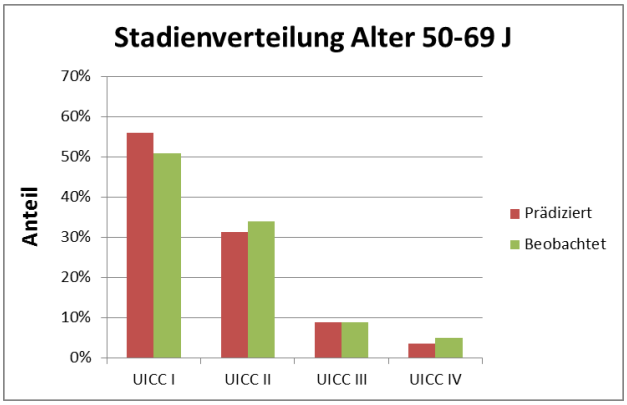


Abbildung A. 1. Appendix: Zusätzliche Ergebnisse der externen Validierung: Stadienverteilung bei unterschiedlichen Teilnahmeraten.

9.2. Sensitivitätsanalysen: zusätzliche Ergebnisse

In den folgenden Tabellen werden zusätzliche Ergebnisse von Sensitivitätsanalysen zu den Lebensqualitätsdaten und Testgüte der Mammographie, die im Modell verwendet wurden, berichtet.

Tabelle A. 2 Appendix: Sensitivitätsanalyse zur Variation der Lebensqualitätsdaten für alle Alterskategorien pro 100 Frauen.

				Lebensqualität in allen Alterskategorien um relative 5%			
	Basisfallanalyse			erhöht		reduziert	
Strategie	**Lebensjahre**	**QALY**	**Rel. Diff. (%)**	**QALY**	**Rel. Diff. (%)**	**QALY**	**Rel. Diff. (%)**
Kein Screening	3847,7	3310,0	13,98	3475,5	9,67	3144,5	18,28
Screening: Alter 50-69 J, 2 J (Ref)	3870,0	3322,5	14,15	3488,7	9,86	3156,3	18,44
Screening: Alter 50-69 J, 2 J; 70-74 J, 3 J	3871,2	3322,9	14,16	3489,1	9,87	3156,8	18,46
Screening: Alter 50-74 J, 2 J	3872,5	3323,1	14,19	3489,3	9,90	3156,9	18,48
Screening: Alter 50-69 J, 2 J; 70-79 J, 3 J	3872,6	3323,2	14,19	3489,5	9,90	3157,0	18,48
Screening: Alter 50-79 J, 2 J	3873,4	3322,9	14,21	3489,1	9,92	3156,7	18,50
Screening: Alter 45-69 J, 2 J	3876,0	3325,6	14,20	3491,9	9,91	3159,3	18,49
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J	3877,1	3325,0	14,24	3491,3	9,95	3158,7	18,53
Screening: Alter 45-74 J, 2 J	3877,7	3326,0	14,23	3492,3	9,94	3159,6	18,52
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-74 J, 3 J	3878,1	3325,4	14,25	3491,7	9,96	3159,0	18,54
Screening: Alter 45-49 J, 1 J; 50-74 J, 2 J	3878,7	3325,4	14,27	3491,7	9,98	3159,0	18,55
Screening: Alter 45-79 J, 2 J	3879,0	3325,8	14,26	3492,2	9,97	3159,4	18,55
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-79 J, 3 J	3879,4	3325,6	14,28	3491,9	9,99	3159,2	18,56
Screening: Alter 45-49 J, 1 J; 50-79 J, 2 J	3880,0	3325,2	14,30	3491,5	10,01	3158,9	18,59

DCIS: ductales Carcinoma in situ, QALY: quality-adjusted life year, Ref.: Referenzstrategie, Rel. Diff.: relative Differenz

Tabelle A. 3 Appendix: Sensitivitätsanalyse zur Variation der Lebensqualitätsdekremente aufgrund von Screening pro 100 Frauen.

				Lebensqualitätsdekrement um relative 5%			
	Basisfallanalyse			erhöht		reduziert	
Strategie	Lebensjahre	QALY	Rel. Diff. (%)	QALY	Rel. Diff. (%)	QALY	Rel. Diff. (%)
Kein Screening	3847,7	3310,0	13,98	3310,0	13,98	3310,0	13,98
Screening: Alter 50-69 J, 2 J (Ref)	3870,0	3322,5	14,15	3322,2	14,16	3322,8	14,14
Screening: Alter 50-69 J, 2 J; 70-74 J, 3 J	3871,2	3322,9	14,16	3322,6	14,17	3323,3	14,15
Screening: Alter 50-74 J, 2 J	3872,5	3323,1	14,19	3322,7	14,20	3323,4	14,18
Screening: Alter 50-69 J, 2 J; 70-79 J, 3 J	3872,6	3323,2	14,19	3322,9	14,20	3323,6	14,18
Screening: Alter 50-79 J, 2 J	3873,4	3322,9	14,21	3322,5	14,22	3323,3	14,20
Screening: Alter 45-69 J, 2 J	3876,0	3325,6	14,20	3325,2	14,21	3326,0	14,19
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J	3877,1	3325,0	14,24	3324,5	14,25	3325,5	14,23
Screening: Alter 45-74 J, 2 J	3877,7	3326,0	14,23	3325,5	14,24	3326,4	14,22
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-74 J, 3 J	3878,1	3325,4	14,25	3324,9	14,27	3325,9	14,24
Screening: Alter 45-49 J, 1 J; 50-74 J, 2 J	3878,7	3325,4	14,27	3324,9	14,28	3325,9	14,25
Screening: Alter 45-79 J, 2 J	3879,0	3325,8	14,26	3325,3	14,28	3326,3	14,25
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-79 J, 3 J	3879,4	3325,6	14,28	3325,0	14,29	3326,1	14,26
Screening: Alter 45-49 J, 1 J; 50-79 J, 2 J	3880,0	3325,2	14,30	3324,6	14,32	3325,8	14,29

DCIS: ductales Carcinoma in situ, QALY: quality-adjusted life year, Ref.: Referenzstrategie, Rel. Diff.: relative Differenz

Tabelle A. 4 Appendix: Sensitivitätsanalyse zur Variation der Lebensqualitätsdekremente aufgrund von positiven Befunden pro 100 Frauen.

				Lebensqualitätsdekrement um relative 5%			
	Basisfallanalyse			erhöht		reduziert	
Strategie	Lebensjahre	QALY	Rel. Diff. (%)	QALY	Rel. Diff. (%)	QALY	Rel. Diff. (%)
Kein Screening	3847,7	3310,0	13,98	3310,0	13,975	3310,0	13,98
Screening: Alter 50-69 J, 2 J (Ref)	3870,0	3322,5	14,15	3322,4	14,150	3322,5	14,15
Screening: Alter 50-69 J, 2 J; 70-74 J, 3 J	3871,2	3322,9	14,16	3322,9	14,163	3323,0	14,16
Screening: Alter 50-74 J, 2 J	3872,5	3323,1	14,19	3323,0	14,189	3323,1	14,19
Screening: Alter 50-69 J, 2 J; 70-79 J, 3 J	3872,6	3323,2	14,19	3323,2	14,188	3323,3	14,19
Screening: Alter 50-79 J, 2 J	3873,4	3322,9	14,21	3322,9	14,212	3323,0	14,21
Screening: Alter 45-69 J, 2 J	3876,0	3325,6	14,20	3325,5	14,203	3325,7	14,20
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J	3877,1	3325,0	14,24	3324,9	14,241	3325,1	14,24
Screening: Alter 45-74 J, 2 J	3877,7	3326,0	14,23	3325,9	14,229	3326,1	14,23
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-74 J, 3 J	3878,1	3325,4	14,25	3325,3	14,254	3325,5	14,25
Screening: Alter 45-49 J, 1 J; 50-74 J, 2 J	3878,7	3325,4	14,27	3325,3	14,267	3325,5	14,26
Screening: Alter 45-79 J, 2 J	3879,0	3325,8	14,26	3325,7	14,264	3325,9	14,26
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-79 J, 3 J	3879,4	3325,6	14,28	3325,5	14,278	3325,6	14,27
Screening: Alter 45-49 J, 1 J; 50-79 J, 2 J	3880,0	3325,2	14,30	3325,1	14,302	3325,3	14,30

DCIS: ductales Carcinoma in situ, QALY: quality-adjusted life year, Ref.: Referenzstrategie, Rel. Diff.: relative Differenz

Tabelle A. 5 Appendix: Sensitivitätsanalyse zur Variation der relativen Lebensqualitätsdekremente aufgrund von entdeckter und behandelter DCIS und Brustkrebs pro 100 Frauen.

				Lebensqualitätsdekrement um relative 5%			
	Basisfallanalyse			erhöht		reduziert	
Strategie	Lebensjahre	QALY	Rel. Diff. (%)	QALY	Rel. Diff. (%)	QALY	Rel. Diff. (%)
Kein Screening	3847,7	3310,0	13,98	3314,7	13,85	3302,3	14,18
Screening: Alter 50-69 J, 2 J (Ref)	3870,0	3322,5	14,15	3327,3	14,02	3305,4	14,59
Screening: Alter 50-69 J, 2 J; 70-74 J, 3 J	3871,2	3322,9	14,16	3327,8	14,04	3305,8	14,61
Screening: Alter 50-74 J, 2 J	3872,5	3323,1	14,19	3327,9	14,06	3305,8	14,63
Screening: Alter 50-69 J, 2 J; 70-79 J, 3 J	3872,6	3323,2	14,19	3328,1	14,06	3305,9	14,64
Screening: Alter 50-79 J, 2 J	3873,4	3322,9	14,21	3327,8	14,09	3305,5	14,66
Screening: Alter 45-69 J, 2 J	3876,0	3325,6	14,20	3330,4	14,08	3306,4	14,70
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J	3877,1	3325,0	14,24	3329,8	14,12	3305,6	14,74
Screening: Alter 45-74 J, 2 J	3877,7	3326,0	14,23	3330,8	14,10	3306,6	14,73
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-74 J, 3 J	3878,1	3325,4	14,25	3330,2	14,13	3305,9	14,75
Screening: Alter 45-49 J, 1 J; 50-74 J, 2 J	3878,7	3325,4	14,27	3330,2	14,14	3305,9	14,77
Screening: Alter 45-79 J, 2 J	3879,0	3325,8	14,26	3330,7	14,14	3306,2	14,77
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-79 J, 3 J	3879,4	3325,6	14,28	3330,4	14,15	3305,9	14,78
Screening: Alter 45-49 J, 1 J; 50-79 J, 2 J	3880,0	3325,2	14,30	3330,0	14,18	3305,5	14,81

DCIS: ductales Carcinoma in situ, QALY: quality-adjusted life year, Ref.: Referenzstrategie, Rel. Diff.: relative Differenz

Tabelle A. 6 Appendix: Sensitivitätsanalyse Szenario 1: Lebenszeitrisiko für diagnostizierten Brustkrebs, DCIS und Tod durch diagnostizierten Brustkrebs bei einer gleichzeitigen Erhöhung der Sensitivität und Reduktion der Spezifität um absolut 5%.

	Basisfallanalyse			Szenario 1 (Sensitivität+5%, Spezifität-5%)					
Strategie	Lebens-zeit Risiko: diagn. BK (%)	Lebens-zeit Risiko: diagn. DCIS (%)	Lebenszeit Risiko: Tod durch BK (%)	Lebenszeit Risiko: diagn. BK (%)	Lebenszeit Risiko: diagn. DCIS (%)	Lebenszeit Risiko: Tod durch BK (%)	Rel. Diff. BK (%)	Rel. Diff. DCIS (%)	Rel. Diff. Tod BK (%)
Kein Screening	11,12	1,01	4,28	11,12	1,01	4,28	0,000	0,000	0,000
Screening: Alter 50-69 J, 2 J (Ref)	10,78	7,65	3,05	10,77	7,71	3,01	-0,013	0,743	-1,374
Screening: Alter 50-69 J, 2 J; 70-74 J, 3 J	10,80	7,67	2,94	10,79	7,73	2,89	-0,004	0,744	-1,588
Screening: Alter 50-74 J, 2 J	10,83	7,71	2,80	10,83	7,76	2,75	0,006	0,746	-1,759
Screening: Alter 50-69 J, 2 J; 70-79 J, 3 J	10,86	7,76	2,76	10,86	7,82	2,71	0,023	0,772	-1,956
Screening: Alter 50-79 J, 2 J	10,89	7,79	2,68	10,89	7,85	2,63	0,026	0,763	-1,966
Screening: Alter 45-69 J, 2 J	10,71	8,60	2,80	10,70	8,67	2,75	-0,018	0,760	-1,817
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J	10,70	8,67	2,77	10,70	8,73	2,72	-0,010	0,626	-1,750
Screening: Alter 45-74 J, 2 J	10,74	8,65	2,64	10,74	8,71	2,58	-0,005	0,769	-2,122
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-74 J, 3 J	10,72	8,70	2,66	10,72	8,76	2,61	0,000	0,638	-1,998
Screening: Alter 45-49 J, 1 J; 50-74 J, 2 J	10,74	8,71	2,60	10,74	8,77	2,55	0,003	0,636	-2,055
Screening: Alter 45-79 J, 2 J	10,83	8,76	2,45	10,83	8,83	2,39	0,026	0,783	-2,489
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-79 J, 3 J	10,79	8,80	2,49	10,79	8,86	2,43	0,030	0,665	-2,403
Screening: Alter 45-49 J, 1 J; 50-79 J, 2 J	10,82	8,83	2,41	10,83	8,88	2,36	0,034	0,651	-2,422

DCIS: ductales Carcinoma in situ, QALY: quality-adjusted life year, Ref.: Referenzstrategie, Rel. Diff.: relative Differenz; BK: Brustkrebs

Tabelle A. 7 Appendix: Sensitivitätsanalyse Szenario 2: Lebenszeitrisiko für diagnostizierten Brustkrebs, DCIS und Tod durch diagnostizierten Brustkrebs bei einer gleichzeitigen Reduktion der Sensitivität und Erhöhung der Spezifität um absolut 5%.

	Basisfall			Szenario 2 (Sensitivität-5%, Spezifität+5%)					
Strategie	Lebens zeit Risiko: diagn. BK (%)	Lebens zeit Risiko: diagn. DCIS (%)	Lebens zeit Risiko: Tod durch BK (%)	Lebens zeit Risiko: diagn. BK (%)	Lebens zeit Risiko: diagn. DCIS (%)	Lebens zeit Risiko: Tod durch BK (%)	Rel. Diff. BK (%)	Rel. Diff. DCIS (%)	Rel. Diff. Tod BK (%)
Kein Screening	11,12	1,01	4,28	11,12	1,01	4,28	0,000	0,000	0,000
Screening: Alter 50-69 J, 2 J (Ref)	10,78	7,65	3,05	10,78	7,59	3,09	0,019	-0,877	1,456
Screening: Alter 50-69 J, 2 J; 70-74 J, 3 J	10,80	7,67	2,94	10,80	7,61	2,99	0,009	-0,871	1,678
Screening: Alter 50-74 J, 2 J	10,83	7,71	2,80	10,83	7,64	2,85	-0,002	-0,870	1,867
Screening: Alter 50-69 J, 2 J; 70-79 J, 3 J	10,86	7,76	2,76	10,85	7,69	2,82	-0,018	-0,898	2,063
Screening: Alter 50-79 J, 2 J	10,89	7,79	2,68	10,89	7,72	2,73	-0,023	-0,888	2,090
Screening: Alter 45-69 J, 2 J	10,71	8,60	2,80	10,71	8,53	2,86	0,024	-0,871	1,927
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J	10,70	8,67	2,77	10,70	8,61	2,82	0,014	-0,710	1,867
Screening: Alter 45-74 J, 2 J	10,74	8,65	2,64	10,74	8,57	2,70	0,010	-0,877	2,253
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-74 J, 3 J	10,72	8,70	2,66	10,72	8,64	2,72	0,004	-0,720	2,124
Screening: Alter 45-49 J, 1 J; 50-74 J, 2 J	10,74	8,71	2,60	10,74	8,65	2,66	0,000	-0,718	2,194
Screening: Alter 45-79 J, 2 J	10,83	8,76	2,45	10,83	8,68	2,52	-0,022	-0,892	2,648
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-79 J, 3 J	10,79	8,80	2,49	10,79	8,73	2,56	-0,027	-0,747	2,550
Screening: Alter 45-49 J, 1 J; 50-79 J, 2 J	10,82	8,83	2,41	10,82	8,76	2,48	-0,032	-0,734	2,590

DCIS: ductales Carcinoma in situ, QALY: quality-adjusted life year, Ref.: Referenzstrategie, Rel. Diff.: relative Differenz, BK: Brustkrebs

Tabelle A. 8 Appendix: Sensitivitätsanalyse: Verhinderte Brustkrebstodesfälle und positive Screeningbefunde bei einer Erhöhung oder Reduktion der falsch-negativen Rate (1-Sensitivität) um relativ 5% des Ausgangswertes in der Basisfallanalyse.

	Basisfallanalyse		Falsch-negative Rate um relativ 5%							
			erhöht				reduziert			
Strategie	Verh. Todesfälle pro 100 Frauen (n)	Positive Befunde pro 100 Frauen (n)	Verh. Todesfälle pro 100 Frauen (n)	Positive Befunde pro 100 Frauen (n)	Rel. Diff. Verh. Todes-fälle (%)	Rel. Diff. Positive Befunde (%)	Verh. Todesfälle pro 100 Frauen (n)	Positive Befunde pro 100 Frauen (n)	Rel. Diff. Verh. Todes-fälle (%)	Rel. Diff. Positive Befunde (%)
Kein Screening	0,00	0	0,00	0	0,00	0,00	0,00	0	0,00	0,00
Screening: Alter 50-69 J, 2 J (Ref)	1,23	87	1,22	98	-0,69	13,37	1,24	98	0,68	13,45
Screening: Alter 50-69 J, 2 J; 70-74 J, 3 J	1,34	93	1,33	105	-0,69	13,08	1,34	105	0,68	13,16
Screening: Alter 50-74 J, 2 J	1,47	105	1,46	118	-0,66	12,31	1,48	118	0,65	12,39
Screening: Alter 50-69 J, 2 J; 70-79 J, 3 J	1,51	104	1,50	117	-0,69	12,76	1,52	117	0,68	12,85
Screening: Alter 50-79 J, 2 J	1,60	115	1,59	129	-0,64	11,92	1,61	129	0,64	12,00
Screening: Alter 45-69 J, 2 J	1,47	123	1,46	136	-0,77	10,88	1,48	137	0,76	10,96
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J	1,51	147	1,50	161	-0,73	9,21	1,52	161	0,72	9,27
Screening: Alter 45-74 J, 2 J	1,64	135	1,62	149	-0,75	10,56	1,65	149	0,74	10,64
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-74 J, 3 J	1,62	153	1,60	168	-0,73	9,21	1,63	168	0,72	9,28
Screening: Alter 45-49 J, 1 J; 50-74 J, 2 J	1,68	159	1,66	174	-0,71	9,06	1,69	174	0,70	9,12
Screening: Alter 45-79 J, 2 J	1,82	151	1,81	167	-0,72	10,29	1,83	167	0,71	10,37
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-79 J, 3 J	1,78	164	1,77	179	-0,73	9,29	1,80	180	0,72	9,35
Screening: Alter 45-49 J, 1 J; 50-79 J, 2 J	1,86	175	1,85	191	-0,68	8,97	1,87	191	0,67	9,03

Rel. Diff.: Relative Differenz; Verh.: Verhinderte

Tabelle A. 9 Appendix: Sensitivitätsanalyse: Verhinderte Brustkrebstodesfälle und positive Screeningbefunde bei einer Erhöhung oder Reduktion der falsch-positiven Rate (1-Sensitivität) um relativ 5% des Ausgangswertes in der Basisfallanalyse.

	Basisfallanalyse		Falsch-positiven Rate um relativ 5%							
			erhöht				reduziert			
Strategie	Verh. Todesfälle pro 100 Frauen (n)	Positive Befunde pro 100 Frauen (n)	Verh. Todesfälle pro 100 Frauen (n)	Positive Befunde pro 100 Frauen (n)	Rel. Diff. Verh. Todes-fälle (%)	Rel. Diff. Positive Befunde (%)	Verh. Todesfälle pro 100 Frauen (n)	Positive Befunde pro 100 Frauen (n)	Rel. Diff. Verh. Todes-fälle (%)	Rel. Diff. Positive Befunde (%)
Kein Screening	0,00	0	0,00	0			0,00	0		
Screening: Alter 50-69 J, 2 J (Ref)	1,23	98	1,23	103	0,00	4,41	1,23	94	0,00	-4,41
Screening: Alter 50-69 J, 2 J; 70-74 J, 3 J	1,34	105	1,34	110	0,00	4,42	1,34	100	0,00	-4,42
Screening: Alter 50-74 J, 2 J	1,47	118	1,47	123	0,00	4,45	1,47	112	0,00	-4,45
Screening: Alter 50-69 J, 2 J; 70-79 J, 3 J	1,51	117	1,51	122	0,00	4,43	1,51	112	0,00	-4,43
Screening: Alter 50-79 J, 2 J	1,60	129	1,60	135	0,00	4,47	1,60	123	0,00	-4,47
Screening: Alter 45-69 J, 2 J	1,47	137	1,47	143	0,00	4,51	1,47	130	0,00	-4,51
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J	1,51	161	1,51	168	0,00	4,58	1,51	154	0,00	-4,58
Screening: Alter 45-74 J, 2 J	1,64	149	1,64	156	0,00	4,52	1,64	143	0,00	-4,52
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-74 J, 3 J	1,62	168	1,62	175	0,00	4,58	1,62	160	0,00	-4,58
Screening: Alter 45-49 J, 1 J; 50-74 J, 2 J	1,68	174	1,68	182	0,00	4,58	1,68	166	0,00	-4,58
Screening: Alter 45-79 J, 2 J	1,82	167	1,82	174	0,00	4,53	1,82	159	0,00	-4,53
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-79 J, 3 J	1,78	180	1,78	188	0,00	4,57	1,78	171	0,00	-4,57
Screening: Alter 45-49 J, 1 J; 50-79 J, 2 J	1,86	191	1,86	200	0,00	4,59	1,86	182	0,00	-4,59

Rel. Diff.: Relative Differenz; Verh.: Verhinderte

A5 Kommentare

A5.1 Bericht im Vergleich zu anderen systematischen Übersichten

Im Rahmen der systematischen Literaturrecherche wurden 8 systematische Übersichten [7,29,33,39,102-105] identifiziert: 7 Übersichtsarbeiten enthalten nach Alterssubgruppen differenzierte Auswertungen bzw. Schlussfolgerungen, und 1 Arbeit [105] betrachtet ausschließlich Frauen unter 50 Jahren. Obwohl sich die systematischen Übersichtsarbeiten (genauso wie die vorliegende Nutzenbewertung) fast ausschließlich mit denselben Primärstudien beschäftigten, kommen die Arbeiten in Bezug auf Frauen unter 50 Jahren und über 69 Jahren zu unterschiedlichen Schlussfolgerungen, sowohl in Bezug auf die Reduktion der brustkrebsspezifischen Mortalität als auch in Bezug auf die Nutzen-Schaden-Abwägung insgesamt. Dies stellen auch Mandrik et al. [29] fest, in deren systematische Übersichtarbeit ausschließlich systematische Übersichtarbeiten einbezogen wurden (Review of Reviews).

Für die Frauen unter 50 Jahren und ab 70 Jahren stellen die Autorinnen und Autoren aller Arbeiten mit eigener metaanalytischer Zusammenfassung übereinstimmend eine statistisch nicht signifikante Reduktion der brustkrebsspezifischen Mortalität für beide Altersgruppen fest [33,39,103,104]. Dabei sind die Ergebnisse bei den Frauen ab 70 Jahren wegen der begrenzten Datenbasis mit größerer Schätzunsicherheit behaftet, d. h., es finden sich breitere Konfidenzintervalle.

In den Schlussfolgerungen werden die Ergebnisse in manchen Arbeiten als Vorteil interpretiert [103,104], in anderen hingegen als nicht ausreichend, um eine Empfehlung für ein allgemeines Screening dieser Altersgruppen [105] bzw. generell für alle Altersgruppen [33] auszusprechen. Aus 2 systematischen Übersichtsarbeiten lässt sich keine klare Positionierung oder Empfehlung für oder gegen ein allgemeines Screening in diesen Altersgruppen ableiten [39,102]. Bei 2 weiteren ist dies wegen des Fokus dieser Arbeiten auch nicht zu erwarten gewesen [7,29]. Mandrik et al. [29], deren Zielsetzung darin bestand, die Variabilität zwischen den systematischen Reviews zu untersuchen und zu erklären, können keinen Faktor ausmachen, der die Unterschiede zwischen den Schlussfolgerungen der 58 eingeschlossenen Arbeiten erklären könnte.

A5.2 Bericht im Vergleich zu internationalen Leitlinien

Die deutsche S3-Leitlinie für die Früherkennung, Diagnostik, Therapie und Nachsorge des Mammakarzinoms [3] empfiehlt die Teilnahme am 2-jährigen nationalen Mammografie-Screening-Programm für Frauen zwischen 50 und 69 Jahren. Für Frauen ab einem Alter von 70 Jahren wird die Teilnahme an Früherkennungsmaßnahmen unter Berücksichtigung des individuellen Risikoprofils und des Gesundheitsstatus sowie einer mehr als 10-jährigen Lebenserwartung empfohlen. Für Frauen im Alter zwischen 40 und 49 Jahren empfiehlt die Leitlinie die Teilnahme nach individueller Risikoanalyse, einer Nutzen-Risiko-Abwägung und unter Berücksichtigung der Präferenzen und der Einwände der Frau. Angaben zum empfohlenen Intervall des Screenings bei Frauen zwischen 40 und 49 Jahren sowie ≥ 70 Jahren

enthält die Leitlinie nicht. Sie verweist allerdings darauf, dass Forschungsbedarf hinsichtlich der Optimierung des Screeningprogramms bezüglich der Altersgrenzen und der Nutzen-Risiko-Relation betreffend eines Screenings zwischen 45 und 49 Jahren sowie ≥ 70 Jahren besteht.

Die Leitlinie der European Commission Initiative on Breast Cancer (ECIBC) empfiehlt für Frauen im Alter zwischen 45 und 49 Jahren und im Alter zwischen 70 und 74 Jahren jeweils die Teilnahme an einem 2- oder 3-jährigem Screening [106]. Als Datenbasis für diese Empfehlungen wurden im Wesentlichen die auch diesem Bericht zugrunde liegenden RCTs herangezogen.

Für Frauen im Alter zwischen 45 und 49 Jahren wurde kein Vorteil eines Screenings in Bezug auf die Endpunkte Gesamtmortalität, brustkrebsspezifische Mortalität, Mastektomien, Konsequenzen aus falsch-positiven Befunden und Überdiagnosen festgestellt. Allerdings war das Risiko für das Auftreten von Tumoren der Stadien IIA und höher bei Frauen mit Screening geringer als für Frauen ohne Screening. Basierend darauf empfehlen die Autorinnen und Autoren ein Screening für Frauen im Alter zwischen 45 und 49 Jahren, was im Wesentlichen der Empfehlung des vorliegenden Berichts entspricht.

Für Frauen im Alter zwischen 70 und 74 Jahren stellt die ECIBC in Übereinstimmung mit dem vorliegenden Bericht keinen statistisch signifikanten Effekt des Screenings hinsichtlich der Endpunkte Gesamtmortalität und brustkrebsspezifische Mortalität fest. Mit der Begründung, dass sich der relative Effekt für diese Altersgruppe nicht wesentlich von dem in anderen Altersgruppen (50–69 Jahre) unterscheidet, leiten die Autorinnen und Autoren an dieser Stelle einen Vorteil des Screenings ab. Der in der Leitlinie angegebene Effektschätzer (relatives Risiko [RR] 0,77; 95 %-KI: [0,54; 1,09]) unterscheidet sich allerdings deutlich von dem der vorliegenden Metaanalyse (IDR 0,91; 95 %-KI: [0,61; 1,36], siehe Abbildung 10). Der Unterschied erklärt sich daraus, dass für die gepoolte Analyse aus 2 Studien aus der Swedish-Two-County-Studie – die diese Analyse wegen der deutlich größeren Fallzahl stark dominiert – für die ECIBC-Leitlinie Daten zum Zeitpunkt 20 Jahre verwendet wurden (siehe Tabelle 36). Im Unterschied dazu wurden in der vorliegenden Metaanalyse Ergebnisse eines früheren Zeitpunkts aus dieser Studie gewählt (10,8 Jahre). Letztlich zeigt sich zu beiden Zeitpunkten – unabhängig davon, welcher gegebenenfalls als aussagekräftiger betrachtet wird – eine geringe Präzision der Schätzung und kein statistisch signifikanter Effekt. Abgesehen davon ist die Übertragung des Screeningeffekts von einer jüngeren Altersgruppe auf Frauen ab 70 Jahren kritisch zu sehen (siehe Kapitel 5). Ferner wurden in der ECIBC-Leitlinie für die Endpunkte Mastektomien, Konsequenzen aus falsch-positiven Befunden und Überdiagnosen aufgrund der nicht ausreichenden Datenlage Beobachtungsstudien herangezogen oder Effekte aus anderen Altersgruppen (50–69 Jahre) extrapoliert. Aufgrund dieser Vorgehensweise empfehlen die Autorinnen und Autoren im Unterschied zum vorliegenden Bericht ein Screening für Frauen im Alter zwischen 70 und 74 Jahren.

Aufgrund der auch in dieser Leitlinie festgestellten teilweise nicht ausreichenden Datenlage hinsichtlich beider Fragestellungen der vorliegenden Bewertung wurden vor allem für die

Empfehlungen zum Untersuchungsintervall des Screenings die Daten von Beobachtungsstudien und Modellierungsstudien herangezogen [106]. Die sowohl in der ECIBC-Leitlinie als auch für den vorliegenden Bericht herangezogenen RCTs haben eine maximale Screeningintervalllänge von 24 Monaten. Die Empfehlung für ein 3-jähriges Screeningintervall resultiert ausschließlich aus Daten von Beobachtungsstudien oder Modellierungen [106]. Insgesamt raten die Autorinnen und Autoren der Leitlinie bei Festlegung des Screeningintervalls für Frauen zwischen 45 und 49 Jahren sowie zwischen 70 und 74 Jahren dazu, die bestehenden nationalen Screeningprogramme, die national teilweise sehr unterschiedlichen Möglichkeiten des Zugangs für Frauen zu den Screeningprogrammen sowie die Präferenzen der Frauen und der Entscheidungsträger zu berücksichtigen.

A5.3 Kritische Reflexion des Vorgehens

A5.3.1 Betrachtete Studientypen

Es wurden für die vorliegende Bewertung RCTs herangezogen sowie Studien, die von den Autorinnen und Autoren als randomisiert bezeichnet, jedoch im Rahmen der vorliegenden Bewertung als quasirandomisierte Studien aufgefasst wurden. Obwohl die so gewonnene Datenbasis nicht ausreichte für eine abschließende Aussage zum Nutzen und Schaden für Fragestellung 2 (Frauen ab 70 Jahren), erschien es in dieser Bewertung nicht sinnvoll, die Einschlusskriterien bezüglich der Studientypen auf Studien niedrigerer Evidenzstufe auszuweiten, z. B. prospektive vergleichende Kohortenstudien. Denn es ist fraglich, ob solche Studien einen Effekt des Screenings auf die Gesamtmortalität oder die brustkrebsspezifische Mortalität in einer Größenordnung gezeigt hätten, die Schlussfolgerungen im Sinne einer Kausalität erlaubt hätten.

Mit der Stockholm-Studie wurde auch eine als quasirandomisiert eingeschätzte Studie eingeschlossen. Der in dieser Studie angewandte aktive Zuteilungsmechanismus, die Zuteilung zu den Gruppen anhand des Geburtsdatums, stellt keine adäquate Randomisierungsmethode dar. Denn bei diesem Vorgehen ist es möglich, Einfluss auf die Zuteilung zu den Gruppen zu nehmen. So könnte – falls den zuteilenden Personen das Geburtsdatum bekannt ist – die potenzielle Teilnehmerin gar nicht erst in die Studie eingeschlossen werden, wenn sie nicht der „gewünschten“ Gruppe zugeteilt würde. Aus diesem Grund wird bei einer solchen Studie von einer geringen Ergebnissicherheit ausgegangen. In der Stockholm-Studie kann jedoch von einem Einschluss (nahezu) vollständiger Alterskohorten der Region ausgegangen werden, was das Risiko der aktiven Beeinflussung der Gruppenzuteilung reduziert. Der Unterschied in der Ergebnissicherheit der Stockholm-Studie zu den 6 Studien mit mäßiger Ergebnissicherheit erschien daher nicht sehr groß. Aus diesem Grund erschien es adäquat, die Stockholm-Studie in der vorliegenden Bewertung zu berücksichtigen.

A6 Literatur

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 [Zugriff: 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. Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft DK, AWMF). Interdisziplinäre S3-Leitlinie für die Früherkennung, Diagnostik, Therapie und Nachsorge des Mammakarzinoms [online]. 2020 [Zugriff: 01.04.2021]. URL: https://www.leitlinienprogramm-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 [Zugriff: 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 [Zugriff: 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. 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.

13. 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.

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

15. 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.

16. 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.

17. 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.

18. 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.

19. 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.

20. 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.

21. 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.

22. 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.

23. Deutsches Mammographie-Screening-Programm. Jahresbericht Evaluation 2019 [online]. 2021 [Zugriff: 03.12.2021]. URL: https://fachservice.mammo-programm.de/download/evaluationsberichte/neu_KOOPMAMMO_Jahresbericht_Eval_2019_20211112_web-Einzelseite.pdf.

24. Statistisches Bundesamt. Tabelle 2.1.2: Die drei häufigsten Todesursachen nach Geschlecht und Altersgruppen [Gesundheitliche Lage der Männer in Deutschland, 2014] [online]. [Zugriff: 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.

25. 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.

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

27. 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.

28. 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.

29. Mandrik O, Zielonke N, Meheus F et al. Systematic reviews as a 'lens of evidence': Determinants of benefits and harms of breast cancer screening. Int J Cancer 2019; 145(4): 994-1006. https://dx.doi.org/10.1002/ijc.32211.

30. 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.

31. 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.

32. 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.

33. 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.

34. 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.

35. 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. S. 55-91.

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

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

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

39. Canelo-Aybar C, Ferreira DS, Ballesteros M et al. Benefits and harms of breast cancer mammography screening for women at average risk of breast cancer: A systematic review for the European Commission Initiative on Breast Cancer. J Med Screen 2021: 969141321993866. https://dx.doi.org/10.1177/0969141321993866.

40. Duffy SW, Smith RA. A note on the design of cancer screening trials. J Med Screen 2015; 22(2): 65-68. https://dx.doi.org/10.1177/0969141315577847.

41. 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.

42. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden: Version 6.0 [online]. 2020 [Zugriff: 21.01.2021]. URL: https://www.iqwig.de/methoden/allgemeine-methoden_version-6-0.pdf.

43. ICH Expert Working Group. ICH harmonised tripartite guideline: structure and content of clinical study reports; E3 [online]. 1995 [Zugriff: 25.11.2020]. URL: https://database.ich.org/sites/default/files/E3_Guideline.pdf.

44. Moher D, Hopewell S, Schulz KF et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c869. https://dx.doi.org/10.1136/bmj.c869.

45. Des Jarlais DC, Lyles C, Crepaz N et al. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Public Health 2004; 94(3): 361-366. https://dx.doi.org/10.2105/ajph.94.3.361.

46. Von Elm E, Altman DG, Egger M et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007; 147(8): 573-577. https://dx.doi.org/10.7326/0003-4819-147-8-200710160-00010.

47. Sutton AJ, Abrams KR, Jones DR et al. Methods for meta-analysis in medical research. Chichester: Wiley; 2000.

48. Veroniki AA, Jackson D, Viechtbauer W et al. Recommendations for quantifying the uncertainty in the summary intervention effect and estimating the between-study heterogeneity variance in random-effects meta-analysis. Cochrane Database Syst Rev 2015; (Suppl 1): 25-27.

49. Kuss O. Statistical methods for meta-analyses including information from studies without any events-add nothing to nothing and succeed nevertheless. Stat Med 2015; 34(7): 1097-1116. https://dx.doi.org/10.1002/sim.6383.

50. Caro JJ, Briggs AH, Siebert U et al. Modeling good research practices--overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--1. Value Health 2012; 15(6): 796-803. https://dx.doi.org/10.1016/j.jval.2012.06.012.

51. Briggs AH, Weinstein MC, Fenwick EA et al. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-6. Med Decis Making 2012; 32(5): 722-732. https://dx.doi.org/10.1177/0272989X12458348.

52. Siebert U. Entscheidungsanalytische Modelle zur Sicherung der Übertragbarkeit internationaler Evidenz von HTA auf den Kontext des deutschen Gesundheitssystems; ein Methodenbeitrag zu HTA [online]. 2005 [Zugriff: 02.06.2021]. URL: https://portal.dimdi.de/de/hta/hta_berichte/hta099_bericht_de.pdf.

53. Eddy DM, Hollingworth W, Caro JJ et al. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. Med Decis Making 2012; 32(5): 733-743. https://dx.doi.org/10.1177/0272989X12454579.

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

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

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

57. 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.

58. 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.

59. 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.

60. 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.

61. 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.

62. 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.

63. 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.

64. Andersson I. Radiographic screening for breast carcinoma. I. Program 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.

65. 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.

66. 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.

67. 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.

68. 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.

69. 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.

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

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

72. 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.

73. 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.

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

75. 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.

76. 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.

77. 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.

78. Shaevitch D, Taghipour S, Miller AB et al. Tumor 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.

79. 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.

80. 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.

81. 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.

82. 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.

83. 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.

84. 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.

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

86. 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.

87. 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.

88. 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.

89. 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.

90. 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 end point. J Med Screen 2002; 9(4): 159-162. https://dx.doi.org/10.1136/jms.9.4.159.

91. 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.

92. 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.

93. 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.

94. 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.

95. 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.

96. 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.

97. 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.

98. University of Oxford. Extending the National Health Service (NHS) Breast Screening Age Range (AgeX) [online]. 2020 [Zugriff: 24.08.2021]. URL: https://ClinicalTrials.gov/show/NCT01081288.

99. University of Oxford. Nationwide cluster-randomised trial of extending the NHS breast screening age range in England: AgeX trial protocol [online]. 2018 [Zugriff: 22.11.2021]. URL: http://www.agex.uk/File+download/18/AgeX+Protocol+-+v5.1+September+2018+%28MREC+Approved%29.pdf.

100. NHS Breast Screening Programme. Completion of randomisation into the AgeX trial, with follow-up continuing throughout the 2020s [online]. 2020 [Zugriff: 22.11.2021]. URL: http://www.agex.uk/File+download/34/AgeX+completion+of+randomisation.pdf.

101. Martín Andrés A, Silva Mato A. Choosing the optimal unconditioned test for comparing two independent proportions. Computat Stat Data Anal 1994; 17(5): 555-574. https://dx.doi.org/10.1016/0167-9473(94)90148-1.

102. Myers ER, Moorman P, Gierisch JM et al. Benefits and Harms of Breast Cancer Screening: A Systematic Review. JAMA 2015; 314(15): 1615-1634. https://dx.doi.org/10.1001/jama.2015.13183.

103. 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 [Zugriff: 11.06.2021]. URL: https://www.ncbi.nlm.nih.gov/books/n/es124/pdf/.

104. Nelson HD, Fu R, Cantor A et al. Effectiveness of Breast Cancer Screening: Systematic Review and Meta-analysis to Update the 2009 U.S. Preventive Services Task Force Recommendation. Ann Intern Med 2016; 164(4): 244-255. https://dx.doi.org/10.7326/m15-0969.

105. van den Ende C, Oordt-Speets AM, Vroling H et al. Benefits and harms of breast cancer screening with mammography in women aged 40-49 years: A systematic review. Int J Cancer 2017; 141(7): 1295-1306. https://dx.doi.org/10.1002/ijc.30794.

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

107. Shea BJ, Grimshaw JM, Wells GA et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007; 7: 10. https://dx.doi.org/10.1186/1471-2288-7-10.

108. 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 [Zugriff: 27.05.2021]. URL: https://training.cochrane.org/handbook/version-6.1/chapter-4-tech-suppl.

109. 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.

A7 Studienlisten

A7.1 Liste der gesichteten systematischen Übersichten

1. Canelo-Aybar C, Ferreira DS, Ballesteros M et al. Benefits and harms of breast cancer mammography screening for women at average risk of breast cancer: A systematic review for the European Commission Initiative on Breast Cancer. J Med Screen 2021: 969141321993866. https://dx.doi.org/10.1177/0969141321993866.

2. 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.

3. Mandrik O, Zielonke N, Meheus F et al. Systematic reviews as a 'lens of evidence': Determinants of benefits and harms of breast cancer screening. Int J Cancer 2019; 145(4): 994-1006. https://dx.doi.org/10.1002/ijc.32211.

4. Myers ER, Moorman P, Gierisch JM et al. Benefits and Harms of Breast Cancer Screening: A Systematic Review. JAMA 2015; 314(15): 1615-1634. https://dx.doi.org/10.1001/jama.2015.13183.

5. 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 [Zugriff: 11.06.2021]. URL: https://www.ncbi.nlm.nih.gov/books/n/es124/pdf/.

6. Nelson HD, Fu R, Cantor A et al. Effectiveness of Breast Cancer Screening: Systematic Review and Meta-analysis to Update the 2009 U.S. Preventive Services Task Force Recommendation. Ann Intern Med 2016; 164(4): 244-255. https://dx.doi.org/10.7326/m15-0969.

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. van den Ende C, Oordt-Speets AM, Vroling H et al. Benefits and harms of breast cancer screening with mammography in women aged 40-49 years: A systematic review. Int J Cancer 2017; 141(7): 1295-1306. https://dx.doi.org/10.1002/ijc.30794.

A7.2 Liste der ausgeschlossenen Publikationen aus der fokussierten Informationsbeschaffung

Nachfolgend aufgeführt finden sich aus der aus der Basis-SÜ Canelo-Aybar 2021 (siehe Abschnitt A3.1.1) identifizierte Publikationen, die relevanten Studien zugeordnet sind, aber die Einschlusskriterien des Berichts nicht erfüllen.

Nicht E1

1. Miller AB, Baines CJ, To T et al. Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years. CMAJ 1992; 147(10): 1477-1488.

2. Miller AB, To T, Baines CJ et al. Canadian National Breast Screening Study-2: 13-year results of a randomized trial in women aged 50-59 years. J Natl Cancer Inst 2000; 92(18): 1490-1499. https://dx.doi.org/10.1093/jnci/92.18.1490.

3. Tabar L, Chen HH, Duffy SW et al. Primary and adjuvant therapy, prognostic factors and survival in 1053 breast cancers diagnosed in a trial of mammography screening. Jpn J Clin Oncol 1999; 29(12): 608-616. https://dx.doi.org/10.1093/jjco/29.12.608.

Nicht E5

1. Shapiro S. Screening: assessment of current studies. Cancer 1994; 74(1 Suppl): 231-238. https://dx.doi.org/10.1002/cncr.2820741306.

A7.3 Liste der ausgeschlossenen Publikationen aus der bibliografischen Recherche mit Ausschlussgründen

Nicht E1

1. Murillo R, Diaz S, Perry F et al. Increased breast cancer screening and downstaging in Colombian women: a randomized trial of opportunistic breast-screening. Int J Cancer 2016; 138(3): 705-713. https://dx.doi.org/10.1002/ijc.29801.

Nicht E2

1. Lee Lin F, Nguyen T, Pedhiwala N et al. A Longitudinal Examination of Stages of Change Model Applied to Mammography Screening. West J Nurs Res 2016; 38(4): 441-458. https://dx.doi.org/10.1177/0193945915618398.

Nicht E4

1. Olsson A, Garne JP, Tengrup I et al. Overweight in relation to tumour size and axillary lymph node involvement in postmenopausal breast cancer patients-differences between women invited to vs. not invited to mammography in a randomized screening trial. Cancer Epidemiol 2009; 33(1): 9-15. https://dx.doi.org/10.1016/j.canep.2009.04.008.

Nicht E5

1. Correction to Lancet Oncol 2015; 16: 1127. Lancet Oncol 2015; 16(9): e427. https://dx.doi.org/10.1016/s1470-2045%2815%2900240-5.

2. Bewley S, Blennerhassett M, Payne M. Cost of extending the NHS breast screening age range in England. BMJ 2019; 365: l1293. https://dx.doi.org/10.1136/bmj.l1293.

3. Boyd NF. The review of randomization in the Canadian National Breast Screening Study; Is the debate over? CMAJ 1997; 156(2): 207-209.

4. Bryant H. The review of randomization in the Canadian National Breast Screening Study; What does the verdict mean for clinicians? CMAJ 1997; 156(2): 213-215.

5. Duffy SW, McCann J, Godward S et al. Some issues in screening for breast and other cancers. J Med Screen 2006; 13(Suppl 1): S28-S34.

6. Duffy SW, Tabar L, Smith RA. Screening for breast cancer with mammography. Lancet 2001; 358(9299): 2166. https://dx.doi.org/10.1016/S0140-6736(01)07194-X.

7. Fletcher SW, Elmore JG. Mammographic screening for breast cancer. N Engl J Med 2003; 348(17): 1672-1680. https://dx.doi.org/10.1056/NEJMcp021804.

8. Garcia-Albeniz X, Hernan MA, Logan RW et al. Continuation of Annual Screening Mammography and Breast Cancer Mortality in Women Older Than 70 Years. Ann Intern Med 2020; 172(6): 381-389. https://dx.doi.org/10.7326/m18-1199.

9. Gøtzsche PC, Olsen O, Ponzone R et al. Beyond randomized controlled trials: organized mammographic screening substantially reduces breast carcinoma mortality (multiple letters). Cancer 2002; 94(2): 578-583.

10. Hakama M, Malila N. Estimation of overdiagnosis by detection to incidence ratio in cancer screening. J Med Screen 2019; 26(3): 135-138. https://dx.doi.org/10.1177/0969141319842910.

11. Jaffe RB. [Commentary on]Effect on mammographic screening from age 40 on breast cancer mortality at 10 years' follow-up: a randomized controlled trial. Obstet Gynecol Surv 2007; 62(5): 314-315.

12. 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 Research &Treatment 2017; 165(1): 9-15. https://dx.doi.org/10.1007/s10549-017-4302-9.

13. Kopans DB. Major failings of trial procedures and quality of screening fatally compromise the results of the Canadian National Breast Screening Studies. J Med Screen 2021: 969141320986186. https://dx.doi.org/10.1177/0969141320986186.

14. Kopans DB, Halpern E, Hulka CA. Mammography screening for breast cancer. Reply to the commentaries. Cancer 1994; 74(4): 1212-1216. https://dx.doi.org/10.1002/1097-0142(19940815)74:4<1212::aid-cncr2820740406>3.0.co;2-5.

15. Le TTT, Adler FR. Is mammography screening beneficial: An individual-based stochastic model for breast cancer incidence and mortality. PLoS Comput Biol 2020; 16(7): e1008036. https://dx.doi.org/10.1371/journal.pcbi.1008036.

16. Lee JH, Zuckerman D. Screening for breast cancer with mammography. Lancet 2001; 358(9299): 2164-2165. https://dx.doi.org/10.1016/s0140-6736(01)07190-2.

17. Miller AB. Screening for breast cancer with mammography. Lancet 2001; 358(9299): 2164. https://dx.doi.org/10.1016/S0140-6736(01)07189-6.

18. Nystrom L, Bjurstam N, Jonsson H et al. Reduced breast cancer mortality after 20+ years of follow-up in the Swedish randomized controlled mammography trials in Malmo, Stockholm, and Goteborg. J Med Screen 2017; 24(1): 34-42. https://dx.doi.org/10.1177/0969141316648987.

19. Ryser MD, Gulati R, Eisenberg MC et al. Identification of the Fraction of Indolent Tumors and Associated Overdiagnosis in Breast Cancer Screening Trials. Am J Epidemiol 2019; 188(1): 197-205. https://dx.doi.org/10.1093/aje/kwy214.

20. Senn S. Screening for breast cancer with mammography. Lancet 2001; 358(9299): 2165. https://dx.doi.org/10.1016/S0140-6736(01)07192-6.

21. Tabar L, Chen TH, Hsu CY et al. Evaluation issues in the Swedish Two-County Trial of breast cancer screening: An historical review. J Med Screen 2017; 24(1): 27-33. https://dx.doi.org/10.1177/0969141316631375.

22. Thornton H. Screening for breast cancer with mammography. Lancet 2001; 358(9299): 2165. https://dx.doi.org/10.1016/S0140-6736(01)07193-8.

23. van den Broek JJ, van Ravesteyn NT, Mandelblatt JS et al. Comparing CISNET Breast Cancer Incidence and Mortality Predictions to Observed Clinical Trial Results of Mammography Screening from Ages 40 to 49. Med Decis Making 2018; 38(1_suppl): 140S-150S. https://dx.doi.org/10.1177/0272989x17718168.

Nicht E7

1. Batur P. In women aged 40 to 48 y, annual mammography vs. usual care reduced breast cancer mortality at 10 but not 23 y. Ann Intern Med 2021; 174(2): JC18. https://dx.doi.org/10.7326/ACPJ202102160-018.

A7.4 Liste der ausgeschlossenen Dokumente aus den durch den G-BA übermittelten Dokumenten mit Ausschlussgründen

Nicht E5

1. Duffy SW, Tabár L, Yen AM et al. Mammography screening reduces rates of advanced and fatal breast cancers: results in 549,091 women. Cancer 2020; 126(13): 2971-2979. https://dx.doi.org/10.1002/cncr.32859.

2. Duffy SW, Tabár L, Yen AM et al. Beneficial effect of consecutive screening mammography examinations on mortality from breast cancer: a prospective study. Radiology 2021; 299(3): 541-547. https://dx.doi.org/10.1148/radiol.2021203935.

3. Hellquist BN, Duffy SW, Abdsaleh S et al. Effectiveness of population-based service screening with mammography for women ages 40 to 49 years: evaluation of the Swedish Mammography Screening in Young Women (SCRY) cohort. Cancer 2011; 117(4): 714-722. https://dx.doi.org/10.1002/cncr.25650.

4. Hellquist BN, Duffy SW, Nystrom L et al. Overdiagnosis in the population-based service screening programme with mammography for women aged 40 to 49 years in Sweden. J Med Screen 2012; 19(1): 14-19. https://dx.doi.org/10.1258/jms.2012.011104.

5. Khil L, Heidrich J, Wellmann I et al. Incidence of advanced-stage breast cancer in regular participants of a mammography screening program: a prospective register-based study. BMC Cancer 2020; 20(1): 174. https://dx.doi.org/10.1186/s12885-020-6646-5.

6. Monticciolo DL, Newell MS, Hendrick RE et al. Breast cancer screening for average-risk women: recommendations from the ACR Commission on breast imaging. J Am Coll Radiol 2017; 14(9): 1137-1143. https://dx.doi.org/10.1016/j.jacr.2017.06.001.

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

8. Tabár L, Yen AM, Wu WY et al. Insights from the breast cancer screening trials: how screening affects the natural history of breast cancer and implications for evaluating service screening programs. Breast J 2015; 21(1): 13-20. https://dx.doi.org/10.1111/tbj.12354.

9. Weigel S, Heindel W, Heidrich J et al. Digital mammography screening: sensitivity of the programme dependent on breast density. Eur Radiol 2017; 27(7): 2744-2751. https://dx.doi.org/10.1007/s00330-016-4636-4.

A8 Bewertung der Qualität der berücksichtigten systematischen Übersichten aus der fokussierten Recherche

Die Bewertung der Qualität der Informationsbeschaffung erfolgte in Anlehnung an AMSTAR, Item 3 [107].

Tabelle 41: Bewertung der Informationsbeschaffung zur systematischen Übersicht Canelo-Aybar 2021

Canelo-Aybar 2021 [39]	**Bewertung**	**Erläuterung**
1. Wurden mindestens 2 verschiedene Informationsquellen durchsucht (z. B. bibliografische Datenbanken und Studienregister)?	ja	1. Bibliografische Datenbanken 2. Sichtung von Referenzlisten 3. Anfragen an ExpertInnen
2. Wurden mindestens 2 verschiedene bibliografische Datenbanken durchsucht?	ja	PubMed, EMBASE and Cochrane Library
3. Werden Suchzeitraum oder Suchdatum genannt?	ja	We searched MEDLINE (19/04/2016), EMBASE (19/04/2016) and CENTRAL (19/04/2016) databases using pre-defined algorithms for both systematic reviews and individual studies.
4. Werden zumindest die wichtigsten Freitextbegriffe oder Schlagwörter der Suchstrategie genannt?	ja	Darstellung der Suchstrategien im Supplement
Frage 1.–4. mit ja: umfassend; mindestens 1 Frage mit nein: fraglich	umfassend	Beobachtungsstudien nur über Sichtung der Referenzlisten von SR identifiziert

A9 Suchstrategien

A9.1 Bibliografische Datenbanken

Suche nach systematischen Übersichten

1. PubMed

Suchoberfläche: NLM

#	Searches
1	mammography screening
2	systematic [sb]
3	#1 and #2
4	#3 AND 2011:2020[DP]

2. Health Technology Assessment Database

Suchoberfläche: INAHTA

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

Suche nach Primärstudien

1. MEDLINE

Suchoberfläche: Ovid

- Ovid MEDLINE(R) 1946 to May 14, 2021

Es wurde folgender Filter übernommen:

- RCT: Lefebvre [108] – 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 tumor* 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).

Suchoberfläche: Ovid

- Ovid MEDLINE(R) Epub Ahead of Print and In-Process, In-Data-Review & Other Non-Indexed Citations May 14, 2021

#	**Searches**
1	(breast* adj3 (cancer* or carcinoma* or tumor* 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

Suchoberfläche: Ovid

- Embase 1974 to 2021 May 14

Es wurde folgender Filter übernommen:

- RCT: Wong [109] – Strategy minimizing difference between sensitivity and specificity

#	**Searches**
1	exp Breast Cancer/
2	(breast* adj3 (cancer* or carcinoma* or tumor* 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

Suchoberfläche: Wiley

- Cochrane Central Register of Controlled Trials: Issue 4 of 12, April 2021

#	**Searches**
#1	[mh "Breast Neoplasms"]
#2	(breast* NEAR/3 (cancer* or carcinoma* or tumor* 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

A9.2 Studienregister

1. ClinicalTrials.gov

Anbieter: U.S. National Institutes of Health

- URL: http://www.clinicaltrials.gov
- Eingabeoberfläche: Expert Search

Suchstrategie
mammography AND screening AND breast

2. International Clinical Trials Registry Platform Search Portal

Anbieter: World Health Organization

- URL: https://trialsearch.who.int
- Eingabeoberfläche: Standard Search

Suchstrategie
mammogr* AND screen* AND breast OR mammography AND screening AND breast