

**Biomarkerbasierte Tests zur Entscheidung für
oder gegen eine adjuvante systemische
Chemotherapie beim primären
Mammakarzinom mit Befall von 1 bis 3
Lymphknoten und bei prämenopausalen
Patientinnen ohne Lymphknotenbefall**

A horizontal bar composed of 18 rectangular segments of varying shades of blue and grey. The word 'BERICHTSPLAN' is centered in white text on a dark blue segment.

BERICHTSPLAN

Projekt: D23-01A Version: 1.0 Stand: 27.11.2023

D23-01B

Impressum

Herausgeber

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema

Biomarkerbasierte Tests zur Entscheidung für oder gegen eine adjuvante systemische Chemotherapie beim primären Mammakarzinom mit Befall von 1 bis 3 Lymphknoten und bei prämenopausalen Patientinnen ohne Lymphknotenbefall

Auftraggeber

Gemeinsamer Bundesausschuss

Datum des Auftrags

12.05.2023 und 21.09.2023

Interne Projektnummer

D23-01A | D23-01B

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

Schlagwörter

Tumormarker – Biologische, Mammatumoren – Mensch, Lymphatische Metastasierung, Nutzenbewertung

Keywords

Tumor Markers – Biological, Breast Neoplasms, Lymphatic Metastasis, Benefit Assessment

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	v
Abkürzungsverzeichnis.....	vi
1 Hintergrund.....	1
2 Fragestellungen	4
2.1 Mammakarzinompatientinnen mit 1 bis 3 befallenen Lymphknoten (Project D23-01A).....	4
2.2 Prämenopausale Mammakarzinompatientinnen ohne Lymphnotenbefall (Project D23-01B)	4
3 Projektverlauf.....	5
4 Methoden	7
4.1 Kriterien für den Einschluss von Studien in die Untersuchung.....	7
4.1.1 Population	7
4.1.2 Prüf- und Vergleichsintervention	8
4.1.3 Patientenrelevante Endpunkte	8
4.1.4 Studientypen	9
4.1.5 Studiendauer	9
4.1.6 Publikationssprache	9
4.1.7 Tabellarische Darstellung der Kriterien für den Studieneinschluss	9
4.1.8 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen	10
4.2 Informationsbeschaffung.....	10
4.2.1 Fokussierte Informationsbeschaffung von systematischen Übersichten	11
4.2.2 Umfassende Informationsbeschaffung von Studien.....	11
4.2.3 Anwendung von Limitierungen auf Datenbankebene	13
4.2.4 Selektion relevanter Studien aus der umfassenden Informationsbeschaffung...	13
4.3 Informationsbewertung und -synthese	14
4.3.1 Darstellung der Einzelstudien.....	14
4.3.2 Bewertung des Verzerrungspotenzials der Ergebnisse.....	14
4.3.3 Metaanalysen	15
4.3.4 Prüfung auf Nichtunterlegenheit	16
4.3.5 Sensitivitätsanalysen	17
4.3.6 Subgruppenmerkmale und andere Effektmodifikatoren.....	17

4.3.7	Aussagen zur Beleglage	18
4.4	Ergänzende Darstellung von Prognosestudien für Mammakarzinompatientinnen mit 1 bis 3 befallenen Lymphknoten (Projekt D23-01A)	19
4.4.1	Kriterien für den Einschluss von Prognosestudien für die ergänzende Darstellung	19
4.4.1.1	Population	20
4.4.1.2	Biomarker	20
4.4.1.3	Patientenrelevante Endpunkte	20
4.4.1.4	Studientypen	20
4.4.1.5	Studiendauer	21
4.4.1.6	Publikationssprache	21
4.4.1.7	Studiengröße	21
4.4.1.8	Tabellarische Darstellung der Kriterien für den Studieneinschluss	21
4.4.1.9	Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen	21
4.4.2	Informationsbeschaffung	22
4.4.3	Bewertung des Verzerrungspotenzials der Ergebnisse	22
4.4.4	Darstellung der Einzelstudien	23
5	Literatur	24
Anhang A	Suchstrategien in bibliografischen Datenbanken	27

Tabellenverzeichnis

	Seite
Tabelle 1: Übersicht über die Kriterien für den Studieneinschluss (RCTs)	10
Tabelle 2: Regelmäßig abgeleitete Aussagesicherheiten für verschiedene Evidenzsituationen beim Vorliegen von Studien derselben qualitativen Ergebnissicherheit	19
Tabelle 3: Übersicht über die Kriterien für den Studieneinschluss im Projekt D23-01A (Prognosestudien).....	21

Abkürzungsverzeichnis

Abkürzung	Bedeutung
ASCO	American Society of Clinical Oncology
AWMF	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e. V.
AHRQ	Agency for Healthcare Research and Quality
DKG	Deutsche Krebsgesellschaft e. V.
DKH	Deutsche Krebshilfe
G-BA	Gemeinsamer Bundesausschuss
HER2/neu; HER2	human epidermal growth factor receptor 2 (Rezeptor des humanen epidermalen Wachstumsfaktors 2)
HTA	Health Technology Assessment
ICD-10	International Statistical Classification of Diseases and Related Health Problems, Revision 10 (Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme, 10. Revision)
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
NICE	National Institute for Health and Care Excellence
PAI-1	Plasminogen-Aktivator-Inhibitor 1
RO-Primäroperation	Resektion eines Tumors im Gesunden
RCT	Randomized controlled Trial (randomisierte kontrollierte Studie)
TNM	TNM Classification of Malignant Tumours (TNM-Klassifikation)
UEs	unerwünschte Ereignisse
uPA	Urokinase-Typ Plasminogen-Aktivator

1 Hintergrund

Der Gemeinsame Bundesausschuss (G-BA) hat am 12.05.2023 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Bewertung von biomarkerbasierten Tests zur Entscheidung für oder gegen eine adjuvante systemische Chemotherapie beim primären Mammakarzinom mit Befall von 1 bis 3 Lymphknoten beauftragt (Projekt D23-01A). Der G-BA hat am 21.09.2023 das IQWiG mit einer zusätzlichen Bewertung von biomarkerbasierten Tests zur Entscheidung für oder gegen eine adjuvante systemische Chemotherapie bei prämenopausalen Patientinnen mit primärem Mammakarzinom ohne Lymphknotenbefall beauftragt (Projekt D23-01B). Die Bearbeitung beider Aufträge wird gemäß der Auftragskonkretisierung vom IQWiG zeitlich synchronisiert und soll in einem gemeinsamen Berichtsplan, Vorbericht und Abschlussbericht vorgelegt werden. In bisherigen, thematisch verwandten Berichten wie der Nutzenbewertung D14-01 [1], dem Addendum D18-01 [2] und dem Rapid Report D19-01 [3] wurden Patientinnen mit positivem und / oder negativem Nodalstatus (Lymphknotenbefall und / oder kein Lymphknotenbefall) gleichermaßen untersucht. Die geplanten Nutzenbewertungen der Projekte D23-01A und D23-01B sehen jeweils die Recherche, Darstellung und Bewertung des aktuellen Wissensstandes zum Einsatz von biomarkerbasierten Tests zur Entscheidung für oder gegen eine adjuvante systemische Chemotherapie bei den oben genannten Populationen von Brustkrebspatientinnen vor. Die detaillierte Projekthistorie zur Beauftragung dieser Projekte findet sich in Kapitel 3.

Definition des Krankheitsbildes

Der Begriff Primärtumor beschreibt den ursprünglichen Ort des ersten Auftretens eines Tumors im Körper. Primärer Brustkrebs (Mammakarzinom, ICD-10 [4] C50) ist eine von der Brustdrüse ausgehende bösartige Neubildung. Bei der lokoregional begrenzten Ersterkrankung beschränkt sich die Krankheit auf einen begrenzten Bereich in der Brustdrüse, gegebenenfalls mit einer Ausdehnung auf wenige regionäre Lymphknoten. Bei lokal fortgeschrittenem Brustkrebs hat sich der Krebs auf große Teile der Brust und / oder auf die Brustwand oder Haut ausgebreitet, eine Metastasierung liegt noch nicht vor. Das Wiederauftreten des Mammakarzinoms nach einem Zeitraum, in dem der Krebs durch eine vorangegangene Behandlung nicht mehr festgestellt werden konnte, wird als Rezidiv bezeichnet. Ein lokales beziehungsweise lokoregionales Rezidiv kann u. a. in der Brust und an der Thoraxwand sowie in den regionalen Lymphknoten der Axilla auftreten. Darüber hinaus können Fernmetastasen auftreten, bei denen sich der Krebs vom Primärtumor auf entfernt liegende Lymphknoten oder in andere Organsysteme wie z. B. das Gehirn ausgebreitet hat [5].

Mammakarzinompatientinnen mit 1 bis 3 befallenen Lymphknoten (Projekt D23-01A)

Im Fokus des Projekts D23-01A stehen Patientinnen mit primärem Hormonrezeptor-positivem, Rezeptor des humanen epidermalen Wachstumsfaktors-2

(HER2)/neu-negativem Mammakarzinom (entspricht den intrinsischen Subtypen Luminal A und Luminal B, HER2/neu-negativ [5]) mit 1 bis 3 befallenen Lymphknoten. Bei diesem Projekt handelt es sich um die Bewertung einer neuen Untersuchungs- und Behandlungsmethode in der vertragsärztlichen Versorgung (siehe Kapitel 3).

Prämenopausale Mammakarzinompatientinnen ohne Lymphnotenbefall (Projekt D23-01B)

Im Fokus des Projekts D23-01B stehen prämenopausale Patientinnen mit primärem Hormonrezeptor-positivem, HER2/neu-negativem Mammakarzinom (entspricht den intrinsischen Subtypen Luminal A und Luminal B, HER2/neu-negativ [5]) ohne Lymphknotenbefall. Bei diesem Projekt handelt es sich um die Überprüfung einer zu Lasten der Krankenkassen erbringbaren vertragsärztlichen Leistung (siehe Kapitel 3).

Definition eines Biomarkers

Der Begriff „Biomarker“ setzt sich aus den Wörtern „biologisch“ und „Marker“ zusammen. Dabei bezieht sich der Begriff auf eine ausgedehnte Unterkategorie von medizinischen Zeichen. Gemeint sind damit Anzeichen des Gesundheitszustands z. B. einer Patientin. Diese Anzeichen können von außen betrachtet sowie genau und reproduzierbar gemessen werden. Medizinische Zeichen sind als Gegensatz zu medizinischen Symptomen zu sehen, da sich die Symptome eher auf z. B. von der Patientin subjektiv wahrgenommene Anzeichen von Krankheit oder Gesundheit beschränken [6].

Bei Biomarkern ist es nötig zwischen diagnostischen, Überwachungs-, pharmakodynamischen, responsiven, prognostischen oder prädiktiven Biomarkern zu unterscheiden. Ein diagnostischer Biomarker identifiziert oder attestiert dabei das Vorhandensein einer spezifischen Krankheit oder ermittelt beispielsweise eine Patientin mit einem Subtyp einer Krankheit. Es ist auch möglich, Biomarker seriell zu messen, um beispielsweise eine Krankheit oder einen medizinischen Zustand zu beurteilen. Solche Biomarker werden als Überwachungsbiomarker bezeichnet. Ein pharmakodynamischer oder responsiver Biomarker ist als eine Konzentrationsänderung als Reaktion auf einen Umweltfaktor oder ein Arzneimittel zu sehen. Ein prognostischer Biomarker wird verwendet, um z. B. die Wahrscheinlichkeit eines Krankheitsverlaufs oder -rückfalls zu ermitteln. Mittels eines prädiktiven Biomarkers kann eine Aussage zum Ansprechen auf eine bestimmte Behandlung getroffen werden [7]. Im Fall der 2 vorliegenden Fragestellungen geht es um den Nutzen von prädiktiven Biomarkern zur Entscheidung für oder gegen eine adjuvante systemische Chemotherapie.

Für die Patientinnen, die der vorliegende Bericht betrachtet, ist unklar, ob ihnen eine adjuvante systemische Chemotherapie, die nach den etablierten klinisch-pathologischen Kriterien angezeigt sein könnte, von Nutzen ist oder ob ihnen eine solche Chemotherapie auf Basis des biomarkerbasierten Testresultats erspart werden kann. Da ein großer Teil dieser

Patientinnen von einer Chemotherapie möglicherweise nicht profitiert, ist es von besonderem Interesse, über einen prädiktiven Marker diejenigen Patientinnen zu identifizieren, die mit hoher Wahrscheinlichkeit einen Nutzen von der Chemotherapie haben bzw. die, die mit hoher Wahrscheinlichkeit nicht davon profitieren werden.

Weiterführende Informationen z. B. zu intrinsischen Mammakarzinom-Subtypen, zur Rolle prädiktiver Marker in der Therapieentscheidung, zur Bestimmung von Biomarkern bei Patientinnen mit Mammakarzinom sowie zur Behandlung der Erkrankung nach einer Primäroperation sind im Abschlussbericht D14-01 [1] genauer beschrieben.

2 Fragestellungen

2.1 Mammakarzinompatientinnen mit 1 bis 3 befallenen Lymphknoten (Projekt D23-01A)

Das Ziel der vorliegenden Untersuchung ist

- die Nutzenbewertung von biomarkerbasierten Tests zur Entscheidung für oder gegen eine adjuvante systemische Chemotherapie im Vergleich zu einer biomarkerunabhängigen oder einer 2. biomarkerbasierten Entscheidungsstrategie

bei Patientinnen mit primärem Hormonrezeptor-positivem, Rezeptor des humanen epidermalen Wachstumsfaktors 2 (HER2)/neu-negativem Mammakarzinom mit Befall von 1 bis 3 Lymphknoten.

2.2 Prämenopausale Mammakarzinompatientinnen ohne Lymphnotenbefall (Projekt D23-01B)

Das Ziel der vorliegenden Untersuchung ist

- die Nutzenbewertung von biomarkerbasierten Tests zur Entscheidung für oder gegen eine adjuvante systemische Chemotherapie im Vergleich zu einer biomarkerunabhängigen oder einer 2. biomarkerbasierten Entscheidungsstrategie

bei prämenopausalen Patientinnen mit primärem Hormonrezeptor-positivem, HER2/neu-negativem Mammakarzinom ohne Lymphnotenbefall.

3 Projektverlauf

Der G-BA hat am 12.05.2023 das IQWiG mit der Bewertung von biomarkerbasierten Tests zur Entscheidung für oder gegen eine adjuvante systemische Chemotherapie beim primären Mammakarzinom mit Befall von 1 bis 3 Lymphknoten beauftragt (Projekt D23-01A).

Der G-BA hat am 21.09.2023 das IQWiG mit der Bewertung von biomarkerbasierten Tests zur Entscheidung für oder gegen eine adjuvante systemische Chemotherapie bei prämenopausalen Patientinnen mit primärem Mammakarzinom ohne Lymphknotenbefall beauftragt (Projekt D23-01B).

Es wurde auf die Einbindung einer oder eines externen Sachverständigen verzichtet. Um im Rahmen der Projektbearbeitung einen Eindruck zu gewinnen, wie Patientinnen (oder ihre Angehörigen) die Erkrankung erleben, welche Therapieerfahrungen sie gemacht haben und was sie sich von einer Behandlung wünschen, wurden am 14.06.2023 Betroffene eingebunden.

Auf Basis des gemeinsamen Berichtsplans werden die vorläufigen Bewertungen vorgenommen. Diese werden in einem gemeinsamen Vorbericht veröffentlicht, der zur Anhörung gestellt wird. Im Anschluss an die Anhörung zum Vorbericht erstellt das IQWiG einen gemeinsamen Abschlussbericht. Dieser Bericht wird an den G-BA übermittelt und grundsätzlich 4 Wochen später auf der Website des IQWiG veröffentlicht. An selber Stelle wird auch die Dokumentation der Anhörung zum Vorbericht veröffentlicht.

Der Zeitplan für alle Arbeitsschritte der Berichterstellung ist auf der Website des IQWiG dargelegt.

Bisherige Untersuchungen des IQWiG zu Biomarkern beim Mammakarzinom und resultierende G-BA-Beschlüsse

Gemäß Auftrag des G-BA vom 03.01.2013 untersuchte das IQWiG im Projekt D13-02 [8] eine Urokinase-Typ-Plasminogen-Aktivator(uPA)- und Plasminogen-Aktivator-Inhibitor-1(PAI-1)-gestützte Strategie zur Entscheidung für oder gegen eine adjuvante Systemtherapie im Vergleich zu einer Entscheidungsstrategie unabhängig von uPA und PAI-1, jeweils bei Patientinnen mit invasivem Mammakarzinom und intermediärem Rezidivrisiko nach RO-Primäroperation hinsichtlich patientenrelevanter Endpunkte [8].

Der G-BA hat am 22.04.2014 das IQWiG mit der Nutzenbewertung D14-01 [1] von biomarkerbasierten Tests zur Entscheidung für oder gegen eine adjuvante systemische Chemotherapie beim primären Hormonrezeptor-positivem, HER2/neu-negativem Mammakarzinom mit und ohne Lymphknotenbefall (0 bis 3 befallene Lymphknoten) beauftragt.

Am 06.07.2018 wurde das IQWiG durch den G-BA daraufhin mit einem Addendum D18-01 [2] zum Abschlussbericht D14-01 beauftragt zur Bewertung von 2 damals aktuell publizierten Studien.

Am 12.07.2019 hat der G-BA das IQWiG mit der Recherche, Darstellung und Bewertung des aktuellen Wissensstandes zu biomarkerbasierten Tests zur Entscheidung für oder gegen eine adjuvante systemische Chemotherapie beim primären Mammakarzinom in Form des Rapid Reports D19-01 [3] beauftragt als Aktualisierung zum Auftrag D14-01.

Unter Berücksichtigung der Berichte des IQWiG hat der G-BA am 20.06.2019 und 15.10.2020 beschlossen [9,10], für die Anwendung bei Patientinnen mit einem primären Hormonrezeptor-positiven, HER2/neu-negativen, nodalnegativen und nicht metastasierten Mammakarzinom unabhängig vom Menopausenstatus die folgenden 4 biomarkerbasierten Tests in die Richtlinie zu Untersuchungs- und Behandlungsmethoden in der vertragsärztlichen Versorgung aufzunehmen: Oncotype DX Breast Recurrence Score, EndoPredict, MammaPrint und Prosigna.

4 Methoden

Die Bewertungen der Projekte D23-01A und D23-01B werden auf Grundlage der Allgemeinen Methoden 7.0 [11] erstellt.

Der Nutzen wird in beiden Projekten jeweils auf Basis von randomisierten kontrollierten Studien (RCTs) bewertet. Außerhalb der Nutzenbewertung werden ausschließlich für das Projekt D23-01A zur Information des G-BA auch Prognosestudien (prospektiv geplante Kohortenstudien) ergänzend dargestellt, siehe Abschnitt 4.4.

Die Zielpopulation in der bisherigen Nutzenbewertung D14-01 [1] und dem Rapid Report D19-01 [3] waren sowohl Patientinnen mit Befall der Lymphknoten (1 bis 3) als auch Patientinnen ohne Befall der Lymphknoten. Somit stellen diese früheren Berichte des IQWiG die Grundlage der Informationsbeschaffung für die aktuellen Nutzenbewertungen der Projekte D23-01A und D23-01B dar. Zugleich finden die projektspezifischen methodischen Festlegungen der vorherigen Biomarker-Berichte [1,3,12] auch hier Berücksichtigung.

In dem Rapid Report D19-01 [3] wurden keine Biomarker identifiziert, die ein so ähnliches molekularbiologisches Testprinzip (bspw. Auswahl und Anzahl der Zielgene) zeigten, dass eine Konkordanz zu erwarten wäre. Dies wurde durch die Ergebnisse dieses Rapid Reports bestätigt, in dem lediglich ein geringes Ausmaß an Konkordanz zwischen den untersuchten Biomarkern festgestellt wurde. Das Ausmaß an Konkordanz lag je nach Biomarkerpaar zwischen 43 % und 74 %. Somit sind für die Projekte D23-01A und D23-01B keine Erkenntnisse aus Konkordanzstudien¹ zu erwarten, die die Übertragung eines ggf. festgestellten Nutzens von einem Biomarker auf einen weiteren ermöglichen könnten. Daher wird in den vorliegenden Nutzenbewertungen jeweils auf die Identifizierung und Darstellung von Konkordanzstudien verzichtet.

4.1 Kriterien für den Einschluss von Studien in die Untersuchung

Die genannten Einschlusskriterien gelten für die beiden Projekte D23-01A und D23-01B gleichermaßen, sofern es nicht anders angegeben wird.

4.1.1 Population

In die Bewertung werden Studien mit Patientinnen mit primärem Hormonrezeptor-positivem, HER2/neu-negativem und nicht metastasiertem Mammakarzinom und 1 bis 3 befallenen Lymphknoten aufgenommen (Projekt D23-01A).

¹ Diese Studien untersuchen die Übereinstimmung der diagnostischen Aussagen der in RCTs validierten Biomarker mit jenen weiterer noch nicht validierter Biomarker [12].

In die Bewertung werden Studien mit prämenopausalen Patientinnen mit primärem Hormonrezeptor-positivem, HER2/neu-negativem Mammakarzinom ohne Lymphknotenbefall aufgenommen (Projekt D23-01B).

In beiden Projekten handelt es sich bei der Population jeweils um Patientinnen nach RO-Primäroperation. Für beide Projekte müssen die Patientinnen mit adjuvanter endokriner Therapie behandelt werden, wie es dem derzeitigen Therapiestandard des Hormonrezeptor-positiven Mammakarzinoms gemäß aktuellen Leitlinienempfehlungen der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e. V. (AWMF), Deutschen Krebsgesellschaft e. V. (DKG) und Deutschen Krebshilfe (DKH) [5] entspricht.

4.1.2 Prüf- und Vergleichsintervention

Die Prüfindervention ist ein biomarkerbasierter Test zur Entscheidung für oder gegen eine adjuvante systemische Chemotherapie.

Als relevante Biomarker in diesem Kontext werden jegliche Tests oder Testkombinationen betrachtet, die molekularbiologische Informationen aus Tumorgewebebeurteilungen beinhalten, die nicht bereits in der aktuellen Risikoklassifikation (Östrogen-, Progesteron-, HER2/neu-Rezeptorstatus, Classification of Malignant Tumours(TNM)-Status etc.) enthalten sind [5]. Der Auswertungsalgorithmus und auch die Grenzwerte der Tests müssen prospektiv festgelegt worden sein. Sowohl Testentwicklung als auch -validierung müssen an einer unabhängigen Stichprobe durchgeführt worden sein.

Die Vergleichsintervention ist eine biomarkerunabhängige Entscheidungsstrategie oder eine 2. biomarkerbasierte Entscheidungsstrategie.

4.1.3 Patientenrelevante Endpunkte

Für die Untersuchung werden folgende patientenrelevante Endpunkte betrachtet:

- Mortalität und Morbidität
 - Gesamtüberleben
 - krankheitsfreies Überleben (die Patientenrelevanz dieses Endpunkts wird anhand der konkreten Operationalisierung in den jeweiligen Studien überprüft; siehe unten)
- gesundheitsbezogene Lebensqualität
- Nebenwirkungen
 - unerwünschte Ereignisse (UEs) sowohl infolge des diagnostischen Tests als auch infolge der sich anschließenden Maßnahmen

Da Patientinnen in der jeweils relevanten Population mit einem kurativen Therapieansatz behandelt werden und in den meisten Fällen davon auszugehen ist, dass sie nach Abschluss der Primärtherapie tatsächlich in Remission sind, stellt ein Rezidiv, das im Rahmen von regulären Nachsorgeuntersuchungen entdeckt wird, ein patientenrelevantes Ereignis dar.

4.1.4 Studientypen

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 4.1.2 genannten Interventionen und alle in 4.1.3 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 Nutzenbewertungen einfließen.

In Anhang A des Berichtsplans von D14-01 [12] wurden bereits geeignete RCT-Studiendesigns wie z. B. das Interaktionsdesign, Strategiedesign oder Anreicherungsdesign detailliert dargestellt, auf welche an dieser Stelle verwiesen wird.

4.1.5 Studiendauer

Eingeschlossen werden Studien mit einer Mindestdauer der Nachbeobachtung von 1 Jahr. Für die Endpunkte Gesamtmortalität und krankheitsfreies Überleben wird eine Nachbeobachtungsdauer von idealerweise 10 Jahren festgelegt. Analog dem Vorgehen im Bericht D14-01 werden für den Fall, dass keine Daten zu einer Nachbeobachtungsdauer von ca. 10 Jahren vorliegen, auch Zwischenauswertungen laufender RCTs von mindestens 5 Jahren dargestellt, um eine Abschätzung der zu erwartenden Ergebnisse für den relevanten 10-Jahres-Zeitraum vorzunehmen.

4.1.6 Publikationssprache

Die Publikation muss in deutscher oder englischer Sprache verfasst sein.

4.1.7 Tabellarische Darstellung der Kriterien für den Studieneinschluss

In der folgenden Tabelle 1 sind die Kriterien aufgelistet, die Studien erfüllen müssen, um in die Bewertung eingeschlossen zu werden.

Tabelle 1: Übersicht über die Kriterien für den Studieneinschluss (RCTs)

Einschlusskriterien	
E1-A	Population Projekt D23-01A: Patientinnen mit primärem Hormonrezeptor-positivem, HER2/neu-negativem und nicht metastasiertem Mammakarzinom und 1 bis 3 befallenen Lymphknoten (siehe auch Abschnitt 4.1.1)
E1-B	Population Projekt D23-01B: prämenopausale Patientinnen mit primärem Hormonrezeptor-positivem, HER2/neu-negativem Mammakarzinom ohne Lymphknotenbefall (siehe auch Abschnitt 4.1.1)
E2	Prüfintervention: biomarkerbasierter Test zur Entscheidung für oder gegen eine adjuvante systemische Chemotherapie (siehe auch Abschnitt 4.1.2)
E3	Vergleichsintervention: jegliche biomarkerunabhängige oder eine 2. biomarkerbasierte Entscheidungsstrategie für oder gegen eine adjuvante systemische Chemotherapie (siehe auch Abschnitt 4.1.2)
E4	patientenrelevante Endpunkte wie in Abschnitt 4.1.3 formuliert
E5	Studientyp: RCTs, die den Nutzen und ggf. den Zusatznutzen einer biomarkerbasierten Therapieentscheidung evaluieren können, also z. B. Studien im Interaktions-, Strategie- oder Anreicherungsdesign beziehungsweise geeignete Modifikationen (siehe auch Abschnitt 4.1.4)
E6	Studiendauer: mindestens 1 Jahr bzw. 10 Jahre (siehe auch Abschnitt 4.1.5)
E7	Publikationssprache: Deutsch oder Englisch
E8	Vollpublikation verfügbar ^a
<p>a. Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [13] oder ein Bericht über die Studie, der den Kriterien des CONSORT-Statements [14] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zur Studienmethodik und zu den Studienergebnissen nicht vertraulich sind.</p> <p>CONSORT: Consolidated Standards of Reporting Trials; E: Einschlusskriterium für RCTs; HER2/neu: Rezeptor des humanen epidermalen Wachstumsfaktors 2; ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; RCT: Randomized controlled Trial (randomisierte kontrollierte Studie)</p>	

4.1.8 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen

Für die Einschlusskriterien E1-A bzw. E1-B (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 E1-A bzw. E1-B, E2 und E3 bei weniger als 80 % erfüllt sind, werden nur dann eingeschlossen, wenn Subgruppenanalysen für Patientinnen vorliegen, die die Einschlusskriterien erfüllen.

4.2 Informationsbeschaffung

Der Abschnitt 4.2 gilt gleichermaßen für beide Projekte D23-01A und D23-01B.

4.2.1 Fokussierte Informationsbeschaffung von systematischen Übersichten

Parallel zur Erstellung des Berichtsplans erfolgte eine Recherche nach systematischen Übersichten in MEDLINE (umfasst auch die Cochrane Database of Systematic Reviews), der International Health Technology Assessment (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 fand am 27.04.2023 statt. Die Suchstrategien für die Suche in bibliografischen Datenbanken finden sich in Anhang A. Die Selektion erfolgte durch 1 Person und wurde anschließend von einer 2. Person überprüft. Diskrepanzen wurden durch Diskussion zwischen beiden aufgelöst.

Die Referenzlisten der identifizierten systematischen Übersichten werden hinsichtlich relevanter Primärstudien gesichtet (siehe Abschnitt 4.2.2).

4.2.2 Umfassende Informationsbeschaffung von Studien

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

Die Informationsbeschaffung knüpft an diejenige der vorangegangenen Nutzenbewertung D14-01 [1] und des Rapid Reports D19-01 [3] an. Dieser Teil der Informationsbeschaffung wird ergänzt um eine systematische Recherche nach relevanten Studien beziehungsweise Dokumenten für den in den vorherigen Berichten nicht abgedeckten Zeitraum.

Da sich die Projekte D23-01A und D23-01B nicht auf bestimmte biomarkerbasierte Tests beschränken, werden alle Publikationen, die in der Bewertung D19-01 das Einschlusskriterium E2 (Prüfintervention bzw. Biomarker) nicht erfüllten, gesichtet.

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
 - European Medicines Agency. EU Clinical Trials Register

Aufgrund der separaten Suchen in ClinicalTrials.gov und im EU Clinical Trials Registry, werden aus dem Suchergebnis des International Clinical Trials Registry Platform Search Portal Einträge dieser beiden Register entfernt.

- Herstelleranfragen

Die Herstelleranfragen ermöglichen den Überblick über alle von Herstellern durchgeführten Studien unabhängig vom Publikationsstatus. Für potenziell relevante Studien aus den Herstellerangaben werden vollständige Studienunterlagen (i. d. R. vollständige Studienberichte) angefordert und in die Bewertungen einbezogen. Alle Informationen zur Methodik und zu Ergebnissen, die in die Nutzenbewertungen eingehen, werden im Bericht des IQWiG veröffentlicht. Angefragt wurden folgende Hersteller:

- Agendia NV
- BioMedica Diagnostics, Inc.
- Biotheranostics, Inc. | A Hologic Company
- Exact Sciences Deutschland GmbH
- Janssen-Cilag GmbH
- Myriad Genetics, Inc.
- NanoString Technologies, Inc.
- NeoGenomics Laboratories, Inc.
- QIAGEN GmbH
- Sekisui Diagnostics LLC
- Veracyte, Inc.

Weitere Informationsquellen und Suchtechniken

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

4.2.3 Anwendung von Limitierungen auf Datenbankebene

Fokussierte Informationsbeschaffung von systematischen Übersichten

Die Suchen wurden auf das Publikationsdatum ab Juni 2019 eingeschränkt. Die MEDLINE-Suchstrategie enthält Limitierungen auf deutsch- und englischsprachige Publikationen sowie auf Humanstudien.

Umfassende Informationsbeschaffung von Studien

Aufgrund der vorangegangenen Nutzenbewertungen werden die Suchen auf das Publikationsdatum ab August 2019 eingeschränkt.

Mit den Suchstrategien werden folgende Publikationstypen ausgeschlossen: Kommentare (MEDLINE) und Editorials (MEDLINE, Embase), da diese i. d. R. keine Studien enthalten [15], sowie Conference Abstract und Conference Review (Embase). Außerdem enthalten die Suchstrategien Limitierungen auf deutsch- und englischsprachige Publikationen sowie auf Humanstudien (MEDLINE, Embase). In der Embase-Suche werden MEDLINE-Datensätze und in der Suche im Cochrane Central Register of Controlled Trials Einträge aus Studienregistern ausgeschlossen.

4.2.4 Selektion relevanter Studien aus der umfassenden Informationsbeschaffung

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

Duplikate werden mit Hilfe des Literaturverwaltungsprogrammes EndNote entfernt. 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 1 und Tabelle 3) 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 gesichtet. Die identifizierten Studien werden dann auf ihre Relevanz geprüft. Der gesamte Prozess wird anschließend von einer 2. Person überprüft. Diskrepanzen werden durch Diskussion zwischen den beiden aufgelöst.

4.3 Informationsbewertung und -synthese

Der Abschnitt 4.3 gilt gleichermaßen für beide Projekte D23-01A und D23-01B.

4.3.1 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 beschrieben.

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 4.3.3 bis 4.3.6 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.

4.3.2 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 Intention-to-treat-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.

Für alle Studien wird zunächst die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung bewertet. Ergibt sich daraus ein endpunktübergreifend hohes Verzerrungspotenzial für die Ergebnisse, so werden die restlichen endpunktübergreifenden Kriterien nur noch in Hinblick darauf überprüft, ob so gravierende Probleme bestehen, dass die Ergebnisse nicht verwertbar sind. Sind die Ergebnisse bezüglich beider Kriterien (Randomisierungssequenz, verdeckte Gruppenzuteilung) potenziell niedrig verzerrt, werden auch die restlichen endpunktübergreifenden Kriterien für die Studien bewertet und das Prüfergebnis für alle Kriterien dokumentiert. Im Falle subjektiver oder subjektiv erhobener Endpunkte ergibt sich aus einer fehlenden Patientenverblindung bereits ein endpunktübergreifend hohes Verzerrungspotenzial, sodass für die Bewertung der weiteren Kriterien des endpunktspezifischen Verzerrungspotenzials der Ergebnisse in diesem Fall wie zuvor beschrieben verfahren wird.

4.3.3 Metaanalysen

Die geschätzten Effekte und Konfidenzintervalle aus den Studien werden mittels Forest Plots dargestellt. Die Heterogenität zwischen den Studien wird mithilfe des statistischen Tests auf Vorliegen von Heterogenität [16] untersucht. Es wird außerdem untersucht, welche Faktoren eine vorhandene Heterogenität möglicherweise verursachen. Dazu zählen methodische

Faktoren (siehe Abschnitt 4.3.5) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.3.6). Falls vorhandene Heterogenität durch solche Faktoren zumindest zum Teil erklärt werden kann, so wird der Studienpool nach diesen Faktoren aufgespalten und die weiteren Berechnungen erfolgen in den getrennten Studienpools.

Wenn das Vorhandensein von Heterogenität nicht ausreichend sicher ausgeschlossen werden kann, wird für die Gesamteffektschätzung ein Modell mit zufälligen Effekten verwendet. Dabei erfolgt die Metaanalyse mithilfe des Knapp-Hartung-Verfahrens. Ist das Konfidenzintervall nach Knapp-Hartung schmäler als das Konfidenzintervall nach DerSimonian-Laird, wird die Knapp-Hartung-Schätzung mit Ad-hoc-Varianzkorrektur weiter betrachtet. Ansonsten wird die Schätzung ohne Ad-hoc-Varianzkorrektur betrachtet. Die Schätzung des Heterogenitätsparameters erfolgt nach Paule-Mandel [17]. Als Ergebnis wird der gemeinsame Effekt mit Konfidenzintervall im Forest Plot dargestellt, wenn die Schätzung informativ ist, also z. B. das Konfidenzintervall des gepoolten Effekts vollständig in der Vereinigung der Konfidenzintervalle der Einzelstudien enthalten ist. Zusätzlich wird ggf. zur Darstellung der Heterogenität das Prädiktionsintervall angegeben.

In Situationen, in denen die Schätzung von Knapp-Hartung nicht informativ ist oder bei statistisch nachgewiesener Heterogenität, ist eine gemeinsame Effektschätzung nicht sinnvoll. In diesen Fällen erfolgt eine qualitative Zusammenfassung der Studienergebnisse. Bei 4 oder mehr Studien wird hierzu das Prädiktionsintervall im Forest Plot mit dargestellt.

Da die Heterogenität im Fall sehr weniger (2 bis 4) Studien nicht verlässlich geschätzt werden kann, wird im Fall von 2 Studien aus pragmatischen Gründen ein Modell mit festem Effekt verwendet, sofern keine deutlichen Gründe dagegensprechen. Im Fall von mehr als 2 Studien muss die Anwendung eines Modells mit festem Effekt explizit begründet werden. Ist die Knapp-Hartung-Schätzung (ggf. mit Ad-hoc-Varianzkorrektur) informativ, so wird das Knapp-Hartung-Konfidenzintervall zur Ableitung einer Nutzensaussage herangezogen, wenn das Ergebnis statistisch signifikant ist. Liegt kein statistisch signifikantes Ergebnis vor, so wird zusätzlich geprüft, ob das DerSimonian-Laird-Verfahren ein statistisch signifikantes Ergebnis liefert. Ist dies der Fall, werden die Studienergebnisse qualitativ zusammengefasst. Ist das Ergebnis nach DerSimonian-Laird nicht statistisch signifikant, wird die Knapp-Hartung-Schätzung (ggf. mit Ad-hoc-Varianzkorrektur) verwendet. Ist die Schätzung nach Knapp und Hartung (ggf. mit Ad-hoc-Varianzkorrektur) nicht informativ, werden die Studienergebnisse qualitativ zusammengefasst.

4.3.4 Prüfung auf Nichtunterlegenheit

Die in der vorliegenden Bewertung untersuchten biomarkerbasierten Tests sollen der Entscheidungsunterstützung für oder gegen eine adjuvante systemische Chemotherapie beim primären Mammakarzinom dienen. Patientinnen, die der Empfehlung eines Biomarkertests

folgen und auf eine Chemotherapie verzichten, setzen sich einem potenziell erhöhten Risiko für ein Rezidiv aus im Vergleich zu Patientinnen, die entsprechend einer biomarkerunabhängigen Strategie mit einer Chemotherapie behandelt werden. Dieses Rezidivrisiko unter einer biomarkerbasierten Behandlungsstrategie sollte nur so geringfügig erhöht sein, dass es durch den Vorteil der vermiedenen Chemotherapie-Nebenwirkungen aufgewogen wird. Methodisch wird diese Anforderung über die Formulierung einer Nichtunterlegenheitsfrage in Hinblick auf das krankheitsfreie Überleben bearbeitet.

Im Folgenden wird die entsprechende projektspezifische Methodik zur Ableitung einer Nutzensaussage dargestellt. Als Schwelle für die Bewertung der Nichtunterlegenheit für die Risikodifferenz beim krankheitsfreien Überleben nach 10 Jahren werden 3 Prozentpunkte (obere Grenze des 95 %-Konfidenzintervalls) festgelegt (vergleiche D18-01). Diese Schwelle von 3 Prozentpunkten entspricht der Größenordnung, die von der American Society of Clinical Oncology (ASCO) als akzeptable Schwelle für die Risikodifferenz genannt wird [18]. Liegen Kaplan-Meier-Schätzungen für das krankheitsfreie Überleben vor, so basieren die Interpretationen der Ergebnisse auf den gegebenenfalls daraus ableitbaren Risikodifferenzen. Alternativ kann das 95 %-Konfidenzintervall für das Hazard Ratio zur Prüfung der Nichtunterlegenheit verwendet werden. Eine Herleitung einer Nichtunterlegenheitsgrenze für das Hazard Ratio hängt vom Basisrisiko der relevanten Population ab. Die Methodik der Herleitung einer Nichtunterlegenheitsgrenze für das Hazard Ratio in Abhängigkeit vom Basisrisiko ist in D18-01 Anhang B [2] erläutert.

4.3.5 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 eingestufteffekt kann z. B. dazu führen, dass nur ein Hinweis auf anstelle eines Belegs für einen (höheren) Nutzen attestiert wird (zur Ableitung von Aussagen zur Beleglage siehe Abschnitt 4.3.7).

4.3.6 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- bzw. 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 Meta-regressionen 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,
- Projekt D23-01A: Menopausenstatus (prä- / postmenopausal),
- prognostische Faktoren für ein Rezidiv (z. B. Tumorgröße) [12].

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 4.3.7).

4.3.7 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 2 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.

Tabelle 2: Regelhaft abgeleitete Aussagesicherheiten für verschiedene Evidenzsituationen beim Vorliegen von Studien derselben qualitativen Ergebnissicherheit

		Anzahl Studien				
		1 (mit statistisch signifikantem Effekt)	≥ 2			
			gemeinsame Effektschätzung sinnvoll	gemeinsame Effektschätzung nicht sinnvoll		
			Metaanalyse statistisch signifikant	konkludente Effekte ^a		
			deutlich	mäßig	nein	
Qualitative Ergebnis- sicherheit	hoch	Hinweis	Beleg	Beleg	Hinweis	–
	mäßig	Anhaltspunkt	Hinweis	Hinweis	Anhaltspunkt	–
	gering	–	Anhaltspunkt	Anhaltspunkt	–	–
a. Unter konkludenten Effekten wird eine Datensituation verstanden, in der es möglich ist, einen Effekt im Sinne der Fragestellung abzuleiten, obwohl eine gemeinsame Effektschätzung nicht sinnvoll möglich ist (siehe Abschnitt 3.1.4 der Allgemeinen Methoden [19]).						

Abschließend erfolgt jeweils separat für die beiden Projekte eine endpunktübergreifende Bewertung des Nutzens. In diese übergreifende Bewertung wird auch die Datenvollständigkeit und die sich daraus möglicherweise ergebende Verzerrung aufgrund von Publication Bias oder Outcome Reporting Bias einbezogen.

4.4 Ergänzende Darstellung von Prognosestudien für Mammakarzinompatientinnen mit 1 bis 3 befallenen Lymphknoten (Projekt D23-01A)

Der G-BA hat am 15.10.2020 beschlossen [10], auf Basis der Darstellung von Prognosestudien im Rapid Report D19-01 neben Oncotype DX Breast Recurrence Score auch die biomarkerbasierten Tests EndoPredict, MammaPrint und Prosigna im Anwendungsgebiet nodalnegativer Patientinnen in die vertragsärztliche Versorgung aufzunehmen. Gemäß Auftrag des G-BA wurden die Daten im Rapid Report D19-01 nicht im Sinne einer Nutzenbewertung ausgewertet. Im vorliegenden Bericht werden für das Projekt D23-01A (nodalpositive Patientinnen) die Daten zu Prognosestudien ebenfalls nur ergänzend und rein deskriptiv dargestellt, und zwar hinsichtlich des Risikos testnegativer Patientinnen, bei Verzicht auf eine adjuvante systemische Chemotherapie im Verlauf von mindestens 5 Jahren ein Rezidiv zu erleiden oder zu versterben.

4.4.1 Kriterien für den Einschluss von Prognosestudien für die ergänzende Darstellung

In folgenden Punkten entspricht die hier angewendete Methodik dem G-BA-Auftrag für D19-01, welche teilweise von dem Bericht D14-01 abgewichen ist. (Die ursprüngliche Methodik aus dem Abschlussbericht D14-01 ist daher zusätzlich in Klammern dargestellt.)

- Prognosestudien werden eingeschlossen, wenn die Nachbeobachtung mindestens 5 Jahre betrug (statt 10 Jahre).

- Bei einer gemeinsamen Auswertung mehrerer Kohorten in einer Prognosestudie führt eine fehlende Berücksichtigung des Studieneffekts in der Modellierung nicht zum Ausschluss.
- Es werden zusätzlich die Endpunkte Gesamtüberleben und rezidivfreies Überleben betrachtet (statt nur krankheitsfreies Überleben).

4.4.1.1 Population

Die Population zu Prognosestudien entspricht der gleichen Population wie in RCTs des Projekts D23-01A und ist in dem Abschnitt 4.1.1 beschrieben.

4.4.1.2 Biomarker

Als relevante Biomarker werden jegliche Tests oder Testkombinationen betrachtet, die molekularbiologische Informationen aus Tumorgewebeproben beinhalten, die nicht bereits in der aktuellen Risikoklassifikation enthalten sind. Darunter wird jeglicher biomarkerbasierte Test verstanden, mit dem man prognostische Aussagen über Gesamtüberleben, krankheitsfreies Überleben oder das Rezidivrisiko der Zielpopulation treffen kann. Der Auswertungsalgorithmus und auch die Grenzwerte des Tests müssen prospektiv festgelegt worden sein. Sowohl Testentwicklung als auch -validierung müssen an einer unabhängigen Stichprobe durchgeführt worden sein.

4.4.1.3 Patientenrelevante Endpunkte

Für die Untersuchung werden folgende patientenrelevante Endpunkte betrachtet:

- Mortalität und Morbidität
 - Gesamtüberleben
 - krankheitsfreies Überleben
 - rezidivfreies Überleben

Dabei sind diese Endpunkte sowohl bei Testpositiven als auch Testnegativen relevant und werden daher vergleichend betrachtet.

4.4.1.4 Studientypen

Es werden prospektiv geplante Kohortenstudien mit nachvollziehbarem Patientenfluss betrachtet, aus denen sich das Gesamtüberleben, krankheitsfreie Überleben oder Risiko für Rezidive ableiten lässt.

Auch geeignet sind prospektiv geplante Prognosestudien, in denen konserviertes Tumormaterial retrospektiv mit einem Biomarker untersucht wurde.

4.4.1.5 Studiendauer

Eingeschlossen werden Prognosestudien mit einer Mindestdauer der Nachbeobachtung von 5 Jahren.

4.4.1.6 Publikationssprache

Die Publikation muss in deutscher oder englischer Sprache verfasst sein.

4.4.1.7 Studiengröße

Aufgrund der nur ergänzenden Darstellung werden Prognosestudien erst ab einer Stichprobengröße der relevanten Population von 200 Patientinnen eingeschlossen.

4.4.1.8 Tabellarische Darstellung der Kriterien für den Studieneinschluss

In der folgenden Tabelle 3 sind die Kriterien aufgelistet, die Studien erfüllen müssen, um in die ergänzende Darstellung eingeschlossen zu werden.

Tabelle 3: Übersicht über die Kriterien für den Studieneinschluss im Projekt D23-01A (Prognosestudien)

Einschlusskriterien	
P1	Population: Patientinnen mit primärem Hormonrezeptor-positivem, HER2/neu-negativem und nicht metastasiertem Mammakarzinom und 1 bis 3 befallenen Lymphknoten (siehe auch Abschnitt 4.4.1.1)
P2	Biomarker: biomarkerbasierter Test gemäß Abschnitt 4.4.1.2
P3	Endpunkte: Gesamtüberleben, krankheitsfreies Überleben und rezidivfreies Überleben (siehe auch Abschnitt 4.4.1.3)
P4	Studientyp: Prognosestudien bzw. prospektiv geplante Kohortenstudien (siehe auch Abschnitt 4.4.1.4)
P5	Studiendauer: mindestens 5 Jahre (siehe auch Abschnitt 4.4.1.5)
P6	Publikationssprache: Deutsch oder Englisch
P7	Studiengröße: mindestens 200 Patientinnen (siehe auch Abschnitt 4.4.1.7)
P8	Vollpublikation verfügbar ^a
<p>a. Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [13] oder ein Bericht über die Studie, der den Kriterien des CONSORT- [14], oder STROBE-Statements [20] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zur Studienmethodik und zu den Studienergebnissen nicht vertraulich sind.</p> <p>CONSORT: Consolidated Standards of Reporting Trials; HER2/neu: Rezeptor des humanen epidermalen Wachstumsfaktors 2; ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; P: Einschlusskriterium für Prognosestudien zur ergänzenden Darstellung; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology</p>	

4.4.1.9 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen

Für das Einschlusskriterium P1 (Population) reicht es aus, wenn bei mindestens 80 % der eingeschlossenen Patientinnen dieses Kriterium erfüllt ist. Liegen für solche Studien Subgruppenanalysen für Patientinnen vor, die das Einschlusskriterium erfüllen, wird auf diese

Analysen zurückgegriffen. Studien, bei denen das Einschlusskriterium P1 bei weniger als 80 % erfüllt ist, werden nur dann eingeschlossen, wenn Subgruppenanalysen für Patientinnen vorliegen, die das Einschlusskriterium erfüllen. Ebenfalls eingeschlossen werden Studien, die zu mindestens 80 % das Einschlusskriterium P2 erfüllen (Biomarker).

4.4.2 Informationsbeschaffung

Die Informationsbeschaffung für Prognosestudien entspricht der für RCTs (siehe Abschnitt 4.2).

Da Prognosestudien im Projekt D23-01A von nachrangiger Bedeutung sind, wurde auf eine Suche nach diesem Studientyp in Studienregistern verzichtet.

4.4.3 Bewertung des Verzerrungspotenzials der Ergebnisse

Die zur Einschätzung einer Aussagesicherheit notwendige qualitative Ergebnissicherheit wird ausschließlich anhand des Anteils der in der Auswertung berücksichtigten Patientinnen beurteilt. Ein Berücksichtigungsanteil von unter 70 % ergibt eine geringe, einer von 70 % bis 90 % eine mäßige qualitative Ergebnissicherheit. Ein Berücksichtigungsanteil von über 90 % wird als hohe qualitative Ergebnissicherheit bewertet.

Bei diesem Anteil handelt es sich um das Verhältnis der in der jeweiligen Auswertung berücksichtigten Patientinnen zu den Patientinnen, die in der Auswertung eigentlich einzuschließen wären;

$\frac{\text{in der Auswertung berücksichtigte Patientinnen}}{\text{in der Auswertung einzuschließende Patientinnen}} \geq \text{z. B. } 70 \% \text{ Berücksichtigungsanteil.}$

In die Berechnung des Anteils der in der Auswertung berücksichtigten Patientinnen (Tumorproben) gehen nur diejenigen Tumorproben ein, die für die Bewertung im Projekt D23-01A relevant sind. Bleiben in der Auswertung Tumorproben von Patientinnen unberücksichtigt, die für die Fragestellung des Projekts D23-01A nicht relevant sind (z. B. Patientinnen, die Chemotherapie erhielten oder Hormonrezeptor-negativ sind), dann hat das keine Konsequenzen für die Berechnung des Anteils berücksichtigter Patientinnen. Ebenso wenig werden fehlende Tumorproben berücksichtigt, bei denen man davon ausgehen kann, dass ihre Daten zufallsbedingt und krankheitsunabhängig fehlen (z. B. wegen einer nicht erfolgreichen Bestimmung der Markerausprägung). Wenn allerdings Tumorproben fehlen oder nicht ausreichend Tumormaterial im Gewebekblock zur Verfügung steht, wird angenommen, dass nur 1 Drittel zufällig fehlt und die verbleibenden 2 Drittel krankheitsbedingt, also nicht zufällig fehlen (z. B. wegen eines sehr kleinen Tumors). Diese werden als Tumorproben von Patientinnen betrachtet, die in der Auswertung einzuschließen wären.

4.4.4 Darstellung der Einzelstudien

Prognosestudien mit einer geringen qualitativen Ergebnissicherheit (siehe Abschnitt 4.4.3) werden nicht ergänzend dargestellt.

Ergebnisse aus Prognosestudien, die für die Fragestellung des Projekts D23-01A relevant sind und die bereits für die Berichte D14-01 und D19-01 extrahiert wurden (Auswertungen zur Population mit Befall von 1 bis 3 Lymphknoten), werden im Projekt D23-01A nicht erneut dargestellt. Solche Ergebnisse sind im Abschlussbericht D14-01 im Abschnitt A4.3.2 in der Tabelle 26 (N1 bei Gnant 2014 [21,22]) und in D19-01 im Abschnitt 5.2.3 in den Tabellen 15, 17, 18, 19, 21, 23, 24 und 25 zu finden.

Alle Ergebnisse aus der aktuellen Informationsbeschaffung werden rein deskriptiv und in verkürzter Form dargestellt.

Die Darstellung orientiert sich dabei an den Tabellen aus den Berichten D14-01 und D19-01. Dies umfasst die Charakteristika des Studiendesigns und der Studienpopulation sowie die Ergebnisse zu den Endpunkten.

Die Daten werden zusätzlich zu einer biomarkerübergreifenden tabellarischen Übersicht zusammengefasst, um den Vergleich der Evidenz zu den einzelnen Biomarkertests zu ermöglichen.

5 Literatur

1. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Biomarkerbasierte Tests zur Entscheidung für oder gegen eine adjuvante systemische Chemotherapie beim primären Mammakarzinom; Abschlussbericht [online]. 2016 [Zugriff: 12.05.2023]. URL: https://www.iqwig.de/download/d14-01_abschlussbericht_biomarker-bei-mammakarzinom.pdf.
2. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Biomarkerbasierte Tests zur Entscheidung für oder gegen eine adjuvante systemische Chemotherapie beim primären Mammakarzinom; Addendum zum Auftrag D14-01 [online]. 2018 [Zugriff: 12.05.2023]. URL: https://www.iqwig.de/download/d18-01_biomarker-bei-mammakarzinom_addendum-zum-auftrag-d14-01_v1-1.pdf.
3. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Biomarkerbasierte Tests zur Entscheidung für oder gegen eine adjuvante systemische Chemotherapie beim primären Mammakarzinom; Rapid Report [online]. 2020 [Zugriff: 12.05.2023]. URL: https://www.iqwig.de/download/d19-01_biomarker-bei-mammakarzinom_rapid-report_v1-1.pdf.
4. Bundesinstitut für Arzneimittel und Medizinprodukte. ICD-10-GM Version 2023; Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme; 10. Revision; German Modification; Version 2023 [online]. 2022 [Zugriff: 11.05.2023]. URL: <https://www.dimdi.de/static/de/klassifikationen/icd/icd-10-gm/kode-suche/htmlgm2023/index.htm>.
5. Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft, Deutsche Krebshilfe, AWMF). Interdisziplinäre S3-Leitlinie für die Diagnostik, Therapie und Nachsorge des Mammakarzinoms; Kurzversion 4.3 [online]. 2021 [Zugriff: 12.05.2023]. URL: https://register.awmf.org/assets/guidelines/032-045Olk_S3_Mammakarzinom_2021-07_1.pdf.
6. Strimbu K, Tavel JA. What are biomarkers? Curr Opin HIV AIDS 2010; 5(6): 463-466. <https://dx.doi.org/10.1097/COH.0b013e32833ed177>.
7. Califf RM. Biomarker definitions and their applications. Exp Biol Med (Maywood) 2018; 243(3): 213-221. <https://dx.doi.org/10.1177/1535370217750088>.
8. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Bestimmung der Antigenexpressionslevel von uPA und PAI-1 beim primären Mammakarzinom mit intermediärem Rückfallrisiko nach R0-Primäroperation; Abschlussbericht [online]. 2014 [Zugriff: 12.05.2023]. URL: https://www.iqwig.de/download/d13-02_abschlussbericht_antigenexpression-bei-mammakarzinom.pdf.

9. Gemeinsamer Bundesausschuss. Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Richtlinie Methoden vertragsärztliche Versorgung: Biomarkerbasierte Tests zur Entscheidung für oder gegen eine adjuvante systemische Chemotherapie beim primären Mammakarzinom [online]. 2019 [Zugriff: 25.05.2023]. URL: https://www.g-ba.de/downloads/39-261-3809/2019-06-20_MVV-RL_Biomarker-Test-prim-Mamma_BAnz.pdf.
10. Gemeinsamer Bundesausschuss. Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Richtlinie Methoden vertragsärztliche Versorgung: Biomarkerbasierte Tests zur Entscheidung für oder gegen eine adjuvante systemische Chemotherapie beim primären Mammakarzinom [online]. 2020 [Zugriff: 04.05.2023]. URL: https://www.g-ba.de/downloads/39-261-4524/2020-10-15_MVV-RL_Biomarker-Tests_BAnz.pdf.
11. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden; Version 7.0 [online]. 2023 [Zugriff: 21.09.2023]. URL: https://www.iqwig.de/methoden/allgemeine-methoden_version-7-0.pdf.
12. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Biomarkerbasierte Tests zur Entscheidung für oder gegen eine adjuvante systemische Chemotherapie beim primären Mammakarzinom; Berichtsplan [online]. 2015 [Zugriff: 12.05.2023]. URL: https://www.iqwig.de/download/d14-01_berichtsplan_biomarker-bei-mammakarzinom.pdf.
13. 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.
14. 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>.
15. Waffenschmidt S, Navarro-Ruan T, Hobson N et al. Development and validation of study filters for identifying controlled non-randomized studies in PubMed and Ovid MEDLINE. *Res Synth Methods* 2020; 11(5): 617-626. <https://dx.doi.org/10.1002/jrsm.1425>.
16. Sutton AJ, Abrams KR, Jones DR et al. *Methods for meta-analysis in medical research*. Chichester: Wiley; 2000.
17. 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.

18. Harris LN, Ismaila N, McShane LM et al. Use of biomarkers to guide decisions on adjuvant systemic therapy for women with early-stage invasive breast cancer: American Society of Clinical Oncology clinical practice guideline. *J Clin Oncol* 2016; 34(10): 1134-1150. <https://dx.doi.org/10.1200/jco.2015.65.2289>.
19. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden; Version 6.1 [online]. 2022 [Zugriff: 27.01.2022]. URL: <https://www.iqwig.de/methoden/allgemeine-methoden-v6-1.pdf>.
20. 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>.
21. Gnant M, Filipits M, Greil R et al. Predicting distant recurrence in receptor-positive breast cancer patients with limited clinicopathological risk: using the PAM50 Risk of Recurrence score in 1478 postmenopausal patients of the ABCSG-8 trial treated with adjuvant endocrine therapy alone. *Ann Oncol* 2014; 25(2): 339-345. <https://dx.doi.org/10.1093/annonc/mdt494>.
22. Filipits M, Nielsen TO, Rudas M et al. The PAM50 risk-of-recurrence score predicts risk for late distant recurrence after endocrine therapy in postmenopausal women with endocrine-responsive early breast cancer. *Clin Cancer Res* 2014; 20(5): 1298-1305. <https://dx.doi.org/10.1158/1078-0432.CCR-13-1845>.
23. 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.

Anhang A Suchstrategien in bibliografischen Datenbanken

Suche nach systematischen Übersichten

1. MEDLINE

Suchoberfläche: Ovid

- Ovid MEDLINE(R) ALL 1946 to April 21, 2023

Es wurde folgender Filter übernommen:

- Systematische Übersicht: Wong [23] – High specificity strategy

#	Searches
1	Gene Expression Profiling/
2	(gene* adj3 expression* adj5 (signature* or profil*)).ab,ti.
3	(immunohistochemi* adj3 (marker* or biomarker* or test* or assay*)).ab,ti.
4	Urokinase-Type Plasminogen Activator/
5	Plasminogen Activator Inhibitor 1/
6	(urokinase adj3 plasminogen activat*).ab,ti.
7	plasminogen activator inhibitor*.ab,ti.
8	(upa or pai-1).ab,ti.
9	Mammaprint*.ab,ti.
10	("70" adj1 gene* adj5 (signatur* or profil*)).ab,ti.
11	(Oncotype* or 21 gene* or recurrence score).ab,ti.
12	Endopredict*.ab,ti.
13	Breast Cancer Index*.ab,ti.
14	(pam50 or pam 50 or 50 gene* or Breast Cancer Prognostic Gene Signature*).ab,ti.
15	(76 gene* adj5 signatur*).ab,ti.
16	(Randox or "Breast Cancer Array").ab,ti.
17	((gene* or genomic*) adj3 grade index).ab,ti.
18	Mammostrat*.ab,ti.
19	((Nottingham prognostic index or NPI) adj1 plus).ab,ti.
20	or/1-19
21	(breast adj3 cancer*).ab,ti.
22	Breast Neoplasms/
23	or/21-22
24	Cochrane database of systematic reviews.jn.
25	(search or MEDLINE or systematic review).tw.
26	meta analysis.pt.
27	or/24-26
28	and/20,23,27
29	28 and (english or german or multilingual or undetermined).lg.
30	29 and 201906:3000.(dt).

2. International HTA Database

Suchoberfläche: INAHTA

#	Searches
1	Gene Expression Profiling[mh]
2	(gene* AND expression* AND (signature* OR profil*))
3	immunohistochemi* AND (marker* OR biomarker* OR test* OR assay*)
4	Urokinase-Type Plasminogen Activator[mh]
5	Plasminogen Activator Inhibitor 1[mh]
6	urokinase AND plasminogen activat*
7	plasminogen activator inhibitor*
8	upa OR pai-1
9	Mammaprint*
10	gene* AND (signatur* or profil*)
11	Oncotype* OR "21 gene*" OR "recurrence score"
12	Endopredict*
13	Breast Cancer Index*
14	pam50 OR "pam 50" OR "50 gene*" OR "Breast Cancer Prognostic Gene Signature*"
15	76 gene* AND signatur*
16	Randex OR "Breast Cancer Array"
17	(gene* OR genomic*) AND grade index
18	Mammostrat*
19	(Nottingham prognostic index OR NPI) AND plus
20	#19 OR #18 OR #17 OR #16 OR #15 OR #14 OR #13 OR #12 OR #11 OR #10 OR #9 OR #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1
21	Breast Neoplasms[mh]
22	breast AND cancer*
23	#22 OR #21
24	#23 AND #20
25	(*) FROM 2019 TO 2023
26	#25 AND #24