

**Bestimmung der
Antigenexpressionslevel von
uPA und PAI-1 beim
primären Mammakarzinom
mit intermediärem
Rückfallrisiko nach R0-
Primäroperation**

**Dokumentation und Würdigung der
Anhörung zum Vorbericht**

Auftrag: D13-02
Version: 1.0
Stand: 22.08.2014

Impressum

Herausgeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema:

Bestimmung der Antigenexpressionslevel von uPA und PAI-1 beim primären Mammakarzinom mit intermediärem Rückfallrisiko nach R0-Primäroperation

Auftraggeber:

Gemeinsamer Bundesausschuss

Datum des Auftrags:

03.01.2013

Interne Auftragsnummer:

D13-02

Anschrift des Herausgebers:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Im Mediapark 8 (KölnTurm)
50670 Köln

Tel.: +49 (0)221 – 35685-0

Fax: +49 (0)221 – 35685-1

E-Mail: berichte@iqwig.de

Internet: www.iqwig.de

Inhaltsverzeichnis

	Seite
Abkürzungsverzeichnis.....	iii
1 Dokumentation der Anhörung	1
2 Würdigung der Anhörung	2
3 Offenlegung potenzieller Interessenkonflikte	3
3.1 Potenzielle Interessenkonflikte von Stellungnehmenden aus Organisationen, Institutionen und Firmen	3
3.2 Potenzielle Interessenkonflikte von weiteren Teilnehmern der wissenschaftlichen Erörterung (externe Sachverständige, externe Reviewer).....	4
4 Dokumentation der wissenschaftlichen Erörterung – Teilnehmerliste, Tagesordnung und Protokoll.....	7
4.1 Teilnehmerliste der wissenschaftlichen Erörterung	7
4.2 Tagesordnung der wissenschaftlichen Erörterung	7
4.3 Protokoll der wissenschaftlichen Erörterung.....	7
4.3.1 Begrüßung und Einleitung.....	7
4.3.2 Tagesordnungspunkt 1: Rolle nicht-randomisierter klinischer Studien	9
4.3.3 Tagesordnungspunkt 2: Rolle einer Per-Protokoll-Analyse.....	22
4.3.4 Tagesordnungspunkt 3: Prognostische Evidenz im Projekt	27
4.3.5 Tagesordnungspunkt 4: Verschiedenes	40
Anhang – Dokumentation der Stellungnahmen	42

Abkürzungsverzeichnis

Abkürzung	Bedeutung
ELISA	enzyme linked immunosorbent assay
G-BA	Gemeinsamer Bundesausschuss
HER2/neu	Rezeptor des humanen epidermalen Wachstumsfaktors 2
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
RCT	randomized controlled trial (randomisierte kontrollierte Studie)
uPA	Urokinase-Typ Plasminogen-Aktivator
PAI-1	Plasminogen-Aktivator-Inhibitor 1

1 Dokumentation der Anhörung

Am 28.03.2014 wurde der Vorbericht in der Version Nr. 1.0 vom 21.03.2014 veröffentlicht und zur Anhörung gestellt. Bis zum 29.04.2014 konnten schriftliche Stellungnahmen eingereicht werden. Insgesamt wurden 3 Stellungnahmen form- und fristgerecht abgegeben. Diese Stellungnahmen sind im Anhang abgebildet.

Unklare Aspekte in den schriftlichen Stellungnahmen wurden in einer wissenschaftlichen Erörterung am 12.06.2014 im IQWiG diskutiert. Das Wortprotokoll der Erörterung befindet sich in Kapitel 4.

Eine Würdigung der in der Anhörung vorgebrachten Aspekte befindet sich im Kapitel „Diskussion“ des Abschlussberichts. Im Abschlussbericht sind darüber hinaus Änderungen, die sich durch die Anhörung ergeben haben, zusammenfassend dargestellt. Der Abschlussbericht ist auf der Website des IQWiG unter www.iqwig.de veröffentlicht.

2 Würdigung der Anhörung

Die im Rahmen der Anhörung vorgebrachten Aspekte wurden hinsichtlich valider wissenschaftlicher Argumente für eine Änderung des Vorberichts überprüft. Die wesentlichen Argumente wurden im Kapitel „Diskussion“ des Abschlussberichts gewürdigt. Der Abschlussbericht ist auf der Website des IQWiG unter www.iqwig.de veröffentlicht.

3 Offenlegung potenzieller Interessenkonflikte

Im Folgenden sind die potenziellen Interessenkonflikte der Stellungnehmenden sowie weiterer Teilnehmer an der wissenschaftlichen Erörterung zusammenfassend dargestellt. Alle Informationen beruhen auf Selbstangabe der einzelnen Personen anhand des „Formblatts zur Offenlegung potenzieller Interessenkonflikte“. Das Formblatt ist unter www.iqwig.de abrufbar. Die in diesem Formblatt aufgeführten Fragen finden sich im Anschluss an diese Zusammenfassung.

3.1 Potenzielle Interessenkonflikte von Stellungnehmenden aus Organisationen, Institutionen und Firmen

Organisation/Institution	Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6
American Diagnostica GmbH	Klöpinger, Martina ¹	ja	nein	nein	nein	nein	nein
	Olenik, Claudia ¹	ja	nein	nein	nein	nein	nein
Deutsche Krankenhausgesellschaft e. V.	Bielecki, Eva ¹	ja	nein	nein	nein	nein	nein
	Rohrig, Angela ¹	ja	nein	nein	nein	nein	nein
Deutsche Krebsgesellschaft e. V.	Bruns, Johannes ¹	nein	ja	ja	nein	nein	nein
	Fehm, Tanja ¹	nein	ja	ja	ja	ja	nein
	Scharl, Anton ¹	ja	ja	ja	nein	nein	nein
	Thomssen, Christoph ¹	nein	ja	ja	nein	ja	nein

¹ Version 11/2013

3.2 Potenzielle Interessenkonflikte von weiteren Teilnehmern der wissenschaftlichen Erörterung (externe Sachverständige, externe Reviewer)

Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6
Kreienberg, Rolf ²	nein	ja	nein	nein	nein	nein
Rutjes, Anne ²	nein	ja	nein	nein	ja	nein

² Version 12/2011

Im „Formblatt zur Offenlegung potenzieller Interessenkonflikte“ wurden folgende 6 Fragen gestellt (Version 12/2011):

Frage 1: Sind oder waren Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor angestellt bei einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere bei einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband?

Frage 2: Beraten Sie oder haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor ein Unternehmen, eine Institution oder einen Interessenverband im Gesundheitswesen, insbesondere ein pharmazeutisches Unternehmen, einen Hersteller von Medizinprodukten oder einen industriellen Interessenverband, direkt oder indirekt beraten?

Frage 3: Haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor direkt oder indirekt von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband, Honorare erhalten für Vorträge, Stellungnahmen oder Artikel?

Frage 4: Haben Sie und / oder hat die Einrichtung¹, die Sie vertreten, abseits einer Anstellung oder Beratungstätigkeit innerhalb des laufenden Jahres und der 3 Kalenderjahre davor von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband, finanzielle Unterstützung für Forschungsaktivitäten, andere wissenschaftliche Leistungen oder Patentanmeldungen erhalten?

Frage 5: Haben Sie und / oder hat die Einrichtung¹, bei der Sie angestellt sind bzw. die Sie vertreten, innerhalb des laufenden Jahres und der 3 Kalenderjahre davor sonstige finanzielle oder geldwerte Zuwendungen (z. B. Ausrüstung, Personal, Unterstützung bei der Ausrichtung einer Veranstaltung, Übernahme von Reisekosten oder Teilnahmegebühren ohne wissenschaftliche Gegenleistung) erhalten von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere von einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband?

Frage 6: Besitzen Sie Aktien, Optionsscheine oder sonstige Geschäftsanteile eines Unternehmens oder einer anderweitigen Institution, insbesondere von einem pharmazeutischen Unternehmen oder einem Hersteller von Medizinprodukten? Besitzen Sie Anteile eines „Branchenfonds“, der auf pharmazeutische Unternehmen oder Hersteller von Medizinprodukten ausgerichtet ist?

¹ Sofern Sie in einer ausgedehnten Institution tätig sind, genügen Angaben zu Ihrer Arbeitseinheit, zum Beispiel Klinikabteilung, Forschungsgruppe etc.

Im „Formblatt zur Offenlegung potenzieller Interessenkonflikte“ wurden folgende 6 Fragen gestellt (Version 11/2013):

Frage 1: Sind oder waren Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor angestellt bei einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere bei einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband?

Frage 2: Beraten Sie oder haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor ein Unternehmen, eine Institution oder einen Interessenverband im Gesundheitswesen, insbesondere ein pharmazeutisches Unternehmen, einen Hersteller von Medizinprodukten oder einen industriellen Interessenverband, direkt oder indirekt beraten?

Frage 3: Haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor direkt oder indirekt von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband, Honorare erhalten für Vorträge, Stellungnahmen oder Artikel?

Frage 4: Haben Sie und / oder hat die Einrichtung¹, für die Sie tätig sind, abseits einer Anstellung oder Beratungstätigkeit innerhalb des laufenden Jahres und der 3 Kalenderjahre davor von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband, finanzielle Unterstützung für Forschungsaktivitäten, andere wissenschaftliche Leistungen oder Patentanmeldungen erhalten?

Frage 5: Haben Sie und / oder hat die Einrichtung¹, für die Sie tätig sind, innerhalb des laufenden Jahres und der 3 Kalenderjahre davor sonstige finanzielle oder geldwerte Zuwendungen (z. B. Ausrüstung, Personal, Unterstützung bei der Ausrichtung einer Veranstaltung, Übernahme von Reisekosten oder Teilnahmegebühren ohne wissenschaftliche Gegenleistung) erhalten von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere von einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband?

Frage 6: Besitzen Sie Aktien, Optionsscheine oder sonstige Geschäftsanteile eines Unternehmens oder einer anderweitigen Institution, insbesondere von einem pharmazeutischen Unternehmen oder einem Hersteller von Medizinprodukten? Besitzen Sie Anteile eines „Branchenfonds“, der auf pharmazeutische Unternehmen oder Hersteller von Medizinprodukten ausgerichtet ist?

¹ Sofern Sie in einer ausgedehnten Institution tätig sind, genügen Angaben zu Ihrer Arbeitseinheit, zum Beispiel Klinikabteilung, Forschungsgruppe etc.

4 Dokumentation der wissenschaftlichen Erörterung – Teilnehmerliste, Tagesordnung und Protokoll

4.1 Teilnehmerliste der wissenschaftlichen Erörterung

Name	Organisation / Institution / Firma / privat
Ralf Bender	IQWiG
Eva Bielecki	Deutsche Krankenhausgesellschaft e. V.
Tanja Fehm	Arbeitsgemeinschaft für Gynäkologische Onkologie e. V.
Martina Klöppinger	American Diagnostica GmbH
Rolf Kreienberg	privat
Martina Markes	IQWiG
Alexander Mensch	IQWiG
Sandra Molnar	IQWiG
Claudia Olenik	American Diagnostica GmbH
Angela Rohrig	Deutsche Krankenhausgesellschaft e. V.
Thilo Rörtgen	Protokollant Sitzungsprotokollarischer Dienst des Landtags NRW
Anne Rutjes	Institut für Sozial- und Präventivmedizin, Universität Bern
Stefan Sauerland	IQWiG
Fülöp Scheibler	IQWiG
Christoph Thomssen	Deutsche Krebsgesellschaft e. V.
Jürgen Windeler	IQWiG

4.2 Tagesordnung der wissenschaftlichen Erörterung

	Begrüßung und Einleitung
TOP 1	Rolle nicht randomisierter klinischer Studien
TOP 2	Rolle einer per-Protokoll-Analyse
TOP 3	Prognostische Evidenz im Projekt
TOP 4	Verschiedenes

4.3 Protokoll der wissenschaftlichen Erörterung

Datum: 12.06.2014, 11:00 bis 12:50 Uhr

Ort: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG),
Im Mediapark 8, 50670 Köln

Moderation: Jürgen Windeler

4.3.1 Begrüßung und Einleitung

Moderator Jürgen Windeler: Ich freue mich, dass Sie gekommen sind. Die Bedingungen sind im Moment etwas schwierig, was den Bahnverkehr in Nordrhein-Westfalen angeht. Und

da sich dann alle Leute auf das Auto stürzen, ist das Autofahren auch nicht viel besser. Aber Sie haben es offenbar alle geschafft und auch zeitgerecht. Ich begrüße Sie ganz herzlich zu der Erörterung zu dem Vorbericht D13-02. Ich freue mich auf einen bestimmt spannenden Austausch. Ich beginne mit meinen notwendigen einleitenden Bemerkungen. Diejenigen, die das nicht zum ersten Mal machen, werden das schon kennen.

Der erste und wichtigste Hinweis ist, dass die Erörterung aufgezeichnet wird. Deswegen stehen die Mikrofone hier; deswegen gibt es einen Protokollanten. Die Aufzeichnung wird veröffentlicht. Indem Sie jetzt hier erschienen sind, sind Sie damit einverstanden, dass das aufgezeichnet und veröffentlicht wird. Falls Sie Ihre Meinung jetzt noch ändern sollten, dann müsste ich Sie bitten, von der Teilnahme an dieser Erörterung Abstand zu nehmen. Aber ich sehe jetzt keinen, der solche Meinungen kundtut. Der daran anschließende Hinweis ist, dass Sie bitte wegen der Aufzeichnung bei jeder und nicht nur wie manchmal üblich bei der ersten Äußerung Ihren Namen sagen, damit der Protokollant und das Tonband das alles zuordnen können. Das waren die ganz wesentlichen Punkte.

Der nächste Punkt ist der, dass ich auf das Ziel der Erörterung hinweise. Wir machen hier keine Anhörung im üblichen Sinne, wie Sie das vielleicht aus der Politik oder anderen Stellungnahmeverfahren kennen, wo man also das, was man schon schriftlich zu Papier gebracht hat, noch einmal vorliest. Wir haben das alles bekommen und gelesen. Sie können davon ausgehen, dass wir das, was wir jetzt nicht zur Sprache bringen, auch meinen, verstanden zu haben. Es geht jetzt hier darum, dass wir, wie Sie auch der Tagesordnung entnehmen können, drei konkrete Punkte zur Sprache bringen wollen, wo wir noch Unklarheiten, Fragen, Diskussionsbedarf haben. Das bedeutet - ich möchte diese Bitte anschließen -, dass Sie hier nicht schriftliche Stellungnahmen verlesen - das ist hier nicht der Platz dafür -, sondern wir dies bitte als wissenschaftliche Erörterung - so ist das auch gedacht - verstehen und uns austauschen über das, was zu klären ist.

Unter dem Punkt „Verschiedenes“, der in der vorliegenden Tagesordnung noch fehlt, den es aber wie üblich gibt, besteht Gelegenheit, von Ihrer Seite Dinge anzusprechen, die in der Tagesordnung nicht explizit erwähnt worden sind.

Das waren die wesentlichen Punkte. Sie haben eine Tagesordnung bekommen. Gibt es zu dieser Tagesordnung Fragen, Streichungswünsche, Ergänzungswünsche? Gibt es zum allgemeinen Vorgehen noch Fragen? - Das sehe ich auch nicht.

Die Veranstaltung ist von 11 bis 14 Uhr terminiert. Ich gehe im Moment davon aus, dass wir eher fertig werden. Sollte sich die Diskussion hinziehen, wird es einen Imbiss zwischendurch und nicht hinterher geben.

Wir steigen nun in die Tagesordnung ein.

4.3.2 Tagesordnungspunkt 1: Rolle nicht-randomisierter klinischer Studien

Martina Markes: In einer Stellungnahme werden die verwendeten Einschlusskriterien für Studientypen mit denen aus anderen Berichten verglichen. Dort werden laut Berichtsplan gegebenenfalls zusätzlich zu den RCTs auch CCTs eingeschlossen. Wir möchten Sie fragen: Aus welchen Gründen sollten in diesen Bericht auch CCTs eingeschlossen werden, obwohl RCTs zu der Fragestellung des Projekts möglich sind und auch von uns gefunden und eingeschlossen wurden? Die Frage war von American Diagnostica.

Claudia Olenik: Bezugnehmend auf die Anmerkungen der Deutschen Krankenhausgesellschaft haben wir gesehen, dass in anderen Projekten sehr wohl auch andere Studiendesigns akzeptiert werden. Uns ist klar mit der Chemo-N0-Studie, die die einzige Studie war, die eingeschlossen wurde, dass es sich hier um einen RCT handelt, allerdings - das haben Sie selber in Ihrem Vorbericht dargestellt aus verschiedenen Gründen - sind in die Studie letztendlich doch sehr viel weniger Patienten eingeflossen, als ursprünglich geplant. Das hatte verschiedene Gründe. Zum einen wurde die Studie ja vorzeitig beendet aufgrund der Änderung der Leitlinienempfehlungen. Es war ethisch nicht mehr vertretbar, dass die Hochrisikogruppe überhaupt keine Therapie erhält. Die neuen Leitlinien haben dann doch einen Einsatz der endokrinen und der Chemotherapie empfohlen. Ein großer Teil der Gruppe, die randomisiert wurde, der Hochrisikogruppe, und zwar ein Drittel, hat vor der Randomisierung beschlossen, sie möchten doch nicht in die Randomisierung hinein. Die Patientinnen haben also selbst für sich die Entscheidung getroffen. Das war vorwiegend die Entscheidung zugunsten einer Nichttherapie, weil zu diesem Zeitpunkt für diese Art der Patienten auch keine Therapie vorgesehen war, also weder endokrin noch Chemotherapie.

Es hat dann im weiteren Verlauf dieser Studie, also bei Patienten, die randomisiert wurden, teilweise Wechsler gegeben. Und was bei dieser Art der Studie war: Es wurde eine ältere Chemotherapie verwendet, die damalige CMF-Therapie, die vielleicht nicht ganz so in der Wirkung ist wie heutige Chemotherapien.

Aus diesem Grunde denken wir, dass bei diesem RCT-Ansatz ... Man sieht ja in der Intention-to-treat-Analyse eine Tendenz. Diese Tendenz wird - das ist ja der zweite Punkt - in der Per-Protokoll-Analyse signifikant noch einmal unterstützt, und man sieht einen Nutzen der eingesetzten Chemotherapie, was Disease-free Survival angeht.

Sie stellen jetzt in Ihrem Bericht dar, dass in dieser Studie kein Nutzen, kein Schaden für die Patientin zu sehen ist. Ich denke, wenn Sie eine einzige Studie einschließen, die zwar ein RCT ist, aber die Ergebnisse dann nicht ausreichend sind, um den Schaden bzw. Nutzen zu belegen, dann frage ich mich, warum nicht weitere Studien herangezogen werden, wie es in anderen Fällen ist. Das war uns eben nicht ganz klar, weil Sie doch für das andere Projekt sagen: Wenn die Ergebnisse ausreichend sind aus den RCTs, um den Schaden oder Nutzen zu belegen, dann werden nach Bedarf auch andere Studiendesigns herangezogen.

Christoph Thomssen: Ich möchte kurz darauf hinweisen: Die Überschrift heißt ja „Bestimmung des Testes zur Abklärung des intermediären Rückfallrisikos nach der Primäroperation“. Das heißt, es geht hier vor allen Dingen um das Zielkriterium Prognose und nicht um das Zielkriterium Prädiktion. Wenn wir das Zielkriterium Prognose nehmen, dann sind auch im Internationalen nach Oxford Levels of Evidence zum Beispiel gar keine randomisierten Studien gefordert, sondern da reichen Inceptive Cohort Studys, also prospektive Kohortenstudien. Das ist für diese Faktoren klar erfüllt mit der Publikation Jänicke 1993 (veröffentlicht in Breast Cancer Research); diese Studie war eine Kohortenstudie, bei der die prognostische Evidenz gezeigt werden konnte, die mit der Chemo-N0-Studie als Feldstudie ja noch einmal bestätigt werden konnte.

Wenn Sie das Zielkriterium Prädiktion nehmen wollen - da ist ja bereits vieles von der Vorrednerin gesagt worden -: Randomisierte Studien wären natürlich ein wunderbarer Gold-Standard. Das ist - das muss man ganz klar sagen - bei solchen komplexen Fragestellungen nicht leicht durchführbar. Das zeigen auch die aktuellen Studien der anderen Tests wie MINDACT oder TAILORx, die ähnliche Schwierigkeiten hatten, bis ein Drittel Verweigerer in der Randomisierung. Deswegen haben sich viele Forscher überlegt, wie man es anders macht, und setzen Level of Evidence niedriger an. Schon dann, wenn wir retrospektiv Datensammlung aus randomisierten Studien neu analysieren, würde das den Level of Evidence 1 erfüllen.

So etwas Ähnliches haben wir letztlich auch noch mal gemacht als eine Art Anwendungsbeobachtung mit der Meta-Analyse zur Frage, ob die Chemotherapie einen prädiktiven Wert hat, Frau Harbeck 2002 im Cancer Research, wo sehr schön aus zwei Kollektiven - Kohorten - von 1988 bis 1997 - München und Rotterdam - herausgearbeitet werden konnten, dass uPA und PAI, wenn es hoch ist, sehr wohl einen starken prädiktiven Wert hat für eine Chemotherapie, hier sogar auch anthracyclinhaltige Therapien dabei.

Moderator Jürgen Windeler: Ich bitte darum, dass wir den TOP 3, wo es um den Stellenwert der prognostischen Daten geht, nicht zu sehr mit dieser Thematik vermischen. Dass das etwas miteinander zu tun hat, ist klar, aber wir sollten uns auf TOP 1 konzentrieren.

Stefan Sauerland: Sie haben sehr klar gesagt, dass es sich um eine recht komplexe Fragestellung handelt, wo also Diagnose und Therapie ineinandergreifen müssen, damit insgesamt ein Nutzen entsteht. Ich glaube, das ist ein wesentliches Argument, warum man sich in diesem Feld jetzt primär mit RCTs beschäftigen sollte. Die grundsätzliche Tatsache, dass RCTs möglich sind, ist ja nicht allein beim uPA/PAI-Test erkennbar, sondern es ist auch bei anderen Tests erkennbar. Auch die Möglichkeit, dass aus bereits in der Vergangenheit durchgeführten randomisierten Studien retrospektive Analysen auf der Basis von archivierten Proben gemacht werden, ist eine Möglichkeit, der wir uns nicht verwehren würden. Das Entscheidende ist jetzt: Was könnte man mit CCTs, also nicht-randomisierten Studien, gewinnen, die jetzt nicht zu prognostischen Aussagen Daten liefern, sondern Studien, die irgendwie einen Vergleich von Diagnose und Therapie in einem Kollektiv, wo der Test

angewendet wird, machen mit einem anderen Kollektiv, wo dieser Test nicht eingesetzt wurde? Das muss dann auch noch in einer entsprechend hohen Qualität durchgeführt werden. Da haben wir große Zweifel, dass ein solches Studiendesign tatsächlich aussagekräftig genug ist, um in diesem Feld wirklich eine adäquate Aussagesicherheit zu ermöglichen.

Christoph Thomssen: Rein theoretisch gebe ich Ihnen in den meisten Sachen recht. Sie müssen nur daran denken:

Erstens. Retrospektiv das durchzuführen, ist immer schwierig, weil wir Frischgewebe brauchen. Das gibt es einfach von großen Studien nicht.

Zweitens. Eine randomisierte Studie, wie es auch in dem Gutachten von Frau Zimmer gefordert ist, im eigentlichen Sinne für die Analyse des prädiktiven Wertes müsste ja das Design „Faktor vorhanden“ oder „Faktor nicht vorhanden“ haben und in jedem der beiden Arme noch mal randomisieren. Das haben wir uns ja sehr genau überlegt bei dem Design damals. Das Problem ist: In der Kohortenstudie (Pilotstudie zu Chemo-N0) lag das Risiko für die Patientinnen, nach fünf Jahren ein Rezidiv zu bekommen, bei unter 5 %. So einer Patientin können Sie schlecht vorschreiben, ich mache eventuell randomisiert Chemotherapie. Das geht aus ethischen Gründen nicht. Statistisch ist es auch schwierig, weil Sie dann extrem große Zahlen brauchen. Von daher haben wir uns damals beschränkt und gesagt: Wir können aus praktikablen Gründen tatsächlich nur in dem Hochrisikoarm nach Chemotherapie randomisieren. Sie müssen daran denken, 1993 war Chemotherapie bei nodalnegativem Mammakarzinom nicht vorgesehen.

Stefan Sauerland: Wir haben diesen TOP jetzt bezogen auf die nicht-randomisierten Studien. Die Frage ist: Wie könnte so eine Studie aussehen, die nicht-randomisiert durchgeführt wird, die aber in den wesentlichen Eigenschaften diesem von Ihnen skizzierten Design entspräche und die tatsächlich auch in ihren Vergleichbarkeiten zwischen den beiden Gruppen so sichere Daten liefern könnte, dass man tatsächlich daraus etwas ableiten könnte? Es ist jetzt nicht einfach so, dass man hier eine Klinik nehmen kann, München, und eine andere Klinik nehmen kann, Rotterdam. In einer Klinik wird der Test durchgeführt und man guckt mal, was passiert mit den Frauen, und in der anderen Klinik wird dieser Test nicht durchgeführt. Dann entstehen natürlich enorme Verzerrungspotentiale, alleine dadurch, dass die Frauen unterschiedlich in die Klinik hineinkommen, in dem gesamten Land ein unterschiedliches diagnostisches Vorgehen, vielleicht ein unterschiedliches Screening vorhanden ist, sodass sich diese Effekte auch in einer adjustierten Analyse nicht mehr werden korrigieren lassen. Da habe ich große Sorge, gerade wenn es hier um kleine Unterschiede geht, dass man solche kleinen Unterschiede nicht mit einer adäquaten Sicherheit wird charakterisieren können, wenn man denn auf der anderen Seite diese ganzen Verzerrungspotentiale mit in der Studie drin hat.

Christoph Thomssen: Um auf die Analyse Rotterdam zurückzukommen: Sowohl in München als auch in Rotterdam ist der Test an Kohorten durchgeführt worden. Es waren Kohorten von Patienten von 1988 bis 1999, weil in beiden Zentren sich jemand für dieses

Thema interessiert hat. In Rotterdam war der Kollege John Foekens verantwortlich. Der hatte für die Estrogen-Rezeptor-Analyse damals die Zytosole sammeln müssen. Daher hatte er das Material als Kohorte. Deswegen sind das zwei sehr wertvolle Kollektive. Sowohl das Münchener Kollektiv als auch das Rotterdamer Kollektiv sind echte Patientenkohorten. Diese beiden sind zusammengeworfen worden. Da ist einfach untersucht worden: Wie haben die sich verhalten im Verlauf unter Chemotherapie oder Nicht-Chemotherapie? Dies ist natürlich nicht im Rahmen einer randomisierten Studie entschieden worden, sondern nach klinischem Gusto. Da - das hatte ich als Zeichnung in meinen schriftlichen Einträgen dargestellt - hat sich wirklich sehr konsistent für alle Subgruppen herausgestellt, dass der hohe uPA-Wert einen klaren prädiktiven Wert hat für den Nutzen einer Chemotherapie. Die Hazard Ratio lag bei etwa 0,5, was ja konsistent ist auch mit den Daten der Chemo-N0-Analyse, während für uPA-niedrig oder für Homotherapie zeigte sich die Hazard Ratio nur bei 0,8, also deutlich schlechter. Das ist eine Art Anwendungsbeobachtung.

Fueloep Scheibler: Wenn ist es richtig verstanden habe, ist es so, dass in beiden Kollektiven sowohl Patientinnen eine Chemotherapie erhalten haben als auch andere Patientinnen, die die nicht erhalten haben. Es ist nicht so, dass in Rotterdam keine Chemotherapie gemacht wurde.

Christoph Thomssen: Es ist ein gemischtes Kollektiv. 1988 bis 1999 war es ja so, dass bei mehr als vier Lymphknoten teilweise gesagt wurde, machen wir lieber keine Chemotherapie, weil die Bonadonna-Daten hatten damals gezeigt, dass Chemotherapie bei mehr Lymphknoten nichts nutzt. Von daher ist es ein gemischtes Kollektiv von Patienten, die nach klinischem Empfinden in diesen beiden Kliniken adjuvante Chemotherapie bekommen haben oder nicht oder adjuvante Hormontherapie bekommen haben oder nicht. Die Guidelines waren damals langsam wechselnd. Es war anfangs so, dass man bei nodalnegativ grundsätzlich gar keine Therapie gemacht hat, und später hat man sogar mit Anthracyclinen behandelt, sodass - das ist der große Vorteil davon - dieses Kollektiv - sogar etwa ein Viertel - anthracyclinhaltige Chemotherapien erhalten hat.

Fueloep Scheibler: Dann habe ich den Eindruck, dass wir die Studie nur unterschiedlich bezeichnen. Für uns wäre ein CCT eine Interventionsstudie.

Christoph Thomssen: Können Sie den Begriff „CCT“ erklären?

Fueloep Scheibler: Controlled clinical trial. Das wäre für uns eine Studie, in der die Patientinnen vom Versuchsleiter in die beiden Gruppen eingeteilt werden. Für uns besteht der Unterschied zwischen CCT und RCT eigentlich nur noch darin, dass diese Einteilung einmal randomisiert erfolgt und einmal nach anderen Kriterien, die eben nicht randomisiert sind. Aber so wie ich es jetzt verstehe, handelt es sich um zwei große Kohortenstudien, deren Daten zusammengeführt wurden, aber das ist eine Beobachtungsstudie.

Christoph Thomssen: Korrekt, weil wir aus den beiden vorhergehenden Studien gesehen haben: uPA/PAI ist ein starker prognostischer Faktor. Wir haben aus der ITT-Analyse auch

den Hinweis, dass es zumindest tendenzmäßig - 2002 hatten wir noch nicht alle Daten ausgewertet - einen prädiktiven Wert hat, sodass wir gesagt haben, das wollen wir jetzt auch in einer Anwendungs-, Beobachtungs- oder großen Kohortenstudie überprüfen, so wie das ja auch von Arzneimitteln verlangt wird heutzutage, dass man Phase-IV-Studien macht, um das noch einmal zu belegen, dass das alles passt.

Moderator Jürgen Windeler: Ich habe einen Punkt noch nicht verstanden. Bei den beiden Kohorten ist der Marker gemessen worden. Ist jetzt der Marker in die Entscheidung mit eingegangen, oder ist er ausdrücklich nicht in die Entscheidung mit eingegangen?

Christoph Thomssen: Nicht eingegangen.

Moderator Jürgen Windeler: Das heißt, er war gar nicht bekannt?

Christoph Thomssen: Es ist sozusagen post-hoc bestimmt worden.

Moderator Jürgen Windeler: Es sind also Entscheidungen getroffen worden aufgrund der Klinik- und sonstiger Befunde. Und der Marker ist sozusagen hinterher ermittelt worden, und es ist sozusagen eine Korrelation zwischen den Entscheidungen und den Markerergebnissen hergestellt worden?

Christoph Thomssen: Genau.

Fueloep Scheibler: Ich glaube tatsächlich, dann sind wir eigentlich eher im Bereich der Prognosestudien. Das sind nach unserer Begriffsdefinition keine CCTs, sondern das sind einarmige Beobachtungsstudien, aus denen man die prognostische Fähigkeit dieses Markers ableiten kann. Darüber würden wir dann unter Tagesordnungspunkt 3 diskutieren wollen. Aber, wie gesagt, das ist für uns keine Interventionsstudie.

Christoph Thomssen: Die Intervention ist ja aufgrund klinischer Entscheidungen gefällt worden. Man hat damit schon einen Unterschied, dass man erkennen kann, welche Patienten Chemotherapie bekommen haben und welche nicht. Und man hat aufgrund des Biomarkers dann auch verschiedene Prognosen festgestellt, die auf eine Wirksamkeit der Chemotherapie rückschließen lassen. Insofern erfüllt dies für uns schon die Begrifflichkeit des prädiktiven Wertes.

Moderator Jürgen Windeler: Ich muss noch einmal nachfragen, weil ich es verstehen möchte: Sie haben jetzt eine Korrelation zwischen dem Biomarker, post-hoc gemessen, und der Entscheidung für oder gegen eine Chemotherapie hergestellt?

Christoph Thomssen: Wir haben eine Korrelation hergestellt zwischen dem Biomarker und der Prognose, die gemessen worden ist, und dann in Abhängigkeit von der Durchführung einer Chemotherapie.

Moderator Jürgen Windeler: Okay.

Stefan Sauerland: Der entscheidende Unterschied zwischen einer randomisierten Studie und dieser Studie ist doch der, dass die Entscheidung für eine Chemotherapie oder gegen eine Chemotherapie nach den damaligen Konsensus-Kriterien erfolgt ist und damit eben nicht randomisiert ist. Das heißt, wenn Sie diese beiden Gruppen vergleichen und da irgendwelche Effekte in der Hochrisiko- und in der Niedrigrisikogruppe ermitteln, dann werden diese Effekte entscheidend dadurch bestimmt sein, ob eine Frau andere Risikofaktoren hat. Das heißt, die eigentlichen Hazard Ratios, die Sie dort vergleichen, sind im engeren Sinne nicht valide, weil die Zuteilung zu dieser Chemotherapie oder eben zu keiner Chemotherapie nicht randomisiert erfolgt ist.

Christoph Thomssen: Grundsätzlich haben Sie recht, aber ich möchte dagegenhalten, erstens ist es eine sehr große Zahl, zweitens aus zwei unterschiedlichen Regionen, die auch unterschiedliche Therapiegewohnheiten hatten - Holland und Deutschland war anders; Süddeutschland war anders als Norddeutschland -. Und in der multivariaten Analyse - das war der Punkt „adjustiert“ - war trotzdem die Chemotherapie und der uPA-Wert ein unabhängiger Faktor. Adjustiert, also im Cox-Modell, war das trotzdem ein unabhängiger Faktor. Das spricht dafür, dass neben dem Nodalstatus, der ja auch die Unabhängigkeit gezeigt hat, eben auch der uPA/PAI-Wert einen klaren prädiktiven Wert hat.

Fueloep Scheibler: Ich glaube trotzdem, dass man diese Art von Studien unterscheiden sollte von kontrollierten Studien, weil hier die Intervention, also ob die Frau eine Chemotherapie erhalten hat oder nicht, nicht einfach von einer Gruppenzuteilung abhängt. Für uns wäre ein CCT beispielsweise, wenn man sagen würde, montags Chemotherapie, dienstags bekommen die Patienten keine, mittwochs wieder Chemotherapie. Dann ist nämlich dieser Zuteilungsmechanismus unabhängig von prognostischen Faktoren. Wenn aber, wie ich das hier vermute, die Zuteilung zu der Frage Chemotherapie ja oder nein genau aufgrund von starken prognostischen Faktoren erfolgt, dann handelt man sich damit einen großen Bias ein, nämlich den Bias by Indication. Es ist also zu vermuten, dass Frauen mit einem höheren Risiko vermehrt in die Chemotherapiegruppe gekommen sind und Frauen mit niedrigem Risiko keine Chemotherapie bekommen haben. Wenn man dann Unterschiede in den Endpunkten sieht, dann waren die möglicherweise schon vorher vorhanden.

Christoph Thomssen: Diese Studie ist ganz klar keine randomisierte kontrollierte Studie, sondern diese Studie ist ein weiterer Baustein, um zu zeigen, dass uPA/PAI einen klaren prädiktiven Wert hat. Wir müssen immer von der Ausgangssituation ausgehen. Die ursprüngliche Intention der Marker war ja nicht, den prädiktiven Wert festzustellen, sondern die Prognose, TOP 3.

Zusätzlich haben wir gesagt: Wenn wir jetzt schon die Patienten charakterisieren, die eine schlechte Prognose haben, dann wollen wir auch wissen, ob die von einer Chemotherapie profitieren. Das Wort „prädiktiv“ haben wir damals noch gar nicht in der stringenten Form, in der wir es heute formulieren - zu der Definition von „prädiktiv“ hat ja unsere Arbeitsgruppe relativ viel beigetragen - verwendet. Damals haben wir ganz einfach gesagt: Wir wollen eben

wissen, ob die Hochrisikopatienten profitieren, definiert nach uPA/PAI. Deswegen haben wir damals randomisiert, und wir sind davon ausgegangen: Wenn wir bei nodalnegativ davon ausgehen, dass ein kleiner Chemotherapiebenefit da sein wird, und wir finden jetzt gerade in der Hochrisikogruppe, dass dieser Effekt da ist, bei der Niedrigrisikogruppe der Effekt gar nicht da sein kann, weil zu niedriges Risiko, dann haben wir schon den ersten Baustein. Der ist mit der Chemo-N0-Studie gegeben.

Der zweite Baustein war praktisch die - ich nenne sie mal so - Meta-Analyse, diese Analyse mit den 3.500 Patienten aus Rotterdam und München, in der gezeigt wird, tatsächlich konsistent, uPA/PAI sagt den Effekt der Chemotherapie besser voraus als für die Hormontherapie. Und gerade weil wir hier den Vergleich zwischen Hormon- und Chemotherapie haben, ist das ein deutlicher Hinweis.

Der dritte Punkt wäre die multivariate Analyse. Ich bin jetzt kein Statistiker, aber meines Erachtens kann man mit einer multivariaten Analyse schon zu einem guten Teil den Bias wieder herausrechnen, den man durch eine solche nichtkontrollierte Form der Studie hineinbaut.

Claudia Olenik: Ich möchte mich dem Vorredner anschließen. Es ging doch jetzt auch um die Frage, warum wir zusätzliche Studien einschließen möchten. Wir haben Studien vorliegen, die eben nicht dem RCT-Design entsprechen, die aber konsistent sind mit den Ergebnisse, die also die Ergebnisse der Chemo-N0-Studie stützen. Wir sind uns im Klaren: Die ITT-Analyse ist nicht signifikant, aber sie zeigt eine Tendenz. Also ich wiederhole: In der Per-Protokoll-Analyse sehen wir einen Nutzen. Das ist doch schon mal als sehr positiv zu bewerten, wenn es in die gleiche Richtung geht und durch andere Studien gestützt werden kann.

Ich möchte jetzt auf einen anderen Punkt kommen. Sie begründen die ausschließliche Akzeptanz der RCTs damit, dass es andere Studien gibt, die aktuell im randomisierten Design die ähnliche Fragestellung untersuchen. Herr Thomssen hat es schon angesprochen, die MINDACT-Studie, die TAILORx-Studie. Jetzt haben wir aber doch hier genau die gleichen Probleme. Mit diesem Projekt, mit dieser Untersuchung soll ja der aktuelle Stand untersucht werden. Die Auswertungen der beiden angesprochenen Studien liegen frühestens 2015 vor. Die Chemo-N0-Studie war eine sauber geplante randomisiert-kontrollierte Studie, zumindest im Hybrid-Design. Wenn wir uns jetzt die anderen Studien anschauen: Natürlich haben wir hier sehr große Patientenzahlen. Aber wir sehen auch bei diesen Studien, dass es sehr schwierig ist, dieses Design umzusetzen. Und erst die Zukunft wird zeigen, ob diese theoretisch geplanten Designs im klinischen Alltag umsetzbar sind und auch die gewünschten Ergebnisse bringen werden. Ich weiß nicht, ob man sich auf Studien beziehen soll, die erst in der Zukunft ausgewertet werden, wenn wir hier einen Parameter haben, für den wir doch aus Ihrer Sicht relativ valide Ergebnisse vorliegen haben. Was machen wir denn in der Zeit bis dahin, bis im Prinzip die anderen Ergebnisse vorliegen?

Ralf Bender: Die Kernfrage war ja: Warum schließen wir nicht andere Studiendesigns ein außer RCT, um das Bild zu vervollständigen?

Claudia Olenik: Genau.

Ralf Bender: Aufgrund der Methodik des Instituts - wir nennen das „Ableitung der Beleglage“ - ist es so, dass wir das Ergebnis eines RCT nicht nivellieren können durch Studien niedrigeren Evidenzlevels. Also selbst wenn wir weitere CCTs hätten - wir haben jetzt gar keine, sondern wir haben eher so etwas wie Beobachtungsstudien, Prognosestudien -, können diese das Ergebnis eines RCT nicht kippen. Und das Ergebnis des RCT ist negativ. Hier wird schon argumentiert mit der Per-Protokoll-Analyse, das sei doch signifikant. Das ist so nicht ganz richtig. Die Analyse des RCT ist die ITT, aber da kommen wir unter TOP 2 noch genauer zu. Das Ergebnis des RCT ist negativ, nicht signifikant, zeigt keinen Nutzen. Dieses Ergebnis kann nicht gekippt werden durch Studien niedrigeren Evidenzlevels. Aus diesem Grund brauchen wir die auch nicht einzuschließen.

Claudia Olenik: Ich weiß nicht, ob es negativ ist. Sie sagen „negativ“, wir sagen, man sieht ja eine Tendenz. Klar, es ist nicht signifikant, auf statistischer Seite ist das Ergebnis damit natürlich negativ.

Ich hatte vorhin ein bisschen ausgeholt. Ich hatte Ihnen ja noch mal dargestellt, welche Gründe oder Faktoren, die ja von uPA/PAI-1 unabhängig sind, möglicherweise dazu geführt haben, dass dieser Nutzeneffekt nicht so deutlich in der Auswertung herausgekommen ist, nämlich dass ja im Prinzip durch die veränderten Leitlinien - Herr Thomssen hat es ja angesprochen - ... Bei den nodalnegativen Patienten waren ja früher überhaupt keine Therapien vorgesehen. Jetzt kommt eine Studie daher und sagt: Okay, wir machen randomisiert Chemotherapie. Plötzlich bekommen die Patienten Angst: Chemotherapie, o Gott. - Wir bekommen die Anrufe zu Hause. Die Patienten rufen an. Da ist eine extreme Angst vorhanden. Und dann entschließen sie sich plötzlich, doch wieder aus der Studie auszusteigen. Wir hatten ursprünglich 640 im High-Risk-Arm, die in die Randomisierung hineingehen sollten. 122 Patienten, also ein ganzes Drittel, sagen auf einmal: Nein, doch keine Randomisierung, lieber gar keine Therapie. - Und selbst nach der Randomisierung immer noch die Unsicherheit. Ich wollte damit nur darstellen und denke, das sollte doch auch berücksichtigt werden, dass wir hier zu einem Zeitpunkt die Studie, die sauber geplant war, aufgesetzt haben, aber natürlich durch die äußeren Gegebenheiten es zu einer starken Verminderung der Patientenzahlen kam. Ursprünglich waren 900 vorgesehen. Nur 647 sind dann aufgenommen worden. Wenn man das jetzt so wegwischt und sagt, die Studie negatives Ergebnis, dann denke ich, sollte man diese Punkte - die haben Sie ja in Ihrem Vorbericht sogar selber mit aufgenommen ... Deshalb finde ich, dass man das nicht so ganz krass wegwischen kann. Deshalb erlauben Sie uns bitte, andere Studienmodelle hinzuzuziehen, die die Ergebnisse stützen. Das sollte, finde ich, immer im Gesamten dargestellt werden.

Christoph Thomssen: Ich denke, das muss man mal aufgreifen, dass wir die Sachen im Gesamtzusammenhang sehen müssen. Die Chemo-N0-Studie war damals im wissenschaftlichen Sinne eine Pilotstudie überhaupt für solche Prognosestudien und prädiktiven Studien. Es gab eine einzige andere Studie von dem Kollegen aus Italien, Paradiso, der mit dem Proliferationsmarker TLI (thymidine labeling index) eine prädiktive Studie aufgelegt hat im ähnlichen Design wie unsere Studie und eben dann Anthracyclin versus nicht Anthracyclin gemessen hat, auch publiziert hat. Ansonsten gab es bisher gar nicht solche Biomarkerstudien. Die MINDACT-Studie ist im Wesentlichen auf unserem Design der Folgestudie, der NNBC-Studie, aufgebaut, weil wir da in der EORTC eng zusammengearbeitet haben. Die MINDACT-Studie hat die gleichen Rekrutierungsprobleme gehabt und auch Complianceprobleme. Ein Drittel - das hat Emil Rutgers auch publiziert - der Patienten hat praktisch die Randomisierung verweigert. Und die MINDACT-Studie wird uns auch nicht genau sagen, Prognose oder Prädiktion, weil dort ja praktisch auch die beiden Zielkriterien gemischt werden in der Randomisierung der heterogenen Gruppe. Sagt Ihnen die MINDACT-Studie was?

Fueloep Scheiber: Ja.

Christoph Thomssen: Und die TAILORx-Studie hat ähnliche Probleme. Die mussten ihre Fallzahl extrem erhöhen auf 11.000 und werden nur auswerten, ob in der Mittelgruppe, wo sie nicht genau wissen, ob der Oncotype-DX einen Wert hat oder nicht, jetzt Chemotherapie besser ist oder nicht. Das heißt, auch da bekommen wir keine sauberen Antworten auf die Frage Prädiktion- oder Prognosewert dieser Faktoren. Deswegen, denke ich, muss man auch im Gesamtumfeld diese Chemo-N0-Studie bewerten und sagen, wenn ich statistisch ganz stringente Kriterien von 2014 nehme, dann hat diese Studie ihre Fehler in der Durchführung und in der Planung. Ich hätte es vielleicht anders gemacht heutzutage. Aber wir haben jetzt nichts anderes, wir werden auch nichts anderes bekommen. Jetzt müssen wir es unter dem Punkt bewerten, was überhaupt klinisch und faktisch möglich ist. Und das zeigt die Geschichte „randomisierte Studien bei prädiktiven Faktoren“, dass das ein ganz schwieriges Feld ist und dass wir glücklich sein können, dass wir überhaupt so weit gekommen sind.

Claudia Olenik: Ich möchte jetzt auch noch einmal das, was Herr Thomssen gesagt hat, ... Das ist nämlich gerade bei dieser TAILORx-Studie von Oncotype-DX ... Die haben ja diese Dreieinteilung der Risikopatienten. Und wenn man jetzt denkt, dass hier randomisiert untersucht wird, dann muss man sich wirklich vor Augen halten, die haben zwei Gruppen, für die ist einfach nach dem Recurrence Score klar, wie sie behandeln. Da vergleichen die nicht mit anderen Faktoren. Und nur in deren Mittlerisikogruppe - und das ist nicht die intermediäre unklare Gruppe, von der wir hier sprechen; das vermischt sich nämlich oft in den Diskussionen - wird dann noch mal randomisiert. Die anderen Gruppen werden entsprechend deren Auswertung nach dem Recurrence Score behandelt.

Stefan Sauerland: Frau Olenik, grundsätzlich ist wichtig, dass es überhaupt erst mal einen Nachweis einer Interaktion gibt. Da ist es fast unerheblich, auf welche Subgruppen sich das

bezieht. Sie sagen jetzt, die geplanten Studien würden auch schon eine Einschränkung vornehmen, weil man in dem Niedrigrisikoarm und dem Hochrisikoarm gar nicht mehr randomisieren könne. Aber selbst wenn dann in dem Mittelrisiko, das mittlere Risiko der Mittelrisikogruppe sozusagen, wenn sich dort eine Interaktion zeigt, das wäre ja zumindest ein Beleg dafür, dass überhaupt der Test sinnvoll verwendbar ist, um über die Entscheidung Chemotherapie ja oder nein eine Aussage treffen zu können.

Das andere, was jetzt mehrfach angesprochen wurde, ist die historische Komponente. Da können wir jetzt auch nicht einfach sagen, wir haben jetzt das Jahr 2014, aber die Studien sind halt in einer Zeit geplant worden, wo man bestimmte Dinge noch nicht wusste. Wir können uns doch nicht dem Stand des heutigen Wissens verschließen. Wir müssen die Daten so interpretieren, wie sie in den heutigen Kontext passen. Da kann man natürlich ein gewisses Verständnis haben für studienplanungstechnische Aspekte, die damals vielleicht völlig richtig waren, aber die heute anders zu sehen sind. Und wenn Sie den Ausblick in die Zukunft wagen wollen, dann wird es ja irgendwann auch dazu kommen, dass man nicht mehr einen Test gegen keinen Test vergleichen wird, sondern wir kommen dann in eine Welt, wo man einen Test gegen einen anderen Test vergleichen muss. Und das ist etwas, wo man dann noch mal ganz neu nachdenken sollte.

Ralf Bender: Ich wollte einen Punkt näher erläutern. Sie haben zu Recht gesagt, wir müssen die Ergebnisse des RCT im Kontext sehen und können diese Ergebnisse nicht einfach so wegwischen. Das ist natürlich richtig. Dass dieses Ergebnis erst einmal negativ ist, darf natürlich nicht in der Weise missverstanden werden, dass wir daraus schlussfolgern würden, es sei nachgewiesen, dass hier kein Nutzen vorliegt. So ist es natürlich nicht. Sie sagen, die ITT-Analyse zeigt einen Trend. Das bedeutet im Klartext, der Effektschätzer ist in der richtigen Richtung. Es ist aber nicht signifikant, und es ist auch nicht ganz knapp nicht signifikant. Der p-Wert ist also nicht dicht bei den 5 %, sondern schon sehr weit davon weg. Die obere Konfidenzgrenze lag bei 1,2 oder 1,3, doch deutlich hoch. Jetzt kann man sagen, hier ist ein Trend in die richtige Richtung, kombiniert mit dem Ergebnis aus der Beobachtungsstudie. Sicher, man kann so etwas wie ein Potential daraus ableiten, eine Hypothese, ob da was dran sein könnte, aber dieser Bericht ist doch dazu da, Belege für einen Nutzen zu liefern. Und die liegen eben zurzeit nicht vor. Das heißt nicht, dass belegt ist, dass kein Nutzen da ist. Sobald eine neue Studie kommt, kann man uns ja neu beauftragen. Vielleicht zeigt sich dann der Nutzen. Das kann sein, das streiten wir nicht ab. Aber die jetzige Datenlage liefert keine Nutzenbelege. Der RCT ist negativ. Und dieses Ergebnis kann nicht nivelliert werden durch ein CCT, sodass es nach unserer Ansicht keinen Sinn macht, hier jetzt noch Studien außerhalb des RCT-Designs einzuschließen, weil es an dem Fazit im Moment sowieso nichts ändern würde. Vielleicht in der Zukunft, wenn andere Studien vorliegen, könnte es sein, dass es tatsächlich in Kombination des alten RCT mit dem neuen, wenn wir vielleicht eine Meta-Analyse machen, vielleicht signifikant ist. Dann hätten wir einen Nutzenbeleg. Aber im Moment haben wir ihn nicht. Und das sagt der Bericht.

Claudia Olenik: Also wenn Sie das so sagen, dass im Prinzip der Nutzen nicht weggewischt wird ... Ich hatte die Pressemitteilung vorliegen. Da war das ein bisschen anders formuliert. Wir sitzen hier zusammen. Ich denke, wir haben alle das wissenschaftliche Know-how, dass uns klar ist, was Ihre Untersuchung besagt, aber man muss sehr, sehr vorsichtig sein, wie das Ganze interpretiert wird. Da sehe ich einfach das Problem, wenn hier groß gesagt wird, kein Nutzen bei intermediärem Risiko, uPA und PAI-1, was das eben auch auslöst.

Martina Klöppinger: Bei den Patientinnen.

Moderator Jürgen Windeler: Dieses Schwarze-Peter-Spiel möchte ich eigentlich ungern mitmachen. Nicht wir sind dafür verantwortlich, dass die Evidenzlage so ist, wie sie ist.

Claudia Olenik: Da hätte vielleicht noch jemand über die Pressemitteilung - ich weiß nicht, wer die geschrieben hat - gucken müssen. Sie formulieren sehr viel deutlicher, klarer ...

Moderator Jürgen Windeler: Ich persönlich. Sie dürfen also mich dafür verantwortlich machen.

Claudia Olenik: Okay. - Die Überschrift war nicht ganz so präzise, wie man es sich gewünscht hätte. Aber wir wollen uns jetzt nicht damit aufhalten.

Ich wollte Ihnen recht geben. Wenn Sie sagen, okay, man kann es nicht wediskutieren. Ich sage ja jetzt auch gar nicht, die hat jetzt einen Nutzen, sondern dass es in die richtige Richtung geht und dass wahrscheinlich möglicherweise aufgrund der zu geringen Fallzahlen, wie auch immer ... Wenn also höhere Fallzahlen dagewesen wären, wäre es wahrscheinlich statistisch signifikant geworden. Dass einfach jetzt von Ihnen im Prinzip zugestanden wird, okay, diese Studie geht in die richtige Richtung, und dass durch Nachfolgestudien im Prinzip dieser Effekt, den uPA und PAI-1, wahrscheinlich hat - wir gehen jetzt einfach davon aus -, dass der auch nachgewiesen werden kann.

Moderator Jürgen Windeler: Ich möchte das gerne ein bisschen auf den Punkt bringen. Unser Vorbericht und nach momentanem Stand absehbar unsere Bewertung heißt, der Nutzen ist nicht belegt. So wird das auch in der Pressemitteilung stehen. Da werden wir auch keine Relativierung machen und sagen, ein bisschen Tendenz oder so etwas, sondern der Nutzen ist nicht belegt. Anforderungen an Belege sind klar und sind auch publiziert.

Das zweite Problem, das ich im Moment sehe, ist, dass wir uns unter bestimmten Bedingungen vorstellen können, CCTs einzuschließen. CCT heißt, Studien, die diese Interaktionsfrage zwischen dem Test und der erfolgreichen Therapie beantworten, ohne Randomisierung. Ich habe im Moment noch keinen Hinweis auf eine solche existierende CCT gehört. Wenn es die gibt, dann wären wir natürlich dankbar, davon zu erfahren. Wir haben auch danach gesucht. Aber vielleicht haben Sie da noch spezielle Anregungen.

Der dritte Punkt: Vor vielen Jahren - das IQWiG ist jetzt zehn Jahre alt -, vor vielleicht sechs Jahren, vor meiner Zeit, sind hier Erörterungen über PET gelaufen mit genau einer ähnlichen Fragestellung, nämlich prädiktive Fragestellung, wie ändert sich mit PET die Therapie. Von Stellungnehmenden, von Wissenschaftlern, von Firmen ist gesagt worden: Solche Studien sind ganz schwierig zu machen. Das kann man eigentlich gar nicht. - Inzwischen werden sie gemacht. Inzwischen - Recherche gibt es dazu - gibt es eine Reihe von Studien international, die genau diese Frage für ein solches Verfahren, im Kern für ein prädiktives Verfahren, untersucht. Ich glaube, dass man sich bei dieser Thematik langsam mal daran gewöhnen sollte, dass wir Studien brauchen, die diese prädiktive Frage beantworten, und zwar die Frage: Welche Konsequenzen hat der Biomarker oder irgendein anderer Marker auf die Therapie, und ist das für die betroffenen Patientinnen von Vorteil oder nicht? Wenn wir diese Frage im Moment nicht positiv beantworten können - so scheint die Beleglage zu sein -, dann liegt - darauf will ich Wert legen; darauf bezog sich meine Bemerkung eben - die Verantwortung dafür bei denjenigen, die diese Evidenz produzieren oder nicht produzieren, nicht bei uns, die die angucken müssen.

Claudia Olenik: Ich möchte noch mal deutlich machen: Es ist TOP 3. TOP 1 und TOP 3 kann man eigentlich gar nicht wirklich auseinanderziehen, wenn es wirklich um den Nutzen für die Patientinnen geht. Dort haben wir dann ja auch die Studien vorliegen, die das zeigen. Ich bin kein Kliniker; wir haben hier aber Kliniker sitzen. Es ist einfach ganz wichtig, dass man prognostische und prädiktive Faktoren gemeinsam betrachtet, zumindest zum derzeitigen Zeitpunkt. Was hier bei der Analyse gemacht wird, das ist ja sauber aus statistischer Sicht. Das ist mir ganz klar. Hier wird Prädiktion abgetrennt von der Prognose. Sie brauchen einfach ein Messinstrument, wie Sie Nutzen messen können, Prädiktion, aber - das werden wir hoffentlich nachher noch in der Diskussion sehen - dass der prognostische Wert eines Markers auch ganz wichtig ist und eben auch einen Nutzen für Patientinnen und Arzt bei der Entscheidungsfindung hat. Wir werden uns nachher - ich will nicht vorgreifen - noch mal ganz genau Ziel und Auftrag anschauen. Aus meinem Verständnis - Sie können es mir dann erklären - steckt nicht nur dann „Wie wirkt jetzt die Chemotherapie?“ dahinter, sondern Auftrag ist eigentlich, auch zu schauen: Hilft uPA/PAI-1 bei einer Empfehlung, bei einer Entscheidung für oder gegen eine adjuvante Chemotherapie? Und da kommt sehr wohl TOP 3 zur Sprache.

Moderator Jürgen Windeler: Mit der Beschreibung Ihrer Frage habe ich gar kein Problem: ob es hilft. Die Frage ist, was die Kriterien für dieses „hilft“ sind. Da kommen wir ja vielleicht bei TOP 3 drauf.

Ich habe den Eindruck, dass TOP 1 zu Ende diskutiert ist. - Herr Thomssen.

Christoph Thomssen: Sie sagen die ganze Zeit: ITT negativ.

Vorab noch mal: Ich sehe es auch so: Die Prognose steht im Vordergrund der Fragestellung, nicht die Prädiktion. Aber hier diskutieren wir die Prädiktion.

Sie sagen, ITT sei negativ, damit sei die Studie im Sinne der Studienfragestellung negativ. Rein statistisch gesehen haben Sie damit recht. Deswegen müssen wir natürlich in der erschwerten Situation, dass es ganz schwer ist, für solcher Marker Evidenz herzubekommen, schauen, welche Argumente sprechen dafür. Und da würde ich eben schon weggehen von der reinen Statistik, sondern dann eben schon zusammenbauen. Was würden wir denn erwarten bei so einer Patientengruppe, wenn wir schon das Hochrisikokollektiv damit ziemlich effektiv herauskristallisieren können? Würden wir denn da erwarten, dass irgendein anderer Effekt kommt, oder würden wir nicht sowieso erwarten, dass es genau diejenigen sind, die von der Chemotherapie profitieren? Dann haben wir die ITT-Analyse, die in der Kappler-Meier-Kurve wunderbar auseinandergeht, deckungsgleich mit der Per-Protokoll-Analyse. Wenn Sie mal klassische Studien nehmen: Die Bonadonna-Studie hatte, glaube ich, 190 Patienten. Damit ist die adjuvante Chemotherapie eingeführt worden. In Italien gab es damals halt keine Wechsler. Da hat die Bonadonna--Studie gesagt: Du bekommst Chemo, und du bekommst keine. Das war direkter. Dadurch sind die Konfidenzintervalle schmaler, und nur die Konfidenzintervalle sind ja letztendlich das, was zu der fehlenden Signifikanz geführt hat.

Und dann haben wir als drittes Argument die Meta-Analyse München/Rotterdam, in der rückgerechnet, weil über eine sehr große Zahl von Patienten sehr schön gezeigt worden ist, dass dieser prädiktive Wert sehr wohl da ist.

Sie haben gefragt: Gibt es CCTs? - Die NNBC-Studie, die wir ja in der Folge mit 4.200 Patienten durchgeführt haben, ist ja eigentlich so eine CCT, weil wir da tatsächlich den Zentren freigelassen haben, zu sagen, ihr dürft nach klinischen Faktoren auswählen, oder die anderen, ihr dürft nach uPA/PAI auswählen. Hier muss man sich nur festlegen. Das war für uns einfacher als zu randomisieren. Randomisieren - das hatte ich auch geschrieben - ist in solchen Situationen schwierig. Das Problem bei der NNBC-Studie ist nur: Wir haben die ja ausgewertet, auch beim ASCO vorgestellt. Die Fragestellung war, Chemotherapie A versus B. Da hat sich jetzt kein Unterschied gezeigt, aber letztes Endes ist das vielleicht auch zu viel erwartet von den Markern, auch noch den Unterschied zwischen zwei Chemotherapien zu erwarten. Und in der Frage Prognose lässt sich momentan mangels Ereignisse noch nicht viel sagen. Trotz 4.200 Patienten haben wir halt noch nicht genügend Evidenz, um da eine Aussage machen zu können. Das weist aber auch darauf hin, dass wir mit den Prognosen gar nicht so schlecht liegen. Das Low-Risk-Kollektiv ist eben so gut, dass wir hier noch keine Aussage machen.

Claudia Olenik: Weil wir ja hier die Trennung zwischen Prognose und Prädiktion haben: Sie beziehen sich ja auch auf Sankt Gallen 2007. Wenn man dem folgt, dann ist es doch tatsächlich so, dass die Einstufung des Rezidivrisikos einer Patientin als hoch bzw. niedrig im Rahmen von den Leitlinien als gleichbedeutend erachtet wird mit der Empfehlung für oder gegen eine adjuvante Chemotherapie. Das heißt, das ist gleichbedeutend. Wir trennen das hier in unserer Erörterung, aber von den Leitlinien ... Ich weiß, dass inzwischen natürlich auch andere Empfehlungen existieren, die mehr die Tumorbiologie aufgreifen, aber Sie beziehen sich ja vor allem auf Sankt Gallen 2007. Wenn Sie da mal hinschauen, wie wird denn

empfohlen, dass in deren Risikogruppen die Chemotherapie gegeben wird, da ist es tatsächlich, dass es als gleich erachtet wird.

Stefan Sauerland: Für die Organisation des Treffens hier, wäre es, glaube ich, hilfreich, wenn wir in den TOPs bleiben würden. Da würde ich sagen: TOP 1 haben wir eigentlich hinreichend diskutiert. Wir haben jetzt verstanden, dass aus Ihrer Sicht primär die Harbeck-Analyse von 2002 und die laufende Auswertung der NNBC-Studie für diesen Bericht eventuell eine Relevanz haben können. Wir werden das prüfen. Ich glaube, damit kann man jetzt den TOP 1 erst mal auf die Seite legen, und man könnte sich mit TOP 2 beschäftigen. Das ist mein Vorschlag.

Moderator Jürgen Windeler: Ich habe TOP 1 ja ohnehin abschließen wollen. Wenn das jetzt unterstützt wird, umso besser. Von daher frage ich noch mal ausdrücklich: Gibt es weitere Wortmeldungen zu TOP 1?

Christoph Thomssen: Um noch mal die Größenordnung klarzumachen: CMF bei der nodalnegativen High-Risk-Situation definiert in der uPA/PAI liegt im absoluten Effekt von etwa 10 % in der gleichen Größenordnung wie Bonadonna-Chemotherapie CMF versus nichts. Das macht die Sache sehr wahrscheinlich, dass das alles passt. Bei Aromatasehemmern haben Sie 20.000 Patienten randomisiert, um einen Effekt von 2,5 % herauszubekommen. Aromatasehemmer sind heutzutage Standard.

Moderator Jürgen Windeler: Gut. Dann schließe ich TOP 1 ab, und wir kommen zum

4.3.3 Tagesordnungspunkt 2: Rolle einer Per-Protokoll-Analyse

Martina Markes: In zwei Stellungnahmen wird vorgeschlagen, die Nutzenbewertung auch auf Basis einer Per-Protokoll-Analyse durchzuführen. Entsprechend unserer Methoden legen wir einer Nutzenbewertung in der Regel die ITT-Auswertung zugrunde. Und wir möchten Sie bitten, zu begründen, warum gerade in diesem Bericht von diesem Auswertungsprinzip abgewichen werden sollte.

Christoph Thomssen: Letztendlich bin ich gefragt. Die Problematik ist ja im vorhergehenden Tagesordnungspunkt schon angesprochen worden mit der Frage der Prädiktion: War es notwendig zu randomisieren in einer Hochrisikosituation nodalnegativ zu einer Zeit, wo bei nodalnegativer Erkrankung im Allgemeinen damals eben keine Chemotherapie empfohlen wurde? Wir haben diese Randomisierung trotzdem durchgesetzt, um genau das klar herauszubekommen, ob bei Patienten, die nach uPA/PAI Hochrisiko sind, eine Chemotherapie etwas nutzt. Es hat sich herausgestellt, dass ein Drittel der Patienten die Randomisierung verweigert hat, von den Ärzten wegen niedrigen Gradings oder so etwas in eine andere Gruppe geschoben worden sind oder die Patientinnen nach der ersten Chemotherapie oder so abgebrochen haben. Das macht natürlich so eine ITT-Analyse extrem schwierig - das sagte ich schon vorher -. Dann werden die Konfidenzintervalle sehr breit, und dann bekommen Sie keine Signifikanz mehr. Deswegen haben wir uns - das sieht man in dem

Consort-Diagramm, in der Zehn-Jahres-Analyse, sehr schön - da noch mal herangemacht und haben genau die Gruppe B1 und B2, also die, die randomisiert worden sind und für Chemotherapie und nicht, ohne die Protokollverletzer ausgewertet, um zu zeigen, wie dann, wenn man die Therapie tatsächlich durchführt, die Effekte sind. Da haben wir ja sehr deutlich die signifikanten Effekte zeigen können.

Ich denke, die ITT-Analyse ist ja a priori Standard deswegen, weil man a) klinische Realität abbilden möchte und b) auf jeden Fall vermeiden möchte, einen Effekt zu überschätzen. Wenn Sie aber ein Drittel der Patienten nicht mehr im richtigen randomisierten Arm drin haben, weil die einfach den Arm wechseln, dann konterkariert das natürlich den Sinn und Zweck einer ITT-Analyse. Aus diesem Grunde haben wir uns zu diesem Kunstgriff gezwungen gesehen.

Grundsätzlich muss man sagen: Die ITT-Analyse - ich sagte es vorhin schon - liegt in ihrer Tendenz eindeutig in der gleichen Richtung wie die signifikante Per-Protokoll-Analyse und wie - das haben wir erwartet - auch alle anderen Analysen, sodass wir in diesem Fall die Per-Protokoll-Analyse schon für eine sehr wichtige Analyse halten. Es greifen ja durchaus auch andere Arbeitsgruppen zu solchen Kunstgriffen. Die Aromatasehemmer waren ja in ihrem Survivalbenefit auch nicht signifikant, weil sehr viele nach Aufdecken der Daten plötzlich noch zum Aromatasehemmer gegriffen hatten. Die BIG-1-98-Studien werden Sie vielleicht kennen, Letrozol versus Tamoxifen fünf Jahre. Nach Aufdeckung der Daten wurde nach zwei oder drei Jahren von den Patienten doch zum Aromatasehemmer sekundär gegriffen, das ja auch effektiv ist. Dann kam eben kein Survivalbenefit mehr heraus in der ITT-Analyse. Dann hat man eine zensierte Analyse gemacht, alle die herausgeworfen, die umgestiegen sind. Dann hat man natürlich eine Überschätzung des Effektes. Und dann haben die sich die Inverse Probability Censoring Weigthing (IPCW) -Analyse als statistischen Kunstgriff zunutze gemacht. Und dann kam man plötzlich zu einem Mittelwert, der dann immer signifikant war, was das Survival betraf. Es ist also offensichtlich durchaus Usus, wenn man methodische Schwierigkeiten hat, statistische Kunstgriffe zu ergreifen, die sozusagen die Realität besser abbilden.

Anne Rutjes: Wenn die Protokollverletzungen eine erhebliche Rate sind und man jetzt begründet, dass man eine Per-Protokoll-Analyse durchführen könnte, dann hat man eigentlich im Prinzip die Annahme gemacht, dass diese Protokollverletzungen completely at random sind, dass die Rezidivrate nichts mit dem Patientenprofil zu tun hat und nichts mit der Prognose. Ich denke, das kann man nicht machen. Was man dann letztendlich hat, ist, dass man ein Selektionsbias einführt. Die RCT kann man dann nicht mehr als RCT ansehen, sondern eher als eine observationelle Studie, wenn man eine Per-Protokoll-Analyse durchführt. Ich sehe das nicht als eine Lösung oder eine Unterstützung. Es bringt mir nicht mehr an Verständnis, wo der wahre Schätzer liegt. Ich sehe schon das Problem, dass die ITT die Tendenz hat, nichts mehr zu zeigen. Nur das Problem mit so einer Anzahl von Protokollverletzungen kann man, glaube ich, nicht beheben, auch nicht mit einem Inverse Probability Censoring Weigthing.

Christoph Thomssen: Die Kaplan-Meier-Kurven sind ja praktisch deckungsgleich zwischen ITT und Per-Protokoll. Das heißt, die ITT zeigt ja jetzt nicht das Gegenteil von dem, was die Per-Protokoll-Analyse zeigt und umgekehrt. Per-Protokoll ergibt jetzt kein anderes Ergebnis als die ITT-Analyse. Insofern sehe ich die zugegebene Gefahr eines Bias für nicht so hoch, zumal in beiden Gruppen, sowohl im Arm für keine Therapie als auch im Arm für Therapie, etwa gleich viele Patienten geswicht sind.

Anne Rutjes: Gibt es denn überhaupt Daten von denen, die die Studien durchgeführt haben, was das Patientenprofil gewesen ist, von denen, die herausgegangen sind?

Christoph Thomssen: Eine war landsmannschaftlich, wenn ich das so sagen darf. In München haben sie halt gesagt: Bleibt mir weg mit dem Gift. Ich will keine Chemotherapie. - Und in Norddeutschland hat man gesagt: Für potentiell 1 % Survivalbenefit möchte ich auf jeden Fall das Gift haben. - Das Zweite war, dass die Ärzte gesagt haben: uPA/PAI ist gut, aber das Grading ist so niedrig, dann machen wir lieber keine Chemotherapie. - Das sind die typischen Argumente gewesen.

Anne Rutjes: Das beruhigt mich nicht.

Christoph Thomssen: Natürlich.

Ralf Bender: Ergänzend dazu: Das Argument, dass die ITT-Analyse vermutlich hier einen verwässerten Effekt zeigt, ist ohne Zweifel richtig, bei so vielen Protokollverletzungen. Aber deshalb zu argumentieren, dann gehen wir auf eine Analyse, die einen starken Selektionsbias besitzt, also verlassen quasi das Design einer randomisierten Studie und gehen wieder auf das Level einer Beobachtungsstudie, dem können wir uns nicht anschließen. Wir bleiben bei der konservativen Analyse, die eben hier keinen Beleg eines Nutzens zeigt. Das ist und bleibt so.

Christoph Thomssen: Die PP-Analyse zeigt eine Hazard Ratio von 0,5, die ITT-Analyse zeigt eine Hazard Ratio von 0,7, also irgendwo in dem Bereich wird der Effekt sein. Das ist ja so ähnlich bei den Industriestudien zu den Aromatasehemmern gezeigt worden. Man hat zwei verschiedene Hazard Ratios angenommen und dazwischen diese zensierte Analyse gesetzt und war dann bei einem Wert, der dazwischenliegt.

Die ITT-Analyse im Allgemeinen ist ja dafür eingeführt worden, wenn aus irgendeinem Grunde die Therapie nicht ganz vollständig oder irgendwie variiert durchgeführt ist, da jetzt die Varianzen, die den Bias enthalten kann, herauszurechnen und dann die ungünstigste Situation anzunehmen. Denn wenn eine ITT-Analyse in dieser Form konterkariert wird, indem praktisch ein Drittel der Patienten nicht randomisierungsgerecht behandelt werden, dann haben wir ein Problem, das auszuwerten.

Das zweite Problem ist - deswegen, finde ich, ist der Zusammenhang auch wichtig -: Das ist die einzige Studie, bei der man überhaupt noch gegen einen negativen Arm, also einen Arm ohne Therapie, hatte randomisieren können. Wir konnten 1999 die Studie nicht mehr

fortführen, weil dann die Chemotherapie Standard geworden ist. Das heißt, das hier war und ist die letzte Möglichkeit, einen prädiktiven Effekt solcher Marker im statistisch sauberen Sinne mittels einer randomisierten Studie durchzuführen. Wenn wir das jetzt nicht anerkennen, weil wir sagen, wir haben die Kriterien zu hoch gesetzt oder die selbst gesetzten Kriterien sind nicht erfüllt, dann ist der Marker, was die Prädiktion betrifft, durch, wenn man keine anderen Argumente mehr anerkennt. Das muss man ganz klar sagen. Aber ich sagte vorhin schon: Arbeitsgruppen aus Amerika wie die von Daniel Hayes und Richard Simon haben dann eben gesagt: „Okay, dann müssen wir den Level of Evidence eben heruntersetzen“, und haben dann den Schwellenwert geändert, was ich auch durchaus kritisch sehe, aber unter dem Aspekt, dass sich die Patientenkollektive und Therapiegewohnheiten ändern, muss man das auch mal mit ins Felde führen.

Moderator Jürgen Windeler: Mir fällt jetzt dazu sehr viel ein, aber ich sage mal nichts.

Ralf Bender: Der Bereich, wo der Effekt in Wahrheit vermutlich liegt, wird nicht gegeben durch die Streubreite der Punktschätzer aus verschiedenen Analysen, sondern durch das Konfidenzintervall, und zwar das Konfidenzintervall durch die ITT-Analyse. Und die zeigt, dass sowohl ein Nutzen im Bereich des Möglichen ist wie auch ein Null-Effekt wie auch ein Schaden im Bereich des Möglichen liegt. Fazit: Kein Beleg eines Nutzens.

Claudia Olenik: Ich möchte mich Herrn Thomssen anschließen. Wir hatten das ja auch mit der Per-Protokoll-Analyse in unserer Stellungnahme geschrieben. Es ging mir nicht darum, dass anstatt der ITT-Auswertung nur die Per-Protokoll-Analyse genutzt werden soll. Mir geht es immer nur darum - da wiederhole ich mich; Sie haben es auch schon aufgegriffen -, dass man das einfach zusammensehen sollte und nicht abgeschnitten, dass man nur ITT und „mmm“, sondern dass ja die Per-Protokoll-Analyse ... Ich möchte jetzt auch gar nicht die Werte zusammen verrechnen in irgendeiner Form - das kann ich nicht; ich bin kein Statistiker -, aber dass man einfach sagt: Wir haben eine Per-Protokoll-Analyse. Wir haben ein mögliches Verzerrungspotential. Wir haben aber einen sauberen Behandlungseffekt, den wir hier erst einmal sehen und tendenziell als stützend sehen. Ich will es nicht ersetzen. Das ist mir klar, dass das nicht in Ordnung ist. Die ITT-Analyse ist ohne Frage die beste Form der Analyse mit dem geringsten Verzerrungspotential. Aber unter den gegebenen Umständen - und deshalb in unserer Stellungnahme noch mal aufgeführt - sollte man die Per-Protokoll-Analyse nicht außen vor lassen. Das war unsere Intention.

Moderator Jürgen Windeler: Wir haben eben keinen sauberen Therapieeffekt - das ist das Problem -, sondern wir haben einen möglicherweise verzerrten Therapieeffekt, weil wir die Randomisierung faktisch zerstört haben. Das ist auch keine randomisierte Studie mehr, die wir in der Per-Protokoll-Analyse angucken. Ich glaube, das ist das große Problem. Sie schreiben in Ihrer Stellungnahme auch Dinge der ITT-Analyse zu, was die Rekrutierung angeht, die mit ITT erst mal nichts zu tun haben. Das hat höchstens etwas mit der Patientenzahl zu tun. Aber wenn ich in diesem konkreten Fall einen Großteil der Patienten aus den randomisierten Gruppen entferne und nicht weiß, was das Ergebnis sein wird, dann ist es

keine randomisierte Studie mehr. Insofern gibt es keinen sauberen Therapieeffekt. Das ist das große Problem.

Christoph Thomssen: Ich möchte noch einmal auf die Meta-Analyse hinweisen. Auch dort ist die Hazard Ratio ja in der gleichen Richtung. Die Meta-Analyse hat insofern einen besonderen Wert, als wir in der Chemo-N0-Studie eine endokrine Therapie gar nicht zugelassen hatten. Das heißt, die Patienten haben wirklich sauber entweder Chemotherapie bekommen oder gar nichts. In der Meta-Analyse sind natürlich auch Mischkollektive von Patienten dabei, die sowohl endokrine Therapie als auch Chemotherapie bekommen haben. Und dennoch lässt sich in der Meta-Analyse sehr schön dieser Effekt für die Chemotherapie herausarbeiten, der sehr deutlich ist, also Hazard Ratio 0,5 für hohes uPA/PAI mit Chemotherapie versus 0,8 für die anderen Arme. Das ist ein deutlicher Effekt, und die Konfidenzintervalle haben sich auch nicht überlappt. Deswegen, denke ich, spielt für die Gesamtbewertung der Prädiktion auch hier die Meta-Analyse noch eine starke Rolle.

Was ist das Ziel unserer Sitzung? - Das Ziel der Sitzung ist doch, zu finden, ob wir mit solchen Markern a) Patienten Chemotherapie ersparen können und b) vielleicht auch vorhersagen können, ob die Chemotherapie einen Effekt hat. Wenn wir es mit randomisierten Studien nicht hinbekommen, weil es heutzutage auch gar nicht mehr gehen wird in einem solchen Design, dann müssen wir uns überlegen, ob wir es aufgrund der bisherigen Daten akzeptieren können, einen solchen Marker für so eine Fragestellung anzunehmen oder nicht. Das ist doch das Ziel vom Ganzen. Es ist doch nicht das Ziel vom Ganzen, uns gegenseitig zu zeigen, wer wissenschaftlich besser argumentieren kann.

Moderator Jürgen Windeler: .. auch, vielleicht.

Der Punkt, den Sie zum Schluss genannt haben - das muss ich gestehen -, beschäftigt mich. Sie sagten, heutzutage geht das nicht mehr. Das hat ja zwei Hintergründe. Der eine Hintergrund kann sein, heute ist alles anders. So ein bisschen schimmert das aus Ihrer Äußerung heraus.

Christoph Thomssen: Die Therapie ist anders.

Moderator Jürgen Windeler: Dann frage ich mich, was uns die Studien von 1999 überhaupt noch sagen.

Der andere Punkt ist - das kann ja auch ein Hintergrund Ihrer Äußerung sein -, dass heute der Stellenwert von uPA/PAI so etabliert ist, dass man da nichts mehr randomisieren oder überprüfen kann. Dann frage ich mich natürlich bei dieser Datenlage, woher diese Wertschätzung von uPA/PAI eigentlich kommt.

Die Frage, das geht heute nicht mehr, wobei ich nicht dafür plädiere, heute eine Studie zu machen, aber wenn wir gemeinsam feststellen, dass die Studienlage so aussagefähig nicht ist

... Ich frage mich eben, was steckt in der Tat hinter der Frage, die so ein bisschen heißt, das kann man heute eigentlich gar nicht mehr fordern.

Christoph Thomssen: Ich meine, man kann keine Studie mehr fordern, bei der wir einen Therapiearm ohne Chemotherapie versus einen Therapiearm mit Chemotherapie vorschlagen, weil das heutzutage nicht mehr machbar ist.

Tanja Fehm: Aus Sicht der AGO ist es so, dass uPA/PAI-1 - das ist ja TOP 3 - eigentlich ein etablierter prognostischer Marker ist, Level of Evidence 1. Und natürlich ist es so, dass in der klinischen Routineversorgung, wenn man sich hinsichtlich der Prognoseeinschätzung unsicher ist, in vielen Bereichen uPA/PAI-1 routinemäßig mitgemacht wird - das muss man ganz klar sagen -, nicht als prädiktiver Marker, aber als prognostischer Marker. Somit wird es schwer sein, so eine Studie aufzulegen, wenn jemand mit einem hohen Risiko nach uPA/PAI-1 dann keine Chemotherapie erhält, wenn das im Sinne einer, was Sie sagen, Studie evaluiert werden soll. Dazu ist er als prognostischer Marker zu sehr etabliert. Und die Prognose bestimmt auch die Indikation zur Chemotherapie. Ganz klar.

Moderator Jürgen Windeler: Weitere Wortmeldungen zu TOP 2? - Dann schließen wir diesen Punkt ab. Wir haben an dieser Stelle die Positionen auch ausgetauscht.

Wir kommen zum

4.3.4 Tagesordnungspunkt 3: Prognostische Evidenz im Projekt

Martina Markes: In verschiedenen Stellungnahmen steht, dass eine Reduktion des Risikos für ein Rezidiv durch Chemotherapie erst ab einem erhöhten Rezidivrisiko klinisch relevant ist und demzufolge Patientinnen mit einem niedrigeren Rezidivrisiko eine Chemotherapie erspart werden könne. Unsere Frage ist: Bei welchem Rezidivrisiko liegt diese Grenze, und gibt es irgendwelche Konventionen oder Leitlinien, die Sie uns nennen können, die Aussagen dazu machen, welches Rezidivrisiko entsprechend gering ist? Wie wird dort gegebenenfalls diese Grenze ermittelt? Es geht bei dieser Frage also um die Grenze.

Christoph Thomssen: Punkt 1: Diese Grenze ist zeitlich schwankend. Die Konvention war 1990 anders als im Jahre 2014.

Punkt 2: Das ist natürlich von Individuum zu Individuum schwankend. Als Ärzte dürfen wir da nicht paternalistisch herangehen, wie Herr Gold schon mal gesagt hat, und vorschreiben, was gemacht wird und welches Risiko akzeptabel ist, sondern das ist natürlich individuell unterschiedlich.

Wenn wir Patienten in Sachen Chemotherapie etwas empfehlen wollen - insofern ist die patient shared decision making trotzdem immer noch ein bisschen eine Beschreibung eines Zustandes, wo der Arzt mit seiner Meinung einen ganz erheblichen Impact hat -, dann werden wir uns schwertun, wenn wir 2 % oder 1 % Survivalbenefit durch eine Chemotherapie erreichen, einer Patientin unbedingt zuzuraten, eine solche Chemotherapie zu machen. Wir

müssen immer daran denken, die Chemotherapie reduziert ja im besten Fall mit ganz modernen toxischen Therapien 36 % bis 40 % des Rezidivrisikos, also etwa zwei Fünftel der Mortalitätsrate auf 15 Jahre. Wenn wir also ein Ausgangsrisiko von unter 5 % haben, dann sind zwei Fünftel genau diese 2 %. Deswegen taten wir uns jedenfalls 1993 schwer, bei einem solch niedrigen Rezidivrisiko überhaupt eine Chemotherapie auch in randomisierter Form in Betracht zu ziehen. Damals waren 10 % Mortalität nach 10 Jahren der Effekt, den man sich überlegt hatte als Chemotherapiebenefit.

Stefan Sauerland: Jetzt muss man sich sehr stark um die exakten Zahlen kümmern und überlegen, ob die Zahlen wirklich in all diesen Aspekten tragfähig sind.

Das eine ist: Was heißt überhaupt Niedrigrisiko, wenn also der uPA/PAI-Test negativ ist?

Der zweite Punkt ist, wie weit sich jetzt wirklich diese Grenze beziehen lässt auf eine Präferenz einer Frau. Sie haben ja gesagt, dass es hier nicht um paternalistische Entscheidung geht, sondern dass die Frau mit einbezogen werden sollte. Es gibt aber eine ganze Reihe von Studien, die zeigen uns, dass, wenn man Frauen, die eine Chemotherapie hinter sich haben, fragt, ob sie denn – als theoretische Überlegung - bereit sind, eine solche Chemotherapie noch mal zu durchlaufen für einen Überlebensvorteil von etwa 1 % oder 5 % oder auch nur 1 Promille, ein großer Teil der Frauen auch für 1 % - das ist die absolute Risikoreduktion beim Überleben - , ganz grob die Hälfte, bereit ist, auch in diesem Fall eine Chemotherapie auf sich zu nehmen. Das würde dagegensprechen, dass es hier eine klare Grenze gibt. Und es würde auch dagegensprechen, dass man bei einer großen Zahl von Frauen mit einem nach uPA/PAI-Test niedrigeren Risiko sagen kann, das ist jetzt für uns entscheidend, weil die Entscheidung liegt offenbar an anderen Faktoren und ist eben nicht in einem so derartig niedrigen Risikobereich, als dass jetzt für die Mehrzahl der Frauen völlig sonnenklar ist, dass hier eine Chemotherapie gar nicht mehr infrage kommt.

Tanja Fehm: Man muss ja ganz klar sagen aus der klinischen Routine heraus: Die Chemotherapie ist ja potentiell eine Therapie, die für die Patientinnen tödlich enden kann, wenn sie eine schwerste Komplikation hat. Das ist ja genau das, worüber wir mit der Patientin im Low-Risk-Bereich sprechen, der fließend ist. Wir haben ja früher Adjuvant! Online genommen, letztendlich um mit den Patienten in diesem intermediären-Risikobereich zu diskutieren. Man muss sagen - den Bereich 2 % bis 4 %, 5 % haben wir oft als Grenze angesetzt -, die Chemotherapie hat 1 % bis 2 % schwerwiegende Komplikationen zu dem Preis von vielleicht 1 % bis 2 % Überlebensbenefit, damit die Patientin die Entscheidung treffen kann. Aber diese Gruppe, die Sie zitieren, ist sicher sehr, sehr gering für 1 %. Ich kenne diese Studie auch. Aber die Patientinnen haben wir häufig unmittelbar nach der Chemotherapie befragt. Wie wäre es denn, wenn man die drei oder vier Jahre hinterher befragt hätte, ob sie dann immer noch sagen würden, sie würden es nicht machen oder sie würden es machen? Deswegen darf man nie vergessen, dass die Chemotherapie - das muss mit in ein Aufklärungsgespräch einfließen - potentiell tödlich sein kann. Das sind genau die 1 % bis 2 % gerade bei den modernen Chemotherapien mit schweren Komplikationen, wo die

Patientin ihre Entscheidung relativiert, wenn sie weiß, dass ihr Rezidivrisiko bei 1 % liegt unter dem Aspekt der Nebenwirkungen der Chemotherapie. Das fließt häufig bei den Diskussionen mit den Patientinnen nicht mit ein.

Stefan Sauerland: Eine kurze Nachfrage dazu: Sie sagen, man hätte die Frauen nicht nach der Chemotherapie befragen sollen, sondern vielleicht nach drei Jahren oder vor der Chemotherapie. Aber gerade wenn doch die Frau die Chemotherapie frisch hinter sich hat, würde man doch erwarten, dass dann die Abneigung der Chemotherapie gegenüber am größten ist, dass also der Überlebensvorteil, den sich eine Frau erhofft oder den eine Frau als akzeptabel erachtet, deutlich höher ist. Also wäre doch der realistische Bereich sogar unter 1 %, vielleicht im Promillebereich, wo eine Frau eine Chemotherapie für sich als akzeptabel empfindet.

Tanja Fehm: Ich denke, dass die Nebenwirkungen, die Komplikationen der Chemotherapie erst später zum Tragen kommen, gerade bei den modernen Chemotherapien.

Ich glaube, dass das gar nicht so sehr die Frage ist, sondern die Frage ist vielmehr, welchen Bereich wir als Low-Risk-Bereich definieren. Das wollte ich damit zum Ausdruck bringen, dass der Übergang natürlich fließend ist und dass es individuell ist und jede Frau für sich selber unterschiedlich definiert, welchen Preis sie für die Chemotherapie letztendlich zahlen müsste. Da kann natürlich das uPA/PAI-1 das noch mal präzisieren für die Patientin, wie ihr Risikobereich ist. Wir können sagen, mit Adjuvant! Online war ja früher ein Bereich, da haben wir noch mal ein zusätzliches Tool, was in Adjuvant! Online nicht mit eingeht, wo man den Low-Risk-Bereich noch mal präzisieren kann und sich die Patientin dadurch noch mal leichter in der Entscheidungsfindung tut.

Rolf Kreienberg: Ich habe eine Frage. Wir haben uns die Studien angesehen hinsichtlich des niedrigen Rezidivrisikos. Das schwankt ja sehr stark. Wo ist denn tatsächlich die Grenze, wenn uPA und PAI-1 niedrig ist? Wo liegt die denn tatsächlich? Denn die schwankt ja in den verschiedenen Studien erheblich zwischen 5 und 14 %.

Christoph Thomssen: In der ersten Pilotstudie in München war das ja nur ein Zentrum. Da lag die Rezidivrate bei 5 % nach fünf Jahren. In der Feldstudie lag sie, glaube ich, bei etwa 12 %. Man muss da allerdings davon ausgehen, dass die Patientinnen kein Tamoxifen bekommen haben. Heutzutage würden die Patientinnen ja Tamoxifen bekommen, womit das Risiko deutlich niedriger wäre.

Was wir auch machen müssen, ist, wir müssen uns natürlich auch das Umfeld anschauen, wenn wir die Tumorgröße nehmen oder den Nodalstatus oder Hormonrezeptor-Status oder HER2-Status als Risikofaktor annehmen. Dann liegen wir auch immer in dem Bereich und werden da nicht besser. Insofern muss man sagen: Mit uPA/PAI ist gerade in dem Kollektiv, was wir untersucht haben, die Risikoabschätzung am besten möglich gewesen. Wenn wir das kombinieren mit einer endokrinen Therapie, was ja in den Studien bisher nicht gelaufen ist,

dann kommen wir in den Bereich, wo es sehr mühsam wird - deutlich unter 10 %, irgendwo bei 5 %, 6 % Rezidivrisiko -, über Chemotherapien zu sprechen. Denken Sie daran: Die Oncotype-Analysen nennen das Niedrigrisikokollektiv dann, wenn sie unter 12 % Fünfjahresrezidivrate haben. Das ist bei den Amerikanern Niedrigrisiko. Wenn Sie sich die Kurven von Oncotype-Score anschauen, dann sehen Sie es bei 12 %. Das ist der Cut-off für Niedrigrisiko. So unterschiedlich wird es auch betrachtet. Wenn wir jetzt mal die anderen Marker auslassen, sondern nur mal die traditionellen Marker nehmen und uns da überlegen, was Niedrigrisiko und Hochrisiko ist, dann muss man sagen, dass wir mit uPA/PAI in einem sehr guten Bereich liegen, was auch mit Grading, G1, möglich ist, was mit Tumorgröße schlechter möglich ist, was mit Nodalstatus noch schlechter möglich ist. Aus dem Grunde haben wir am Schluss gesagt: Eigentlich interessiert uns vor allen Dingen die Gruppe G2, nämlich die Patienten, wo wir mit anderen Markern nicht weiterkommen.

Stefan Sauerland: Herr Thomssen, die Frage ist ja jetzt nicht, ob sich die prognostische Einschätzung verbessern lässt, sondern die Frage ist, ob die prognostische Einschätzung in einen Bereich hineinführen kann bei einer Mehrzahl der Frauen, sodass tatsächlich das Risiko eines Rezidivs so gering ist, dass sich die Frage Chemotherapie ja oder nein von alleine beantwortet. Da haben wir zum einen die Unsicherheit, wie hoch das Rezidivrisiko in der Niedrigrisikogruppe überhaupt ist nach uPA/PAI-Testung - und das schwankt -. Und das Zweite ist die Grenze, die man unterschreiten muss. Und da sagte Frau Fehm, dass das natürlich schwammig ist und dass das natürlich von der individuellen Entscheidungsfindung der Frau abhängt. Das führt dazu, dass wir offenbar keine klare Verbindung zwischen diesen beiden Zahlen herstellen können, die es uns erlaubt, irgendwie regelhaft einen Zusammenhang zwischen einem negativen uPA/PAI-Test und der therapeutischen Konsequenz und dem Effekt einer weggelassenen Chemotherapie herzustellen.

Claudia Olenik: Es ist tatsächlich so, dass wir ungefähr 7.000 Patientinnen pro Jahr mit uPA/PAI-1 testen. Die Empfehlung wird ja entsprechend dann auch von den Klinikern ausgesprochen. Die Patientin kann sich das natürlich noch anders überlegen.

Wie Frau Fehm schon sagte, es wird immer Patientinnen geben, die selbst bei 0,01 % Nutzen sagen, noch eine Chemotherapie drauf. Vielleicht ist das auch nur eine Auswahl der Patientinnen, mit denen ich gesprochen habe. Ich bin nicht in der Klinik. Es gibt aber immer eine extreme Angst vor der Chemotherapie. Frau Fehm nannte auch die Langzeitwirkung. Das ist ja oft gar nicht bekannt, die kognitiven Störungen, die wirklich jahrelang anhalten können, das Fatigue-Syndrom usw. Klar, die unmittelbaren Probleme, Haarverlust usw., während der Chemotherapie, aber das, was Jahre später auch noch kommt, selbst wenn es keinen tödlichen Ausgang hat ... Die Lebensqualität ist teilweise derart heruntergesetzt, weil die Patientinnen so müde, so kraftlos sind. Ein Großteil der Patientinnen - ich kann das nicht in Prozent beziffern - hat eine extreme Angst davor. Die sind wirklich dankbar um die Parameter, und die würden auch viel Geld bezahlen. Das zeigt ja schon, wie viel ihnen das Wert ist. Die zahlen Tausende von Euro für irgendwelche Tests, damit sie ein besseres Hilfsmittel haben für eine bessere Entscheidung, was sie tun sollen. Und die Ärzte nehmen ja dankend den

uPA/PAI-1-Test an. Sonst würde er ja nicht so weitläufig in Deutschland eingesetzt. Und auch in Frankreich wird er eingesetzt.

Moderator Jürgen Windeler: Ich glaube, wir sollten von solchen Vermarktungsstrategien Abstand nehmen. Wir machen hier wissenschaftliche Erörterungen. Ich glaube, die Frage, ob Menschen in höchster Not etwas bezahlen würden, als Kriterium dafür, dass es etwas taugt, heranzuziehen, ...

Claudia Olenik: Damit war nicht unser Test gemeint.

Moderator Jürgen Windeler: Das haben Sie aber gesagt.

Claudia Olenik: Nein. Das möchte ich jetzt richtigstellen.

Moderator Jürgen Windeler: Dann machen Sie das.

Claudia Olenik: Die Not ist so groß, dass die Patientinnen sogar hingehen, Tests, die noch nicht empfohlen werden von Leitlinien ... Das sind nicht die Tests, von denen wir hier sprechen, die hier erörtert werden. Die haben so ein dringendes Bedürfnis, keine Chemotherapie zu erhalten, dass sie sich sogar an Hersteller wenden, die noch nicht mal empfohlen sind.

Moderator Jürgen Windeler: Danke für die Richtigstellung.

Der zweite Punkt ist, dass ich mit Interesse zur Kenntnis nehme, dass die Frage der Nebenwirkungsbewertung dort, wo es um Arzneimittel geht, die Nebenwirkungen haben, ganz anders läuft als in dieser Sitzung hier. Dort, wo es um Arzneimittel geht, etwa im Zusammenhang mit den frühen Nutzenbewertungen, wo Nebenwirkungen, und zwar gravierende Nebenwirkungen, auftreten, wird von den Klinikern und von Experten in der Regel gesagt: Die sind beherrschbar. Die sind nicht so schlimm. Das geht alles vorbei. Das hat keine relevanten Auswirkungen auf den Haupteffekt, also etwa Veränderungen der Lebensqualität. - Ich sehe das, was die Einschränkung der Lebensqualität angeht durch Chemotherapie, genauso wie Sie. Daran gibt es wenig zu deuteln. Insofern macht es, glaube ich, Sinn bei dem, was wir gerade diskutieren, abzuwägen, ob man die Chemotherapie wirklich braucht oder nicht.

Ich würde jetzt gerne auf die prognostische Evidenz kommen. Die prognostische Evidenz könnte für uns - wenn überhaupt - so etwas sein wie eine Hilfskonstruktion im Zusammenhang mit randomisierten Studien, weil man natürlich in einer Situation, wo man zuverlässig - darauf zielen auch die Fragen von Frau Markes und Herrn Sauerland - sagen würde, da kommt eine Chemotherapie - ich sage es etwas überspitzt - nach menschlichem Ermessen nicht mehr infrage, über die Frage der randomisierten Studien und darüber, ob es vielleicht doch geht, keine Gedanken mehr machen muss, aber eben nur dann, wenn man wirklich sicher ist, die kommt nicht mehr infrage. Deswegen ist auch die Frage der Grenze

und die Frage der Zuverlässigkeit der Grenze und die Frage, ob ab einer bestimmten Grenze diese Chemotherapie wirklich nicht mehr infrage kommt - über subjektive Patientenreferenzen, die dann vielleicht doch zu wollen, reden wir im Moment nicht - , von sehr zentraler Bedeutung, weil sie darüber entscheidet, ob wir sagen würden, jawohl, für diese Situation braucht man wirklich keine Studien mehr oder man braucht sie doch.

Rolf Kreienberg: Herr Windeler, Sie haben eigentlich alles vorweggenommen, was ich sagen wollte. Das finde ich gut.

Man sollte es nicht mit subjektiven Argumenten unterstützen, sondern es geht um die Objektivität und wirklich um die Frage der unteren Grenze: Wann kann man Chemotherapie weglassen, und ist das Testsystem in der Lage, das mit einer gewissen Sicherheit zu definieren? Alles andere ist unwichtig. Dass Patienten das natürlich wollen und nach jedem Strohalm greifen, den sie bekommen können, ist völlig klar. Darum geht es hier aber nicht, sondern es geht darum, ob man ein Testsystem aussucht und empfiehlt, das die untere Grenze mit einer an Sicherheit grenzenden Wahrscheinlichkeit definiert. Das ist die richtige Frage.

Die andere Frage ist dann umgekehrt: Die anderen bekommen dann Chemotherapie obendrauf. Und in welchem Setting geht das? Darum geht es. Das war vorhin meine Frage: Welcher Grenzwert gibt uns aus den bisher vorliegenden Studien diese Aussage?

Moderator Jürgen Windeler: Ich möchte das ergänzen, was Herr Kreienberg gesagt hat. Diese an Sicherheit grenzende Wahrscheinlichkeit bezieht sich darauf, ob Studien gemacht werden müssen oder nicht, nicht darauf, was in der Praxis tatsächlich passiert. Das letztere ist jetzt nicht die Frage hier, sondern die Frage ist, ob wir in eine Situation kommen, wo die Erwägung, sollte man Chemotherapie machen oder nicht, keine Rolle mehr spielt. Denn da würden wir sagen: Dann brauchen wir auch keine Studien zu fordern. Es geht nicht darum, ob das in der Praxis genauso gemacht wird.

Christoph Thomssen: Zu dem ersten Punkt, bei Arzneimitteln werden die Nebenwirkungen kleingeschrieben: Da sprechen Sie natürlich jetzt den Falschen an, weil wir ja da gerade mit der wissenschaftlichen Intention herangegangen sind, wie wir Therapien ersparen können. Ich kann nur sagen: In unserer wissenschaftlichen Diskussion - Frau Fehm wird mir da recht geben - diskutieren wir eigentlich nicht - das sagen wir den Firmen auch - über die Grad-3-Nebenwirkungen, weil das die Nebenwirkungen sind, weswegen die Patienten ins Krankenhaus müssen, sondern wir reden über die Grad-2-Nebenwirkungen, weswegen die Patienten Beschwerden haben. Das muss man ganz klar feststellen. Aber ich denke, dass die Nebenwirkungen hier jetzt nicht Thema sind, sondern es geht um die Frage der tatsächlichen und erreichbaren Überlebensraten.

In den großen deutschen Brustzentren haben wir derzeit 7 % Todesfälle nach fünf Jahren, also 93 % überleben das nicht-metastasierte Mammakarzinom. Ich glaube, das ist eine ziemlich gute Zahl. Es sind, glaube ich, 87 % nach zehn Jahren, die das überleben. Das weiß ich von

Heidelberg, von Hamburg, von unserem eigenen Kollektiv. Bei euch ist das, glaube ich, auch so.

In der Chemo-N0-Studie haben wir ohne Tamoxifen in der Low-Risk-Gruppe 12 bis 13 % Rezidive nach zehn Jahren und knapp 10 % Todesfälle über alles, da sind auch die Herzinfarkte dabei. Das heißt, wir liegen in der Low-Risk-Gruppe ohne Therapie schon sehr in der Nähe der Gruppe von Patienten, die eine maximale Therapie bekommen haben. Es ist die Frage, ob man das weiter steigern kann. Wenn wir bei der Low-Risk-Gruppe noch Tamoxifen hypothetisch dazugegeben hätten, kann man sicher davon ausgehen, dass man auf zehn Jahre noch mal 40, 50 % der Todesfälle hätte einsparen können. Dann lägen wir irgendwo bei 94 % Survival. Das ist auch in den Kurven gut sichtbar, weil die Todesfälle erst ab dem Jahr acht aufgetreten sind. Das sind die typischen Zeitpunkte, wo wir bei hormonrezeptorpositiven Patienten die Ereignisse erwarten würden. Die bekommen ihre Rezidive und ihre Todesfälle sehr spät. Das passt auch eigentlich ganz gut dazu. Nach acht Jahren liegt die Survival-Rate bei 98 %.

Was ist das, was Patienten von uns erwarten abgesehen von ewigem Leben? - Ich denke, dazu ist schon einiges gesagt worden. Patienten haben dazu ihre subjektive Meinung. Ansonsten muss man sich auf Konventionen einlassen, und die liegen irgendwo im Bereich von 3 bis 5 % über dem natürlichen Überleben wahrscheinlich. Weit darunter werden die Effekte wirklich sehr gering. Da liegen wir bei 1 % bis 2 % absoluter Benefit zusätzlich. Da werden Sie in der Community wenige Leute finden, die da noch empfehlen, eine Chemotherapie zu machen.

Tanja Fehm: Das ist genau das, was auch ich gesagt habe. Sobald die tödlichen Nebenwirkungskomplikationen der Chemotherapie den Benefit nicht ausgleichen, empfehlen wir auch keine Chemotherapie. Bei 1 bis 2 % würden wir definitiv keine Chemotherapie empfehlen. Häufig nehmen die Brustzentren einen Wert um 5 %, wo man das als Option für die Patientinnen diskutiert. Diese Grenzwerte würden Sie nirgendwo so scharf finden. Das sind einfach Konventionen, wo man sagt: Okay, ab dem Bereich sind die Nebenwirkungen der Chemotherapie so stark, die auftreten können, dass der Nutzen nicht mehr gegeben ist. Und dann gibt es einen Graubereich, wo man es optional diskutieren kann. Die meisten ab 5 % sprechen es dann etwas konkreter an. Aber das sind keine hart definierten Grenzen. Das muss man bei dem Punkt auch sagen.

Christoph Thomssen: Eine Ergänzung: Wir haben ja damals CMF geprüft. CMF ist ja in der Therapie ein bisschen lästig, aber danach sind die Nebenwirkungen vergessen. Heutzutage geben wir in der Regel Anthracycline und Taxane. Wenn wir Chemotherapie sagen, dann machen wir nicht ein bisschen Chemotherapie, sondern richtig Chemotherapie, weil wir auch den richtigen Effekt haben wollen. Taxane und Anthracycline machen aber Langzeitnebenwirkungen, Anthracycline in einem sehr hohen Prozentsatz im späteren Leben schwere Herzinsuffizienz, die nicht reversibel sind - das hatten wir gerade im Ärzteblatt publiziert -, und Taxane machen eine schwere Polyneuropathie in vielen Fällen, die

subklinisch andauert - die geht nicht mehr weg; das ist schlimmer als eine diabetische Neuropathie; daran muss man denken -.

Und wenn wir von Chemotherapie sprechen ... Ich kann mich, Herr Sauerland, nicht mehr an die Studie erinnern, welche Chemotherapien die bei den Patienten untersucht haben. Ich kenne die Daten auch. Ich weiß jetzt gar nicht, welche Chemotherapien die untersucht hatten. Wenn Sie daran denken, welche Langzeitnebenwirkungen letztendlich da entstehen, dann muss man schon sagen: Die Zahlen, die Frau Fehm und auch ich hinten angegeben haben, liegen im einstelligen Prozentbereich. Viel mehr Effekte können wir den Frauen nicht zumuten.

Stefan Sauerland: Es sind ja nicht nur diese beiden Zahlen, die fraglich sind, also wo die Grenze ist und wie hoch das Risiko in der Niedrigrisikogruppe ist, sondern Sie brauchen für Ihre Kalkulation, die Sie gerade gemacht haben, immer auch noch die Annahme, dass das Hazard Ratio, also die relative Reduktion eines Rezidives, in den beiden Gruppen gleich ist. Wenn Sie von 5 % Rezidivrisiko ausgehen und dann daraus berechnen, dass sich das absolute Risiko um 1 bis 2 % verändern wird, dann setzen Sie implizit voraus, dass das Hazard Ratio, was Sie in der Gesamtgruppe bei uPA/PAI hoch, uPA/PAI niedrig gesehen haben, über alle Strata konstant ist. Und da sind wir auch ein wenig am zweifeln, ob denn dafür überhaupt Daten vorliegen, gerade wenn Sie sagen, dass sich die Chemotherapieregime im Laufe der Zeit natürlich geändert haben und dann natürlich noch eine weitere Komponente der Variabilität mit hineinkommt.

Christoph Thomssen: Wir reden ja vom Proportional Hazard Ratio, was über die Zeit konstant ist, und nicht von time-varying Hazard Ratios. Insofern müssen wir mal von der Annahme ausgehen, dass wir das Risiko über die Zeit konstant reduzieren. Wir haben natürlich auch Daten zur zeitlichen Variierung des Risikos. Zur Risikoreduktion haben wir keine. Da wissen wir nur aus den allgemeinen Chemotherapiestudien, frühe Rezidive werden besser verhindert als späte Rezidive. Das gilt mutmaßlich auch für die uPA/PAI-Patienten. Aber insgesamt denke ich, dass wir immer ein Log-Rank Test über das Gesamte machen.

Ralf Bender: Nein, es geht nicht um die Variation über die Zeit, sondern es geht um die Variation zwischen verschiedenen Subgruppen. Ob da wirklich das gleiche Hazard Ratio für alle Subgruppen gilt, das wissen wir nicht. Man nimmt es an, weil man keine Daten hat, das Gegenteil zu belegen. Aber Stefan Sauerland hat nur darauf hingewiesen, ein Bestandteil der Argumentationskette ist, dass dieses Hazard Ratio auch in der Niedrigrisikogruppe gilt. Die Argumentation würde sofort zusammenbrechen. Wenn da - dafür haben wir natürlich auch keine Daten, keine Belege, nur hypothetisch - der Relativeffekt sogar größer wäre in der Niedrigrisikogruppe, dann könnte man diese ganze Argumentation nicht fahren.

Christoph Thomssen: Drei Punkte:

Erstens. Wir reden hier über die Prognose und nicht über die Prädiktion.

Zweitens. Wenn wir uns jetzt schon auf die Prädiktionsnummer einlassen: In der Oxford Meta-Analyse, Richard Peto, Chemotherapie usw., über 100.000 Patienten, konnte sehr schön gezeigt werden, dass die Wirkung von der Chemotherapie ziemlich unabhängig ist von den Risikogruppen, was die relative Risikoreduktion betrifft. Wir haben ja die Hypothese für unseren Marker, dass wir für uPA/PAI hoch eine höhere relative Risikoreduktion haben als für uPA/PAI niedrig. Das ergibt sich zum Beispiel aus der schon mehrfach zitierten Meta-Analyse von Frau Harbeck München/Rotterdam. Da sehen wir sehr deutlich: Hazard Ratio bei uPA/PAI niedrig für Chemotherapie bei 0,8, Hazard Ratio für uPA/PAI hoch bei 0,5. Wenn wir diese Daten als die günstigsten Daten zugrunde legen, dann liegt der Effekt für die Reduktion der Mortalität oder des disease-free survival in der Niedrigrisikogruppe sogar noch niedriger, als wir im besten Falle annehmen würden. Wenn wir es nach der Oxford Meta-Analyse annehmen würden, dann würden wir sagen, etwa 30 bis 40 % Reduktion über 15 Jahre, wenn wir es nach der Meta-Analyse Harbeck ansehen, dann würden wir sagen, wahrscheinlich sogar nur 10 bis 20 % Reduktion. Da ist der Benefit sogar noch geringer einzuschätzen. Aus diesem Grunde - deswegen haben Sie mir eine wunderbare Vorlage gegeben - sind wir der Meinung, dass es eben keinen Sinn macht, einer Low-Risk-Gruppe in irgendeiner Form eine adjuvante Chemotherapie anzubieten.

Moderator Jürgen Windeler: Wenn Sie so wollen, haben Sie mir gerade eine Vorlage gegeben. Ich glaube, dass Sie jetzt noch mal sehr schön gezeigt haben, dass die Trennung in Prädiktion und Prognose überhaupt keinen Sinn macht. Sie argumentieren ja die ganze Zeit mit den Therapieeffekten in dieser Niedrigrisikogruppe, das heißt natürlich über die Prädiktion, dass man bei dieser Gruppe nicht mehr erfolgreich eine Therapie anwenden sollte. Es geht da wieder um die Therapieeffekte, nicht nur um die Prognose.

Ich würde an dieser Stelle noch mal sagen, was ich vorhin schon mal gesagt habe: Diese prognostische Evidenz an dieser Stelle wäre für uns eine Möglichkeit, unter bestimmten Rahmenbedingungen - das betrifft die Schwelle, das betrifft die Größe des Therapieeffektes - zu sagen -, es gibt einen Bereich, wo die Frage, und zwar möglichst empirisch gestützt, sollte man diese Frauen therapieren oder nicht, faktisch entschieden ist, keine Rolle mehr spielt. Dann würde sich unsere Forderung, darüber müssten wir jetzt Studiendaten haben, in diesem Bereich erledigen. Prognostische Daten alleine machen für uns, wenn man sozusagen die Entscheidung davon abkoppelt, überhaupt keinen Sinn, und die würden wir auch nicht betrachten wollen. Wenn man sie betrachtet, dann kommen Sie aber sofort in den Bereich der Prädiktion.

Christoph Thomssen: Die ursprüngliche Hypothese war ja mit Zahlen von 1993, dass wir gesagt haben: Bei nodalnegativem Mammakarzinom sind 70 bis 80 % der Patienten alleine durch die Operation geheilt. Das heißt, wenn wir Chemotherapie machen, würden diese 70 % der Patientinnen die Chemotherapie überflüssig bekommen.

Zweitens wissen wir, dass von den restlichen 30 % nur 10 % tatsächlich profitieren werden. Das heißt, weitere 20 % absolut würden die Therapie nicht überflüssig, sondern vergeblich

bekommen. Das ist ja die Hypothese, von der wir ausgehen. Deswegen haben wir gesagt: Da wir mit unseren bisherigen Prognosefaktoren diese natürlichen Krankheitsverläufe sehr schlecht abbilden können, brauchen wir neue Prognosefaktoren. Und mit dem uPA/PAI sind wir immerhin so weit gekommen, dass wir sagen können: 50 % der Patienten könnten wir tatsächlich dem Low-Risk-Kollektiv zuordnen, sodass wir sagen können: Das sind Patienten, denen mutmaßlich die Chemotherapie erspart werden kann, weil sie eben alleine durch die Operation geheilt sind. Für die anderen 50 % haben wir jetzt noch keinen besseren Marker, aber für die nehmen wir schon mal an, dass das höhere Risiko aufgrund der Daten, die wir haben, da ist. Das ist sozusagen der Versuch, näher zu kommen. Da muss man sagen: Es gibt keinen anderen Faktor, der dem natürlichen Verlauf so gut nahegekommen ist in der Abbildung.

Was die Therapieeffekte betrifft - das war ja ein bisschen von Ihnen in die Diskussion eingebracht worden -: Wir sagen ja, wenn wir 50 % der Patienten aufgrund ihrer guten Prognose die Therapie ersparen können, dann müssen wir uns auch nicht über Therapieeffekte unterhalten. Und die Therapieeffekte diskutieren wir nur, um noch mal klarzumachen, wenn wir eine gute Prognose haben, dann sind die Therapieeffekte eben so tiny, dass wir angesichts der Nebenwirkungen, auch wenn der Marker nicht uPA/PAI heißen würde, sondern Nodalstatus oder Grading oder so etwas, keine Chemotherapie empfehlen. Und wenn ich es sagen darf: Für das Grading haben wir ja Daten, wo wir sagen, zumindest in der Studie von immerhin 13 Zentren, sprich 13 verschiedenen Pathologen, dass Grading so ein starker Faktor war, dass wir sagen müssen: Das war der stärkere Faktor von allen Faktoren. G1-Patienten laufen so gut, dass wir keine Chemotherapie brauchen - es waren nur 2 Rezidive aufgetaucht -, und G3-Patienten laufen so schlecht, dass wir für diese beiden Gruppen eigentlich einen Prognosefaktor nicht mehr brauchen. Da ist die Entscheidung klar. G3 laufen so schlecht wie nodalpositive, bekommen eine Chemo oder eine Therapie in irgendeiner Form, und G1 laufen so gut, dass wir sagen, da passt es nicht mit der Chemotherapie. Das heißt, es bleiben in der Studie noch 70 %, heutzutage mit einem optimierten Grading vielleicht 50 bis 60 % der Patienten übrig, wo tatsächlich so eine Zusatzfrage notwendig ist.

Moderator Jürgen Windeler: Weitere Wortmeldungen?

Claudia Olenik: Nur einen Hinweis: Sie hatten sich, glaube ich, die Peto-Arbeit aufgeschrieben. Die ist sogar im Vorbericht zitiert.

Stefan Sauerland: Die ist bekannt.

Christoph Thomssen: Die dürfte Herrn Sauerland bekannt sein. Das ist sozusagen unsere Muttermilch.

Stefan Sauerland: Die wurde quasi durchgeführt zu einer Zeit, als es den uPA/PAI-Test noch nicht gab. Deswegen sind diese Daten nicht so einfach anwendbar.

Claudia Olenik: Weil Sie das Gesicht ein bisschen verzogen haben, habe ich gedacht, Sie ...

Stefan Sauerland: Nein, nein.

Claudia Olenik: Okay.

Moderator Jürgen Windeler: Weitere Wortmeldungen?

Fueloep Scheibler: Ich habe doch noch eine Frage: Auf welche Datenbasis würden Sie die Prognose insbesondere in diesem Kollektiv stützen wollen? Wenn wir sagen, wir würden tatsächlich Prognose akzeptieren, unter einer gewissen Prognose ist die Chemotherapie nicht mehr sinnvoll - ob das jetzt 2 %, 5 % sind, sei dahingestellt -: Auf welche Studie stützt denn jetzt Ihre Annahme, dass es tatsächlich so ist, dass bei uPA/PAI-negativen Frauen die Prognose unterhalb von ...

Christoph Thomssen: Wir haben zwei prospektive Studien. Wir haben die prospektive Studie München publiziert von Jänicke 1993, Breast Cancer Research und Treatment. Und wir haben die Chemo-N0-Studie. In der Vorpublikation haben wir sehr schön gezeigt, dass da die Kurven praktisch sehr kongruent laufen. Das heißt, zwei Studien bestätigen das. Eine Studie, eine Pilotstudie, hat das evaluiert prospektiv, und die zweite Studie hat das Ganze prospektiv im Feldansatz bestätigt, das heißt multizentrisch. Das sind die beiden Hauptstudien, auf die wir das stützen. Mit einer Studie alleine hätten wir nach Oxford schon Level 1B erreicht. Die brauchen nämlich eine inceptive cohort study, um einen prognostischen Marker als hochevidenzbewiesen anzusehen.

Wir haben das ja noch weiter ausgeführt - ich glaube, ich hatte es einmal dargestellt -, weil wir nach den Kriterien von Gary Clark von 1990 aus dem Prognosemekka San Antonio ... Er hat ein paar Kriterien aufgeführt, die damit losgehen: Ein Prognosefaktor muss ein biologisches Modell haben. Ein Prognosefaktor muss einfach zu bestimmen sein. Er muss in der Pilotstudie prospektiv evaluiert sein. Er muss in der prospektiven Studie auch in der multivariaten Analyse nicht korrelieren, sondern einen unabhängigen Effekt haben. Diese Punkte hatten wir ja alles mit München 1 abgearbeitet. Hinzu kam: Es muss bestätigt sein in einer großen multizentrischen Studie. Und der Therapieeffekt sollte irgendwann gezeigt werden. Und dann sollte noch die Meta-Analyse sein. Praktisch haben wir jedenfalls unserer Meinung nach alle diese Punkte abgearbeitet.

Es gibt noch weitere Studien, die von John Foekens aus Rotterdam, die in der gleichen Zeit, Anfang der 90er-Jahre, über uPA/PAI gezeigt worden sind mit jeweils 650 Patienten, die genau das Gleiche gezeigt haben. Und die Meta-Analysen von 15 oder 16 Studiengruppen aus Europa haben das letztendlich auch gezeigt, wobei da nicht sehr viele retrospektive Analysen hineingekommen sind.

Was die absoluten Zahlen betrifft, würde ich mich auf München 1 und Chemo-N0 im Wesentlichen beziehen wollen, und dann natürlich auf die NNBC3-Studie, wo wir versucht haben, das mit 4.500 Patienten zu analysieren, wo die Rezidivrate noch zu niedrig ist, um überhaupt etwas sagen zu können.

Fueloep Scheibler: Noch eine kurze Rückfrage. Bei der Kohorte in München waren wir uns nicht ganz sicher. Kann man tatsächlich die Patientencharakteristika so fein auseinanderdröseln, dass man sagen kann, wie der prognostische Wert für die Patientinnen ist, die uns eigentlich interessieren? Aufgrund der Publikationen hatten wir den Eindruck, man kann die Population nicht so genau ...

Christoph Thomssen: Das wird natürlich bei 374 Patienten oder wie viele darin waren, wo auch nodalpositive Patienten dabei waren, irgendwann fraglich, wenn die Gruppen sehr klein werden, dass man da noch etwas herausbekommt. Aber es hat sich in der Studie, zumindest in der multivariaten Analyse, auch gegenüber S-Phase, auch gegenüber Cathepsin D oder den traditionellen Faktoren, sehr klar gezeigt, dass uPA/PAI der stärkste unabhängige Faktor ist in der Subgruppe der Nodalnegativen.

Wenn wir jetzt Ihre Frage weiter spinnen, dann haben Sie natürlich recht. Vieles geht ins Pragmatische. Zunächst ist die Frage, warum wir uns auf G2 beschränkt haben. Das ergibt sich aus der Chemo-N0-Studie, weil das eben die Mittelgruppe ist. Das ist ja auch das, was wir letztendlich in der NNBC-Studie praktisch als Algorithmus vorgegeben haben. Der Algorithmus ist ja auch publiziert. Und wenn wir jetzt moderne Daten heranziehen, dann müssen wir sagen, dann brauchen wir nur noch die Entscheidung Chemotherapie ja oder nein. Und dann fallen die Triple-Negativen raus, weil die sowieso Chemotherapie bekommen. Es fallen die HER2-Positiven heraus, weil die sowieso Anti-HER2-Therapie mit Chemotherapie bekommen. Es bleiben dann die Rezeptorpositiven übrig, wo es um die Frage geht: Kann man unterscheiden zwischen luminal A und luminal B Typ, also zwischen rezeptorpositiv low-risk und rezeptorpositiv high-risk? Dazu wissen wir von der NNBC-Studie, auch publiziert dass wir zumindest in dieser Gruppe etwa halbieren können mit uPA/PAI. Da haben wir das Follow-up noch nicht. Das lag ja wegen der Ereignisse noch nicht vor. Und aus einer Kohortenstudie, die wir in Halle durchführen mit fünf anderen Zentren, wissen wir, dass wir das Risikokollektiv etwa halbieren können. Der Nutzen von damals passt also auch weiterhin.

Fueloep Scheibler: Noch eine Rückfrage: Bedeutet das, was Sie gerade geschildert haben, dass für diese Kohorte in München mit diesen 316 Patienten ja nur ein sehr geringer Teil letztendlich die Frauen sind, für die jetzt die ... Die rezeptorpositive Subgruppe ist ja dann wahrscheinlich nur ein Drittel dieser Kohorte.

Christoph Thomssen: Heutzutage würde ich sagen, für ungefähr 15.000, 20.000 Patienten, haben wir ausgerechnet, wäre es in Deutschland relevant, also etwa ein Drittel.

Fueloep Scheibler: Ist es möglich, diese Daten für dieses eine Drittel noch mal zu re-analysieren, also ob der prognostische Wert, der im gesamten Kollektiv gefunden wurde, sich auch zeigt für dieses Drittel, das heute eigentlich vor der Frage steht?

Christoph Thomssen: Jein. Wir sind dabei, die pathologischen Blöcke von der Chemo-N0-Studie zusammenzubekommen, aber das erweist sich als etwas schwierig, weil wir die HER2-

Bestimmung damals nicht gemacht haben. Und wenn wir sie gemacht hätten, wäre sie qualitätsmäßig auch nicht gut genug, als dass wir sie heute hätten verwenden können. Sie wissen ja: Damals haben wir 30 % HER2-Positive gehabt. Bei den Rezidiven liegen wir bei 15 %. Diese Daten müssten sozusagen erst geschaffen werden. München 1, da war das HER2 gerade erst erfunden. Da sind die Blöcke wohl auch nicht mehr vorhanden.

Für die aktuellen Studien, die wir in Halle als Kohortenstudien extra mit diesem Zweck haben laufen lassen, wissen wir, da haben wir den HER2-Status zum Beispiel. Da haben wir leider noch nicht die Prognose. Aber da wissen wir, dass wir solche Gruppeneinteilungen durchaus vornehmen können. Deswegen kenne ich die Zahlen auch, dass wir bei Rezeptorpositiven, HER2-Negativen etwa halbieren können.

Fueloep Scheibler: Da liegen die Daten schon vor?

Christoph Thomssen: Die sind ausgewertet, die haben wir nur noch nicht publiziert.

Fueloep Scheibler: Das ist für uns eine wichtige Information.

Christoph Thomssen: Das ist insofern auch wichtig, weil wir durchaus auch den Vorwurf und die Argumente bekommen haben, mit uPA/PAI würden wir möglicherweise nur die Triple-Negativen und die HER2-Positiven erkennen, und dann wäre das kein Zusatzeffekt. Das zeigt sich in dieser Kohortenstudie sehr deutlich, und auch in der NNBC-Studie zeigt es sich sehr deutlich, dass es über die Gruppen unterschiedlich ist.

Fueloep Scheibler: Wann erwarten Sie bei der NNBC-Studie die Ergebnisse, ganz konkret auf diese Fragestellung bezogen?

Christoph Thomssen: Da warten wir auf die Finanzierung für das nochmalige Follow-up, weil das ursprünglich nicht vorgesehen war. Da haben wir nicht damit gerechnet, dass wir so wenige Ereignisse haben im nodalnegativen Low-Risk-Kollektiv. Das ist das Problem. Wir haben jetzt etwa vier Jahre Follow-up, und jetzt müssen wir es etwa verdoppeln.

Fueloep Scheibler: Noch eine Frage: Wie groß ist die Hallenser Studie?

Christoph Thomssen: 1.200 Patienten.

Fueloep Scheibler: Und davon die Rezeptorpositiven? Oder sind das die 1.200?

Christoph Thomssen: 1.200 Patienten insgesamt. Davon sind ungefähr 40 % nodalpositiv.

Fueloep Scheibler: Okay.

Christoph Thomssen: Aus fünf Zentren. Somit auch wieder eine gewisse Abbildung der Realität.

Moderator Jürgen Windeler: Gibt es noch Wortmeldungen zu TOP 3? - Das scheint nicht der Fall zu sein. - Dann kommen wir zu dem ergänzten Punkt, nämlich

4.3.5 Tagesordnungspunkt 4: Verschiedenes

Wortmeldungen, Fragen, die Sie in diesem Zusammenhang mitgebracht haben, die Sie noch interessieren?

Claudia Olenik: Ich nehme Bezug auf den Vorbericht. Hinten wird ausgeführt, welche Studien ausgeschlossen wurden aufgrund zum Beispiel Studientyp. Da sind dann Leitlinien ausgeschlossen worden. Ich weiß nicht, ob das ein Versehen ist. Ich habe mich nur ein bisschen gewundert. Vielleicht können Sie mir das erklären. Auf Seite 50 ist ja „Nicht E5“, also Studientyp. Deshalb wurde es ausgeschlossen. Dann unter Nr. 13 sind die ASCO-Leitlinien ausgeschlossen worden. Das ist mir aufgefallen. Ich bin nicht alle Details durchgegangen.

Martina Markes: In unserem ersten Screeningschritt tauchte die mit auf, also im Titel-Abstract-Screening. Und wir wussten nicht genau, ob es sich um eine Leitlinie handelt oder ob innerhalb der Leitlinie vielleicht systematische Auswertungen oder weitere Studien drin sind. Im Volltextscreening haben wir dann erkannt, dass das keine Studie ist, die wir im Sinne unserer Einschlusskriterien in den Bericht einschließen würden und haben sie deshalb ausgeschlossen. Also als Leitlinie ist sie nach wie vor relevant, nur ist es keine Studie, die wir im Sinne der Einschlusskriterien in diesen Bericht einschließen konnten.

Claudia Olenik: So ist es zustande gekommen?

Martina Markes: Ja.

Claudia Olenik: Okay.

Moderator Jürgen Windeler: Weitere Wortmeldungen?

Christoph Thomssen: Vielleicht ein weiches Kriterium: Wir freuen uns ja, dass wir in Deutschland mal so einen Marker auf die Reihe bekommen haben. Und wir würden uns freuen, wenn wir uns in Deutschland nicht nur auf niederländische oder amerikanische oder kalifornische Marker beschränken müssten. Aber das ist ein weiches Argument.

Der wichtigere Punkt ist: Natürlich würde ich mich freuen oder würden wir erwarten, dass wir aufgrund der guten Daten zur prognostischen Unterscheidung eine Zulassung bekämen oder eine Zustimmung oder ein entsprechendes Zertifikat von Ihnen für die routinemäßige Durchführung von Urokinase und PAI-1 bei den entsprechenden Patientengruppen. Wenn das nicht durchführbar wäre, wäre uns natürlich trotzdem daran gelegen, diese Fragestellungen, die von Ihnen vorgeschlagen worden sind, weiter abzarbeiten. Dann gibt es doch meines Erachtens einen Kompromissweg, dass man sagt, dass die Kassen solche

Anwendungsbeobachtungen oder protokollgerechten Bestimmungen auch noch mitfinanzieren. Die Frage ist, ob man so einen Weg dann gehen könnte.

Moderator Jürgen Windeler: Das ist natürlich jetzt kein Thema für diesen konkreten Auftrag. Wenn man sich jetzt den gesetzlich hierfür vorgesehenen Weg vorstellen würde, dann wäre eine Option, die aber jetzt nicht in unserem Auftrag enthalten sein wird, die 137e-Regelung, also die Regelung, dass man mit einem solchen Bericht ein Potential feststellt und dann möglicherweise auf dieser Ebene Studien macht, die entweder ganz oder teilweise von den GKV finanziert werden könnten. Das wäre ein Weg, der da möglich wäre, wobei der Weg dann weniger uns betrifft, sondern eben eine Abwägung und Entscheidung des G-BA wäre, und man sich sicherlich auch die Frage stellen müsste, ob auf die Frage, die der G-BA dann hätte, nämlich von einem Potential zu einem Nutzen zu kommen, der G-BA dann sagen müsste, welche Studien er dafür haben wolle. Da wäre meine Prognose, dass dafür Anwendungsbeobachtungen nicht reichen werden. Was immer man da genau macht, ist eine andere Diskussion. Das wäre jetzt ein Weg, den ich mir im Rahmen des G-BA vorstellen könnte. Ob einzelne Kassen noch andere Wege auch im Rahmen von Modellvorhaben meinen gehen zu wollen, ist eine andere Diskussion. Die hat mit dem G-BA nichts zu tun. Aber der Weg über die 137e-Regelung mit der Bewertung zusammen wäre einer, den man sich im Prinzip vorstellen kann.

Christoph Thomssen: Das würden wir begrüßen, aber wir würden es noch mehr begrüßen, wenn Sie mit unserer Argumentation zur prognostischen Wertigkeit dieser Marker mitgehen könnten und das entsprechend dem G-BA empfehlen könnten.

Moderator Jürgen Windeler: Wir haben das ja heute ausführlich zur Kenntnis genommen und gewürdigt. Wir werden das jetzt noch mal diskutieren und erwägen und sehen, was wir aus den Argumenten machen können.

Ich habe den Eindruck, dass niemandem mehr etwas auf der Seele brennt. Vielen Dank, dass Sie da waren. Vielen Dank für Ihre Anmerkungen, Kommentare, Argumente.

Ich schließe hiermit die Sitzung.

Anhang – Dokumentation der Stellungnahmen

Inhaltsverzeichnis

	Seite
A.1 – Stellungnahmen von Organisationen, Institutionen und Firmen	A 2
A.1.1 – American Diagnostica GmbH.....	A 2
A.1.2 – Deutsche Krankenhausgesellschaft e. V.....	A 11
A.1.3 – Deutsche Krebsgesellschaft e. V.....	A 14

A.1 – Stellungnahmen von Organisationen, Institutionen und Firmen

A.1.1 – American Diagnostica GmbH

Autoren:

- Klöppinger, Martina
- Olenik, Claudia

Stellungnahme
von
American Diagnostica GmbH
zum
Vorbericht D13-02

„Bestimmung der Antigenexpressionslevel von
uPA und PAI-1 beim primären Mammakarzinom
mit intermediärem Rückfallrisiko nach R0-Primäroperation“

Auftrag: D13-02

Version: 1.0

Stand: 21.03.2014

Sehr geehrter Herr Prof. Windeler,
Sehr geehrte Damen und Herren,

vielen Dank für die Möglichkeit, zu oben genannten Vorbericht Stellung nehmen zu können.

Der Gemeinsame Bundesausschuss (G-BA) hat das IQWiG gemäß § 139a Abs. 3, Nr. 1 SGB V am 20.12.2012 mit der Recherche, Darstellung und Bewertung des aktuellen medizinischen Wissenstandes zur Bestimmung der Antigenexpressionslevel von uPA und PAI-1 beim primären Mammakarzinom mit intermediärem Rückfallrisiko nach R0-Primäroperation beauftragt.

Der spätere Abschlussbericht des IQWiG soll eine Grundlage für die Beratungen des G-BA zur Bewertung der Bestimmung der Antigenexpressionslevel von uPA und PAI-1 beim primären Mammakarzinom mit intermediärem Rückfallrisiko nach R0-Primärooperation gemäß § 135 Abs. 1 (Methodenbewertung ambulant) und § 137c SGB V (Methodenbewertung stationär) darstellen.

Bei der Durchsicht des Vorberichts sind uns einige Punkte aufgefallen, zu denen die American Diagnostica GmbH wie folgt Stellung nehmen möchte:

Anmerkungen zur projektspezifischen Methodik

1) Bei der Identifikation relevanter Studien wurden im vorliegenden Bewertungsverfahren ausschließlich randomisierte kontrollierte Studien (RCTs) für die Nutzenbewertung berücksichtigt. Dadurch wurde nur eine einzige Studie – die Chemo N0 Studie – in die Bewertung eingeschlossen. Weitere prospektiv geplante Studien mit niedrigerem Evidenzgrad wurden für die Bewertung der Biomarker uPA und PAI-1 nicht berücksichtigt. RCTs weisen zweifelsfrei die bestmögliche Evidenz zur Bewertung des Nutzens eines diagnostischen Verfahrens auf. Allerdings halten wir es für geboten, auch Studien mit geringerem Evidenzlevel einzubeziehen, wenn die auf RCTs basierende Datenlage nicht hinreichend ist, um den patientenrelevanten Nutzen und Schaden mit ausreichender Ergebnissicherheit schätzen zu können, dafür aber andere, methodisch gut validierte Ergebnisse aus prospektiven Studien vorliegen. Ihre methodische Einschränkung hat in dem vorliegenden Bewertungsverfahren dazu geführt, dass ein Großteil der wissenschaftlichen Erkenntnisse und Publikationen ausgeschlossen wurden, die auch wesentliche Informationen zur Bewertung des klinischen Nutzens von uPA und PAI-1 enthalten.

Wir möchten hier besonders auf Informationen und Publikationen zu uPA und PAI-1 verweisen, die Ihrem Institut im Rahmen von verschiedenen Stellungnahmen zu diesem Projekt zugesendet wurden bzw. Ihnen vom G-BA übermittelt wurden. Die ausschließliche Beschränkung auf RCTs in diesem Verfahren ist vor allem deshalb nicht nachvollziehbar, da in ähnlichen Verfahren sehr wohl auch Studien mit geringerem Evidenzlevel akzeptiert werden (siehe auch Punkt 2).

- 2) In einem anderen aktuellen Bewertungsverfahren einer diagnostischen Methode (Projekt D13-01) berücksichtigen Sie laut Berichtsplan auch Studientypen mit geringerem Evidenzlevel.

Auszug aus entsprechendem Berichtsplan D13-01: *„Ist die auf RCTs basierende Datenlage nicht hinreichend, um den patientenrelevanten Nutzen und Schaden der Proteomanalyse mit ausreichender Ergebnissicherheit schätzen zu können (zum Beispiel kleine Fallzahl, wenige Ereignisse [geringe Präzision], hohes Verzerrungspotenzial), werden gegebenenfalls auch verfügbare eindeutig prospektiv geplante, jedoch nicht randomisierte Interventionsstudien mit zeitlich parallelen Kontrollgruppen herangezogen.....“).*

Aus unserer Sicht ist nicht nachvollziehbar, warum von Ihrer Seite in den oben genannten Verfahren (D13-01 & D13-02) unterschiedliche Anforderungen an das Studiendesign gestellt werden. Dieser Punkt wurde bereits in der Stellungnahme der Deutschen Krankenhausgesellschaft zum aktuellen Berichtsplan zum uPA und PAI-1 Verfahren (D13-02) kritisiert.

Dass im vorliegenden Verfahren ausschließlich RCTs in die Nutzenbewertung eingeschlossen werden, begründen Sie damit, dass mittlerweile weitere Biomarker entwickelt wurden (z.B. Oncotype DX, Mammaprint), die aktuell mit vergleichbarer Fragestellung in großen randomisierten Studien untersucht werden. Angesichts der zu erwartenden Ergebnisse für andere Biomarker sei daher eine Nutzenbewertung von uPA und PAI-1 auf der Basis von Studien mit niedrigerem Evidenzgrad nicht zielführend. Dieser Argumentation stimmen wir nicht zu, vor allem da die endgültigen Auswertungen der von Ihnen benannten Studien (z.B. MINDACT, TAILORx) aktuell noch nicht vorliegen. Es wird sich damit erst in der Zukunft zeigen, wie und ob die geplanten randomisierten Studiendesigns in der Realität durchsetzbar sind.

Bitte beachten Sie, dass laut G-BA Auftrag das IQWiG mit der Recherche, Darstellung und Bewertung des aktuellen medizinischen Wissenstandes von uPA und PAI-1 beauftragt wurde. Da aktuell keine bessere Evidenz zu uPA und PAI-1 und auch nicht zu anderen Biomarkerstudien vorliegt, sollten wie in anderen Verfahren auch (z.B. D13-01), für die Nutzenbewertung nicht ausschließlich RCTs, sondern bei Bedarf zusätzlich Studientypen mit geringerem Evidenzlevel herangezogen werden.

- 3) In den Vorbericht wurden ausschließlich Studien aufgenommen, die den prädiktiven Wert von uPA und PAI-1 belegen. Dies wurde von Ihnen damit begründet, dass nur der prädiktive und nicht der prognostische Wert „Auftragsgegenstand“ sei. Wie bereits in verschiedenen Stellungnahmen zu diesem Projekt angemerkt wurde, ist es aus **klinischer Sicht** wenig zielführend, wenn ausschließlich der prädiktive, aber nicht der prognostische Wert für die Nutzenbewertung berücksichtigt wird. Bei Einbeziehung des prognostischen Wertes erweitert sich die Perspektive der Nutzenbewertung auch in die Richtung, dass ein Marker mit prognostischem Wert dazu beiträgt, Patientinnen eine Chemotherapie zu ersparen.

Wegen der Abhängigkeit des absoluten Nutzens von der grundsätzlichen Wahrscheinlichkeit eines Rezidivs (in einer durch Prognosefaktoren definierten Patientengruppe), fordert die ak-

tuelle S3-Leitlinie auch eine Betrachtung aller validierten Prognosefaktoren. So ergibt sich laut S3-Leitlinie der Deutschen Krebsgesellschaft (S3-Leitlinie, Version 2012) eine Chemotherapie-Indikation beim primären Mammakarzinom immer dann, „*wenn der individuelle zu erwartende Nutzen höher ist als mögliche Nebenwirkungen und Spätschäden*“.

Aus klinischer Sicht liefert damit nicht nur der prädiktive sondern auch der prognostische Wert einen wichtigen Beitrag zur Einschätzung dieses individuellen Nutzens und damit zur Therapieplanung (in diesem Fall einer adjuvanten Chemotherapie).

- **Prognostische Marker** erlauben Aussagen über den voraussichtlichen Krankheitsverlauf unabhängig von der Behandlung – bezogen auf relevante klinische Endpunkte (wie krankheitsfreies Überleben oder Gesamtüberleben) – zu treffen (Mandrekar & Sargent, 2009), (Oldenhuis, Oosting, Gietema, & Vries, 2008). Prognostische Marker können darüber hinaus therapeutisch relevant sein und Therapieentscheidungen beeinflussen, wenn sie die Identifizierung von Patientinnen mit einer guten Prognose ohne Therapie erlauben, so dass diesen Patientinnen die adjuvante Systemtherapie erspart werden kann (Simon, 2010) (Freidlin, McShane, & Korn, 2010).
- **Prädiktive Marker** erlauben Aussagen über den voraussichtlichen Behandlungserfolg und erleichtern damit die Auswahl der individuell besten Therapie (Mandrekar & Sargent, 2009), (Oldenhuis, Oosting, Gietema, & Vries, 2008).

Wir möchten an dieser Stelle nochmals betonen, dass in der aktuellen Versorgungssituation prognostische Faktoren sehr wohl klinisch und therapeutisch relevant sind und für eine Therapieentscheidung – für oder gegen eine adjuvante Chemotherapie – verwendet werden können. Insbesondere bei der Patientengruppe mit „intermediärem“ Rezidivrisiko und ohne Befall der Lymphknoten ist die Abschätzung des Rezidivrisikos (Prognose) von therapeutischer Bedeutung. Bei dieser Patientengruppe ist eine genaue Abwägung von Nutzen und Risiko wichtig, da ein Großteil dieser Patientinnen nach R0-Resektion (auch ohne adjuvante Chemotherapie) kein Rezidiv erleiden wird. Die Erkenntnis, dass in einer Gruppe von Patientinnen das grundsätzliche Rezidivrisiko oder Sterberisiko (ohne Chemotherapie) hinreichend gering ist, führt zu der Schlussfolgerung, dass keine Indikation zur adjuvanten Chemotherapie besteht.

Die Bedeutung der Risikoabschätzung (Prognose) ist auch in der Richtlinie des Gemeinsamen Bundesausschusses zur Regelung von Anforderungen an die Ausgestaltung von strukturierten Behandlungsprogrammen verankert. Dort ist niedergelegt, dass die Therapie nach individueller Risikoabschätzung erfolgen soll (DMP-Richtlinie, Version 16. Februar 2012).

Bitte berücksichtigen Sie in diesem Zusammenhang auch unsere ausführliche Stellungnahme, die wir anlässlich der Bekanntmachung des Bewertungsauftrages anhand eines umfangreichen Fragenkatalogs erstellt haben. Die Stellungnahme wurde dem G-BA am 25.01.2013 übermittelt und sollte Ihrem Institut vorliegen. Weitere ausführliche Informationen zu diesem Punkt sind außerdem in den Stellungnahmen der Deutschen Krebsgesellschaft und der American Diagnostica GmbH zum vorläufigen Berichtsplan enthalten.

Bewertung und Interpretation der im Vorbericht eingeschlossenen Studien

Für die vorläufige Nutzenbewertung der Biomarker uPA und PAI-1 wurde eine einzige Studie – die Chemo N0 Studie – als relevant für die Fragestellung identifiziert (siehe dazu Punkt 5.2. und 5.3. im Vorbericht) (Jänicke, et al., 2001) (Harbeck, et al., 2013). Die Beschränkung auf eine einzige Studie erfolgte dadurch, dass für die Nutzenbewertung ausschließlich RCTs vom IQWiG berücksichtigt wurden (siehe auch obenstehende „Anmerkungen zu projektspezifischer Methodik“). Bei der Chemo N0 Studie handelt sich um eine prospektiv randomisierte kontrollierte Studie (RCT), die im Hybriddesign durchgeführt wurde.

1) Prognostischer Wert von uPA und PAI-1: Die prognostische Fragestellung, die wichtiger Gegenstand der Chemo N0 Studie war, wird bei der vorliegenden Nutzenbewertung durch Sie nicht berücksichtigt, da sich Ihr Institut bei seiner Untersuchung ausschließlich auf den prädiktiven Wert von uPA und PAI-1 beschränkt. Dies ist aus klinischer Sicht wenig zielführend, da speziell bei Patientinnen mit intermediärem, d.h. klinisch unklarem Rezidivrisiko, neben uPA und PAI-1 bisher auch keine anderen prädiktiven Faktoren existieren, die den Nutzen einer adjuvanten Chemotherapie vorhersagen könnten (S3-Leitlinie, Version 2012) (AGO-Leitlinie, Version 2014.1D). Selbst modernere Tests wie z.B. Genexpressionsanalysen scheinen dieses Problem derzeit nicht wirklich zufriedenstellend lösen zu können (Schmidt & Untch, 2014). Zum aktuellen Zeitpunkt ist es daher umso wichtiger validierte Prognosefaktoren verfügbar zu haben, die eine zuverlässige Abschätzung des Rezidivrisikos ermöglichen. Mit Hilfe dieser Faktoren können Patientinnen identifiziert werden, die ein so geringes Rezidivrisiko tragen, dass Ihnen eine Chemotherapie erspart werden könnte. Der uPA PAI-1 Test ermöglicht speziell in der Gruppe der Patientinnen mit unklarem (= intermediärem) Rezidivrisiko, eine präzisere Prognosebestimmung als bisherige klinisch pathologische Faktoren und hilft damit eine verbesserte Therapieentscheidung zu treffen.

Die ausschließliche Berücksichtigung des prädiktiven Wertes von uPA und PAI-1 ist aus oben genannten Gründen und unter Berücksichtigung von aktuellen Leitlinienempfehlungen und Versorgungssituation beim primären Mammakarzinom, speziell bei der Gruppe der Patientinnen mit intermediärem Rezidivrisiko daher nicht zu verstehen (siehe dazu auch obenstehende „Anmerkungen zu projektspezifischer Methodik“).

2) Prädiktiver Wert von uPA und PAI-1: Im Rahmen des Vorberichts wurde nur ein Teilaspekt der Chemo N0 Studie – die prädiktive Fragestellung – ausgewertet (Jänicke, et al., 2001) (Harbeck, et al., 2013). In dem entsprechenden Behandlungsarm wurden Patientinnen mit hohem Rezidivrisiko (d.h. hoher Konzentrationen von uPA und/oder PAI-1) randomisiert – entweder wurde ihnen eine Behandlung mit Chemotherapie (CMF) zugewiesen oder sie wurden nur nachbeobachtet (Randomisierungsarm).

Patientinnen mit hohem Rezidivrisiko (hohe uPA und/oder PAI-1 Werte), die einer Randomisierung nicht zustimmten, wurden separat ausgewertet („Beobachtungsstudie im Hochrisikoarm“). Laut Studienprotokoll konnten Patientinnen ohne Angabe von Gründen, die Randomisierung ablehnen (siehe dazu auch Studiendesign 5.2.1 im Vorbericht). Dies führ-

te dazu, dass ein großer Teil an Hochrisikopatientinnen (N=122) für die ITT Auswertung nicht zur Verfügung standen. Ein anderes Vorgehen wäre ethisch auch nicht zu rechtfertigen gewesen.

Für die im Vorbericht ausführlich dokumentierte ITT (*intention-to-treat*)-Analyse wurden alle Patientinnen aus dem Randomisierungsarm eingeschlossen. Da zum Zeitpunkt der Chemo N0 Studie die Wertigkeit einer adjuvanten Chemotherapie vor allem bei nodal-negativen Patientinnen noch nicht nachgewiesen war, wählte eine große Anzahl der Studienteilnehmerinnen trotz Randomisierung eine andere Therapie bzw. keine Therapie. Aus diesem Grund sind im „CMF Arm“ Patientinnen enthalten, die keine Therapie erhielten und im „Kontrollarm“ Patientinnen, die eine CMF Therapie erhielten. Bei der ITT-Auswertung werden Patienten entsprechend der ihnen zugewiesenen Behandlung ausgewertet und nicht entsprechend der tatsächlichen Behandlung nach einem eventuellen „Wechsel“. Deshalb wird der wahre Therapieeffekt in der ITT-Analyse eher unterschätzt.

Patientinnen, die nach Randomisierung die „Behandlung wechselten“ wurden in der *per-protocol Analyse* ausgeschlossen. In der *per protocol population* zeigte sich in der Gruppe mit hohen Konzentrationen von uPA und/oder PAI-1 (hohes Rezidivrisiko) ein statistisch signifikanter Unterschied im krankheitsfreien Überleben (10 Jahres-Überlebensrate) zwischen der Gruppe mit CMF-Chemotherapie und der nicht therapierten Kontrollgruppe. Das Hazard Ratio (HR) betrug 0,48 (95%CI: [0,26-0,88]; p=0,019) (Harbeck, et al., 2013). Die Ergebnisse der Chemo N0 Studie zeigen damit deutlich in der *per protocol Analyse*, dass Patientinnen mit hohen uPA und/oder PAI-1 Werten von einer CMF-Chemotherapie profitierten und ein signifikant verlängertes rezidivfreies Überleben (DFS) aufwiesen.

Bitte berücksichtigen Sie in diesem Zusammenhang auch unsere ausführliche Stellungnahme, die wir anlässlich der Bekanntmachung des Bewertungsauftrages erstellt haben. Die Stellungnahme wurde dem G-BA am 25.01.2013 übermittelt und sollte Ihrem Institut vorliegen.

Mögliche geeignete Studiendesigns

Die einzige von Ihnen eingeschlossene Studie – die Chemo N0 Studie – untersuchte den Nutzen einer adjuvanten Chemotherapie in der Gruppe der Patientinnen mit hohen Konzentrationen von uPA und/oder PAI-1 (RCT im Hybriddesign). Sie haben bei Ihrer Bewertung nur die *intention to treat population* (ITT) aus dem Randomisierungsarm (RCT im Hybriddesign) der Chemo N0 Studie berücksichtigt. In der ITT-Analyse konnte weder für das krankheitsfreie Überleben (DFS) noch für das Gesamtüberleben (OS) ein signifikanter Effekt der Chemotherapie in der Teilpopulation der Patientinnen mit hohen uPA und/oder PAI-1 Werten nachgewiesen werden. Die *per protocol Analyse* zeigte dagegen für das krankheitsfreie Überleben einen signifikanten Effekt einer Chemotherapie in dieser Gruppe. Diese Auswertungen wurden für die vorliegende Nutzenbewertung jedoch nicht berücksichtigt.

Dass in der ITT-Analyse der Chemo N0 Studie kein statistisch signifikanter Effekt einer Chemotherapie in der Gruppe der Patientinnen mit hohen uPA und/oder PAI-1 Werten zu beobachten ist, kann verschiedene Gründe haben, die von den Biomarkern uPA und PAI-1 unabhängig sind (dies wurde zum Teil auch im Vorbericht diskutiert), z.B.:

- zu wenige Patienten wurden in die Studie eingeschlossen (aufgrund neuer Behandlungsleitlinien wurde die Studie vorzeitig abgebrochen)
- zu viele Hochrisikopatienten haben Randomisierung abgelehnt (dies war laut Studienprotokoll ohne Begründung möglich)
- zu viele Hochrisikopatienten haben nach Randomisierung die Behandlung gewechselt (Unterschätzung des ITT Effekts)
- CMF Wirkung zu gering

Aus Ihrer Sicht ist der patientenrelevante Nutzen bzw. Schaden einer uPA- und PAI-1-gestützten Strategie zur Entscheidung für oder gegen eine adjuvante Chemotherapie beim primären Mammakarzinom mit intermediärem Rezidivrisiko nach R0-Primäroperation aufgrund fehlender geeigneter Studien unklar. Da aber bereits Vorschläge zu denkbaren Studiendesigns gemacht werden, um den gefundenen Mängeln abzuweichen, gehen wir davon aus, dass Sie (das IQWiG) an dieser Stelle einen möglichen Nutzen in der Testanwendung sehen, der unter dem neu geschaffenen Begriff des „Potenzials“ im Sinne des § 137e SGB V subsummiert werden kann. Im Rahmen einer solchen Erprobung (§137 e SGB V) müssten dann, falls ausschließlich RCTs akzeptiert werden, also Studien mit höherer Evidenz (RCTs) gefordert werden, neue Studien entsprechend der aktuellen Anforderungen an Biomarkerstudien geplant und durchgeführt werden.

Im Rahmen einer solchen Erprobung sollten dann alle relevanten, zur Verfügung stehenden Tests einbezogen werden. Hierbei sollte insbesondere ein Studiendesign gewählt werden, das den Nutznachweis bei allen Tests ermöglicht und auch deren Budget Impact betrachtet.

Fazit

Aus den oben aufgeführten Gründen möchten wir das IQWiG bitten, den **Vorbericht** unter besonderer Berücksichtigung der folgenden Punkte zu überarbeiten und zu vervollständigen:

- 1) Neben den im Vorbericht ausschließlich berücksichtigten RCTs sollten auch Studientypen mit niedrigerem Evidenzgrad eingeschlossen werden (siehe Einschlusskriterien, Tab.1 im Vorbericht). Diese Forderung ergibt sich aus unserer Sicht aus folgenden Gründen:
 - a) Bei anderen Projekten, die das IQWiG aktuell bearbeitet, werden bei Bedarf Studien mit niedrigerem Evidenzgrad eingeschlossen (z.B. IQWiG Projekt D13-01)
 - b) Aktuell sind für andere Biomarker (z.B. Oncotype DX, MammaPrint, etc.) keine prospektiven Daten mit ausreichender Evidenz verfügbar. Evidenz aus prospektiven Studien (RCTs) werden erst in den kommenden Jahren erwartet. Vor diesem Hintergrund,

sollten für die vorliegende Nutzenbewertung auch Studien mit niedrigerem Evidenzgrad eingeschlossen werden.

- 2) Die klinische Bedeutung des prognostischen Wertes von uPA und PAI-1 für die Therapieentscheidung im aktuellen Versorgungskontext, z. B. Empfehlungen der Arbeitsgemeinschaft Gynäkologische Onkologie (AGO-Leitlinie, Version 2014.1D) oder der interdisziplinären S3-Leitlinie (S3-Leitlinie, Version 2012) muss berücksichtigt werden. Zusätzlich wird der Test von US-amerikanischen Leitlinien (ASCO) auf allerhöchstem Evidenzniveau (LoE-1a) für die Prognoseeinschätzung bei nodal-negativen Mammakarzinompatientinnen empfohlen (Harris, et al., 2007) (Sturgeon, et al., 2008).

Literatur

- AGO-Leitlinie. (Version 2014.1D). *AGO Leitlinien der Kommission Mamma*. http://www.ago-online.de/fileadmin/downloads/leitlinien/mamma/maerz2014/de/2014D_05_Prognostische_und_praediktive_Faktoren.pdf
- DMP-Richtlinie. (Version 16. Februar 2012). *Richtlinie des Gemeinsamen Bundesausschusses zur Regelung von Anforderungen an die Ausgestaltung von Strukturierten Behandlungsprogrammen nach §137f Abs. 2 SGB V in der Fassung vom 16. Februar 2012*. www.g-ba.de/downloads/62-492-623/DMP-RL_2012-02-16.pdf
- Freidlin, B., McShane, L. M., & Korn, E. L. (2010). Randomized Clinical Trials With Biomarkers: Design Issues. *JNCI*, 102, S. 152-160.
- Harbeck, N., Schmitt, M., Meisner, C., Friedel, C., Untch, M., Schmid, M., et al. (2013). Ten-year analysis of the prospective multicenter Chemo-N0 trial validates ASCO-recommended biomarkers uPA and PAI-1. *Eur J Cancer*, 49(8), S. 1825-35.
- Harris, L., Fritsche, H., Mennel, R., Norton, L., Ravdin, P., Taube, S., et al. (2007). American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *Journal of Clinical Oncology*, 25(33), S. 5287-5312.
- Jänicke, F., Prechtel, A., Thomssen, C., Harbeck, N., Meisner, C., Untch, M., et al. (2001). Randomized Adjuvant Chemotherapy Trial in High-Risk, Lymph Node-Negative Breast Cancer Patients Identified by Urokinase-Type Plasminogen Activator and Plasminogen Activator Inhibitor Type 1. *Journal of the National Cancer Institute*, 93 (12), S. 913-920.
- Mandrekar, S. J., & Sargent, D. J. (2009). Clinical Trial Designs for Predictive Biomarker Validation: Theoretical Considerations and Practical Challenges. *Journal of Clinical Oncology*, 4, S. 4027-4034.
- Oldenhuis, C., Oosting, J., Gietema, J., & Vries, E. (2008). Prognostic versus predictive value of biomarkers in oncology. *European Journal of Cancer*, 44, S. 946-953.
- S3-Leitlinie. (Version 2012). *Interdisziplinäre S3-Leitlinie für die Diagnostik, Therapie und Nachsorge des Mammakarzinoms*, Langversion 3.0. http://www.krebsgesellschaft.de/download/S3_Brustkrebs_Update_2012_OL_Langversion.pdf
- Schmidt, M., & Untch, M. (2014). Prediction of benefit from chemotherapy in ER-positive/HER2-negative breast cancer - a problem still to be solved. *Annals of Oncology*, 25, S. 754-755.
- Simon, R. (2010). Clinical trial designs for evaluating the medical utility of prognostic and predictive biomarkers in oncology. *Personalized Medicine*, 7(1), S. 33-47.
- Sturgeon, C., Duffy, M., Stenman, U.-H., Lilja, H., Brünnler, N., Chan, D., et al. (2008). National Academy of Clinical Biochemistry Laboratory Medicine Practice Guidelines for Use of Tumor Markers in Testicular, Prostate, Colorectal, Breast, and Ovarian Cancers. *Clinical Chemistry*, 54(12), S. e11-e79.

A.1.2 – Deutsche Krankenhausgesellschaft e. V.

Autoren:

- Bielecki, Eva
- Rohrig, Angela
- Schlottmann, Nicole

Vorab per E-Mail und Fax

Institut für Qualität und Wirtschaftlichkeit im
Gesundheitswesen
- Stellungnahme zum Vorbericht < D13-02 >
Prof. Dr. med. Jürgen Windeler
Im Mediapark 8 (KölnTurm)
50670 Köln

**Stellungnahme der Deutschen Krankenhausgesellschaft zum Vorbericht D13-02
„Bestimmung der Antigenexpressionslevel von uPA und PAI-1 beim primären
Mammakarzinom mit intermediärem Rückfallrisiko nach R0-Primäroperation“**

Sehr geehrter Herr Professor Windeler,
sehr geehrte Damen und Herren,

der Gemeinsame Bundesausschuss (G-BA) hat das Institut gem. § 139a Abs. 3, 1. SGB V am
03.01.2013 mit der Recherche, Darstellung und Bewertung des aktuellen medizinischen
Wissenstandes zur Bestimmung der Antigenexpressionslevel von uPA und PAI-1 beim
primären Mammakarzinom mit intermediärem Rückfallrisiko nach R0-Primäroperation beauf-
tragt.

Der spätere Abschlussbericht soll eine Grundlage für die Beratungen des G-BA zur Bewertung
der Bestimmung der Antigenexpressionslevel von uPA und PAI-1 beim primären
Mammakarzinom mit intermediärem Rückfallrisiko nach R0-Primäroperation nach
§ 137c SGB V (Methodenbewertung stationär) und § 135 SGB V (Methodenbewertung
ambulant) darstellen.

Gemäß § 137c SGB V hat der G-BA zu prüfen, ob Untersuchungs- oder Behandlungsmethoden,
die zu Lasten der gesetzlichen Krankenkassen im Rahmen einer Krankenhausbehandlung
angewendet werden oder angewendet werden sollen, für eine ausreichende, zweckmäßige und
wirtschaftliche Versorgung der Versicherten unter Berücksichtigung des allgemein anerkannten
Standes der medizinischen Erkenntnisse erforderlich sind.

Die Deutsche Krankenhausgesellschaft (DKG) als eine Trägerorganisation des G-BA nimmt
zum o. g. Berichtsplan des Instituts wie folgt Stellung:

Wir haben uns bei der Durchsicht des Berichts sehr gefreut, dass eine Einordnung des Themas
in den aktuellen Stand des medizinischen Wissens im Kapitel 6.4 stattgefunden hat. Durch den
Vergleich mit anderen HTAs und die Darstellung der aktuellen Leitlinienempfehlungen wird eine
Bewertung des Themas im Rahmen der Arbeitsgruppe deutlich erleichtert. Hierfür möchten wir
uns an dieser Stelle ausdrücklich bedanken.

Allerdings bedauern wir es, dass auf Grund der Eingrenzung des Studientyps auf RCTs im Rahmen der Einschlusskriterien für den Studieneinschluss bereits im Berichtsplan eine starke Einengung der Sichtweise auf die vorliegende Literatur entstanden ist.

Wie in Ihrem Methodenpapier Version 4.1 auf Seite 5 genannt, ist es „die Hauptaufgabe des Instituts [...], die vorgegebene Frage seiner Auftraggeber, ob ein Nutzen oder Schaden durch eine Maßnahme nachzuweisen ist, möglichst zuverlässig zu beantworten.“

Das IQWiG versucht, diese Aufgabenstellung mit Studien der höchsten Evidenzstufe zu beantworten, da diese mit der geringsten Ergebnisunsicherheit behaftet sind. Im vorliegenden Vorbericht musste nun aber festgestellt werden, dass sich bei den festgelegten Ein- und Ausschlusskriterien lediglich ein RCT finden lässt, was zudem nicht mehr dem aktuellen Therapiestandard entspricht und außerdem durch das Studiendesign nur einen Teil der Fragestellung beantworten kann.

Wir hätten uns daher gewünscht, dass im Sinne der evidenzbasierten Medizin, nach der auf der Basis der besten zur Verfügung stehenden Daten entschieden werden soll, auch andere vorliegende Studien zumindest dargestellt und dann in ihrer Ergebnissicherheit bewertet werden. Dieses Vorgehen ist unseres Erachtens nach auch in Ihrem Methodenpapier so dargelegt: „Falls sich herausstellt, dass Studien mit der nötigen Qualität und Präzision generell fehlen, ist es die Kernaufgabe des Instituts, den Sachverhalt zu beschreiben mit der Schlussfolgerung, dass auf Basis der „gegenwärtigen besten“ Evidenz keine sicheren Empfehlungen möglich sind.“ (Methodenpapier Version 4.1 Seite 7)

Mit der Darstellung aller verfügbaren Evidenz und nicht nur eines Teils der vorliegenden Daten zur Fragestellung, hätten sich evtl. auch aktuellere und besser auf die Fragestellung passende Informationen finden lassen, was die weitere Arbeit in der Arbeitsgruppe deutlich erleichtert hätte.

Wir hoffen, dass wir Ihnen mit unseren Anmerkungen einige Anregungen für die Überarbeitung des Vorberichts und die weitere Auftragsbearbeitung geben konnten. Im Übrigen behalten wir uns vor, ggf. ergänzende Aspekte in das weitere Verfahren einzubringen.

Mit freundlichen Grüßen
Der Hauptgeschäftsführer
Im Auftrag



Dr. med. N. Schlottmann
Geschäftsführerin Dezernat Medizin



Angela Röhrig
Referentin Dezernat Medizin



Dr. med. E. Bielecki, MPH
Referentin Dezernat Medizin

A.1.3 – Deutsche Krebsgesellschaft e. V.

Autoren:

- Bruns, Johannes
- Fehm, Tanja
- Thomssen, Christoph
- Scharl, Anton

Anhang zum

„Formblatt zur schriftlichen Stellungnahme zum Vorbericht D13-02“ der DKG e.V.

Stellungnahme der Arbeitsgemeinschaft für Gynäkologische Onkologie e.V. (AGO) und der Deutschen Krebsgesellschaft e.V. (DKG) zum Vorbericht (vorläufige Nutzenbewertung) D13-02, Version 1.0 Stand 21.03.2014

zu

Bestimmung der Antigenexpressionslevel von uPA und PAI-1 durch ELISA-Test in Tumorgewebeextrakten zur Entscheidung für oder gegen eine adjuvante Systemtherapie beim primären Mammakarzinom, welches einem intermediären Rückfallrisiko nach R0-Primäroperation zugeordnet wird

Herausgeber: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)

Auftraggeber: Gemeinsamer Bundesausschuss, vom 20.12.2012/03.01.2013

Berlin, den 29.04.2014

Sehr geehrter Herr Hecken,

sehr geehrte Damen und Herren,

wir möchten die Gelegenheit wahrnehmen und zu dem oben genannten Vorbericht D13-02 vom 21.03.2013, der die Grundlage für die Beratungen des G-BA zum genannten Thema sein wird, Stellung zu nehmen. Zunächst möchten wir wiederholen, dass wir die Auftragserteilung des G-BA begrüßen und unterstützen.

Das Mammakarzinom wird aufgrund seiner vielfältigen histologischen Ausprägungen und biologischen Eigenschaften als heterogene Tumorerkrankung angesehen. Zum Zeitpunkt der Diagnose befindet sich das Karzinom in unterschiedlichen Entwicklungsstadien und weist eine unterschiedliche intrinsische Aggressivität und Aktivität auf. Eine eindeutige Differenzierung der Brustkrebspatientinnen ohne Lymphknotenbefall in die Gruppe mit einem niedrigem bzw. einem hohem Metastasierungsrisiko ist nicht zufriedenstellend möglich. Daher wird aus Sicherheitsgründen einem Großteil der Patientinnen häufig eine Chemotherapie empfohlen, obwohl sie möglicherweise nicht von der Therapie profitieren.

Das individuelle Rückfallrisiko einer Brustkrebspatientin kann anhand der tumorbiologischen Marker uPA/PAI-1 (Plasminogen Aktivator vom Urokinasetyp und sein Inhibitor) deutlich genauer vorhergesagt werden als mit traditionellen klinischen und pathologischen Markern.

Ziel der Bestimmung der uPA und PAI-1-Konzentrationen im Primärtumor von Brustkrebspatientinnen ist somit die Vermeidung von unnötiger adjuvanter Chemotherapie. Die Patientinnen sollen ihrem Rückfallrisiko entsprechend behandelt werden; überflüssige, potentiell toxische Chemotherapien sollen eingespart werden. Der Test ermittelt für Brustkrebs im Frühstadium das individuelle Risiko eines Erkrankungsrückfalls innerhalb der nächsten 10 Jahre und unterstützt die Entscheidung, ob eine adjuvante Chemotherapie nach Operation sinnvoll und hilfreich ist. Neben dem offensichtlichen Vorteil für die Patientinnen konnte in einer Wirtschaftlichkeitsanalyse auch ein günstiges Kosten-Nutzen-Verhältnis gezeigt werden: Durch Anwendung dieses Tests werden Chemotherapie-Kosten eingespart (Jacobs et al. 2013).

Auftrag des Gemeinsamen Bundesausschusses (auszugsweise)

Das IQWiG sollte gemäß § 139a Absatz 3 Nummer 1 SGB V unter Berücksichtigung der Auftragskonkretisierung des G-BA die Recherche, Darstellung und Bewertung des aktuellen medizinischen Wissenstandes zur „Bestimmung der Antigenexpressionslevel von uPA und PAI-1 durch ELISA-Test in Tumorgewebeextrakten zur Entscheidung für oder gegen eine adjuvante Systemtherapie beim primären Mammakarzinom, welches einem intermediären Rückfallrisiko nach R0-Primäroperation zugeordnet wird“, durchführen.

Die Bewertung soll, soweit medizinisch-fachlich geboten und aufgrund der Erkenntnislage möglich, differenziert nach relevanten Patienten- und Interventionscharakteristika in Bezug auf patientenrelevante Zielgrößen und auf der Grundlage von Studien der Evidenzstufe I (2. Kapitel § 11 Absatz 2 der Verfahrensordnung des G-BA) durchgeführt werden.

Insbesondere sollen bei der Bewertung folgende Aspekte berücksichtigt werden:

- Patientencharakteristika: z. B. Alter, Rezeptorstatus (Östrogenrezeptor, Progesteron-rezeptor, HER2), Lymphknotenstatus,
- Vergleichsbehandlung: z. B. andere Methoden zur Entscheidungsfindung zur adjuvanten Chemotherapie,
- patientenrelevante Zielgrößen: z. B. Gesamtüberleben, krankheitsfreies Überleben, metastasenfreies Überleben und eventuell andere Morbiditätsoutcomes (z. B. negative Folgen der Chemotherapie) sowie Lebensqualität.

Das IQWiG hat am 21.03.2014 den vorläufigen Bericht zur Bewertung des klinischen Nutzen der uPA/PAI-1-Bestimmung im Tumorgewebe beim nodalnegativen Mammakarzinom vorgelegt. Wir erlauben uns, diesen im Folgenden zu kommentieren:

1 Bewertung und Interpretation der im Vorbericht eingeschlossenen Studien

1.1 Auftrag

Uns fällt auf, dass die Themenformulierung des IQWiG sich von der Auftragsformulierung des GBA sprachlich und inhaltlich unterscheidet:

G-BA:

Bewertung des „uPA und PAI-1 ELISA-Test“ zur Bestimmung der Antigenexpressionslevel in Tumorgewebeextrakten zur Entscheidung für oder gegen eine adjuvante Systemtherapie beim primären Mammakarzinom, welches einem intermediären Rückfallrisiko nach RO- Primäroperation zugeordnet wird gemäß § 135 Abs.1 sowie gemäß § 137c SGB V (Beschluss des G-BA, 20.20.2012, Seite 1)

IQWiG:

Thema: Bestimmung der Antigenexpressionslevel von uPA und PAI-1 beim primären Mammakarzinom mit intermediärem Rückfallrisiko nach RO-Primäroperation (Vorbericht, 21.0.3.2014, Seite i Impressum)

Der GBA fragt nach einer Bewertung des uPA/PAI1-Testung zur Entscheidung für oder gegen eine adjuvante Systemtherapie auf der Grundlage *des aktuellen medizinischen Wissensstandes*, das IQWiG dagegen reduziert diese Fragestellung auf den prädiktiven Wert des uPA/PAI1-Status. In unserer Stellungnahme zum Berichtsplan hatten wir bereits auf die Gefahr dieser Einschränkung der Fragestellung aufmerksam gemacht, dieser Forderung wurde seitens des IQWiGs offensichtlich bewusst nicht gefolgt (2.1.3 & 2.1.4 DWA zum IQWiG-Berichtsplan D13-02). Die Daten zur Prognoseabschätzung sind ebenfalls Teil des aktuellen medizinischen Wissensstandes.

Im Folgenden ist mit dem Begriff „uPA/PAI1-Status“ die Kombination beider Proteinkonzentrationen im Primärtumor gemeint, die für den klinischen Algorithmus zur Risikoabschätzung ausschlaggebend ist. Hoher uPA/PAI1-Status: mindestens eine der beiden Proteinkonzentration, gemessen mittels ELISA, liegt oberhalb des Schwellenwertes; niedriger uPA/PAI1-Status: beide Proteinkonzentrationen sind unterhalb der Schwellenwerte.

1.2. Studienauswahl

Zur Bewertung hat das IQWiG ausschließlich die aktuelle Publikation der ChemoN0-Studie (Harbeck et al. 2013) untersucht.

Wir meinen, dass dadurch wesentliche Informationen nicht in die Beurteilung einfließen konnten. Sowohl die Erstpublikation der ChemoN0-Studie (Interimanalyse; Jänicke et al. 2001) als auch die vorgeschaltete Pilotstudie (Jänicke et al. 1993) und die Metaanalysen zur prognostischen (Look et al. 2003) und prädiktiven (Harbeck et al. 2002) Bedeutung des uPA/PAI1-Status sollten berücksichtigt werden.

1.3 Klinische Relevanz des uPA/PAI1-Status (Prognose und Prädiktion)

Im klinischen Alltag steht für die Patientin die Frage im Vordergrund, ob eine adjuvante Chemotherapie notwendig ist oder nicht.

Die Indikation zur adjuvanten Chemotherapie ergibt sich zunächst aus der zugrundeliegenden Risikoabschätzung (Prognose): bei 8% Mortalitätsrisiko besteht gemäß Oxford-Overview weniger als 3% absoluter Benefit aus der Chemotherapie (relative Reduktion der 10-Jahres-Mortalität um 36%; EBCTCG 2012). Nur bei erhöhtem Risiko kann eine klinisch relevante absolute Risikoreduktion erreicht werden; Beispiel: Bei 40% Rückfallrisiko ist statistisch mit wenigstens 14% absolutem Benefit zu rechnen)

Dieser Aspekt wird in dem IQWiG-Vorbericht vollständig unterschlagen (siehe 2.1.3 Ziel der Untersuchung DWA zum Berichtsplan D13-02 des IQWiG). Schon in der prospektiven Pilotstudie (Jänicke et al. 2001) konnte gezeigt werden, dass Patientinnen mit niedrigem uPA/PAI1-Status ein sehr niedriges Rückfallrisiko (weniger als 5% Rückfallrisiko nach 5 Jahren; Jänicke et al. 1991) vorhergesagt werden kann; diesen Patientinnen kann die adjuvante Chemotherapie mit ihren potentiellen Nebenwirkungen erspart werden. Umgekehrt hatten nodalnegative Patientinnen mit hohem uPA/PAI1-Status ein deutlich erhöhtes Rückfallrisiko, ähnlich hoch wie bei undifferenzierten Karzinomen (G3) oder bei axillärem Lymphknotenbefall.

Ursprüngliche Intention und Ziel der ChemoN0-Studie war es, diese Daten im Feldversuch zu validieren und damit die Datenkonsistenz und die klinische Praktikabilität dieser Marker zu zeigen. Sekundär sollte in der ChemoN0-Studie gezeigt werden, dass auch die Identifikation von Hochrisikopatientinnen eine sinnvolle therapeutische Konsequenz nach sich ziehen kann durch Nachweis der Wirksamkeit einer adjuvanten Chemotherapie in diesem damals generell nicht behandelten Patientinnenkollektiv.

Durch die Feldstudie ChemoN0 mit verschiedenen Zentren und mehreren Laboratorien, in denen die Konzentrationsbestimmung der Proteine uPA und PAI-1 durchgeführt wurde, konnte die prognostische Diskriminierungsfähigkeit des uPA/PAI1-Status der unizentrischen Pilotstudie eindrucksvoll bestätigt werden (Abb. 1). Nodalnegative Patientinnen mit niedrigem uPA/PAI1-Status hatten in der 10 Jahres-Analyse der ChemoN0-Studie ein Rückfallrisiko von 12,9% (ohne irgendeine adjuvante Therapie, auch nicht endokrin, Sterberisiko 10,2%). Bei hohem uPA/PAI1-Status ergab sich ein signifikant höheres Rückfallrisiko von 23,0% und auch ein signifikant ähnlich höheres Sterberisiko von 20,9% (multivariate Analyse: $HR_{DFS}=1,84$; 95%-K.I. 1,1-3,0; $p=0,017$; $HR_{OS}=1,85$; 95%-K.I. 1,13-3,1; $p=0,02$). Der uPA/PAI1-Status war neben dem Grading der einzige unabhängige, starke und signifikante Prognosefaktor in dieser Analyse.

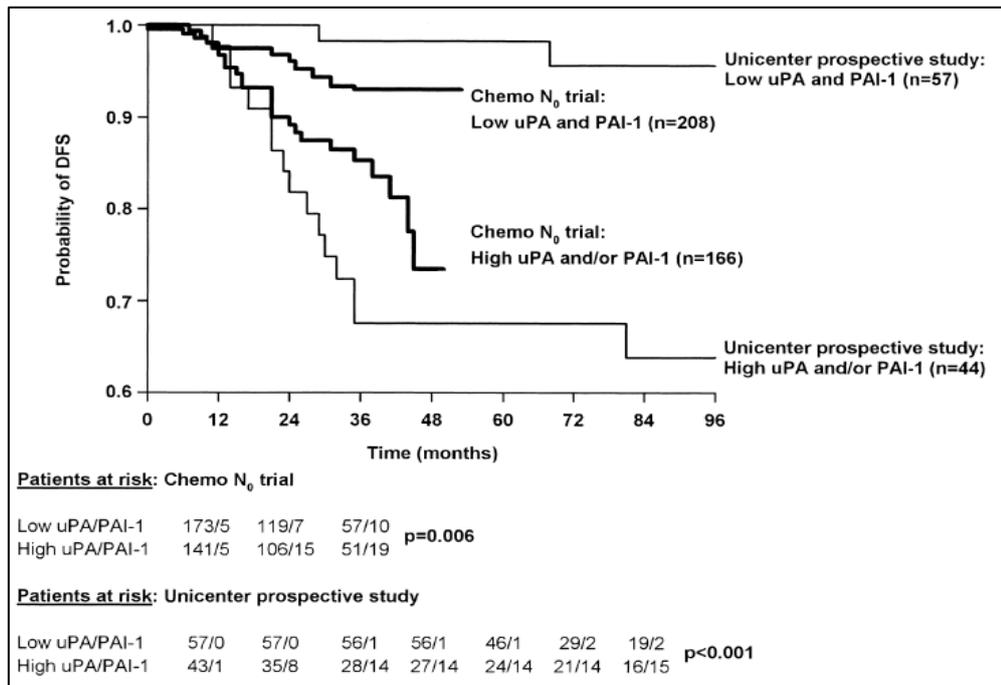


Abbildung 1. Bestätigung der Daten der unizentrischen Pilotstudie (dünne Linien) durch die multizentrische Feldstudie ChemoN0 (dicke Linien); aus: Jänicke et al. J Natl Cancer Inst 2001;93: 913–20).

Auch die Metaanalyse der EORTC Receptor and Biomarker Group (Look et al. 2002) bestätigt die unabhängige und signifikante prognostische Bedeutung des uPA/PAI1-Status beim nodalnegativen Mammakarzinom. Aus diesen Daten (Pilotstudie, Validierungsstudie und Metaanalyse) ergibt sich zusammenfassend die Begründung, Patientinnen mit niedrigem uPA/PAI1-Status eine adjuvante Chemotherapie zu ersparen.

Es gibt keinen anderen Prognosefaktoren oder Biomarker, deren klinische Wertigkeit basierend auf einem biologischen Modell so konsequent durch prospektive Pilotstudie, multizentrische Validierungsstudie und Metaanalyse bestätigt worden ist. Dies ist im Übrigen auch die Begründung, warum sowohl in den amerikanischen Empfehlungen (Harris et al. 2007) als auch in den deutschen Leitlinien der uPA/PAI1-Status mit einem sehr hohen Evidenzniveau bewertet wurde.

Beachtet werden sollte auch, dass die prognostische Bedeutung des uPA/PAI1 Status sich nicht nur auf das krankheitsfreie Überleben bezieht, sondern auch deutlich im Sterberisiko abgebildet wird (Abb. 2; Harbeck et al. 2013).

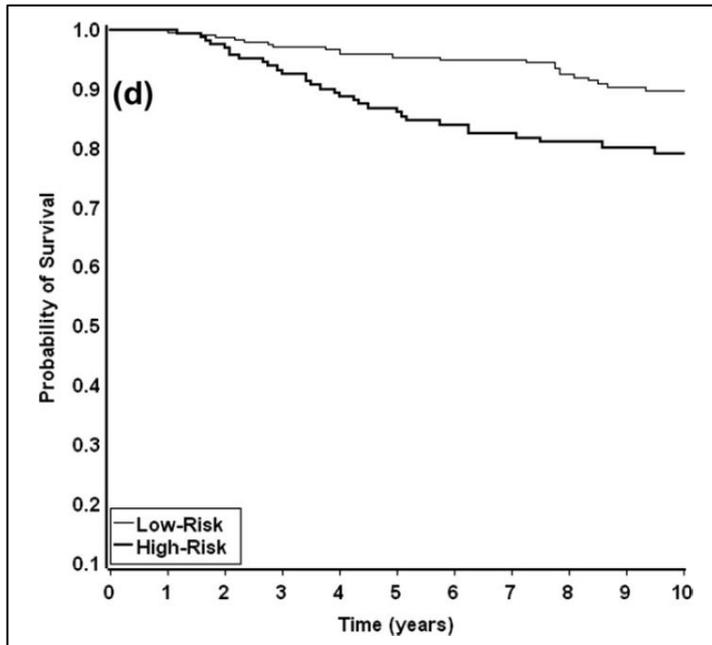


Abbildung 2. Wahrscheinlichkeit des Gesamtüberlebens bei nodalnegativen Patientinnen ohne adjuvante Therapie (keine Chemotherapie, keine endokrine Therapie) in Abhängigkeit vom uPA/PAI1-Status des Primärtumors (aus: Harbeck et al. Eur J Cancer 2013;49(8):1825-35)

Die zweite Fragestellung, prädiktive Bedeutung des uPA/PAI1-Status, wird im IQWiG-Vorbericht ausschließlich unter Bezugnahme auf die aktuell publizierten Daten der ChemoNO-Studie diskutiert. Es muss darauf hingewiesen werden, dass aufgrund der Situation in den 90er Jahren viele Patientinnen nach der Randomisation den experimentellen Arm der Chemotherapie abgelehnt hatten, so dass wir davon ausgehen, dass der Effekt in der intention to treat-Analyse (ITT) unterschätzt wird (HR= 0.74; 95%CI=0.44-1.27; n.s.); aus diesem Grunde haben wir eine per protocol-Analyse (PP) durchgeführt, die den signifikanten Effekt der adjuvanten Chemotherapie klar belegt (adjustierte HR=0,48; 95%-K.I. 0,26-0,88; p=0,019). Der Chemotherapieeffekt in der Hochrisikogruppe wurde in ähnlicher Größenordnung bereits bei der Erstpublikation beschrieben (Jänicke et al. 2001).

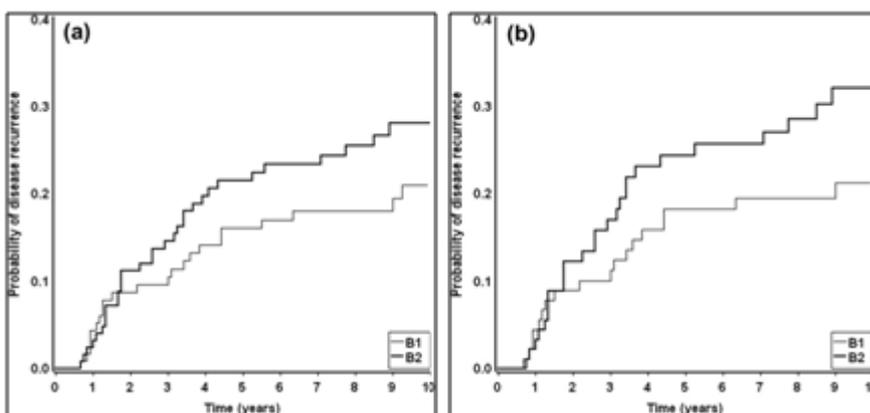


Abbildung 3. Links ITT-Analyse (a) und rechts PP-Analyse (b) des Chemotherapie-Effektes in der Hochrisikogruppe (uPA/PAI1-Status hoch). (a) ITT; CMF (n=117) 10-J-DFS 21,0% (95% CI 14,2-30,4);keine CT (n=125)10-J-DFS 28,1% (95% CI 20,6-37,6); HR=0,74 (95% CI = 0,44–1,27); n.s.(b) PP; CMF (n=91) 10-J-DFS 21,3% (95% CI 13,9-31,9); keine CT (n=90) 10-J-DFS 32,1% (95% CI 22,9-43,8); HR=0,48 (95% CI = 0,26–0,88); p = 0,019)

Rechnerisch und formal im IQWiG-Bericht korrekt dargestellt, ist die ITT-Analyse negativ gewesen. Wenn man aber die Kurvenverläufe der ITT-Analyse und die der PP-Analyse vergleicht, sieht man grundsätzlich vergleichbare Ergebnisse (Abb.3; Harbeck et al. 2013); mit der PP-Analyse wird kein Widerspruch zur ITT-Analyse formuliert, sondern nur CMF-Effekt bei bereinigten Daten besser herausgearbeitet.

Die Voraussage der Sensitivität des Tumors (Prädiktion) gegenüber einer Chemotherapie ist durch die Analyse des uPA/PAI1-Status ebenfalls gegeben. Zusätzlich zur ChemoN0-Studie existieren Ergebnisse aus einer hoch publizierten Metaanalyse mit insgesamt 3.424 Patientinnen aus mehreren Datensets. Der Chemotherapie-Benefit für Patientinnen mit hohem uPA/PAI1-Status ist in dieser Metaanalyse signifikant erhöht (Abb. 4; Harbeck et al. 2002)

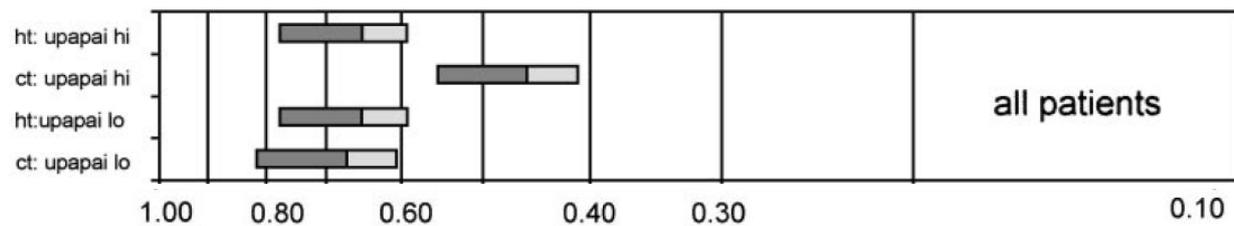


Abbildung 4. Prädiktive Bedeutung des uPA /PAI1-Status (*upapai hi*): hohe Werte ergeben einen hohen Benefit von einer adjuvanten Chemotherapie (*ct*), nicht aber endokrinen Therapie (*ht*); bei niedrigem uPA/PAI-1 Status (*upapai lo*) kein Unterschied zwischen den Effekten der endokrinen Therapie und der Chemotherapie (aus: Harbeck et al. Cancer Research 62: 4617–4622; 2002)

Evidenzbewertung von Prognosefaktoren und prädiktiven Faktoren

Bezüglich der Bewertung des Evidenzniveaus sei angemerkt, dass für die Entwicklung von klinisch relevanten Biomarkern (i.e. Prognosefaktoren, prädiktive Faktoren) sich auf internationalem Niveau inzwischen die Anforderungen deutlich geändert haben. Da es klar geworden ist, dass aufgrund methodischer und ethischer Einschränkungen sich ein prospektives Studiendesign, wie bei Chemo-N0 noch durchgeführt, heutzutage verbietet, werden nach neuer Evidenztabelle auch Daten zu Biomarkern, die retrospektiv aus Teilkollektiven prospektiv-randomisierter Studien mit nicht-Biomarker-fokussiertem Studienziel gewonnen wurden, noch der höchste Evidenzgrad zugestanden wird (Simon et al. 2009). Nach Anwendung dieses Bewertungsschemas erreicht der uPA/PAI1-Status mit prospektiver RCT und Metaanalyse als Basis genauso wie moderne Genexpressionstests ohne prospektive Studien als Basis das höchste Evidenzniveau (LoE I; Abb. 1 & 2; AGO Empfehlungen 2013).

Methodische und ethische Einschränkungen können sich daraus ergeben, dass Patientinnen mit niedrigem Rückfall-Risiko wegen der potentiellen Nebenwirkungen nicht (auch nicht randomisiert im Rahmen einer klinischen Studie) mit einer adjuvanten Chemotherapie behandelt werden können; somit kann und konnte diese Forderung des IQWiG-Vorberichtes a priori nicht erfüllt werden (siehe 5.3.2. Krankheitsfreies Überleben im IQWiG-Vorbericht). Gerade bei Prognosefaktoren mit hoher Güte, die das Niedrigrisiko mit hoher Spezifität erkennen, wird dies zum methodischen Problem: Beim Design der ChemoN0-Studie lagen der Arbeitsgruppe nur die Daten aus der unizentrischen Münchner Pilotstudie vor, dabei ergab sich für Patientinnen mit niedrigem uPA/PAI1-Status ein 5-Jahres-

Rückfallrisiko von nur 5%, aus der theoretisch möglichen relativen Risikoreduktion von 25% durch die damals üblichen Chemotherapien (CMF) hätte maximal ein absoluter Benefit von 1,25% weniger Rückfälle erreicht werden können. Bei nahezu 99% Übertherapie verbietet sich ein solches Studiendesign.

Weitere methodische und ethische Einschränkungen ergeben sich daraus, dass heutzutage so gut wie alle Patientinnen eine adjuvante systemische Therapie (mindestens Tamoxifen) erhalten, somit es nicht mehr möglich ist, gegen eine völlig unbehandelte Kontrollgruppe zu randomisieren, wie es in den 90er Jahren (z.B. in der ChemoNO-Studie) noch möglich war; eine Aussage zur reinen prognostischen Wertigkeit eines Biomarkers ohne Therapieeinfluss kann daher mit heutigen Patientinnenkollektiven nicht mehr getätigt werden.

Fazit

Prospektive klinische Studien belegen, dass aufgrund des Testergebnisses im Gesamtkollektiv nodal-negativer Patientinnen bis zu 55%¹ der Patientinnen eine Chemotherapie und die damit verbundenen Nebenwirkungen erspart werden kann (Jänicke et al. 2001). Darüber hinaus kann mittels des uPA- und PAI-1-Werte auch der Effekt einer adjuvanten Chemotherapie (hohe Werte – stärkerer Effekt) abgeschätzt werden. Die Methode ist hinsichtlich ihrer prognostischen Bedeutung in prospektiven, retrospektiven Studien und in einer Metaanalyse validiert worden. Damit wurde sie als einzige neue Methode zur Prognoseabschätzung auf höchstem Evidenzniveau (Oxford-Level LoE 1a) validiert.

Uns erscheint es wichtig zu betonen, dass die Fragestellung und Forschung in Bezug auf die Verbesserung der Abschätzung des Rückfallrisikos und der individualisierten Therapie seit nahezu 25 Jahren national und international die Therapeuten und Wissenschaftler beschäftigt. Neben den traditionellen Faktoren wie axillärer Lymphknotenstatus, Grading, Tumorgröße, Alter der Patientin, Steroidhormonrezeptorstatus (ER, PgR) und dem prädiktiven Faktor HER2 haben nur die Prognosefaktoren uPA/PAI-1 einen hohen Grad an wissenschaftlicher Evidenz erreicht. Intention für die Entwicklung neuer Methoden der Prognoseabschätzung ist es, Chemotherapie gezielter einzusetzen und Patientinnen vor Schaden zu bewahren.

Die Ergebnisse der fachlichen Auseinandersetzung und Diskussion zu den Prognosemarkern uPA/PAI-1 erfordert die klinische Umsetzung in den Routinealltag. Die Gesundheit und Genesung der Patientinnen, denen eine Chemotherapie erspart werden kann, sollte die sektorübergreifende Abrechnung in unserem Krankenkassensystem möglich machen.

Aufgrund der aktuellen Form der Prognoseabschätzung und neuer gezielt eingesetzter Medikamente (antihormonelle Therapie, anti-HER2-Therapie) begrenzt sich heute die Fragestellung auf Patientinnen mit hormonrezeptorpositiven Tumoren mit unklarem Risikoprofil (in der Regel intermediäres Risiko). Patientinnen mit HER2-überexprimierenden Tumoren erhalten in der Regel eine Chemotherapie in Kombination mit anti-HER2-Substanzen (S3-Leitlinie 2012), Patientinnen mit sogenannten tripelnegativen Tumoren (bezogen auf ER, PgR, HER2) erhalten fast immer eine Chemotherapie, u. U. sogar mit Einsatz von Platinkomplexen (GeparSixtoStudie, 2013; CALGB #40603, 2013).

¹ Beobachtung aus der Pilotstudie bei Betrachtung einer Gesamtkohorte (Jänicke et al. 1991); in der klinischen Realität, durch Begrenzung des Einsatzes auf ER pos. HER2 neg. G2-Tumoren ist dieser Wert niedriger anzusetzen.

1.4 Weitere Kommentare zur Analyse der ChemoNO-Studie im IQWiG Vorbericht

1.4.1 ad Zusammenfassung des Vorberichtes

Es ist klar, wenn man sich a priori auf den prädiktiven Nutzen eines Biomarkers beschränkt und den prognostischen Zusatznutzen vollständig außer Acht lässt, dass dann nur eine Studie zur Bewertung zur Verfügung steht. Wir halten auch die prognostische Aussage für klinisch relevant und für die Patientinnen nützlich und möchten vorschlagen, dass auch dieser Nutzen bewertet wird. Dafür können zahlreiche prospektive und retrospektive Kohortenstudien herangezogen werden, die im vorliegenden Bericht unterschlagen wurden.

Erstes Zielkriterium der ChemoNO-Studie war die prognostische Fragestellung (Welche Patientinnen haben so ein niedriges Rezidivrisiko, dass ihnen eine adjuvante Chemotherapie erspart werden kann?) Gerade das Hybriddesign (Anreicherungsdesign mit zusätzlicher Nachbeobachtung der nicht randomisierten Teilpopulation) ist zur Evaluation der prognostischen Bedeutung geeignet.

Zusammenfassend wird der uPA/PAI1-Bestimmung der Zusatznutzen abgesprochen. Im Gegensatz zum Fazit des IQWiG-Berichtes sehen wir den Nutzen für die Patientinnen ganz eindeutig darin, dass durch Einsatz dieser Marker bis zu 55% der Patientinnen die Chemotherapie erspart werden kann (bei niedrigem uPA/PAI1-Status). Es ist zu kurz gegriffen, nur die Prädiktion zu betrachten.

1.4.2 ad xii Schlagwörter des Vorberichts

Es handelt sich mitnichten um einen „Urin-Plasminogenaktivator“. Korrekt wäre die Bezeichnung: „Plasminogenaktivator vom Urokinasetyp“.

1.4.3 ad Seite 3: uPA/PAI-1 als prädiktive Marker

Patientinnen mit intermediärem Risiko sind eine heterogene Gruppe, die sich aus high risk- und low risk-Patientinnen zusammensetzt. Es war erklärte primäre Intention der Forschung, mittels der uPA/PAI1-Testung, die low risk-Patientinnen zu identifizieren, um ihnen eine adjuvante Chemotherapie zu ersparen. Der hohe Nutzen dieser Testung in Bezug auf diese Fragestellung konnte in der Münchner Pilotstudie und in der confirmierenden ChemoNO-Studie und in der EORTC-Metaanalyse bewiesen werden.

Die Fragestellung zur Prädiktion ist in diesem Zusammenhang eine zweite wissenschaftliche Frage. Wir sind der Meinung, dass durch die Konsistenz der Ergebnisse der zwei Analysen der ChemoNO-Studie zusammen mit der oben zitierten Metaanalyse zum prädiktiven Effekt (Harbeck et al. 2002) ebenfalls überzeugend gezeigt werden konnte, dass ein besonderer Nutzen der Chemotherapie bei Patientinnen mit hohem uPA/PAI1-Status gegeben ist.

Im Übrigen wären andere Ergebnisse auch schwer erklärbar: Wenn im Kollektiv mit niedrigem uPA/PAI1-Status das Rückfallrisiko zu niedrig ist, um überhaupt einen relevanten Effekt der Chemotherapie zu zeigen, muss das high risk-Kollektiv folglich den grundsätzlich immer zu beobachtenden Effekt einer adjuvanten Chemotherapie (auch bei nodalnegativer Situation; EBCTCG 2012) abbilden.

1.4.4 ad Seite 4: 2 Ziel der Untersuchung

Schon ein niedriger uPA/PAI1-Status spricht gegen eine adjuvante Chemotherapie, da bei niedrigem Rezidivrisiko der absolute Effekt einer Chemotherapie a priori als niedrig anzusehen ist. Die prädiktive Aussage eines Markers hinsichtlich der Chemotherapiesensitivität kann somit nicht das einzige Kriterium sein.

1.4.5 ad Seite 7, 4.1.4 Studientypen

Prospektiv randomisierte Studien (RCTs) sind ausschließlich für die Bewertung als prädiktiver Marker notwendig, für die Bewertung der prognostischen Relevanz sind Kohortenstudien (optimal prospektiv) ausreichend. Nach den Vorschlägen des Oxford Centre of Evidence Based Medicine wird der höchste Evidenzlevel vergeben, wenn „Systematic review of inception cohort studies“ vorhanden sind. (OCEBM Levels of Evidence Working Group 2014; download 27.04.2014, <http://www.cebm.net/?o=1025>).

1.4.6 ad Seite 16, 5.1.1 Literaturrecherche

Aufgrund der Einschränkung der Fragestellung auf die potentielle prädiktive Bedeutung des uPA/PAI1-Status und die ausschließliche Suche nach RCTs ist es klar, dass die Literatur für diese Fragestellung erheblich reduziert wurde (es existiert nur eine prospektive randomisierte Studie; für andere Prognosefaktoren existieren bis dato überhaupt keine Ergebnisse aus prospektiv randomisierten Studien (außer für Thymidin labeling Index, Paradiso et al. 2001).

Wenn sich das IQWiG auf die unseres Erachtens wichtigere prognostische Fragestellung eingelassen hätte, wären naturgemäß deutlich mehr Studien in die Bewertung gelangt. Zur prognostischen Relevanz des uPA/PAI1-Status existieren zahlreiche prospektive und retrospektive Kohortenstudien mit weitgehend homogenem Ergebnissen, sowie auch Metaanalysen (Look et al., 2000, 2002).

1.4.7 ad Seite 32, 6.1 Eingeschlossene Studien

An dieser Stelle wird auf die relativ hohe Rate an Rezidiven in der low risk-Gruppe hingewiesen und damit argumentiert, dass die prognostische Information keinen patientenrelevanten Nutzen darstelle. Hierbei wird allerdings außer Acht gelassen, dass die Patientinnen keine adjuvante endokrine Therapie erhalten haben. Hochgerechnet aus dem hohen Anteil von Patientinnen mit Östrogenrezeptor positiven (ER)-positiven Tumoren (90%, s Tab 1A, Harbeck et al. 2013) kann geschätzt werden, dass sich das Rezidivrisiko unter 5 Jahren adjuvante Tamoxifen Therapie von 12,9% um relativ 40% auf 7,5% verringert hätte. Der Chemotherapieeffekt würde sich damit auf unter 2% (absoluter Benefit) schätzen lassen, so dass, wie oben angemerkt, schon aus ethischen Gründen diesen Patientinnen keine adjuvante Chemotherapie angeboten hätte werden dürfen.

1.4.8 ad Seite 34, 6.3 Ergebnisse

Die Reduktion der Bedeutung prognostischer Zusatzinformationen ist den Autoren und Verantwortlichen der Studien bewusst. Wir haben daher in den aktuellen Empfehlungen den Einsatz der uPA/PAI1-Testung auf Patientinnen mit ER-positiven und HER2-negativen G2 Tumoren begrenzt.

In der ChemoNO-Studie wurde tatsächlich ausschließlich CMF eingesetzt; wir wissen aber aus dem EBCTCG Overview, dass der Effekt moderner Therapie (Anthrazykline, Taxane) generell höher ist. Auch wurden in der Metaanalyse zum prädiktiven Wert des uPA/PAI1-Status über 20% Patientinnen mit anthrazyklinhaltigen Therapien eingeschlossen.

1.4.9 Ad Seite 36, 6.6 Vergleich mit anderen Biomarkern

Unverständlich ist, dass im IQWiG-Vorbericht so starke Erwartungen in die potentielle Wertigkeit der Genexpressionsprofile gesetzt werden. Es existieren überhaupt keine prospektiven Daten. Selbst die Pilotstudien für diese Tests wurden anhand retrospektiver selektierter Kollektive gewonnen. Auch die Validierung dieser Tests erfolgte allein auf der Basis retrospektiver Datensets. Für die uPA/PAI1-

Bewertung wurden dagegen vom IQWiG alle Daten aus retrospektiven Analysen, inklusive der Metaanalysen, aus der Bewertung ausgeschlossen.

Die uPA/PAI1-Studien waren im Grunde genommen Pilotstudie für die Evaluation von klinisch relevanten Biomarkern allgemein. Konzept und Design sind in den Studien zu den Gen-Expressionsprofilen aufgenommen und weiterentwickelt worden. Selbst mit aktualisierten Designs zeigen sich ähnliche Schwierigkeiten der Umsetzbarkeit solcher Konzepte.

Auswertungen von Studienkonzepten mit einem RCT Design wie in der MINDACT- und TAILORx- Studie stehen aktuell nicht zur Verfügung, so dass noch keine Rückschlüsse auf deren Machbarkeit und Umsetzbarkeit gezogen werden können. Eine erste Analyse der MINDACT-Studie zur Machbarkeit und Implementierung des Tests zeigt, dass ein Drittel der registrierten Patientinnen im Rahmen der Studie behandelt wurden (Cardoso et al 2008).

Bezugnehmend auf die Aussage der Dokumentation und Würdigung der Anhörung zum vorläufigen Berichtsplan vom 19.12.2013 unter 2.1.2 zu den genannten Multiparameter-Gen Expressionsprofilen ist es somit völlig offen, ob tatsächlich in naher Zukunft positive Aussagen zum Zusatznutzen dieser Test gemacht werden können. Die bisherige Evidenz dieser Profile basiert ausschließlich auf retrospektive Auswertungen, prospektive Daten zu diesen Tests liegen bisher nicht vor.

1.4.10 ad Seite 37, 7 Fazit

Dem Fazit können wir nicht folgen. Wie oben dargestellt, halten wir die sich aus dem uPA/PAI1-Status gewonnene prognostische Information für einen relevanten Zusatznutzen für den klinischen Alltag. Der wahrscheinliche absolute Benefit für die einzelne Patientin kann bereits aus der Prognoseabschätzung geschlossen werden. Darüber hinaus sehen wir auch ausreichend Evidenz, dass Patientinnen mit hohen uPA/PAI1-Werten einen Nutzen aus der adjuvanten Chemotherapie ziehen würden (Jänicke et al. 2001; Harbeck et al. 2013; Harbeck et al. 2002).

2. Benennung von zusätzlichen, im Vorbericht nicht genannten, relevanten Studien,

2.1. siehe oben unter 1.3 dieser Stellungnahme unseren Kommentar zu Metaanalyse und Prädiktion

2.2. Für die Entwicklung von Prognosefaktoren ist die randomisierte kontrollierte Studie (RCT) nicht das geeignete Instrument. Im Allgemeinen werden Analysen für die Bedeutung dieser Faktoren anhand von – im besten Fall prospektiven Kohortenstudien – vorgenommen. Nach den Vorschlägen des Oxford Centre for Evidence-Based Medicine werden Prognosefaktoren ausschließlich auf der Basis von Kohortenstudien und Metaanalysen (*What will happen if we do not add a therapy? (Prognosis): LoE 1 Systematic review of inception cohort studies; LoE 2 Inception cohort studies; LoE 3: Cohort study or control arm of randomized trial.* OCEBM Levels of Evidence (LoE) Working Group 2011) mit dem höchsten Evidenzniveaus bewertet. RCTs sind für die Beurteilung der prognostischen Wertigkeit von Biomarkern auch vom OCEBM nicht vorgeschrieben.

Für die Entwicklung von klinisch relevanten Biomarkern (i.e. Prognosefaktoren, prädiktive Faktoren) haben sich auf internationalem Niveau die Anforderungen deutlich geändert. Da es klar geworden ist, dass aufgrund methodischer und ethischer Einschränkungen sich ein prospektives Studiendesign, wie bei der ChemoNO-Studie noch durchgeführt, heutzutage verbietet, wird nach neuer Evidenztafel

auch Biomarkern, die retrospektiv in Teilkollektiven prospektiv-randomisierter Studien mit nicht-Biomarker-fokussiertem Studienziel evaluiert wurden, noch der höchste Evidenzgrad zugestanden. (Simon et al 2009). Auch bei Anwendung dieses aktuellen Bewertungsschemas erreicht uPA/PAI1 mit prospektiver RCT und Metaanalyse als Basis genauso wie moderne Genexpressionstests ohne prospektive Studien als Basis das höchste Evidenzniveau (LoE I; Abb. 1 & 2; AGO Empfehlungen 2013).

Methodische und ethische Einschränkungen können sich daraus ergeben, dass Patientinnen mit niedrigem Rückfall-Risiko wegen der potentiellen Nebenwirkungen nicht (auch nicht randomisiert im Rahmen einer klinischen Studie) mit einer adjuvanten Chemotherapie behandelt werden können; somit kann und konnte diese Forderung des IQWiG-Vorberichtes a priori nicht erfüllt werden (siehe 5.3.2. Krankheitsfreies Überleben im IQWiG-Vorbericht). Gerade bei Prognosefaktoren mit hoher Güte, die das Niedrigisiko mit hoher Spezifität erkennen, wird dies zum methodischen Problem: Beim Design der ChemoNO-Studie lagen der Arbeitsgruppe nur die Daten aus der unizentrischen Münchner Pilotstudie vor, dabei ergab sich für Patientinnen mit uPA/PAI1-Status ein 5-Jahres-Rückfallrisiko von nur 5%, aus der theoretisch möglichen relativen Risikoreduktion von 25% durch die damals üblichen Chemotherapien (CMF) hätte maximal ein absoluter Benefit von 1,25% weniger Rückfälle erreicht werden können. Bei nahezu 99% Übertherapie verbietet sich ein solches Studiendesign.

Weitere methodische und ethische Einschränkungen ergeben sich daraus, dass heutzutage so gut wie alle Patientinnen eine adjuvante systemische Therapie (mindestens Tamoxifen) erhalten, somit ist es nicht mehr möglich, gegen eine völlig unbehandelte Kontrollgruppe zu randomisieren, wie es in den 90er Jahren (z.B. in der ChemoNO-Studie) noch möglich war. Eine Aussage zur reinen prognostischen Wertigkeit eines Biomarkers ohne Therapieeinfluss kann daher mit heutigen Patientinnenkollektiven nicht mehr getätigt werden.

3. Verweise auf andere qualitativ angemessene Unterlagen

Zusammenfassend geben wir die in Deutschland relevanten Leitlinien zur Therapie und Diagnostik des Mammakarzinom an, die eine Standardisierung der Diagnose, Therapie gewährleisten.

Leitlinien, die die Prognosemarker uPA/PAI-1 empfehlen

- AGO Leitlinien: Die Kommission Mamma der Arbeitsgemeinschaft für Gynäkologische Onkologie (AGO) hat es sich zur Aufgabe gemacht, klinisch relevante Fragen nach Analyse der Evidenzlage zu bewerten und eine Empfehlung für den klinischen Einsatz auszusprechen. Die Kommission erkennt den uPA/PAI1-Status seit 2002 als Prognosemarker mit der Bewertung Level of Evidence 1a und empfiehlt die Bestimmung (AGO+) beim nodalnegativen Mammakarzinom (www.ago-online.org/, siehe Abb. 5 und 6 im Anhang).
- S3 Leitlinien: Die DKG sieht einen Nutzen dieser Faktoren für manche Patientinnen. In der aktuellen Version der evidenzbasierten S3-Leitlinie der Deutschen Krebsgesellschaft von 2008 steht, dass die Bestimmung weitere prognostische Informationen liefern kann. Eine Aktualisierung dieser Leitlinien ist geplant.
(http://www.krebsgesellschaft.de/index.php?seite=wub_llevidenzbasiert_mammakarzinom).
- ASCO Leitlinien: Die amerikanischen Leitlinien für die Therapie des Mammakarzinoms empfehlen die Bestimmung dieser Biomarker beim nodalnegativen Mammakarzinom neben dem Hormonrezeptorstatus, der HER2-Bestimmung und der Nutzung des PCR-basierten Oncotype DX Recurrence-Scores (American Society of Clinical Oncology 2007 Update of Recommendations for the Use of Tumor Markers in Breast Cancer: Harris L et al. JCO 2007).

- Kopie aus dem Statement der AGO zu den Prognosemarkern uPA/PAI-1:
 „Vorteil der Patientinnen mit verfügbaren uPA/PAI-1-Werten
 Bei den meisten Patientinnen mit einem Mammakarzinom ohne Lymphknotenbefall erfolgt die Risikoabschätzung für ein Wiederauftreten der Krankheit entsprechend der vorhandenen klinisch-pathologischen Datenlage. Dementsprechend gehören nur ca. 20% dieses Kollektivs in die Gruppe der Patientinnen mit einem geringen Rückfallrisiko und erhalten keine Chemotherapie. Der Großteil der nodal-negativen Patientinnen erhält sicherheitshalber prophylaktisch eine Chemotherapie, obwohl nur ca. 25% der Patientinnen tatsächlich von einer Chemotherapie profitieren. Dies bedeutet eine enorme Übertherapie in der Gruppe der nodal-negativen Patientin. In den abgeschlossenen prospektiven Studien, in denen anhand von uPA/PAI-1 klassifiziert wurde, konnte der Anteil an Niedrigrisikopatientinnen auf das Doppelte erhöht werden. Die exzellente Datenlage zu den Prognosemarkern mit 10 Jahres Follow-Up-Daten erlaubt in unserer Verantwortung den Patientinnen gegenüber keinen Aufschub, die Bestimmung in den Alltag der Krankenversorgung einzuführen. Jede Patientin sollte die Möglichkeit haben, dass die Werte der Prognosemarker in die Entscheidungsfindung ihre Krebstherapie Einzug haben, um überflüssige Belastungen durch die Chemotherapie zu verhindern.“
- Die internationalen Empfehlungen „St Gallen Konsensus“ haben 2005 die Invasionsmarker uPA/PAI-1 als Marker insbesondere zur Abschätzung der Resistenz endokriner Therapien anerkannt.

4. Anmerkungen zur projektspezifischen Methodik unter jeweiliger Angabe

Siehe zu diesem Punkt bitte unseren direkten Kommentaren zum Vorbericht

5. Validierung der Prognosemarker im Überblick

Wir möchten nochmals die solide Entwicklung der Biomarker uPA/PAI1 für den Einsatz beim nodalnegativen Mammakarzinom zusammenfassend darstellen:

1. Entdeckung des prognostischen Wertes der uPA-Aktivität beim Mammakarzinom (Duffy et al. 1988)
2. Unizentrische Pilotstudie (Kohortenstudie) an der UFK München, TUM (n=345), (Jänicke et al. 1993) wird der prognostische Wert der Tumor-Konzentration von uPA und seines Inhibitors PAI-1 beschrieben (Messung mittels ELISA).
3. Die optimalen Schwellenwerte wurden mittels Maximum-Likelihood-Methode (logrank-Statistik) ermittelt: Für uPA stehen Werte unter 3 ng/mg Protein im Tumorgewebeextrakt für ein niedriges Rezidivrisiko, für PAI-1 wurde 14 ng/mg Protein als optimaler Schwellenwert ermittelt. Für die klinische Nutzung werden die Werte als sogenannter uPA/PAI-1-Status kombiniert: Als Niedrigrisiko gelten nur Tumore mit niedrigem uPA- und PAI-1-Konzentrationen; wenn einer der beiden Werte oder beide Werte erhöht sind, liegt ein erhöhtes Rezidivrisiko vor. **Etwa 50 – 55% der nodalnegativen Patientinnen könnte auf dieses Weise eine adjuvante Chemotherapie erspart werden.**
4. In einer multizentrischen internationalen Feldstudie (gefördert von der DFG; „ChemoN0“) wird der prognostische Wert beim nodalnegativen Mammakarzinom bestätigt und die Tauglichkeit des Testes bei dezentraler Bestimmung bestätigt. (Jänicke et al. 2001; Schmitt et al. 2010; Harbeck et al. 2013). Patientinnen mit erhöhten uPA und/oder PAI-1 Tumorkonzentrationen profitieren - beim randomisierten Vergleich - von einer adjuvanten Therapie mit CMF durch Halbierung des Rückfallrisikos.

5. In zahlreichen Kohortenstudien (national und international) wird die prognostische Wertigkeit von uPA und PAI-1 bestätigt. Interessant ist die tumorentitätsübergreifende Bedeutung dieser Faktoren. Prognostische Bedeutung hat der uPA/PAI-1-Test bei Karzinomen wie Nierenzellkarzinom, Magenkarzinom, Kolonkarzinom und anderen.
6. In einer Metaanalyse wird die unabhängige und starke prognostische Bedeutung von uPA und PAI-1 bestätigt (Look et al. 2002)
7. In einer Metaanalyse kann auch der prädiktive Wert von uPA/PAI-1 herausgearbeitet werden (Harbeck et al. 2003). Patientinnen mit Mammakarzinom, die ein hohe uPA und/oder PAI-1 Konzentration aufweisen, profitieren mehr als andere von einer adjuvanten Chemotherapie (Harbeck et al. 2002).
8. Die Qualitätssicherung mittels standardisierter Kontrollproben und regelmäßigen Ringversuchen belegt die technische Güte des Testverfahrens (Sweep et al. 1998). Es wurden auch genaue Parameter für die Präanalytik festgelegt (Schmitt et al. 2007). Robustheit, Lagerungsresistenz, Temperatur-Resistenz des Testverfahrens werden ebenfalls gezeigt (Promotion Bauerfeind U. 2008).
9. In einer vergleichenden Studie an Stanzbiopsie aus dem Tumor (core needle biopsies) wird die verlässliche Verwendbarkeit des Testes auch bei geringen Gewebemengen (im Vergleich zu den üblichen Mengen aus dem Resektat) gezeigt (Thomssen et al. 2009).
10. In einer der größten internationalen Studie zum nodalnegativen Mammakarzinom (n=4.147) wird die Machbarkeit der routinemäßige Nutzung des ELISA-Verfahrens bestätigt (Kantelhardt et al. 2011).

Literatur:

1. **Health economic impact of risk group selection according to ASCO-recommended biomarkers uPA/PAI-1 in node-negative primary breast cancer.**
Jacobs VR, Kates RE, Kantelhardt E, et al.
Breast Cancer Res Treat. 2013 Apr;138(3):839-50
2. **Ten-year analysis of the prospective multicentre Chemo-N0 trial validates American Society of Clinical Oncology (ASCO)-recommended biomarkers uPA and PAI-1 for therapy decision making in node-negative breast cancer patients.**
Harbeck N, Schmitt M, Meisner C, et al.
Eur J Cancer 2013; 49(8):1825-35
3. **Randomized adjuvant chemotherapy trial in high-risk, lymph node-negative breast cancer patients identified by urokinase-type plasminogen activator and plasminogen activator inhibitor type 1.**
Jänicke F, Prechtel A, Thomssen C, et al. N0 Study Group.
J Natl Cancer Inst. 2001; (93) 913-20
Urokinase (uPA) and its inhibitor PAI-1 are strong and independent prognostic factors in node-negative breast cancer
Jänicke F, Schmitt M, Pache L, et al.
Breast Cancer Res Treat 1993; 24:195–208
4. **Clinical relevance of the urokinase-type and tissue-type plasminogen activators and of their type 1 inhibitor in breast cancer**
Jänicke F, Schmitt M, Graeff H
Semin Thromb Hemost 1991; 17:303–12
5. **Pooled analysis of prognostic impact of uPA and PAI-1 in 8377 breast cancer patients.**
Look M, van Putten W, Duffy M, et al.
JNCI 2002; 94(2): 116-28
6. **Enhanced benefit from adjuvant chemotherapy in breast cancer patients classified high-risk according to urokinase-type plasminogen activator (uPA) and plasminogen activator inhibitor type 1 (n = 3424).**
Harbeck N, Kates RE, Look MP, et al.
Cancer Res. 2002; (62) 4617-22
7. **Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100 000 women in 123 randomised trials**
Early Breast Cancer Trialists' Collaborative Group (EBCTCG)
Lancet 2012; 379: 432–44
8. **American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer**
Harris L, Fritsche H, Mennel R, et al.
J Clin Oncol 2007, 25: 5287-312
9. **Use of Archived Specimens in Evaluation of Prognostic and Predictive Biomarkers**
Simon RM, Paik S, Hayes DF
J Natl Cancer Inst 2009; 101(21) :1446-52
10. OCEBM Levels of Evidence Working Group*. "The Oxford 2011 Levels of Evidence". Oxford Centre for Evidence-Based Medicine. <http://www.cebm.net/index.aspx?o=5653>. Download am 27.04.2014

11. **Randomized clinical trial of adjuvant fluorouracil, epirubicin, and cyclophosphamide chemotherapy for patients with fast-proliferating, node-negative breast cancer.**
Paradiso A, Schittulli F, Cellamare G et al.
J Clin Oncol. 2001 Oct 1;19(19):3929-37.
12. **Clinical application of the 70-gene profile: the MINDACT trial**
Cardoso F, van't Veer, Rutgers E.
J Clin Oncol 2008; 26(5): 729-35
13. **Feasibility of measuring the prognostic factors uPA and PAI-1 in core needle biopsy breast cancer specimens.**
Thomssen C, Harbeck N, Dittmer J et al.
J. Natl. Cancer Inst. 2009: 101, 1028-1029 (2009).
14. **Prospective evaluation of prognostic factors uPA/PAI-1 in node-negative breast cancer: Phase III NNBC3-Europe trial (AGO, GBG, EORTC-PBG) comparing 6 x FEC versus 3 x FEC / 3 x Docetaxel.**
Kantelhardt EJ, Vetter M, Schmidt M, Veyret C, Augustin D, Hanf V, Meisner C, Pöpke D, Schmitt M, Sweep F, von Minckwitz G, Martin PM, Jänicke F, Thomssen C, Harbeck N.
BMC Cancer. 2011 Apr 16;11(1):140

 Prognostic Factors II in Early Breast Cancer					
Factor	LoE ₂₀₀₈	CTS	LoE _{Ox2001}	GR	AGO
➤ER / PgR	II	B	2a	B	+
➤HER2 (IHC, FISH)	II	B	2b	B	+
➤ER/PgR/HER2 as surrogate markers for molecular subtypes	I	B	2b	B	+
➤uPA / PAI-1 (ELISA) in N0	I	A	1a	A	+
➤ Proliferation markers					
➤ Ki-67 before, during or after treatment	II	B	2b	B	+
➤ Mitotic activity Index (MAI)	I	A	1a	A	+

 Prognostic Factors III in Early Breast Cancer					
Factor	LoE ₂₀₀₈	CTS	LoE _{Ox2001}	GR	AGO
➤Tumor cell detection in bone marrow	I	B	1a	B	+/-
➤Circulating tumor cells	I	B	1b	B	+/-
➤Validated multigene arrays in HR+ (Oncotype DX®, EndoPredict®)	I	B	2b	B	+/-
➤Mammprint® in N0-1	II	C	2b	B	+/-
➤Computer guided decision tools (Adjuvant®)	II	C	2b	B	+/-
➤Mammostrat	I	B	2b	B	+/-
➤PAM50	II	B	2b	B	+/-
➤IHC4	I	B	2b	B	+/-

Abbildung 5 und 6: Auszug aus AGO Leitlinien zur Diagnostik Therapie des Mammakarzinoms 2013
 Siehe dazu auch gesondertes Dokument (Auszug_AGO_Leitlinien_2014).