

Mündliche Anhörung



**Gemeinsamer
Bundesausschuss**

gemäß 5. Kapitel § 19 Abs. 2 Verfahrensordnung
des Gemeinsamen Bundesausschusses

hier: Wirkstoff Lenvatinib

Sitzung im Hause des Gemeinsamen Bundesausschusses in Berlin
am 10. November 2015
von 11.26 Uhr bis 12.37 Uhr

– Stenografisches Wortprotokoll –

Angemeldete Teilnehmer der Firma **Eisai GmbH**:

Frau Krapfenbauer
Herr Mehlig
Herr Dr. Schwenke
Herr Dr. Steinert

Angemeldete Teilnehmer der Firma **Bayer Vital GmbH**:

Frau Dr. Hoffmann
Herr Tamoschus

Angemeldete Teilnehmer für die **Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie e. V. (DGHO)**:

Herr Prof. Dr. Fassnacht-Cappeller (nicht anwesend)
Herr Prof. Dr. Wörmann

Angemeldeter Teilnehmer für die **Sektion Schilddrüse der Deutschen Gesellschaft für Endokrinologie e. V.:**

Herr Dr. Kreissl (nicht anwesend)

Angemeldeter Teilnehmer für die **Deutsche Gesellschaft für Nuklearmedizin e. V. (DGN)**:

Herr Prof. Dr. Luster (nicht anwesend)

Angemeldeter Teilnehmer der Firma **medac Gesellschaft für klinische Spezialpräparate mbH**:

Herr Meier

Angemeldete Teilnehmer für den **Verband forschender Arzneimittelhersteller e. V. (vfa)**:

Herr Dr. Rasch
Herr Dr. Werner

Beginn der Anhörung: 11.26 Uhr

(Die angemeldeten Teilnehmer betreten den Raum)

Herr Prof. Hecken (Vorsitzender): Meine sehr verehrten Damen und Herren, herzlich willkommen zur mündlichen Anhörung im Rahmen der frühen Nutzenbewertung. Wir haben es mit der Bewertung eines Orphans zu tun. Basis des heutigen mündlichen Stellungnahmeverfahrens ist die Dossierbewertung des Gemeinsamen Bundesausschusses vom 1. Oktober 2015, die Ihnen allen hinlänglich bekannt ist. Stellung genommen haben im schriftlichen Stellungnahmeverfahren zum einen Eisai GmbH, die DGHO, die Sektion Schilddrüse der Deutschen Gesellschaft für Endokrinologie e. V., die Deutsche Gesellschaft für Nuklearmedizin e. V., Bayer Vital GmbH, medac und der Verband forschender Arzneimittelhersteller.

Ich begrüße – für das Protokoll müssen wir das feststellen – Frau Krapfenbauer, Herrn Mehlig, Herrn Dr. Schwenke und Herrn Dr. Steinert von Eisai, Frau Dr. Hoffmann und Herrn Tamoschus von Bayer, Herrn Professor Dr. Wörmann von der DGHO, Herrn Meier von medac und Herrn Dr. Rasch und Herrn Dr. Werner von vfa. Ist jetzt noch jemand da, der nicht aufgerufen wurde? – Nein, das ist nicht der Fall.

Üblicher Hinweis zum Verfahren: Wir führen wieder Wortprotokoll, deshalb bitte jeweils Namen, entscheidendes Unternehmen oder Institution nennen, wenn Sie sich zu Wort melden.

Wir werden uns heute sicherlich sehr intensiv mit der Frage des Ausmaßes des Zusatznutzens beschäftigen müssen, auf der Basis der doch relativ dünnen Datenqualität, die hier vorliegt, und mit der Frage der Nebenwirkungsprofile des in Rede stehenden Wirkstoffs. Das soll natürlich andere Themen nicht ausschließen. Mein Vorschlag wäre, dass zunächst der pU kurz einleitet, was die aus seiner Sicht wesentlichen Punkte sind. Dann würden wir in die Frage-und-Antwort-Runde eintreten. Wer möchte beginnen? – Herr Mehlig, bitte schön.

Herr Mehlig (Eisai): Vielen Dank, Herr Professor Hecken, für Ihre Einführung und für die Möglichkeit, eine kurze Einleitung geben zu dürfen. Bevor ich das tue, möchte ich gerne meine Kolleginnen und Kollegen vorstellen. Frau Daniela Krapfenbauer ist Gesundheitsökonomin und die bessere Hälfte in unserem deutschen Marktzugangs- und Erstattungssteam. Sie hat hauptverantwortlich die Erstellung des Nutzendossiers für Lenvatinib koordiniert. Herrn Dr. Carsten Schwenke kennen Sie von einer ganzen Reihe von Anhörungen. Er berät uns zu unseren Produkten von Anbeginn der frühen Nutzenbewertung an in methodischen und statistischen Fragestellungen. Herr Dr. Steinert ist der medizinisch-wissenschaftliche Experte in der Geschäftseinheit Onkologie. Einige von Ihnen kennen ihn noch von der letzten Anhörung zu Eribulin bei metastasiertem Brustkrebs. Herr Dr. Steinert ist auch für Lenvatinib von medizinischer Seite verantwortlich und betreut das Präparat. Mein Name ist Hilmar Mehlig. Ich bin Market Access Manager bei Eisai und die andere Hälfte im deutschen Marktzugangs- und Erstattungssteam.

Sehr geehrte Damen und Herren, Herr Professor Hecken hat es bereits gesagt, die Anhörung heute dreht sich um Lenvatinib. Lenvatinib ist ein Orphan Drug zur Behandlung erwachsener Patienten mit progressivem, lokal fortgeschrittenem oder metastasiertem radiojodrefraktärem differenziertem Schilddrüsenkarzinom. Der Gemeinsame Bundesausschuss hat sich in anderen Nutzenbewertungsverfahren bereits mit Schilddrüsenkarzinom beschäftigt, in diesem Fall mit medullärem Schilddrüsenkarzinom. Diese sind abzugrenzen von dem undifferenzierten anaplastischen Schilddrüsenkarzinom und dem für Lenvatinib relevanten differenzierten Schilddrüsenkarzinom. Schilddrüsenkarzinome zählen mit weniger als 2 Prozent pro Jahr an allen Tumorerkrankungen in Deutschland zu den seltenen Tumoren. Das differenzierte Schilddrüsenkarzinom repräsentiert die häufigste Form. Durch das meist

langsame Tumorwachstum, die gute Resezierbarkeit und adäquate Therapiestrategien wie Radiojodtherapie und TSH-Suppression ist die Prognose für die meisten Patienten mit Schilddrüsenkarzinom gut. Wenn ein differenziertes Schilddrüsenkarzinom nicht oder nicht mehr auf eine Radiojodtherapie anspricht, wird die Prognose jedoch deutlich schlechter. Die Patienten haben nach Metastasierung dann noch eine verbleibende Lebenserwartung von etwa drei Jahren. Für Patienten, für die chirurgische Maßnahmen und Radiojodtherapie ausgeschöpft sind, stand vor Beginn der Zulassungsstudie SELECT für das Präparat Lenvatinib keine relevante Behandlungsoption mehr zur Verfügung.

Es gibt unterschiedliche Einschätzungen zu der tatsächlichen Zahl der für Lenvatinib infrage kommenden Patienten in unserem Nutzendossier und dem Bericht des IQWiG dazu. Unstrittig ist dabei aber, dass es sich um eine sehr seltene Erkrankung handelt. Für diese Patienten mit dringendem therapeutischem Handlungsbedarf ist Lenvatinib eine zielgerichtete systemische Therapie, für die der Nachweis einer signifikanten und klinisch relevanten Lebenszeitverlängerung sowie einer deutlichen Verlängerung der Zeit bis zum Auftreten eines Progresses vorliegt.

Was ist und wie wirkt Lenvatinib? Lenvatinib ist ein multipler Rezeptor-Tyrosinkinase-Inhibitor mit einem breiten Inhibitionsprofil. Darüber hinaus zeichnet sich Lenvatinib durch einen neuen Bindungsmechanismus aus, der als Typ V bezeichnet wird. Diese Wirkmechanismen machen Lenvatinib einzigartig in seiner Wirkstoffklasse.

Interessanter noch, als zu wissen, wie Lenvatinib wirkt, ist, was es bewirkt und welchen Nutzen Patienten erfahren können. In der pivotalen Studie SELECT war Overall Survival, das Gesamtüberleben, ein sekundärer, aber sicher ohne jeden Zweifel patientenrelevanter Endpunkt. Es wurden für zwei Datenschnitte sowohl unadjustierte als auch wegen des erlaubten Crossover von Placebo-Patienten zu Lenvatinib adjustierte Analysen zum Gesamtüberleben durchgeführt. Insgesamt zeigen die Analysen zum Gesamtüberleben beider Datenschnitte mit und ohne Adjustierung gleichgerichtete Effekte, die für den zweiten Datenschnitt mit den aussagekräftigeren, reiferen Daten adjustiert statistisch signifikant waren. Über diesen Vorteil im Gesamtüberleben hinaus möchten wir gerne mit Ihnen die Validität und Patientenrelevanz des progressionsfreien Überlebens, des primären Endpunkts der SELECT-Studie, kurz PFS, besprechen.

Wir sind der Überzeugung, dass PFS im Zusammenhang mit dieser schwer zu behandelnden, lebensbedrohlichen Form des Schilddrüsenkarzinoms und den sehr ausgeprägten Behandlungsergebnissen für Lenvatinib patientenrelevant ist. Uns hat ermutigt, dass im Nutzenbewertungsbericht des G-BA, wenn auch nur in Teilen, unserer Argumentation gefolgt wird. Um knapp einige Aspekte zu nennen, die aus unserer Sicht dafür sprechen, dass in diesem konkreten Fall PFS patientenrelevant ist: 99 Prozent der Patienten, die in die Studie eingeschlossen wurden, litten bereits zu Beginn der Studie unter Metastasen. 45 Prozent der Patienten unter Lenvatinib und 48 Prozent der Patienten im Placebo-Arm waren bei Randomisierung gemäß des ECOG-Status symptomatisch.

Wie aggressiv die Tumorbiologie der eingeschlossenen Patienten war, zeigt sich eindrücklich im medianen PFS-Wert von nur 3,6 Monaten in der Placebo-Gruppe, im Gegensatz dazu beträgt der mediane PFS-Wert der Patienten, die mit Lenvatinib behandelt wurden, 18,3 Monate. Es ist für uns schwer vorstellbar, dass das hochsignifikant verlängerte Aufhalten des Progresses der bereits bestehenden und neuen Läsionen um eineinviertel Jahre angesichts einer Gesamtlebenserwartung nach Metastasierung von etwa drei Jahren nicht patientenrelevant sein kann. Unterstützt wird die Patientenrelevanz der PFS-Ergebnisse durch eine objektive Ansprechrates von 64,8 Prozent unter Lenvatinib, verglichen mit 1,5 Prozent unter Placebo. 99 Prozent der Patienten erfuhren unter Lenvatinib eine Tumorreduktion. Eine teilweise oder in vier Patienten sogar vollständige Remission erlebten 70,4 Pro-

zent der Patienten innerhalb der ersten 30 Tage der Therapie. Eine Komplettremission, wie ich sie eben erwähnt habe, die darüber hinaus länger anhaltend war, ist beispiellos in der Literatur.

88 Prozent der Patienten im Placebo-Arm entschieden sich in Abstimmung mit ihrem Arzt bewusst dafür, die Möglichkeit des Crossover nach Auftreten des Progresses wahrzunehmen. Dies drückt für uns neben der zweifelsfrei nachgewiesenen Wirksamkeit erneut die Patientenrelevanz des progressionsfreien Überlebens eindrücklich und unmittelbar aus.

Für die Bewertung der Verträglichkeit wird im Nutzenbewertungsbericht unserer Einschätzung gefolgt, dass die Time-to-first-Event-Analysen des Datenschnitts vom 15. März 2014 herangezogen werden sollten. Allerdings vertreten wir weiterhin die Ansicht, dass der Endpunkt Gesamtzahl unerwünschter Ereignisse, UE, nicht in die Bewertung der Verträglichkeit eingehen sollte, da er eben auch beispielsweise Laborwertveränderungen der CTCAE Grad 1 und 2 beinhaltet, die nicht patientenrelevant erscheinen. Zur Ableitung des Ausmaßes des Zusatznutzens sollte die Gesamtzahl UE also nicht verwendet werden. Die unerwünschten Ereignisse CTCAE Grad 3 und 4, die schwerwiegenden unerwünschten Ereignisse und die unerwünschten Ereignisse, die zum Therapieabbruch führten, sollten natürlich für die Bewertung der Verträglichkeit von Lenvatinib herangezogen werden. Für diese Endpunkte zeigte sich ein deutlich höheres Schadenspotenzial von Lenvatinib, was durch den Vergleich gegenüber Placebo allerdings, wie vom G-BA bestätigt, zu erwarten war. Die unter der Therapie mit Lenvatinib beobachteten unerwünschten Ereignisse wurden als überwiegend kontrollierbar und behandelbar eingestuft. Insgesamt ist das Sicherheitsprofil von Lenvatinib vorhersehbar und beherrschbar. Dem gegenüber stehen nachhaltige, bisher nicht erreichte große Verbesserungen des therapie-relevanten Nutzens bezüglich des Gesamtüberlebens einerseits sowie der stark verlängerten Zeit bis zum Eintritt eines Progresses andererseits, die damit als erheblich einzuschätzen sind. Unter Beachtung der Schwere der Erkrankung sowie der sehr beschränkten Therapieoption, verbunden mit dem dringenden Therapiebedarf, schätzen wir das Ausmaß des Zusatznutzens von Lenvatinib als erheblich ein für Patienten mit progressivem, lokal fortgeschrittenem oder metastasiertem differenziertem Schilddrüsenkarzinom, das nicht auf eine Radiojodtherapie anspricht.

Nochmals vielen Dank für die Möglichkeit dieser Einleitung.

Herr Prof. Hecken (Vorsitzender): Herzlichen Dank, Herr Mehlig. – Fragen? – Frau Teupen, bitte.

Frau Teupen: Warum haben Sie keine Daten zur Lebensqualität erhoben?

Herr Prof. Hecken (Vorsitzender): Herr Mehlig, bitte.

Herr Mehlig (Eisai): Sie treffen einen wunden Punkt, und das gleich mit der ersten Frage. Ich möchte vielleicht einleitend dazu sagen: Als die SELECT-Studie geplant wurde, gab es zu diesem Zeitpunkt keine krankheitsspezifischen Lebensqualitätsinstrumente, die man hätte verwenden können. Die gibt es, soweit ich das weiß, auch heute noch nicht. In der Rückschau hätte man sicherlich die Entscheidung auch anders treffen können und dann immerhin unspezifische Lebensqualitätsinstrumente in die Studie aufnehmen können. Das hat man nicht gemacht, und das ist sicherlich ein Mangel an Taten, den Sie zu Recht ansprechen. Wir haben das intern sehr lange diskutiert. Es gibt einige Initiativen, die ich gerne den Kollegen Steinert ausführen lassen möchte, wenn Ihnen das recht ist, Herr Professor Hecken, um diese Datenlücken zumindest in der Zukunft teilweise zu schließen.

Herr Prof. Hecken (Vorsitzender): Ja, bitte.

Herr Dr. Steinert (Eisai): Wie Herr Mehlig schon angesprochen hat, lagen für die Zulassungsstudie die Lebensqualitätsdaten nicht vor. Dieses Problems ist man sich bewusst. Es werden international im Rahmen einer Beobachtungsstudie Lebensqualitätsdaten erhoben. Es ist eine sehr junge Studie, in die noch relativ wenig Patienten aufgenommen wurden. Hier werden sie mit den Fragebögen FAQ-G, EQ-5D-3L und dem Treatment Satisfaction Questionnaire for Medication erhoben. Gegenwärtig findet eine Rekrutierung für eine Studie statt, die 211 heißt. Im Rahmen dieser 211-Studie werden ebenfalls Lebensqualitätsdaten erhoben.

Herr Prof. Hecken (Vorsitzender): Danke schön. – Frau Dr. Grell.

Frau Dr. Grell: Herr Mehlig, Sie hatten abgehoben auf die Operation als mögliche Option bei diesem Karzinom. Ich finde im Protokoll nicht, dass alle Patienten inoperabel waren. Man kann ja Metastasen Chirurgie machen. Sie hatten in dem Lenvatinib-Arm vier Patienten mit lokalem. Wie kann man sich das vorstellen?

Herr Prof. Hecken (Vorsitzender): Herr Steinert, bitte.

Herr Dr. Steinert (Eisai): Ein zentrales Einschlusskriterium der Studie war, dass die Patienten nachgewiesen radiorefraktär waren und dass sie metastasiert waren und dass eine Resektion keine Option war. Meines Wissens wurde zu 99 Prozent eine Resektion vorgenommen. Demzufolge wurden die Patienten eingeschlossen. Das heißt, eine Resektion als Therapieoption zum Aufhalten des Progresses oder der Erkrankung war für die Patienten, die eingeschlossen waren, keine Option mehr.

Herr Prof. Hecken (Vorsitzender): Nachfrage, Frau Grell.

Frau Dr. Grell: Dann habe ich die Frage: Wo steht das im Protokoll? Ich habe es nämlich nicht gefunden. Sie hatten außer Metastasierten auch noch lokal Fortgeschrittene. Wenn das im Protokoll steht, ist das ja in Ordnung.

Herr Prof. Hecken (Vorsitzender): Ja, bitte, Herr Steinert.

Herr Dr. Steinert (Eisai): Im Rahmen der Studie waren 99 Prozent der Patienten metastasiert. Die Daten zu dem lokal Fortgeschrittenen können wir gerne nachreichen.

Herr Prof. Hecken (Vorsitzender): Frau Grell.

Frau Dr. Grell: Wo steht im Protokoll, dass die nicht operabel waren? Das war meine Frage. Es waren vier Lokale.

Herr Prof. Hecken (Vorsitzender): Herr Steinert.

Herr Dr. Steinert (Eisai): Das können wir gerne nachreichen. Ich kann Ihnen die Frage jetzt nicht konkret beantworten.

Herr Prof. Hecken (Vorsitzender): Okay. – Herr Wörmann.

Herr Prof. Dr. Wörmann (DGHO): Es ist eigentlich nicht meine Aufgabe, die Firmenarbeit zu übernehmen. In der Studie war das für uns für die Bewertung wichtig: 90 Prozent der Patienten hatten Lungenmetastasen, die eingeschlossen waren, und 40 Prozent Knochenmetastasen. Knochenmetastasen würden wir nicht operieren lassen. Lungenmetastasen sind in der Regel multiple. Insofern wä-

re das ein Grund, zu sagen, dass die meisten nicht operabel waren. Aber das muss ich eigentlich nicht beantworten. Zumindest war die Studie so angelegt. Das sind die publizierten Daten dazu.

Herr Prof. Hecken (Vorsitzender): Danke schön, Herr Professor Wörmann. – Herr Mayer.

Herr Dr. Mayer: Ich habe zwei Fragen. Es wurde vorhin angemerkt, dass die Lebensqualitätsdaten fehlen. Was mich auch interessiert: Warum wurden keine Morbiditätsdaten erhoben außer PFS? Ein Schilddrüsenkarzinom macht sicher auch lokal Symptome, aber auch die Metastasen: Schmerzen, Schluckstörungen und Ähnliches, Horner-Syndrom. Das Fehlen von Lebensqualitätsdaten kann ich nach Ihrer Argumentation nachvollziehen; da gab es keine spezifischen Bögen. Aber Symptome hätte man mit erheben können. Das fehlt mir.

Die zweite Frage richtet sich an die Statistiker, wahrscheinlich Herrn Schwenke. Sie haben adjustiert mit dem RPSFT-Modell. Soweit ich darüber Bescheid weiß – ich bin kein Statistiker –, gibt es noch kein allgemein akzeptiertes Adjustierungsverfahren, das immer gültig wäre. Ich meine aber gelesen zu haben, die FDA kritisiert gerade in diesem Fall die Anwendung dieses Modells erheblich, weil die Grundannahme, dass der Überlebensvorteil von Lenvatinib über den gesamten Zeitraum, also nach Randomisierung und auch nach Progress, für die Placebo-Patienten ein gleichbleibender und auch immer gleich großer Effekt ist, von der FDA widerlegt worden ist. Sie sagen, die Placebo-Patienten, die Lenvatinib nach Progress bekommen haben, profitieren statistisch signifikant weniger als die Lenvatinib-Patienten, die von Anfang an mit Lenvatinib behandelt worden sind. Das heißt, die Grundannahme der gleichförmigen Wirksamkeit ist nicht vorhanden. Ich glaube, auch die EMA zieht diese adjustierten Daten in ihrer Begründung nicht wirklich heran. Ich wollte Ihre Einschätzung wissen, wie weit Sie diese Daten des letzten Datenschnittes für die Adjustierung als explorativ oder tatsächlich als beweisend ansehen.

Herr Prof. Hecken (Vorsitzender): Wer macht Teil 1? Wer macht Teil 2? – Herr Mehlig, bitte.

Herr Mehlig (Eisai): Darf ich allgemein auf das Modell antworten und weitergeben, wenn es spezifisch wird? Vielleicht fange ich von hinten an. Die EMA hat, wie das im EPAR zu lesen ist, die Daten als supportiv anerkannt. Sie haben recht: Die EMA hat im EPAR das Modell kritisiert. Es ist auch ihre Aufgabe, die vorgelegten Daten kritisch zu hinterfragen. Herr Dr. Schwenke hat mir das ausführlich erklärt. Es gibt in der statistischen Wissenschaft nicht das eine Modell, wie Sie ganz am Anfang auch gesagt haben, das für alle Adjustierungen geeignet wäre. Das Rank Preserving Structural Failure Time Model ist eines der gängigsten Modelle zur Adjustierung. Es ist in der Literatur, die wir in der Stellungnahme teilweise angeführt haben, als eines der guten Modelle und fallweise – in unserem Fall sehen wir das so – als das bestgeeignete Modell für die Situation dargestellt. Das Modell wird niemals die reine Wahrheit darstellen können. Dafür ist es ein Modell. Es muss für den Crossover adjustiert werden, weil die Verzerrung sonst überhaupt keine Aussagekraft für die Daten mehr zulässt. Den Crossover mussten wir aus ethischen Gründen in diesem Studiendesign anbieten.

So ist das Rank Preserving Structural Failure Time Model schon im Studienprotokoll, also von Anfang an, als das Modell hinterlegt worden, mit dem man die Adjustierung durchführen wollte. Die Adjustierung ist a priori geplant gewesen und so für den zweiten Datenschnitt, der von der EMA eingefordert wurde, auch eingefordert worden. Bei diesem zweiten Datenschnitt mit den aussagekräftigsten und reifen Daten sieht man einen signifikanten Überlebensvorteil für das Präparat.

Vielleicht gebe ich an diesem Punkt für die weiteren Tiefen des RPSFTM an Herrn Dr. Schwenke weiter, der zum Thema Symptomatik noch ein Wort sagen kann.

Herr Prof. Hecken (Vorsitzender): Herr Schwenke.

Herr Dr. Schwenke (Eisai): Vielen Dank. – Vielleicht ganz kurz zu diesem Modell. Letztendlich ist die Annahme, die getroffen wird, um dieses Modell anwenden zu können, die Annahme, dass die Patienten, egal zu welchem Zeitpunkt sie mit der aktiven Substanz starten, immer den gleichen Überlebensvorteil haben. Man kann verschiedene Varianten machen, um diese Annahme zu prüfen. Was nicht geht, ist, einen statistischen Test durchzuführen, um zu schauen, ob die Annahme zutrifft oder nicht. Das heißt, was man machen kann, ist zum Beispiel, grafisch zu schauen, wie sich die Hazards – nicht das Hazard Ratio, sondern die Risiken – über die Zeit verteilen. Dazu haben wir Grafiken im Modul 4, wo man sehr schön sieht, dass die Hazards bis auf gewisse Schwankungen insgesamt relativ konstant sind.

Was wir uns zusätzlich angeschaut hatten, war, wie das second PFS, das heißt die Zeit von der ersten bis zur zweiten Progression, bei den Patienten aussieht, die Crossover hatten. Wir haben gesehen, dass diese Zeit vergleichbar lange ist wie bei den Patienten, die direkt mit Lenvatinib angefangen haben. Wir lagen bei 18 Monaten für die Lenvatinib-Patienten gegenüber knapp 22 Monaten für die Placebo-Patienten, die nach Progression Crossover hatten.

Man muss natürlich berücksichtigen, dass die Patienten, die ein Crossover hatten, eine positive Selektion sind. Das sind ja die Patienten, die nicht gestorben sind, die die Möglichkeit hatten, überhaupt Crossover zu haben. Das heißt, hier erwartet man schon eine gewisse Überschätzung dieser Zeit.

Die FDA und auch die EMA diskutieren diese Modelle sehr kritisch. Das macht auch Sinn; denn, wie Herr Mehlig schon sagte, es gibt kein bestes Modell, es gibt eine Reihe guter Modelle, es gibt eine Reihe schlechter Modelle. Die schlechten Modelle sollte man logischerweise nicht verwenden. Dazu gibt es Simulationsstudien, in denen Wissenschaftler untersucht haben, welche Modelle gar nicht funktionieren und welche Modelle in den allermeisten Fällen funktionieren. Das RPSFT-Modell funktioniert in der Tat in den meisten Fällen. Das heißt, wenn man dieses Modell anwendet und vor allem wenn man prospektiv plant, dieses zu tun, und das dann auch durchführt, dann kann man mit den Ergebnissen etwas anfangen.

Es ist, wie Herr Mehlig auch schon sagte, natürlich ein Modell. Die Wahrheit kann man damit nicht abbilden, aber man kann sich zumindest der Wahrheit nähern. Selbst wenn man davon ausgehen würde, dass man jetzt eine Überschätzung des Effektes hat, dass man auf der anderen Seite, wenn man sich das OS ohne Adjustierung anschaut, eine dramatische Unterschätzung hat, dann muss der mittlere Effekt zumindest tendenziell in Richtung des Ergebnisses des RPSFT-Modells sein. Dementsprechend wären wir immer noch hochgradig signifikant mit den Daten.

Die EMA hat diese Analysen in der Tat nur als supportiv angesehen. Aber der Hintergrund ist hauptsächlich, dass die EMA PFS als primären Endpunkt akzeptiert und deswegen die Zulassung auf Basis des PFS aussprechen kann und OS als solches sowieso nur supportiv ist in diesem ganzen Verfahren. Somit hatte die EMA insgesamt gar nicht die Notwendigkeit, OS besonders zu wichten. Sie war natürlich an den Daten interessiert, logisch, auch interessiert an adjustierten Daten. Aber die Zulassung wurde prinzipiell auf Basis des PFS ausgesprochen, das eben von dem Crossover nicht beeinflusst ist, dass man da eine vernünftige Abschätzung hat, diese 14,7 Monate Vorteil gegenüber Placebo, sodass auf Basis dessen die Zulassung ausgesprochen wird und nicht auf Basis des OS.

Herr Prof. Hecken (Vorsitzender): Nachfrage dazu, Herr Mayer.

Herr Dr. Mayer: Aber Sie gehen davon aus, dass die EMA das anders eingeschätzt hätte als die FDA. Die EMA sagte, man musste gar nicht das OS beurteilen. Die FDA hat ganz klar dieses Modell

hier als eigentlich nicht angezeigt bezeichnet. So habe ich es zumindest nachgelesen. Sie haben zwei Hazard Ratios verwandt, da haben Sie recht. Trotzdem hat man gesagt, es ist auch nur explorativ, aber es gibt Hinweise, dass es nach Progress nicht mehr so wirkt wie zu Beginn eingesetzt. Damit wird das, was beim Placebo – ich sage es einmal ganz platt – als fiktive Überschätzung der Placebo-Gruppe abgezogen wird, überschätzt gewertet. Das heißt, man wird dem Lenvatinib einen größeren Placebo-Effekt zusprechen, als eigentlich vorhanden ist. – Das sei aber dahingestellt.

Ich habe eine konkrete Frage. Soweit ich das gelesen habe, auch zur Methodik, bedarf es einer Adjustierung oder einer Korrektur dieser Unsicherheit über eine Konfidenzintervallkorrektur, die in der Regel immer eine größere Unsicherheit nach sich zieht. Die Konfidenzintervalle werden breiter als in der Primäranalyse. Sie haben aber im zweiten Datenschnitt, in der korrigierten Analyse, noch engere Konfidenzintervalle generiert. Wurde hier ein Korrekturfaktor angewandt? Denn es gibt Analysen, die besagen, den Effektschätzer kann man bestimmen. Der ist entsprechend niedriger und besser. Man hat sozusagen eine doppelte Verbesserung der relativen Überlebenszeit. Aber eine Signifikanz auch über die Konfidenzintervalle ist eigentlich nicht zu bestimmen, jetzt explorativ. Manche adjustieren die Konfidenzintervalle gar nicht oder machen den p-Wert nicht anders. Sie haben ganz konkret einen hoch signifikanten p-Wert berechnet und auch enger werdende Konfidenzintervalle gezeigt. Mich interessiert, wie das methodisch gemacht worden ist, ob Sie adjustiert haben. Wie kann die Unsicherheit, die logischerweise in einem solchen Modell steckt, zu engeren Konfidenzintervallen führen?

Herr Prof. Hecken (Vorsitzender): Herr Schwenke, bitte.

Herr Dr. Schwenke (Eisai): Die Konfidenzintervalle für das Hazard Ratio in dem adjustierten Modell wurden auf dem sogenannten Bootstrap-Verfahren basiert. Bootstrap-Verfahren sind Verfahren, wo man aus den Daten, die man hat, Stichproben zieht mit zurücklegen. Das heißt, man kann Patienten in dieser Stichprobe mehrfach ziehen. Man zieht Stichproben, die genauso groß sind wie die Studie selbst. Man berechnet darüber den Konfidenzbereich.

Dieses Verfahren ist ein übliches Verfahren, wenn es keine numerischen oder keine formelbasierten Verfahren mehr gibt, um Konfidenzintervalle bestimmen zu können. In diesem Fall war es so, dass die Konfidenzintervalle relativ eng waren, weil die Daten sehr homogen sind. Das heißt, wenn ich eine sehr homogene Datenbasis habe, bekomme ich auch relativ enge Konfidenzintervalle. Das ist bei anderen Studien in der Tat anders, wo ich mit Ausreißern eine höhere Streuung in meinem Bootstrap-Verfahren generiere, sodass die Konfidenzintervalle breiter werden. In diesem Fall war es so, dass wir nicht so viele Ausreißer hatten, sodass insgesamt die Konfidenzintervalle sehr präzise geschätzt werden können. Das Konfidenzintervall, das aussagt, dass der wahre Wert mit hoher Sicherheit in diesem Bereich liegt, wird dadurch eng, dass die Daten homogen sind und ich dann davon ausgehen kann, dass das Hazard Ratio in einem relativ engen Bereich liegt. Das ist die Methodik dahinter.

Herr Prof. Hecken (Vorsitzender): Danke schön. – Frau Rozynska, bitte schön.

Frau Rozynska: Ganz kurz zu der Grundannahme dieses Modells, die gerade schon angesprochen wurde, bezüglich der FDA-Kritik. Die FDA ist zu dem Schluss gekommen, dass die Grundannahme des gemeinsamen Behandlungseffekts nicht erfüllt ist. Wie stehen Sie dazu? Wie können Sie das erklären?

Herr Prof. Hecken (Vorsitzender): Herr Schwenke.

Herr Dr. Schwenke (Eisai): Letztendlich kann man dem nur entgegenstellen, was die EMA dazu gesagt hat. Die EMA hat geschrieben, dass das Modell prinzipiell adäquat ist. Auch die FDA hat im Vor-

hinein nicht abgelehnt, dieses Modell in das Protokoll aufzunehmen. Bei der Entwicklung des Protokolls war die FDA beteiligt. Im Nachhinein ist die FDA in der Tat zu diesem Schluss gekommen. Die EMA ist zu einem anderen Schluss gekommen.

(Zuruf: Das stimmt nicht!)

– Die EMA hat nicht gesagt, dass dieses Modell das einzig Wahre ist, ganz sicher. Das Modell ist, wie schon gesagt, eben ein Modell. Aber es ist eine bessere Abschätzung des Effektes, als wenn man die unadjustierten Analysen genommen hätte.

Herr Prof. Hecken (Vorsitzender): Nachfrage.

Frau Rozynska: Soweit ich die Aussage der EMA verstanden habe, kritisiert die EMA grundsätzlich das Modell. Sie sagt, dass es gewisse grundsätzliche Schwächen hat und sie es aus diesem Grund nur unterstützend heranzieht. Akzeptiert hat sie es nach unserer Lesart nicht bezüglich der Daten, die vorgelegt worden sind. Aber die FDA schaut sich konkret die Daten an und überprüft die Hypothese, dass hier ein gemeinsamer Effekt vorliegt, und kommt dann zu dem Schluss, dass dies nicht der Fall ist. Die Frage ist halt: Welche Auswirkung hat das auf die Aussagekraft dieses Endpunktes, der nur durch diese Analyse signifikant wird?

Herr Prof. Hecken (Vorsitzender): Herr Mehlig, bitte.

Herr Mehlig (Eisai): Ich maße mir nicht an, das besser zu wissen als Herr Dr. Schwenke, um Gottes willen. Aber ich würde es gerne auf die Tatsache zurückführen, dass wir sowohl unadjustiert als auch adjustiert Analysen hatten und dass wir gleichgerichtete Effekte gesehen haben, für den ersten Datenschnitt adjustiert und unadjustiert, und gleichgerichtete Effekte für den zweiten Datenschnitt mit den aussagekräftigeren, reiferen Daten, unadjustiert und adjustiert. Sie haben völlig recht, was die EMA in ihrem EPAR geschrieben hat, ist, dass sie das Modell kritisiert wegen der Hypothese, die man letztendlich eben nicht beweisen kann, und dass eine Überschätzung die Folge sein kann. Das steht da drin. Das ist auch im Nutzenbewertungsbericht des G-BA aufgegriffen und zitiert worden. Nichtsdestotrotz haben wir die gleichgerichteten Effekte. Wir haben ein Hazard Ratio von 0,53, und wir haben einen unteren Wert des Konfidenzintervalls für Lenvatinib von 30,9 Monaten, wohingegen wir ein medianes Überleben im Placebo-Arm von 19 Monaten haben. Das ist schon ein deutlicher Unterschied.

Noch einmal: Es ist ein Modell, es nähert sich der Wahrheit an, es ist a priori so definiert worden, dass es verwendet werden soll. Herr Dr. Schwenke hat es schon gesagt. Die Gesundheitsbehörden haben jeweils in der Protokollabstimmung dem ja auch zugestimmt. Das ist das Modell, das vorab festgelegt wurde, das zur Anwendung gekommen ist. Es ist eines der gängigen Modelle, um die Adjustierung durchzuführen. Es ist aus unserer Sicht – wir haben in unserer Stellungnahme die zusätzliche Auswertung hineingebracht – ein sehr gut geeignetes Modell gewesen, um die Adjustierung durchzuführen. Ich denke, wichtig ist, dass wir für alle vier Werte, die wir angeschaut haben, gleichgerichtete Effekte sehen. Bei den unadjustierten Analysen waren die eben nicht signifikant. Für den zweiten Datenschnitt und die größte Zahl der Ereignisse war es adjustiert dann aber sehr wohl und deutlich signifikant mit einem Hazard Ratio von 0,53, womit wir uns eigentlich sehr wohl fühlen.

Herr Prof. Hecken (Vorsitzender): Danke schön. – Ergänzend noch, Herr Steinert.

Herr Dr. Steinert (Eisai): Ich möchte von der stark methodisch geprägten Diskussion den Blick auf die Situation des Patienten richten. Wie wir gesehen haben, ist das ein Patient, der in der Placebo-

Gruppe ein medianes PFS von 3,6 Monaten hat und der mit der Therapie von Lenvatinib eine Verbesserung im medianen PFS von 14,7 Monaten, also eineinviertel Jahren, erreicht hat, im Vergleich zu einer Gesamtlebenserwartung von drei Jahren. Es ist natürlich aus der medizinisch-biologischen Sicht schwer vorstellbar, dass das keinen Einfluss auf das Gesamtüberleben hat, unabhängig davon, ob man das adäquateste Statistikmodell einsetzt.

Herr Prof. Hecken (Vorsitzender): Danke schön. Sie sollten noch im Kopf behalten – vielleicht für später –, dass die Frage von Herrn Mayer noch offen ist, die lautete, wieso Sie nur das PFS angeschaut haben und nicht Dinge, die sich auch in der Morbidität abbilden, Schluckbeschwerden usw., usf. Das müssen wir gleich noch kurz nachfragen. – Frau Müller, bitte.

Frau Dr. Chr. Müller: Ich wollte kurz zurück zur Methodik – das ist ein wichtiger Punkt –; denn wir haben ein Problem: Bei großen Behandlungsunterschieden haben wir fast regelhaft ein Crossover; das ist auch verständlich. Wir müssen in irgendeiner Art und Weise versuchen, mit den Overall-Survival-Daten umzugehen. Ich glaube, es ist schon sehr wichtig, zu verstehen. Ich meine, Sie haben das gemacht; ich finde es sehr gut, dass einmal versucht wurde, zu adjustieren; das wird ja nicht immer versucht. Ich habe Verständnisfragen dazu. Es gibt unterschiedliche Modelle. Es stimmt, das ist eines, das relativ häufig angewandt wird und auch im britischen Raum positiv diskutiert wird. Sie haben darauf abgehoben, Sie hätten das im Studienprotokoll schon „für den Fall, dass“ vorgesehen. Nun kann ich nicht so ganz verstehen: a) weiß ich vorher nicht, ob ich wirklich einen Behandlungsunterschied habe und ein Crossover nötig wird, b) kann ich vorher eigentlich die Voraussetzungen, so wie ich das verstanden habe, nicht prüfen. Ich weiß ja nicht, ob die Wirkung davon abhängig ist, ob ich das zuerst einsetze oder erst nach einem Progress. Die Argumentation ist mir nicht ganz klar. Wenn Sie dazu vielleicht ein, zwei Sätze sagen können.

Kurze Frage vielleicht an Sie, Herr Dr. Schwenke: Welche anderen Modelle haben Sie überhaupt in Erwägung gezogen? Ich habe das bis jetzt so verstanden, dass man sich entscheidet und die Voraussetzung prüft, wenn man bestimmte Informationen hat. Oder haben Sie einfach das am häufigsten verwendete im Studienprotokoll definiert und dann durchgeführt? Ich denke da auch an Weiterentwicklung wie zum Beispiel Branson and Whitehead, die davon ausgehen, dass die Überlebenszeiten parametrische Formen annehmen. Das wäre zum Beispiel eine Möglichkeit, wo man noch etwas anderes drin hätte. Es gibt eben auch andere Modelle, die nicht auf dieser zeitunabhängigen Wirkung basieren. Haben Sie die geprüft, und können Sie dazu noch ein, zwei Worte sagen?

Herr Dr. Schwenke (Eisai): Sehr gerne. – Die erste Frage war, warum genau dieses Modell und kein anderes. Bei der Planung von klinischen Studien hat man bestimmte Annahmen, wie der Behandlungseffekt aussieht. In diesem Fall war es so, dass es aus den Phase-II-Studien sehr gute Hinweise darauf gab, dass eine hohe Wirksamkeit vorliegt. Deswegen wurde unter anderem das Crossover im Studienprotokoll erlaubt und auch mit den Behörden diskutiert, dass die Firma davon ausgeht, dass der Behandlungseffekt so groß ist, dass ein Crossover wahrscheinlich erfolgen wird. Ich weiß nicht, wie es im globalen klinischen Team war, aber dass 88 Prozent der Patienten Crossover haben werden, das ist schon wirklich eine Hausnummer. Damit hatte wahrscheinlich kaum jemand gerechnet, dass wirklich so viele ein Crossover haben.

Auf der anderen Seite ist der Behandlungseffekt von plus 14,7 Monaten in der Progression, den man in der Phase III gesehen hat, eben auch eine Hausnummer, wo man sagt: Auch das konnte man so aus den Phase-II-Studien nicht zwangsläufig erwarten. Das heißt, der Effekt ist schon gigantisch.

Auf der anderen Seite muss man sich in der Planung solcher Studien überlegen: Welches Szenario erwarte ich, und was wäre in diesem Szenario ein gutes Modell? In der Tat wählt man dann bei der

Planung ein Modell, das – erstens – häufig verwendet wird und – zweitens – von dem man aus Simulationsstudien weiß, dass in zumindest der überwiegenden Zahl der Fälle, der Situationen, dieses Modell auch gut funktioniert, nicht zwangsläufig in jeder Situation immer das allerbeste ist, aber zumindest ein gutes Modell, sodass – darauf kommt es an – das Konfidenzintervall so abgeschätzt werden kann, dass man nicht zu optimistisch ist, das heißt, das Konfidenzintervall nicht zu eng wird. Es gibt eben einige Modelle, die das Konfidenzintervall zu schmal machen, die nicht häufig eingesetzt werden, weil sie einfach nicht adäquat sind. Dieses Modell ist eben adäquat in vielen Situationen.

Man darf auch nicht vergessen: Wenn man in das Protokoll 15 verschiedene Adjustierungsmethoden hineinschreibt, ist die Frage: Welches nimmt man? Was ist das beste gerade in Bezug darauf, dass man die Annahmen relativ schlecht testen kann? Man kann die Annahmen wie in diesem Fall nur grafisch sich anschauen, und das haben wir auch gemacht. Das sieht man im Modul 4. Das ist das Einzige, was geht. Die anderen Verfahren sind schön und gut. Ich mag mir die Situation nicht vorstellen, wenn wir hier mit 15 verschiedenen Adjustierungen aufgetreten wären. Denn dann weiß man wieder nicht, welches Modell von den 15 verschiedenen das richtige ist. Was wir im Moment haben, ist zumindest laut FDA und zum Teil auch EMA die nicht adjustierte Variante, die eine absolute Unterschätzung ist, und auf der anderen Seite das adjustierte Modell, das zumindest laut FDA eine gewisse Überschätzung sein kann, sodass sich der wahre Wert wahrscheinlich im oberen Drittel wiederfindet; man weiß nicht genau, wo er liegt. Deswegen wurde bei dieser Studie im Studienprotokoll das eine Modell festgelegt, damit man die eine handfeste konkrete Anweisung hat, wie die Analyse laufen soll. In der Tat wurden danach auch nicht andere Adjustierungsverfahren verwendet und an die Daten angepasst. Das heißt, man hat nicht die 15, 16, 18 verschiedenen Varianten, die theoretisch möglich wären, durchprobiert, sondern ist bei dem vorgeplanten Approach geblieben.

Herr Prof. Hecken (Vorsitzender): Danke schön. – Zur Ergänzung, Herr Wörmann.

Herr Prof. Dr. Wörmann (DGHO): Ich wollte kurz einwenden, dass das schön öfter diskutiert wurde. Ich glaube, wir diskutieren hier bei der Überlebenszeit zwei Fragen, einmal: Was macht Lenvantinib?, und: Macht die frühe Therapie gegenüber einer späteren – ich glaube, es sind 88 Prozent – einen Unterschied? Das heißt, wir mischen hier zwei Punkte. Diese 12 Prozent Unterschied machen jetzt den Überlebensvorteil. Das ist statistisch signifikant und interessant. Aber ich weiß nicht, ob man sich zu sehr daran verbeißen darf. Es wird schwierig sein, einen Überlebensvorteil zu zeigen.

Ich habe einen zweiten Punkt. Wir sind eigentlich der Meinung, dass Placebo als Kontrolle in Deutschland nicht korrekt ist, sondern eigentlich ist Sorafenib die Kontrolle. Wenn wir nach Nutzen und nicht allein nach der Wirkung schauen, dann muss man es gegen Sorafenib sehen, weil beide zugelassen sind. Dann kommen wir, glaube ich, überhaupt nicht mehr auf einen Überlebensvorteil.

Herr Prof. Hecken (Vorsitzender): Herr Mehlig.

Herr Mehlig (Eisai): Wenn ich darauf direkt kurz entgegnen darf. – Herr Professor Wörmann, das ist die Situation heute. Der Grund, warum seinerzeit für die SELECT-Studie Placebo gewählt wurde, war die Tatsache, dass zu dieser Zeit Sorafenib noch lange nicht zugelassen war. Die Studie für unser Präparat ist im Juli 2011 begonnen worden, und Sorafenib ist, glaube ich, im April oder Mai 2014 in Europa zugelassen worden. Somit hatten wir, als wir die Studie design und begonnen hatten, keine relevante Therapiealternative, gegen die wir einen Head-to-Head-Vergleich hätten durchführen können. Dass inzwischen zwei Präparate für diese Erkrankung zugelassen sind, ist die Situation heute, aber eben nicht die Ausgangssituation, als wir unsere Studie planten. Da war Sorafenib selbst noch mitten in seiner eigenen Studie.

Herr Prof. Dr. Wörmann (DGHO): Ich kritisiere nicht das Design der Studie. Ich bin überhaupt begeistert, dass Sie bei einem Orphan Drug mit so einer kleinen Entität eine Studie mit 400 Patienten randomisiert durchgeführt haben. Das ist keine Selbstverständlichkeit. Trotzdem: Wenn wir heute eine Nutzenbewertung machen, müssen wir heute Sorafenib berücksichtigen.

Herr Prof. Hecken (Vorsitzender): Dazu ergänzend, Herr Rimmele.

Herr Rimmele: Die therapeutische Situation sieht noch ganz anders aus: Die Medikamente werden nacheinander gegeben. Ich meine, auch in der Lenvatinib-Studie hatten einige zuvor Sorafenib, und es gibt in der Sorafenib-Studie eine ganze Reihe Patienten, die danach, in der therapeutischen Situation, Lenvatinib bekommen haben. Deswegen ist bei diesen beiden Therapeutika eher die Frage: Was gibt man zuerst?

Herr Prof. Hecken (Vorsitzender): Herr Rodewyk.

Herr Dr. Rodewyk: Nach der Diskussion um die Methodik interessieren mich eigentlich mehr die sechs Todesfälle, die Sie hatten. Sie haben vorhin gesagt, in der Phase-II-Studie hatten Sie tolle Ergebnisse. Gab es da auch schon Todesfälle? 88 Prozent Crossover: Sind die Patienten über die Todesfälle informiert worden? Es gibt ja die Diskussion um die Dosisreduktion. Tut sich da irgendetwas bei den Überlegungen? Da hätte ich gerne noch ein paar Informationen.

Herr Prof. Hecken (Vorsitzender): Wer macht das? – Herr Steinert.

Herr Dr. Steinert (Eisai): Ich versuche beide Fragen aufzugreifen, einmal die Dosisreduktion und was in der Produktentwicklung noch gemacht werden möchte, sowie die Todesfälle. Ich möchte aber einen dezidierten Blick auf die Todesfälle werfen. In der Lenvatinib-Gruppe hatten wir 82 Todesfälle; das sind etwa 31,3 Prozent. Wir hatten 52 Todesfälle in der Placebo-Gruppe; das macht knapp über 40 Prozent aus. Fatale Ereignisse, die im Zusammenhang mit einem unerwünschten Ereignis auftraten, waren etwa 20 Fälle in der Lenvatinib-Gruppe und 6 Fälle in der Placebo-Gruppe. Was hier ganz wichtig ist: Man muss die Behandlungsdauer sehen. Das heißt, wenn wir eine Adjustierung auf die Behandlungsdauer vornehmen, haben wir 0,07 Events pro Patientenjahr für die Lenvatinib-Gruppe und 0,1 Event pro Patientenjahr für die Placebo-Gruppe. Das heißt, wir haben den Placebo-Vergleich; wir haben aber auch die 13,6 Monate. Sie haben natürlich völlig recht: Es sind sechs Todesfälle aufgetreten, die wahrscheinlich oder möglicherweise in einen Zusammenhang mit der Studienmedikation gebracht wurden.

Zu der zweiten Frage. Was die EMA auch gesehen hat, ist, dass die Wirksamkeit von Lenvatinib nicht loszukoppeln ist von der initialen Dosierung von 24 mg – das hat man auch heute schon gehört –, vor allem auch bei den PFS-Events. Was man herausbekommen möchte – das wird im Rahmen der 211-Studie untersucht, die ich vorhin schon erwähnt habe, bei der auch die Lebensqualitätsdaten erhoben werden –, ist, ob man bei einer niedrigeren Dosierung vergleichbare Wirksamkeiten mit einem besseren Nebenwirkungsprofil erreichen kann. Diese Studie wird rekrutiert, ist global aufgesetzt, um diese Frage lösen zu können, durch die entsprechende Generierung von validen Daten.

Herr Prof. Hecken (Vorsitzender): Danke schön. – Frau Wenzel-Seifert.

Frau Wenzel-Seifert: Uns interessiert auch die Safety. Es wurde dargestellt, dass viele dieser Nebenwirkungen, die aufgetreten sind, Wirkstoffklassennebenwirkungen sind, vor allen Dingen – so wurde es erläutert – tritt notwendigerweise Hypertonie auf, die ein Zeichen für das Ansprechen der

Therapie sein soll. Ich hätte gerne einen Kommentar von Herrn Professor Wörmann, ob das wirklich so zu sehen ist. Es sind doch sehr dramatische hypertensive Ereignisse mit neurologischer Symptomatik aufgetreten. Da stellt sich die Frage: Sind die Risiken, die aufgetreten sind, vergleichbar mit denen, die man unter Sorafenib sieht?

Die zweite Frage. Sie haben die Studie 211 angesprochen, die Sie rekrutieren. Wann ist mit Ergebnissen zu rechnen?

Herr Prof. Hecken (Vorsitzender): Erster Teil der Frage an Herrn Wörmann vielleicht. – Bitte schön.

Herr Prof. Dr. Wörmann (DGHO): Wir haben es in unserer Stellungnahme sehr deutlich betont: Wir halten die Nebenwirkungen für erheblich. Das ist eine deutliche Belastung. Die Kollegen sagen, dass es fast keinen Patienten gibt, der langfristig mit der initialen Dosis behandelt wird. Das entspricht aber auch der Zulassungsstudie, wo relativ viele reduziert hatten. Aber obwohl in der Zulassungsstudie zwei Drittel der Patienten mit reduzierter Dosis weitergemacht haben, ist trotzdem der insgesamt hohe Remissionseffekt dabei herausgekommen.

Was Sie eben andeuteten: Ob man aus der Rate oder Schwere der Nebenwirkungen auf den Therapieerfolg alleine schließen kann, sozusagen aus der Motivation: „Halte das aus, weil du dann gut ansprichst!“. Mit solchen Ansätzen haben wir Schwierigkeiten.

Herr Prof. Hecken (Vorsitzender): Zweiter Teil, Herr Steinert, bitte.

Herr Dr. Steinert (Eisai): Grundsätzlich ist zu den Nebenwirkungen zu sagen: Wir haben gegen den Placebo-Arm verglichen. Wir hatten allein 60 Prozent unerwünschte Ereignisse in der Placebo-Gruppe. Daraus kann man schließen, dass die Krankheit an sich und der Progress, den die Patienten erfahren, ein relativ zügiger Progress, dazu führen, dass unerwünschte Ereignisse wahrscheinlich oder möglicherweise kategorisiert werden, aber natürlich keine Kausalität zur Studienmedikation darstellen. Wir haben Nebenwirkungen, und die werden engmaschig monitoriert.

Zu der Hypertonie. Wir wissen – ich möchte nicht das Wort „Klasseneffekt“ benutzen –, dass, wenn der VEGF-Signal-Weg blockiert wird, wir Hypertonien auslösen können. Wichtig zu wissen, ist, dass etwa 56 Prozent aller Patienten, die randomisiert wurden, schon hypertensiv waren. Die Hypertonie ist unter Lenvatinib etwa 2,1 Wochen nach Beginn der Medikation oder Therapie aufgetreten. Das heißt, es war ein relativ früher Event. Es wurden auch als Hypertonien, die im Zusammenhang mit der Studienmedikation auftraten, sogar prähypertensive Werte mit aufgenommen. Das heißt, es wurde relativ frühzeitig eine antihypertensive Therapie eingesetzt. Was wir gesehen haben, ist, dass wir im ersten Therapiezyklus die hypertensiven Events hatten, dass man das durch eine standardisierte antihypertensive Therapie unter Kontrolle bekommen hat, und entsprechend mit Dosismodifikationen, wenn ein persistierender Hypertonus vorlag. Wenn wir eine Gesamtbehandlungsdauer von 13,6 Monaten haben und wir in allen Folgezyklen nach Zyklus 1 diese hypertensiven Events nicht mehr hatten, heißt das, dass man mit einer antihypertensiven Therapie das relativ gut in den Griff bekommen kann.

Herr Prof. Hecken (Vorsitzender): Danke schön. – Herr Mehlig, zur Ergänzung.

Herr Mehlig (Eisai): Wenn ich das ganz kurz ergänzen darf. – Ich denke, die Hypertonien nehmen wir alle sehr ernst, Herr Professor Wörmann. Auch die Ausführungen der Fachinformationen sind sehr ausführlich und aus meiner Sicht erfreulich klar, wie man mit der Hypertonie umzugehen hat, weil es ein ganz offensichtliches Nebenwirkungsproblem des Medikaments ist. Wie Herr Dr. Steinert eben

gesagt hat, hatten in der Studie die Mehrzahl der Patienten eine Hypertonie, als sie in die Studie kamen, und wurden deswegen engmaschig kontrolliert und behandelt, einerseits medikamentös und andererseits mit Dosisreduktion, die in der Fachinformation sehr klar und eindeutig vorgeschlagen wird. Dort steht, wie eine Herunterdosierung vorzunehmen ist, mit dem Wissen, dass Hypertonie unter Lenvatinib häufig beobachtet wird.

Herr Prof. Hecken (Vorsitzender): Danke schön. – Ergänzend dazu, Herr Wörmann.

Herr Prof. Dr. Wörmann (DGHO): Ich will nicht den Eindruck erwecken, dass das ein miserables Präparat ist im Vergleich zu Sorafenib. Sorafenib ist mindestens so unfreundlich. Sorafenib ist dadurch gekennzeichnet, dass es extreme Fatigue macht. Es ist ganz schwierig, unter Sorafenib zu arbeiten. Ich will es nicht schönreden, sodass Lenvatinib so viel schlechter wäre. Diesen Eindruck haben die Kollegen überhaupt nicht, die damit behandeln, sondern eher, dass, gerade was die Fatigue-Symptomatik angeht, Lenvatinib deutlich besser toleriert wird als Sorafenib. – Nur dass nicht der falsche Eindruck entsteht.

Herr Prof. Hecken (Vorsitzender): Danke schön. – Frau Krapfenbauer.

Frau Krapfenbauer (Eisai): Ich wollte ganz kurz etwas nachschieben. Es war nach der Verfügbarkeit der Daten aus der Studie 211 gefragt. Wir gehen momentan davon aus, dass wir Ende August 2020 valide Daten haben werden.

Herr Prof. Hecken (Vorsitzender): Herr Steinert, noch einmal.

Herr Dr. Steinert (Eisai): Ich wollte noch hinzufügen, dass bei der Hypertonie Grad 3, Grad 4 es insgesamt nur bei 1,1 Prozent der Patienten zum Abbruch der Medikation kam. Das heißt, dass das mittels antihypertensiver Standardintervention gut managebar und handhabbar war.

Herr Prof. Hecken (Vorsitzender): Danke schön. – Frau Dr. Grell.

Frau Dr. Grell: Herr Schwenke, ich wollte gerne auf das Overall Survival zurückkommen. Die Grundannahme des Modells RPSFT, das Sie gerechnet haben, geht davon aus, dass der Behandlungszeitpunkt keine Rolle spielt. Wenn ich eine Randomisierung habe und Patienten, die ein Fortschreiten haben und das Crossover machen, muss ich davon ausgehen, dass diese Patienten tendenziell eher ein fortgeschritteneres Stadium haben. Ich würde bei soliden Tumoren eher davon ausgehen, dass sie schlechter ansprechen. Insofern stellt sich die Frage, wie gut dieses Modell für solide Tumore generell ist. Die Frage ist nicht, welches Modell bei vielen vielleicht das richtige Modell oder das schönste Ergebnis ist, sondern die Frage ist, wenn man verschiedene Modelle hat, ob man daraus ein stringentes Bild bekommt. Wenn ich das gerechnete Modell „on treatment“ anschau, das auch Probleme hat – alle Modelle haben Probleme; da stimme ich Ihnen völlig zu –, dann kommt es zu einem Nachteil von Lenvatinib. Deswegen verstehe ich auch die Aussage im EPAR so, die letztlich sagt, die Analyse der Adjustierung nach dem Crossover führt nicht zu Bedenken hinsichtlich eines Schadens des Overall Survivals. So steht es im EPAR korrekt. Das heißt, aus der Summe der verschiedenen Analysen, die im Statistic Review der FDA sehr umfangreich durchgeführt wurden, ergibt sich ein uneinheitliches Bild. Für mich ist überhaupt nicht einschätzbar, wie ein Effekt aus diesen unterschiedlichen Analysen bezüglich des Overall Survivals aussehen könnte, wenn man das zusammennimmt.

Herr Prof. Hecken (Vorsitzender): Noch einmal dazu, Herr Mehlig, bitte.

Herr Mehlig (Eisai): Wenn ich kurz darauf erwidern darf, Frau Dr. Grell, und dann gebe ich an meinen Kollegen weiter, wenn es methodisch relevant wird. – Wir haben gesagt, dass die EMA sagt, aus den Daten ergibt sich kein Nachteil. Das ist genau das, wonach die EMA sucht. Die EMA sucht in den Daten, ob es irgendwelche Hinweise gibt, dass das Präparat zu einem Nachteil beim Gesamtüberleben führen kann. Das ist letztendlich das, wonach geschaut wurde. Das heißt nicht, dass das die Daten relativiert, die vorgelegt wurden, sondern es sagt lediglich, dass das, wonach die EMA schauen muss, nämlich dass das Präparat auf gar keinen Fall einen Hinweis zeigt, dass es sogar das Leben verkürzen kann, statt es zu verlängern, aus den Daten auf keinen Fall herauszulesen ist. Auch ich bin über diesen Satz gestolpert und habe ihn mir von den Kollegen aus der Zulassung erklären lassen.

Weil Sie gesagt haben, verschiedene Modelle und schönstes Ergebnis, schauen, was man hat: Das ist genau der Punkt, den ich vorhin versucht habe anzusprechen. Die Entscheidung für das Modell – danach wurde vorhin auch gefragt – ist zu einem frühen Zeitpunkt getroffen worden. Diese Entscheidung musste getroffen werden. Man muss sie im Studienprotokoll fixieren. Ich muss es a priori festlegen, mit der Vermutung, mit der Erfahrung und dem Wissen, das unsere Statistiker hatten, dass dieses Modell sicherlich geeignet ist. Wenn Sie diese Entscheidung einmal treffen, müssen Sie mit dieser Entscheidung leben. Ich würde ungern vor diesem Gremium sitzen und vielleicht nicht 15, aber zwei oder drei oder vier verschiedene Adjustierungsmodelle verwendet haben und die noch möglichst nach Vorliegen der Daten definiert haben, um zu sagen: Schaut, jetzt haben wir ein Modell gefunden, das positive Ergebnisse zeigt. Die Entscheidung für das Modell sehr früh zu treffen, war aus der Erfahrung und dem Wissen um das Präparat heraus und auch aus den Fallzahlschätzungen heraus getroffen worden. Mit dieser Entscheidung muss man den Weg bis in die Zulassungsverfahren gehen und dann natürlich auch die Ergebnisse genau dieses Modells vorlegen und transparent zur Verfügung stellen.

Herr Prof. Hecken (Vorsitzender): Herr Schwenke, noch einmal, aber bitte nur drei Sätze.

Herr Dr. Schwenke (Eisai): Sehr gerne. – Viel kann ich dem nicht beifügen. Letztendlich geht es darum: Wir sehen grafisch, dass die Hazards einigermaßen stabil sind. Ja, auch wir waren einigermaßen überrascht, dass es wirklich so stabil ist. Auf der anderen Seite sehen wir das zweite PFS, also die Zeit vom ersten zum zweiten PFS, ähnlich lange wie in den Patienten, die direkt Lenvatinib bekommen haben. Somit gehen wir davon aus, dass dieser Effekt wirklich kleiner ist als gedacht, so dass es, wenn es eine Überschätzung ist, eine relativ kleine Überschätzung des Gesamtüberlebens im adjustierten Modell ist.

Herr Prof. Hecken (Vorsitzender): Herr van Poppel, bitte.

Herr van Poppel: Ich habe eine Frage in diesem Zusammenhang. Sie haben mehrmals wiederholt, dass die Adjustierung durch das RPSFT-Modell bereits a priori im Studienprotokoll festgelegt worden war. Ich habe gerade einen Artikel von Schlumberger im *New England Journal of Medicine* dazugenommen. Da findet man im „supplemental material“ zwei Versionen eines statistischen Analyseplans. In der älteren Version ist die RPSFTM-Analyse nicht abgebildet. In der zweiten, der neueren Version ist auch die RPSFTM-Analyse in der Tat abgebildet worden. Aber dieses Dokument ist anscheinend erst nach der Entblindung unterschrieben worden. Stimmt das?

Herr Prof. Hecken (Vorsitzender): Herr Mehlig.

Herr Mehlig (Eisai): Ich habe die beiden Dokumente nicht vorliegen. Ich habe auch vom Studienprotokoll und nicht vom statistischen Analyseplan gesprochen. Im Studienprotokoll – da bin ich sehr züversichtlich; das habe ich geprüft – war das RPSFT-Modell bereits vorab definiert.

Herr Prof. Hecken (Vorsitzender): Reicht das, Herr van Poppel?

Herr van Poppel: Also vor Entblindung, meinen Sie?

Herr Mehlig (Eisai): Das Studienprotokoll wird verabschiedet, bevor ich in die Studie hineingehe. Im Studienprotokoll war es bereits definitiv festgelegt, bevor die Studie begonnen hat und alles andere, was danach kam.

Herr Prof. Hecken (Vorsitzender): Okay. – Ich schaue in die Runde: Gibt es weitere Fragen? – Herr Wörmann, bitte.

Herr Prof. Dr. Wörmann (DGHO): Ich bin ein bisschen beunruhigt über das Image, das Lenvatinib hier hat. Es ist schon eine Substanz, auf die über zwei Drittel der Patienten ansprechen. Das kennen wir sonst überhaupt nicht. Wir haben in unserer Stellungnahme gezeigt, mit Sorafenib gibt es wirklich deutlich höhere Remissionsraten. Der Punkt ist, dass die Remissionsrate hier – das ist sicher ein grober Fehler der Studie – nicht durch direkte Morbiditätsparameter unterlegt worden ist. Wir haben beim progressionsfreie Überleben ein Hazard Ratio von 0,21. Das ist auf der Topliste aller onkologischen Präparate. Das Problem, das wir hier haben, ist, dass wir schon sehen, dass es eine progrediente Erkrankung gewesen ist. So ist das definiert. So sind die Patienten eingeschlossen worden. Dabei kommt dieser Vorteil im progressionsfreien Überleben heraus. Der ist nicht substanziell in der Studie hinterlegt. Was wir erleben, ist, dass die Patienten eine bessere Lebensqualität haben. Aber das ist hier nicht hinterlegt. Ich will nur nicht den Eindruck erwecken, dass das eine Mistsubstanz ist. Klinisch erleben wir schon, dass das eine hochwirksame Substanz ist. Aber ich glaube, die Studien, die durchgeführt werden, hätten noch besser gemacht werden können.

Herr Prof. Hecken (Vorsitzender): Danke schön für diese Ergänzung, Herr Professor Wörmann. Allerdings hatte ich aus den Fragen kein unmittelbares Image abgeleitet. Von Mistsubstanz hatte bislang auch niemand geredet. – Das nur der guten Ordnung halber, damit nicht der Eindruck entsteht, als sei der Eindruck entstanden, als wäre das so. Da muss man doch aufpassen. – Herr Mayer, bitte.

Herr Dr. Mayer: Ich möchte auch noch einmal betonen, dass das keiner von uns gesagt hat. Es ging tatsächlich nur um den apodiktischen Glauben an Zahlen. Es ist klar geworden, dass man so etwas nicht machen kann. Meine Frage, die immer noch nicht beantwortet worden ist, wäre, wie man, auch wenn es eine Phase-II-Studie ist, die Studie mit 400 Patienten durchführen kann, ohne Morbiditäts- und Lebensqualitätsdaten zu erheben. Wenn es eine Studie mit 20 versus 20 explorativ in Phase II wäre, würde ich sagen, man braucht es noch nicht so früh. Aber es ist eine große Studie, eine relevante Studie, wie Herr Wörmann gesagt hat, und trotzdem hat man auf patientenrelevante Endpunkte komplett verzichtet. Warum war das so?

Herr Prof. Hecken (Vorsitzender): Wer macht das: Symptomatik, Schluckbeschwerden, alles, was dazugehört, was man unter diese Begrifflichkeit zählen kann? – Herr Mehlig, bitte.

Herr Mehlig (Eisai): Ich versuche den Punkt nach hinten heraus aufzugreifen, den Sie provozierend gemacht haben. Sie haben gesagt, dass keine patientenrelevanten Endpunkte in der Studie waren. Das – das habe ich auch in meiner Einleitung gesagt – sehen wir anders. Ich denke, das kann man

auch zurückweisen, weil das PFS in diesem konkreten Fall sicherlich als patientenrelevant einzuschätzen ist und das Overall Survival zweifellos, auch wenn wir überraschend intensiv das Modell diskutiert haben, ein patientenrelevanter Endpunkt ist.

Die klinische Symptomatik wäre, hätten wir ein Lebensqualitätsinstrument in der Studie mitgeführt, darüber abgebildet gewesen. Diesen Mangel haben wir schon angesprochen. Da die Lebensqualität nicht gemessen wurde, ist dieser Teil auch nicht gemessen worden. Selbstverständlich haben wir die Verträglichkeit beschrieben. Aber Daten, die wir nicht in der Studie erhoben haben, können wir hier leider nicht nachliefern. Die haben wir nicht.

Herr Prof. Hecken (Vorsitzender): Okay. – Herr Steinert, eine Ergänzung.

Herr Dr. Steinert (Eisai): Ich möchte auf einen Punkt hinweisen, den wir in dieser Studie das erste Mal gesehen haben und der vor allem in dieser Indikation nur wenige Patienten betrifft. Wir haben das erste Mal bei Patienten, die metastasiert waren, eine Komplettremission gesehen; sie war anhaltend. Das war vorher nie der Fall gewesen. – Das möchte ich noch einmal zum Ausdruck bringen.

Herr Prof. Hecken (Vorsitzender): Herzlichen Dank. – Ich sehe keine Wortmeldung mehr. Dann möchte ich Ihnen die Gelegenheit geben, kurz zusammenzufassen, Herr Mehlig, was wir hier diskutiert haben. Bitte schön.

Herr Mehlig (Eisai): Vielen Dank, Herr Professor Hecken. Ich glaube, ich möchte nicht versuchen, die Diskussion umfassend zusammenzuführen.

Herr Prof. Hecken (Vorsitzender): Schade eigentlich.

Herr Mehlig (Eisai): Ich habe mich sehr über die äußerst interessierte und erstaunlich methodenkonzentrierte Diskussion gefreut. Ich hoffe, dass wir zumindest einen guten Teil der Fragen von Ihnen angemessen beantworten konnten. Was wir mitnehmen – das war nicht ganz unerwartet –, ist, dass wir innerhalb unseres bestehenden lernenden Systems in der Firma darauf hinweisen werden, dass das Fehlen von Lebensqualitätsdaten nicht nur in Deutschland hinterfragt werden wird. Das ist routinemäßig. Auch wenn es keine krankheitsspezifischen Instrumente gibt, muss der Versuch unternommen werden, die Lebensqualität zu messen, um die Antworten auf Fragen besser geben zu können, die ansonsten im Raum bleiben.

Ich hoffe, es ist uns gelungen, Ihnen einige unserer Positionen näherzubringen, insbesondere auch über die Relevanz des progressionsfreien Überlebens, das der primäre Endpunkt in dieser Studie war und – Herr Professor Wörmann hat es eben gesagt – mit außergewöhnlichen Werten in dieser Studie gezeigt werden konnte.

Beim Overall Survival nehme ich gerne mit, dass es auch nach heute nicht das eine ideale Modell gibt. Wir werden Diskussionen führen und diese Diskussionen auch bestehen können müssen. Wir fühlen uns aber erneut angesichts der vorliegenden Werte und der Gleichgerichtetheit der Effekte, auch der unadjustierten Analysen, wohl, was die Wirksamkeitsparameter angeht. Das Verträglichkeitsprofil haben wir knapp charakterisiert. Wir haben auf die Hypertonie hingewiesen, darauf, dass damit sehr sorgfältig umgegangen werden muss, dass entsprechend engmaschig monitoriert und kontrolliert werden muss, wenn diese und andere Nebenwirkungen auftreten.

Insgesamt sind wir uns mit der EMA in der Meinung einig, dass wir ein vorhersehbares Verträglichkeitsprofil haben, das in der überwiegenden Zahl der Fälle seitens der Onkologen, die diese Therapie durchführen, beherrschbar ist.

Ich bedanke mich noch einmal auch im Namen meiner Kollegen für die wirklich sehr angeregte Diskussion und gebe gerne an Sie zurück, Herr Professor Hecken.

Herr Prof. Hecken (Vorsitzender): Ganz herzlichen Dank. – Wir danken unsererseits für die Diskussion. Ich will nur noch einmal sagen: Dass Sie von Anfang an jetzt in Ermangelung eines validierten Messinstruments, um in dieser Indikation Lebensqualität zu messen, darauf verzichtet haben, so etwas nicht mitgeführt haben, das haben wir, glaube ich, kapiert. Die Frage von Herrn Mayer war auch nicht darauf bezogen, zu sagen: Wieso haben Sie, nachdem Sie vorher gesagt haben, es gibt kein validiertes Messinstrument, nicht trotzdem gemessen? Die Frage ist vielmehr schlicht und ergreifend: Wäre es nicht möglich gewesen, statt allein auf PFS und andere Dinge zu vertrauen, daneben sich zwei, drei klassische Symptomatiken anzuschauen, die man auch in einer anderen Form hätte abbilden können, woraus man möglicherweise ein bisschen Honig hätte saugen können? Dass Lebensqualität wichtig ist, ist klar. Dass es manchmal an Messinstrumenten mangelt, ist auch klar. Aber das war einfach nicht der Punkt – damit wir hier nicht missverstanden auseinandergehen –, das hätte uns in der Diskussion sicherlich an der einen oder anderen Stelle helfen können.

Danke, dass Sie da waren. Die Anhörung ist beendet.

Schluss der Anhörung: 12.37 Uhr