

Mündliche Anhörung



**Gemeinsamer
Bundesausschuss**

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

hier: Wirkstoff Nivolumab (Nierenzellkarzinom)

Sitzung im Hause des Gemeinsamen Bundesausschusses in Berlin
am 5. September 2016
von 13.00 Uhr bis 14.02 Uhr

– Stenografisches Wortprotokoll –

Angemeldete Teilnehmer der Firma **Bristol-Myers Squibb GmbH & Co. KGaA:**

Frau Dr. Beckers
Frau Dr. Herrmann
Frau Dr. Kupas
Herr Dr. May

Angemeldete Teilnehmer der Firma **Eisai GmbH:**

Herr Mehlig
Frau Stern

Angemeldete Teilnehmer der Firma **Novartis Pharma GmbH:**

Frau Barth
Herr Dr. Guderian

Angemeldete Teilnehmer der Firma **Pfizer GmbH:**

Frau Schmitter
Frau Dr. Strunz

Angemeldete Teilnehmer der Firma **Roche Pharma AG:**

Herr Bleibler
Herr Buhck

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

Herr Prof. Dr. Staehler

Angemeldete Teilnehmer für die **Deutsche Gesellschaft für Urologie (DGU):**

Herr Dr. Johannsen
Herr Prof. Dr. Miller

Angemeldete Teilnehmer für die **Arzneimittelkommission der deutschen Ärzteschaft (AkdÄ):**

Herr Dr. Fetscher
Herr Dr. Spehn

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

Herr Dr. Rasch
Herr Nehls

Beginn der Anhörung: 13.00 Uhr

(Die angemeldeten Teilnehmer betreten den Raum)

Herr Prof. Hecken (Vorsitzender): Meine sehr verehrten Damen und Herren, herzlich willkommen – teilweise waren Sie ja schon da – zur weiteren Anhörung im schriftlichen Stellungnahmeverfahren zu Nivolumab, jetzt zur Behandlung des fortgeschrittenen Nierenzellkarzinoms.

Basis der heutigen mündlichen Anhörung ist die IQWiG-Dossierbewertung vom 28. Juli 2016, bei der das IQWiG drei Gruppen gebildet und in einer Gruppe einen erheblichen, in einer Gruppe einen beträchtlichen und in einer Gruppe einen nicht belegten Zusatznutzen gesehen hatte. Wir haben Stellungnahmen zum einen vom pharmazeutischen Unternehmer, von Bristol-Myers Squibb GmbH & Co. KG, dann von der Arzneimittelkommission der deutschen Ärzteschaft, dann von der DGHO, dann von der DGU, dann von der IAG-N, dann von Eisai GmbH, von Ipsen Pharma GmbH, von Novartis Pharma GmbH, von Pfizer Pharma GmbH, von Roche Pharma GmbH und vom Verband der forschenden Arzneimittelhersteller.

Ich muss zunächst wie üblich wieder die Anwesenheit für das Protokoll prüfen: Für Bristol müssten zunächst Herr Dr. May da sein – jawohl –, dann Frau Dr. Beckers – sie ist auch da, ja –, dann Frau Dr. Kupas – sie war schon da –, dann Frau Dr. Herrmann – sie ist auch da, ja. Dann müsste für die Deutsche Gesellschaft für Urologie Herr Professor Dr. Miller da sein – ja, ich sehe ihn –, dann ebenfalls für die Deutsche Gesellschaft für Urologie Herr Dr. Johannsen – er ist auch da, ja. Für die AkdÄ müssten wieder die Herren Herr Dr. Spehn und Dr. Fetscher da sein – jawohl –, dann für die Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie Herr Professor Dr. Staehler – ja –; Herr Wörmann ist entschuldigt. Herr Mehlig kommt verspätet. Dann ist für Eisai Frau Stern da – ja –, dann für Roche Herr Bleibler – ja –, dann für Roche Herr Buhck – ja –, dann für Pfizer Frau Dr. Strunz – ja –, dann für Pfizer Frau Schmitter – ja –, dann für Novartis Herr Dr. Guderian – ja –, dann für Novartis Frau Barth – ja – und dann für den vfa Herr Dr. Rasch und Herr Nehls. Frage: Wurde jemand nicht aufgerufen? – Das ist nicht der Fall. Dann können wir die Anwesenheit so zu Protokoll nehmen.

Bevor ich dem pharmazeutischen Unternehmer das Wort zu einer Einleitung gebe, kommt der übliche Hinweis auf das Wortprotokoll: Benutzen Sie deshalb bitte jeweils das Mikrofon und nennen Sie Namen und entsendendes Unternehmen oder Fachgesellschaft.

Aus meiner Sicht müssten heute in der Anhörung neben anderen Fragestellungen ganz besonders zwei Punkte diskutiert werden, nämlich zum einen die Frage, inwieweit die Unterteilung in Subgruppen auf Basis der MSKCC-Prognose-Scores gerechtfertigt ist, in deren Folge unterschiedliche Aussagen zum Zusatznutzen abgeleitet werden. Darauf wird der pharmazeutische Unternehmer sicherlich eingehen. In diesem Kontext geht es um die Frage: Welcher der herangezogenen Prognose-Scores ist aus Ihrer Sicht besser geeignet, der MSKCC-Score oder die Heng-Kriterien?

Zum anderen müsste auch darüber gesprochen werden, wie die Gruppe der Patienten, die mit Tamsirolimus vorbehandelt wurden, hinsichtlich ihrer Bedeutung in der Versorgungsrealität und der in Betracht kommenden Therapieoptionen zu bewerten ist. Das sind aber nur zwei Fragestellungen, die ich so einfach einmal in den Raum stellen will.

Als Erstes besteht jetzt für den pharmazeutischen Unternehmer die Möglichkeit, kurz einzuführen. Wer möchte das übernehmen? – Herr May.

Herr Dr. May (BMS): Ich werde das übernehmen.

Herr Prof. Hecken (Vorsitzender): Bitte schön, Herr May, Sie haben das Wort.

Herr Dr. May (BMS): Vielen Dank. – Sehr geehrter Herr Professor Hecken, sehr geehrte Damen und Herren! Vielen Dank, dass wir die Gelegenheit haben, heute hier noch einmal mit Ihnen in den Dialog einzutreten und einige Punkte aus der Stellungnahme zu besprechen.

Das IQWiG hat in seiner Dossierbewertung einen Hinweis auf einen erheblichen bzw. einen beträchtlichen Zusatznutzen von Nivolumab gegenüber Everolimus für einen Großteil der Patienten mit fortgeschrittenem oder metastasiertem Nierenzellkarzinom nach Vortherapie festgestellt. Ich werde mich im Folgenden auf einige wenige Aspekte konzentrieren und nicht durch die komplette Stellungnahme gehen.

Wir würden gern mit Ihnen darüber sprechen, warum wir die Übertragbarkeit des Nutzens für das Anwendungsgebiet Patienten mit fortgeschrittenem Nierenzellkarzinom nach Vortherapie mit Temsirolimus im Wirkmechanismus von Nivolumab begründet sehen, warum wir der Ansicht sind, dass das Merkmal MSKCC-Score als Effektmodifikator für einen Zusatznutzen von Nivolumab als Zufallsbefund zu sehen ist, sobald die Heng-Kriterien hinzugezogen werden, wie wir die gewünschte Analysenzeit bis zur Verschlechterung mit dem MID von 10 Millimeter EQ-5D darstellen und warum wir die Lebensqualität als bewertbar dargestellt sehen.

Bevor wir in die Diskussion einsteigen, möchte ich meine Kolleginnen von BMS vorstellen, die mich heute begleiten. Frau Dr. Beckers zu meiner Linken ist verantwortlich für den Bereich Medizin bei den Urogenitaltumoren, Frau Dr. Kupas ist verantwortlich für die Biostatistik, und Frau Dr. Herrmann zu meiner Rechten aus dem Bereich Market Access ist zuständig für das Dossier. Ich selbst leite den Bereich Medizin in Deutschland.

Für Nivolumab ist das therapeutische Gebiet Nierenzellkarzinom bereits die vierte Indikation, für die es die frühe Nutzenbewertung durchläuft und damit die zweite Indikation, die heute in der Anhörung ist. Das Ergebnis der Nutzenbewertung durch das IQWiG, Herr Professor Hecken, haben Sie ja freundlicherweise schon kurz zusammengefasst. Deswegen möchte ich mich auch kurz fassen und noch einmal betonen, dass die zentrale Nivolumab-Zulassungsstudie nach einer prospektiv definierten Interimsanalyse vorzeitig beendet wurde, da sich beim Gesamtüberleben ein deutlicher Vorteil von Nivolumab gegenüber der zweckmäßigen Vergleichstherapie Everolimus abzeichnete. Diese Vorteile waren auch bei der Morbidität und im Nebenwirkungsprofil vorhanden.

Charakteristisch für das Nierenzellkarzinom ist, dass es oftmals sehr lange nicht diagnostiziert wird und erst bei größerem, fortgeschrittenem Wachstum oder dann entdeckt wird, wenn sich bereits Metastasen an anderen Organen gebildet haben. Sie können dann oftmals nicht mehr operativ entfernt werden. Für Patienten in dieser Phase werden Wirkstoffe wie Nivolumab besonders wichtig, die unabhängig von der Vortherapie ansetzen können, ein eher niedriges Nebenwirkungsspektrum haben und somit die Lebensqualität erhalten oder sogar verbessern können.

Das bringt mich zum ersten Punkt, den ich adressieren möchte – Sie hatten es schon angedeutet –, zur Übertragbarkeit des Nutzens für das Anwendungsgebiet Patienten mit fortgeschrittenem Nierenzellkarzinom nach Vortherapie mit Temsirolimus. Begründet durch den gängigen Therapiealgorithmus, konnten keine Patienten mit einer Temsirolimus-Vorbehandlung in die Studie eingeschlossen werden. Nivolumab hat zudem ein komplett anderes Wirkprinzip als mTOR-Inhibitoren oder antiangiogenetisch wirkende Medikamente. Es aktiviert Abwehrzellen des Immunsystems, um den Tumor zu bekämpfen. Hierin liegt auch begründet, dass Nivolumab unabhängig von den zugelassenen Vortherapien beim Nierenzellkarzinom wirkt. Es gibt daher aus unserer Sicht keinen Grund, warum die Ergebnisse nicht auf eine andere Vortherapie übertragen werden können. Diese Meinung vertritt im Übrigen auch die EMA, die von einer ähnlichen Wirksamkeit von Nivolumab bei Patienten mit antiangiogenetischen Substanzen und mit Temsirolimus in der Vortherapie ausgeht. BMS teilt diese Sicht und sieht aus diesem Grund einen Anhaltspunkt für einen nicht quantifizierbaren Zusatznutzen von Nivolumab.

Ein weiterer Punkt, zu dem wir heute in die Diskussion treten wollen, ist das Bilden von Subgruppen nach MSKCC-Score. Das IQWiG hat, begründet auf einen signifikanten Interaktions-p-Wert, eine getrennte Betrachtung der Subgruppen nach dem MSKCC-Prognose-Score vorgenommen. Diese Interaktion sieht man jedoch nicht konsistent über verschiedene Subgruppen bei anderen Scores, die ebenfalls die Prognose der Patienten abbilden. Insbesondere sieht man beim Heng-Score keinen Hinweis auf Interaktion. Der Heng-Score beinhaltet zusätzlich zu den Komponenten des MSKCC-Scores noch weitere Merkmale und ist somit aus unserer Sicht ein für die Prognose präziserer Score. BMS ist daher der Ansicht, dass das Merkmal des MSKCC-Scores als Effektmodifikator für das Gesamtüberleben als Zufallsbefund zu sehen ist und der erhebliche Zusatznutzen für die Gesamtpopulation gilt. Die Details dazu liegen Ihnen in der Stellungnahme vor. Aus diesem Grund vertritt BMS die Auffassung, der Zusatznutzen ist für die Gesamtpopulation und nicht für einzelne Subgruppen zu bewerten.

Zu meinem dritten Punkt: Die Rücklaufquoten der Fragebögen zur Morbidität und Lebensqualität sind im Dossier korrekt dargestellt, und die Analysen sind vollumfänglich zu bewerten. Zudem regte das IQWiG an, dass für den EQ-5D VAS auch eine Sensitivitätsanalyse mit einer MID von 10 Millimetern zusätzlich dargestellt werden sollte. Dieser Anregung sind wir nachgekommen und können konsistent einen Zusatznutzen von Nivolumab gegenüber Everolimus zeigen. In die Analyse der Zeit bis zur Verschlechterung mit einer MID von 10 Millimetern ergab sich fast eine Verdreifachung der Medianzeit bis zur Verschlechterung und ein absoluter Vorteil von 7,8 Monaten zugunsten von Nivolumab. Damit wird der erhebliche Zusatznutzen in der Gesamtschau eindeutig bekräftigt.

Abschließend möchte ich noch einmal kurz auf die Messung der Lebensqualität eingehen. BMS hat in seiner Zulassungsstudie Daten zur Lebensqualität erhoben, einmal mit einem krankheitsspezifischen Instrument, dem FKSI-DRS, des Weiteren in Form eines multiattributiven generischen Instrumentes zur Messung der gesundheitsbezogenen Lebensqualität, dem EQ-5D. Für den EQ-5D wurden Werte mit VAS sowie der Index dargestellt. Die Analysen der Werte, die mit beiden Instrumenten erhoben wurden, hat das IQWiG in dem Endpunktkomplex Morbidität und Lebensqualität dargestellt. Diese grundsätzliche Wertschätzung der Daten unterstreicht, dass das Unternehmen valide Daten zum Endpunktkomplex Morbidität und Lebensqualität dargestellt hat. Ergebnisse der Analysen der Daten des EQ-5D-Index stimmen stark mit denen des EQ-5D-VAS überein. Wir sehen darin eine weitere Bestätigung des EQ-5D-Index-Wertes als valides Instrument zur Erfassung der Lebensqualität.

Ich habe mich hier auf die wesentlichen Aspekte der Stellungnahme konzentriert, möchte jedoch nicht unerwähnt lassen, dass BMS in sehr vielen Punkten mit der Stellungnahme des IQWiG übereinstimmt. In der Gesamtschau sehen wir jedoch einen erheblichen Zusatznutzen für Nivolumab für alle Patienten, da wir in allen vier Nutzenkategorien überzeugen konnten. Insbesondere das extrem positive Nebenwirkungsprofil im Vergleich zu Everolimus rechtfertigt das Ausmaß „erheblich für alle Patienten unabhängig vom MSKCC-Score“. – Vielen Dank. Gerne treten wir mit Ihnen in den Dialog ein.

Herr Prof. Hecken (Vorsitzender): Ja, ganz herzlichen Dank für diese Einführung, Herr May. – Herr Dr. Eyding hat sich gemeldet. Ich würde gern zunächst nur eine Frage an die DGHO und an die DGU stellen.

Die DGHO sagt in ihrer Stellungnahme, dass die Subgruppe der Patienten nach Vorbehandlung mit Temsirolimus als eigene Untergruppe wenig sinnvoll sei, während ich aus der Stellungnahme der DGU herauslese, dass auch, wenn diese Vortherapie nicht untersucht worden ist, für Patienten nach der Vortherapie in entsprechender Art und Weise doch eine eigenständige Subgruppe durchaus zielführend sein könne. Das ergibt sich mittelbar auch aus der Stellungnahme der AkdÄ, und deshalb sollten wir, bevor wir uns mit Effektmodifikationen beim Gesamtüberleben auf der Basis von MSKCC-Prognose-Scores beschäftigen, zunächst einmal mit der Frage befassen, wie es mit der Gruppe der in der hier beschriebenen Weise vorbehandelten Patienten aussieht und ob dies ein klinisch relevan-

ter Parameter ist oder nicht, damit wir das schon einmal ein Stück weit andiskutiert und abgearbeitet hätten, bevor wir dann eben, wie gesagt, in die filigraneren Diskussionen gehen.

Ich weiß nicht, wer beginnen möchte, vielleicht die DGHO, Herr Professor Staehler, und dann die Urologen. – Bitte schön.

Herr Prof. Dr. Staehler (DGHO): Die Subgruppe der Hochrisikopatienten nach dem MSKCC-Score findet sich in der klinischen Realität nicht wirklich ausgeprägt wieder. So kann man sagen, dass nur ungefähr 4 Prozent aller Patienten, die eine metastasierte Situation haben, in diese Gruppe hineinspielen und behandelt werden. Temsirolimus hat in dieser Gruppe im klinischen Alltag so gut wie keinen Stellenwert mehr. Das heißt, die Substanz wird eigentlich für diese Patienten nicht mehr eingesetzt, insbesondere auch deshalb, weil es keine klaren Empfehlungen für eine weitere Therapie gibt.

Man muss unterscheiden, ob diese Patienten therapiefähig sind – dann werden sie einer Therapie in der Regel auch mit einem Tyrosinkinase-Inhibitor zugeführt – oder nicht; dann erfolgt gar keine Behandlung. Das heißt, Patienten nach einer Temsirolimus-Vorbehandlung finden sich in der Realität so gut wie überhaupt nicht. Wir betreuen in München zum Beispiel etwa 15 Prozent aller metastasierten Patienten der Bundesrepublik, und wir haben in den letzten sechs Jahren keinen einzigen Patienten mit Temsirolimus behandelt. Das unterstreicht die klinisch relevante Situation, dass eben diese Patienten so nicht zu testen sind, geschweige denn in einer Studie oder im klinischen Alltag einzuschließen sind.

Herr Prof. Hecken (Vorsitzender): Danke schön. – Für die Urologen Herr Miller oder Herr Johannsen. Wer möchte?

Herr Prof. Dr. Miller (DGU): Wir würden ergänzend zu dem, was Herr Staehler gerade geäußert hat, sagen, dass vor allem in der Second-Line – unabhängig davon, dass Temsirolimus wenig eingesetzt wird, was wir bestätigen würden – die ursprüngliche Klassifikation nach dem MSKCC-Risiko-Score keine Rolle mehr spielt, weil wir auch wissen, dass sie für die weitere Behandlung einfach irrelevant ist, sodass man in der Second-Line die Entscheidung auf Basis anderer Kriterien trifft, beispielsweise anhand dessen, ob der Patient für die nächste Therapie von seinem Allgemeinzustand her noch behandlungsfähig ist, und solcher Dinge mehr. Das ist aus meiner Sicht kein klinisch relevantes Entscheidungskriterium für die Situation, in der wir Nivolumab einsetzen.

Herr Prof. Hecken (Vorsitzender): Danke schön. – AkdÄ, möchten Sie dazu etwas sagen, oder sollen wir es auf die nächste Fragerunde verschieben?

Herr Dr. Spehn (AkdÄ): Dass Patienten, die Temsirolimus bekommen, nicht als Kontrollgruppe infrage kommen, ergibt sich schon daraus, dass mTOR-Inhibitor in der Regel nicht zweimal hintereinander gegeben wird. Temsirolimus wird wenig eingesetzt. Es war aber vor der Entwicklung von Cabozantinib und von Nivolumab meines Wissens das einzige Medikament, das einen deutlichen Überlebensvorteil gezeigt hatte, und dies eben in der Gruppe der Patienten mit einem schlechten Motzer-Score. Daher kam ja auch der Ansatz, dass der pU den Motzer-Score einbezogen hat, hier zwar nicht als sekundären Endpunkt, aber als präspezifizierte Subgruppenanalyse.

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

Frau Dr. Beckers (BMS): Wir haben sie nach den entsprechenden Risiko-Scores stratifiziert, um eine gleichmäßige, balancierte Verteilung der verschiedenen Risikogruppen in unserer Studie in beiden Behandlungsarmen zu gewährleisten. Die Studie war nicht darauf gepowert, einen Unterschied in der Wirksamkeit von Nivolumab in den unterschiedlichen Risikogruppen einzuschätzen. Es gibt keinerlei Belege, keinerlei valide Leitlinienempfehlung oder Studien, was für Patienten nach Temsirolimus-Vorbehandlung dann in der Zweitlinie gegeben werden sollte.

Herr Prof. Hecken (Vorsitzender): Danke schön. – Herr Dr. Eyding, bitte.

Herr Dr. Eyding: Ich verlasse das Thema jetzt. Ich habe eine Frage zur Übertragbarkeit der Studie. Mir ist aufgefallen, dass das Ersterkrankungsalter nach dem, was in den Leitlinien berichtet worden ist, mit 68 oder 72 Jahren bei Männern und Frauen deutlich höher ist als das, was in der Studie betrachtet worden ist; da beträgt es wohl nur 61 und 62 Jahre, und die über 75-Jährigen machen nur 10 Prozent der Studienpopulation aus. Gleichwohl sieht man eine starke Interaktion auch für das Alter, gerade bei dem Cut von 75 Jahren, wobei man bei den über 75-Jährigen eben keinen Effekt mehr sieht bzw. sogar eine numerische Effekturnkehr beim Gesamtüberleben und auch mehr UEs.

Wie können Sie mich sozusagen beruhigen, dass diese Studie dennoch für die Behandlungssituation repräsentativ ist, wenn das Ersterkrankungsalter eigentlich deutlich höher ist und wir in dieser Altersgruppe in der Studie eigentlich gar keinen Effekt sehen, sondern eher einen ganz schlechten Effekt? Man müsste mir einfach noch einmal erklären, wie es dazu kommt, dass so viel jüngere Patienten in dieser Studie sind, als man von dem Ersterkrankungsaltersschnitt in Deutschland eigentlich erwarten würde.

Herr Prof. Hecken (Vorsitzender): Frau Dr. Beckers, bitte.

Frau Dr. Beckers (BMS): Die Frage der Übertragbarkeit – real world, Studie und vice versa – ist ja ein generelles Problem. Wir müssen einfach die Studie von den Einschlusskriterien und von den Ausschlusskriterien einheitlich gestalten.

Zu Ihrer Frage, wie ich Sie beruhigen kann, dass die Studiendaten übertragbar sind: Natürlich hat sich auch die EMA und haben sich auch die Zulassungsbehörden in ihrem Assessment mit der Frage der Übertragbarkeit und auch mit der Altersgruppe genau befasst, und sie sehen eindeutig, dass die Studie generell übertragbar ist.

Herr Prof. Hecken (Vorsitzender): Nachfrage Herr Eyding.

Herr Dr. Eyding: Was sind denn die Argumente, weswegen es zu dieser Abweichung kommt? Dass die EMA das feststellt, mag ja so sein. Aber was sind die Argumente dafür?

Herr Prof. Hecken (Vorsitzender): Frau Beckers.

Frau Dr. Beckers (BMS): In unserer Studie war das Einschlusskriterium für das Alter mindestens 18 Jahre. Es sind Patienten im Alter von 18 bis 88 Jahren eingeschlossen worden, basierend eben auf diesen Einschlusskriterien, die gefordert wurden – mindestens eine antiangiogenetische Vortherapie –, und am Ende des Tages ist dieses Ergebnis dabei herausgekommen. Über 75 Jahre – da stimme ich Ihnen zu – sind es 9 Prozent der Patienten, die eingeschlossen wurden, was für eine klinische Studie bereits sehr viel ist.

Herr Prof. Hecken (Vorsitzender): Herr Eyding, Nachfrage? – Herr Rasch, bitte.

Herr Dr. Rasch (vfa): Indirekt dazu: In der Nutzenbewertung des IQWiG finde ich keinerlei signifikante Effektmodifikation für das Alter, zumindest für das Stratum, was getestet wurde, über 65, unter 65. Insofern bin ich über die Frage von Herrn Eyding etwas überrascht, nach welcher Effektmodifikation Sie da geschaut haben und an welchen Endpunkten.

Herr Prof. Hecken (Vorsitzender): Herr Eyding, bitte; lindern Sie die Überraschung von Herrn Rasch.

Herr Dr. Eyding: Ich versuche es, ja. – Ich habe bei dem Endpunkt Overall Survival bei dem Cut von 75 Jahren einen Interaktions-p-Wert von 0,12, und er ist auch entsprechend im Dossier so abgebildet. Das wäre auch noch eine Frage gewesen. Ein Teil der Beruhigung könnte darin bestehen, dass man die Effekte für die drei Gruppen hätte, nämlich für kleiner 65, für 65 bis 75 und für größer 75. Da ergibt sich ja auch ein Interaktions-p-Wert von 0,22; das heißt, das verfehlt den Hinweis nur knapp. Wenn es da jetzt gar keinen Trend gäbe oder das bei den 65- bis 75-Jährigen genauso gut aussähe wie bei den unter 65-Jährigen, dann wäre das möglicherweise eine gewisse Beruhigung. Vielleicht haben Sie es im Kopf, wie da die Effekte aussehen. Das wäre ein Punkt, der da ein bisschen lindern könnte. Aber die Interaktion kommt aus diesen 0,12 für die 75-Jährigen.

Herr Prof. Hecken (Vorsitzender): Herr Rasch, dann antworten Sie. Danach würde ich aber gerne dem pharmazeutischen Unternehmer das Wort geben.

Herr Dr. Rasch (vfa): Nur eine Replik dazu: Wenn man sich genügend Strata anschaut, wird man schon irgendwo einen gleichen Hinweis finden. Ein starker Beleg für eine Interaktion sieht für mich anders aus.

Herr Prof. Hecken (Vorsitzender): Ja, gut, okay. – Frau Dr. Kupas.

Frau Dr. Kupas (BMS): Vielleicht noch kurz dies: Wir haben uns vier verschiedene Kategorien des Alters angeschaut, und wir sehen nur bei dieser einen einzigen einen Hinweis auf Interaktion; daher ist die Wahrscheinlichkeit, dass es ein falsch positiver Effekt ist, relativ groß. Gerade wenn man sich den Trennpunkt 65 Jahre anschaut, ist der Interaktions-p-Wert von 0,8455 sehr groß. Von daher gehen wir nicht davon aus, dass da wirklich eine Interaktion stattgefunden hat. Die Subgruppe der 75-Jährigen ist sehr klein. Von daher kann es einfach ein Zufallsbefund sein.

Herr Prof. Hecken (Vorsitzender): Danke schön. – Herr Dr. Kaiser, bitte.

Herr Dr. Kaiser: Ich komme noch einmal auf den MSKCC-Score zurück. Es war ja nicht das IQWiG, das die Subgruppe definiert hat, wie Sie, Herr May, am Anfang geschrieben haben, sondern es ist die Subgruppe, die BMS in seiner Studienplanung, in seiner Auswertung und im Studienbericht definiert hat. Sie als BMS haben also diesem Merkmal eine größere Bedeutung beigemessen. Sie haben das insbesondere eben auch als Stratifizierungsmerkmal genommen, was ja auch eine Gleichverteilung sicherstellen soll. Das heißt, Sie haben dem auch eine inhaltliche Bedeutung gegeben. Diese Stratifizierungsmerkmale sind bei Subgruppenanalysen natürlich dann von besonderer Bedeutung, weil sie eine größere Sicherheit geben; das kann man in der allgemeinen Subgruppenliteratur nachlesen.

Frau Beckers, Sie haben eben gesagt, die Studie sei nicht auf Unterschiede gepowert gewesen. Sie haben hierbei vergessen, dass Sie hier eine starke Interaktion haben, nämlich einen p-Wert unter 0,05 für eine von Ihnen vorab definierte Subgruppe. Es geht also nicht darum, alleine zu sagen, es gibt hier keinen statistisch signifikanten Unterschied in einer Subgruppe, sondern Sie haben auf Basis einer von Ihnen vorab definierten Subgruppenanalyse einen Nachweis für eine Modifikation, und dann ist eine Gesamtaussage eben nicht statthaft. Das ist eine ganz übliche Sache. Wenn Sie da eine andere Power hätten haben wollen, hätten Sie die Studie anders durchführen müssen.

Wenn Sie sagen, die Heng-Kriterien wären jetzt sinnvoller, Sie gleichzeitig aber sagen, Sie haben den MSKCC als Stratifizierungsmerkmal genommen, um inhaltlich eine Gleichverteilung sicherzustellen, was bedeutet das denn dann für die Relevanz Ihrer Studie, wenn Sie jetzt ein falsches Stratifizierungsmerkmal gewählt haben? Ist dann die Studie nicht an sich weniger aussagekräftig? Ich würde so weit nicht gehen; aber irgendwie beißen sich Ihre Argumente immer ein bisschen in den Schwanz: Auf der einen Seite sagen Sie, das ist ein wichtiges Merkmal, und wir wollten aber sicherstellen – –, auf der anderen Seite, wenn Sie dann ein Subgruppenergebnis sehen, das Sie selbst vorab geplant

haben, wollen Sie es dann nicht wahrhaben? Das passt irgendwie nicht zusammen. – Vielleicht können Sie dazu noch einmal Stellung nehmen.

Herr Prof. Hecken (Vorsitzender): Wer möchte beginnen? – Frau Dr. Kupas. – Ich halte nur für das Protokoll fest, dass Herr Mehlig um 13.12 Uhr eingetroffen ist. – Frau Kupas, bitte.

Frau Dr. Kupas (BMS): Wir haben beide Merkmale in der Studie untersucht, und wir haben uns für eines für die Stratifizierung entschieden – Frau Beckers kann vielleicht noch etwas dazu sagen, warum es der MSKCC-Score war –, weil wir einfach einen Prognose-Score brauchen, um die Patienten entsprechend zu balancieren und dann auch auswerten zu können, weil die Ereignisse später eintreten. Der Heng-Score und der MSKCC-Score sind nicht unabhängig, und wir sehen bei den Heng-Kriterien überhaupt gar keine Interaktion. Von daher gehen wir davon aus, dass es ein Zufallseffekt ist, dass hier diese Interaktion aufgetreten ist. Natürlich war es Stratifizierungsmerkmal; trotzdem sind auch bei den Heng-Kriterien die Gruppen relativ balanciert, und da sehen wir wirklich gar nichts – der p-Wert ist relativ groß, größer 0,5 –, und von daher denken wir, es war an dieser Stelle ein falsch positives Ergebnis des Interaktionstests.

Herr Prof. Hecken (Vorsitzender): Ergänzung Frau Beckers, dann Herr Fetscher.

Frau Dr. Beckers (BMS): In Ergänzung zu meiner Kollegin Frau Kupas noch einmal zur Klärung, warum wir uns für zwei Risiko-Scores entschieden haben: Der MSKCC-Score ist weltweit der verbreitetere, er ist älter als der Heng-Score, und er war zu Beginn der Studie eben bereits etabliert. Der Heng-Score wurde später entwickelt, auf der Basis der damals zugelassenen Zweitliniensubstanzen. Er ist validiert worden. Die Studie begann 2012; 2015 war die Validierung des Heng-Scores abgeschlossen. Es liegen Daten für die Validität des Heng-Scores vor, basierend auf allen aktuell zugelassenen Substanzen – das liegt für den Motzer-Score nicht vor –, und der Heng-Score ist auch derjenige, der zum Beispiel von den Leitlinien, gerade der nationalen S3-Leitlinie, für die Risikostratifizierung empfohlen wird.

Herr Prof. Hecken (Vorsitzender): Herr Dr. Fetscher, bitte.

Herr Dr. Fetscher (AkdÄ): Wenn man sich die beiden Scores ansieht, sind sie ja zu 90 Prozent klinisch deckungsgleich: Karnofsky, Zeit seit Diagnose, Hämoglobinwert, Kalzium; das ist ja alles völlig gleich. Der einzige Unterschied ist: Thrombozyten und Neutrophile statt LDH. Als 30 Jahre klinisch tätiger Onkologe kann ich Ihnen sagen: Die beiden Scores haben sicherlich durch verschiedene Zeitpunkte und verschiedene Populationen und Publikationen für Sie ein stark unterschiedliches Gewicht; für Kliniker liegen diese beiden Scores doch sehr nahe beieinander, und deswegen können Sie, glaube ich, auch hier nicht so überzeugend vertreten, dass der eine Score für Sie gut ist und der andere Score für Sie schlecht ist.

Sie sind beide brauchbar. Es gibt eine historische Entwicklung zwischen beiden Scores. Der eine spiegelt ein bisschen angiologische Parameter; die Thrombozyten und die Neutrophilen sind im Motzer-Score nicht erfasst. Aber ich kann den Riesenunterschied, den Sie zwischen diesen beiden Scores machen, als Kliniker nicht nachvollziehen.

Herr Prof. Hecken (Vorsitzender): Frau Beckers, Frau Kupas, Herr Rasch.

Frau Dr. Beckers (BMS): Wir sehen das auch so. Wir sehen keine Unterschiede im Nutzen von Nivolumab je nach Auswertung von Heng-Score oder Motzer-Score. Wir möchten nur betonen, dass der Heng-Score eben für die Patienten validiert ist, die mit einer antiangiogenetischen Therapie vorab behandelt wurden, und das ist die Situation, die in dieser Studie abgebildet wird.

Herr Prof. Hecken (Vorsitzender): Danke. – Frau Kupas.

Frau Dr. Kupas (BMS): Es geht gar nicht darum zu sagen, welcher Score besser oder schlechter ist, sondern wir sehen einfach bei den Heng-Kriterien überhaupt gar keine Interaktion und beim MSKCC eine Interaktion, und wir haben auch andere Faktoren, die die Prognose widerspiegeln, bei denen wir auch keine Interaktion sehen. Deswegen vermuten wir, dass der Interaktionstest an dieser Stelle einfach ein Zufallsbefund war. Es geht gar nicht darum zu sagen, welcher Score der richtigere ist.

Herr Prof. Hecken (Vorsitzender): Herr Rasch.

Herr Dr. Rasch (vfa): Wenn man sich das Ergebnis dieser Subgruppenanalyse anschaut, so ist es, aus der IQWiG-Methodik folgend, ein Beleg; der p-Wert für den Interaktionstest liegt bei 0,48. Wäre dieser p-Wert jetzt bei 0,051 gewesen, wäre das ein Hinweis für eine Effektmodifikation, und dann hätte es für die Subgruppe mit der günstigeren intermediären Prognose da auch einen Anhaltspunkt für nicht ... (akustisch unverständlich) gegeben. Man muss sich auch bei dem Ergebnis anschauen: Es gibt keinerlei Effekturnkehr. Das obere Konfidenzintervall liegt bei 1,02. Das Ergebnis ist also knapp signifikant für diese getrennt betrachtete Subgruppe. Was die IQWiG-Methodik eben in diesem Fall gar nicht berücksichtigt, ist der Power-Verlust für diese Subgruppe. Das sollte man sich vielleicht bei der Betrachtung dieses Ergebnisses noch einmal vor Augen führen. – Danke.

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

Herr Dr. Kaiser: Dazu nur eine Anmerkung; danach habe ich aber eine andere Frage zu einem anderen Komplex. Ich bin gespannt, ob der vfa in den nächsten knappen Situationen, die anders herum sind, auch mit knapper Signifikanz und knappem Konfidenzintervall ober- oder unterhalb argumentiert. Es ist mir neu, dass Sie eine solche methodische Richtung einschlagen, in der Sie plötzlich mit „knapp signifikant“ oder „nicht signifikant“ argumentieren. Ich glaube, da müssen Sie schon konsistent bleiben.

Meine konkrete Frage bezieht sich aber jetzt auf ein anderes Thema, und zwar auf die Rücklaufquoten. Herr May, Sie haben eben beschrieben, dass die Angaben im Dossier korrekt seien. Das nehmen wir jetzt erst einmal so hin. Unser Punkt war ja, dass Ihre Angaben im Dossier inkonsistent zu Ihrem eigenen Studienbericht waren; das war ja unser Problem. Sie haben jetzt eine Erklärung dafür geliefert, und zwar ist das wohl offenbar, so sage ich einmal, durch eine technische Zuordnung geschehen, nämlich dadurch, dass Fragebögen – das sage ich jetzt, damit alle mitgenommen werden –, die zur Base-Line, also zum ersten Erhebungszeitpunkt, erhoben worden sind, aber nachdem das Medikament gegeben worden ist, fälschlicherweise nicht diesem ersten Zeitpunkt zugeordnet worden sind, sondern dem zweiten Zeitpunkt nach vier Wochen. Habe ich das richtig verstanden?

Herr Prof. Hecken (Vorsitzender): Wer macht das? – Frau Kupas, bitte.

Frau Dr. Kupas (BMS): Es ist so, wir haben da eine Diskrepanz zwischen den Rohdaten und den dann daraus entstehenden Analysedaten in ADAM-Datensätzen. In den Rohdaten ist es richtigerweise zu Base-Line zugeordnet worden, und dann ist einfach ein ganz formaler Prozess gelaufen, dass dieser Zeitstempel angeschaut wurde, und in den ADAM-Daten wurde es dann eine Woche später gesetzt, weil eben zum Zeitpunkt der Medikation oder kurz danach der Zeitstempel gesetzt wurde. De facto ist es so: Wir haben da so ein elektronisches System, und da kann das passieren: Wenn der Arzt zu spät auf „Okay“ drückt, dann ist der Zeitstempel nach der Dosis. Aber es ist trotzdem zur Base-Line erhoben worden. Wir haben uns dann an den Rohdaten orientiert, und deswegen haben wir eben mehr Patienten unter Base-Line.

Es gibt noch einen zweiten Grund für die Abweichung: Die Tabellen im CSR zeigen nur die Patienten pro Zeitpunkt, und wir zeigen immer Zeitpunkt plus mindestens einen Nachfolgezeitpunkt, weil wir

Change von Base-Line auswerten. Die Base-Line-Zeile hat also Base-Line und mindestens irgendeinen Nachfolgezeitpunkt.

Herr Prof. Hecken (Vorsitzender): Herr Kaiser.

Herr Dr. Kaiser: Wenn ich darauf eingehen darf: Diese Inkonsistenz war für uns auch nicht das Problem, weil die Frage, welche Grundgesamtheit von Patienten in diesen Auswertungen zu den Rücklaufquoten enthalten ist, genau durch die Methodik erklärt war, aber eben nicht dieser Unterschied zwischen Woche null und Woche vier. Das heißt einfach, das, was im Studienbericht steht, ist durch eine technisch fehlerhafte Zuordnung so nicht richtig, aber das, was im Dossier steht, ist richtig.

Was jetzt auf Basis Ihrer Stellungnahme auffällt, ist Folgendes: Sie haben gesagt, dass bei etwa 50 Patienten – das sind ja nicht so wenige – eine solche Fehlzuordnung technischer Natur stattgefunden hat. Diese Fehlzuordnung hat zum einen ganz offensichtlich nur in dem Nivolumab-Arm stattgefunden – vielleicht können Sie dazu noch sagen, warum –; zum anderen entspricht die Zahl der Patienten, von denen Sie gesagt haben, sie seien fehlzugeordnet worden, etwa 50, nicht der Zahl der Patienten, die dann in Woche null in Ihrer Auswertung im Vergleich zum Studienbericht zusätzlich ausgewertet werden. Letzteres mag dadurch erklärt sein, dass Sie die andere Grundgesamtheit haben, nämlich diejenigen, die auch einen Post-Base-Line-Wert haben. Aber das wäre eben meine Frage. Diskrepanz 50 zu 34 ist die eine Frage, warum nur bei Nivolumab, ist die zweite Frage.

Herr Prof. Hecken (Vorsitzender): Frau Dr. Kupas, bitte.

Frau Dr. Kupas (BMS): Die Diskrepanz lässt sich dadurch erklären, dass es wirklich zwei verschiedene Tabellen sind. Es sind auch nicht nur Patienten im Nivolumab-Arm gewesen. Wir haben in beiden Armen dann natürlich mehr Base-Line-Patienten gehabt, und wenn wir dann die Patienten darstellen, die Base-Line plus mindestens einen Nachfolgezeitpunkt haben, kommen wir zu den Zahlen, die wir im Dossier haben. Ich habe die exakten Zahlen nicht, weil ich natürlich von den globalen Kollegen diese 50 Patienten als Zahl genannt bekommen habe; aber sie sind in beiden Armen aufgetreten, und in beiden Armen ist korrigiert worden.

Herr Prof. Hecken (Vorsitzender): Herr Kaiser.

Herr Dr. Kaiser: Das ist nun nicht ganz unwichtig, weil Sie hierzu ja auch entsprechende Daten und auch noch einmal Sensitivitätsanalysen nachgereicht haben. Das klingt für mich tatsächlich auf Basis der Daten nicht plausibel, dass es in zwei Armen ist. Wie viele Patienten waren denn im Nivolumab-Arm und wie viele in dem anderen Arm fehlzugeordnet zu Base-Line?

Frau Dr. Kupas (BMS): Wie gesagt, die Antwort weiß ich nicht genau. Ich habe nur diese absolute Anzahl von 50. Ich weiß aber, dass es in beiden Armen stattgefunden hat. Dadurch, dass diese Tabellen komplett unterschiedlich aufgesetzt sind, ist es auch schwer nachzuhalten, wo die Patienten genau sind. Ich habe nur exakt nachgefragt, und unsere Tabellen im Dossier entsprechen den Rohdaten, wie sie in der Datenbank hinterlegt sind. So haben wir dann auch ausgewertet, weil das auch für uns klinisch richtig war, wenn es direkt nach der Dosierung war, dass es trotzdem ein Base-Line-Wert ist.

Herr Prof. Hecken (Vorsitzender): Herr Kaiser.

Herr Dr. Kaiser: Vielleicht nur als letzte Anmerkung dazu, weil ich nachvollziehen kann, dass Sie das jetzt so nicht sagen können, noch Folgendes: Die Daten, die Sie geliefert haben, die Studienberichtstabelle und die Tabelle im Dossier sowie Ihre Aussage jetzt, dass dies auch im Vergleichsarm stattgefunden hat, sind mit den Daten so nicht kompatibel.

Herr Prof. Hecken (Vorsitzender): Da müssten Sie also noch einmal nachschauen.

Frau Dr. Kupas (BMS): Dazu kann ich jetzt nichts sagen.

Herr Prof. Hecken (Vorsitzender): Ja, okay. Es wäre nur ganz hübsch, wenn man es vielleicht aufklären könnte. – Frau Müller, bitte.

Frau Dr. Müller: Ich wollte eigentlich noch einmal ganz kurz zurück zu den Scores. Ich hätte da noch einmal eine Frage an die Fachgesellschaften. Ich weiß nicht, ob das okay ist, oder wird es jetzt hier noch thematisiert?

Ich wollte noch einmal fragen: Wie ist das denn, wenn man sie vergleicht? Dann ist ja ein Großteil der Variablen, die erhoben werden, identisch. Die Frage ist: Wenn man jetzt in einem Fall eine Effektmodifikation sieht und in dem anderen nicht – – Ganz kurz einmal: Wo sich das unterscheidet, ist ja entweder der LDH-Wert beim Motzer-Score und sind eben bei den Heng-Kriterien die Granulozyten oder die Thrombozyten, wenn sie erhöht sind. Welche Rolle spielt das in der Klinik? Ist da ein Unterschied, oder wie schätzen Sie das ein, weil es ja ein bisschen ungewöhnlich ist? Wenn sich da einmal eine Effektmodifikation zeigt oder nicht, ist das ein relevanter Unterschied, dass man sozusagen sagt, man guckt jetzt mehr auf das eine und guckt Laktat-Dehydrogenase gar nicht mehr an, oder ist das eigentlich in der Größenordnung vergleichbar, um etwas über das Risiko der Patienten zu sagen? Dann würde das mehr in die Richtung gehen, dass das vielleicht eine zufällige Aufteilung wäre, weil die Variablen, die etwas über den Allgemeinzustand sagen, wie Karnofsky, Zeit bis zum Therapiebeginn usw., ja in beiden Scores genau die gleichen sind. Es geht einfach darum, dass wir es ein bisschen besser verstehen.

Herr Prof. Hecken (Vorsitzender): Ja, ich habe gesehen, Herr Professor Miller und Herr Dr. Johannsen haben offensichtlich die Erkenntnisse und die Absicht, die Frage zu beantworten. – Herr Johannsen, bitte schön.

Herr Dr. Johannsen (DGU): Für uns Kliniker spielen die Heng-Kriterien vor der Erstlinientherapie selbstverständlich eine große Rolle, in den Nachfolgetherapien nicht, weil es sehr schwierig ist, sie verlässlich zu erheben. Wenn noch hinzukommt, dass Patienten Komplikationen hatten, wird es ebenfalls schwierig, weil es natürlich Confounder gibt, zum Beispiel bei der Anämie usw.

Für uns ist entscheidend, wie in der Zweit- oder Drittlinientherapie der allgemeine Zustand des Patienten ist, aber nicht der Heng-Score. Auch der Motzer-Score spielt für uns klinisch dort keine Rolle. Wir würden einem Patienten eine Therapie nicht wegen eines LDH-Wertes vorenthalten. Es gibt sehr viele Dinge, die für uns klinisch von Bedeutung sind, auch von der Paraklinik her; aber bestimmte Befunde müssen wir beim Nierenzellkarzinom anders einordnen als bei vielen anderen Tumoren. Diese Einteilung ist sicherlich für die hiesige Stratifizierung nachvollziehbar wichtig gewesen, um eine gewisse Aussagekraft der Studie zu gewährleisten. Aber für uns Kliniker spielt das in den Nachfolgetherapien keine Rolle.

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

Herr Prof. Dr. Staehler (DGHO): Diese Scores verwenden wir eigentlich nur, um uns in Leitlinien zu orientieren, welche Medikamente wir den Patienten anbieten dürfen. Das heißt, sie dienen der Unterscheidung, ob ein Patient in einer Hochrisikosituation Temsirolimus bekommt oder nicht. Ansonsten sind diese Scores zum einen unbekannt in der „normalen“ Therapielandschaft – das heißt, die meisten Kollegen werden sie klinisch nicht einsetzen können, weil sie sie nicht auswendig kennen –, und zum anderen brauchen wir sie, um Studien vergleichbar zu machen, um zu verstehen, welche Studienpopulationen wir dort vorhanden haben. Übersetzt in den Alltag, heißt das nichts anderes als „fitter Patient – gute Prognose“ bzw. „unfitter Patient – schlechte Prognose“, und daran werden sie sich ori-

entieren und für weitere Therapielinien interessieren, ansonsten gar nicht. Da wird am Ende nur die klinische Situation des Patienten herangezogen: Ist er therapiefähig, oder ist er nicht therapiefähig? Ansonsten sind diese Scores nicht wirklich hilfreich.

Frau Dr. Müller: Okay.

Herr Prof. Hecken (Vorsitzender): Danke schön. – Ich schaue einmal in die Runde. – Herr Eyding, bitte.

Herr Dr. Eyding: Ich habe noch eine Frage, die sich aus der Anhörung zu dem Lungenkarzinom ergibt. Es gibt ja hierbei keine Interaktion mit dem PD-L1-Rezeptorstatus. Meine Frage wäre: Wie viele Proben sind sozusagen anamnestisch aus der First-Line für die Erhebung des PD-L1-Status übertragen worden, und wie viele sind sozusagen an frischen Biopsien für die Zweitliniensituation erzeugt worden?

Herr Prof. Hecken (Vorsitzender): Bitte schön, Frau Beckers.

Frau Dr. Beckers (BMS): Das kann ich für die einzelnen Proben und für die einzelnen Patienten aus der Hüfte heraus geschossen nicht beantworten. Was wir beantworten können, ist, dass für alle Patienten der PD-L1-Status vor Studienbeginn vorliegen musste, und was wir weiterhin beantworten können, ist, dass auch hier gezeigt wurde, dass PD-L1 kein guter prädiktiver Biomarker für die Wirksamkeit von Nivolumab ist.

Herr Prof. Hecken (Vorsitzender): Nachfrage Herr Eyding.

Herr Dr. Eyding: Wir haben ja gehört, dass das irgendwie eine dynamische Situation sein kann, dass sich der Status eben ändern kann, je nach Behandlung und Behandlungslinie und Vorbehandlung. In der vorherigen Anhörung war es so, dass die meisten Proben anamnestisch erhoben worden sind und die geringere Zahl vor der zweiten Linie, also vor Studieneinschluss. Insofern wäre es schon interessant, wie die Situation vergleichbar ist und ob man das irgendwie mit der Abwesenheit dieses Status korrelieren kann, weil das ja schon erst einmal ein ganz plausibler Befund in der Lunge war, dass das PD-L1-Status-abhängig ist, während es hier eben nicht so ist. Insofern wäre es schon eine interessante Information, noch zu wissen, ob die Situation hier vergleichbar war oder ob das hier total anders war. Vielleicht können Sie das wenigstens hier schon sagen.

Herr Prof. Hecken (Vorsitzender): Frau Beckers, ja.

Frau Dr. Beckers (BMS): Ich kann sagen, dass wir oben akzeptiert haben – – In dieser Studie, die sowohl aus dem Primarius stammten als auch aus den Metastasen, gibt es ja auch noch einmal Unterschiede, wie Sie bestimmt vorher auch gehört haben, auch je nachdem, welcher Teil – aus dem Primarius oder aus der Metastase – gerade untersucht wurde. Auch hier kann es Unterschiede geben. Ich kann ferner sagen, dass in dieser Studie Proben akzeptiert wurden, die maximal fünf Jahre alt sein durften.

Herr Prof. Hecken (Vorsitzender): Okay. Also, die Quantifizierung, die wir gern gehabt hätten, Herr Eyding, bekommen wir nicht. Ich glaube, eben waren es so etwa 60 Prozent, die eben älter waren, und dann eben etwa 40 Prozent, die vor Einleitung der neuen Therapie gezogen wurden. – Ich habe jetzt Herrn Dr. Spehn und dann Herrn Dr. Bartmann. – Herr Dr. Spehn.

Herr Dr. Spehn (AkdÄ): Das Thema ist eigentlich schon vorbei. Trotzdem will ich noch eine Frage zum Motzer-Score stellen. Wann haben Sie denn den Heng-Score in die Bewertung hineingenommen? In der Publikation im „New England Journal“ stehen in der Tabelle der Karnofsky, der MSKCC,

der Motzer-Score; da steht aber nichts vom Heng-Score. Ist der Heng-Score nach Abschluss der Studie hineingenommen worden?

Herr Prof. Hecken (Vorsitzender): Frau Dr. Kupas, bitte.

Frau Dr. Kupas (BMS): Nein, das war eine präspezifizierte Subgruppe, von Anfang an. In der Publikation werden nicht immer alle Subgruppen dargestellt. Aber sie war von Anfang an im Studienprotokoll zur Auswertung vorgesehen, und so ist sie auch im Dossier dargestellt.

Herr Prof. Hecken (Vorsitzender): Danke schön. – Herr Dr. Bartmann, bitte.

Herr Dr. Bartmann: Frau Beckers hatte gerade die Metastasierung angesprochen. Da wollte ich einmal nachfragen. Mir ist aus den Unterlagen – es geht hier offensichtlich um fortgeschrittenes Nierenzellkarzinom – der Anteil der nicht metastasierten Tumoren nicht ganz klar geworden. Könnten Sie das noch einmal sagen? Also, ist das bekannt? Also, es ist sicherlich bekannt. Aber ist es auch ersichtlich, wie hoch dieser Anteil dann war?

Dann auch noch eine Anschlussfrage zum Memorial-Sloan-Kettering-Cancer-Score: Ist das nicht nur für metastasierte Patienten entwickelt worden?

Herr Prof. Hecken (Vorsitzender): Frau Beckers.

Frau Dr. Beckers (BMS): In dieser Studie wurden Patienten eingeschlossen, die nicht kurativ operabel waren, also lokal fortgeschritten ohne kurative Operabilität zur R₀-Resektion, bzw. fernmetastasierte Tumoren hatten. Aber es waren nur Patienten, die bereits in dieser Phase waren und schon mindestens eine systemische Vortherapie bekommen haben mussten. Das heißt im Prinzip, vor der systemischen antiangiogenetischen Vortherapie war schon das Kriterium erreicht, dass ein lokales Stadium, das nicht operabel ist, nicht mehr gegeben war. Wie viel Prozent der Patienten, kann ich insofern nicht beantworten. Es waren alle Patienten, die aufgrund ihrer Erkrankung bereits mindestens eine Systemtherapie gebraucht haben. – Ihre zweite Frage habe ich jetzt leider vergessen.

Herr Dr. Bartmann: Das war die Frage noch einmal nach dem Score, danach, ob der Motzer-Score nicht nur für metastasierte Nierenzellkarzinome entwickelt worden ist. Wenn Sie das noch beantworten könnten!

Frau Dr. Beckers (BMS): Der Motzer-Score ist – genauso wie auch in dieser Studie angewandt – für Patienten mit lokal fortgeschrittenen und metastasierten Tumoren entwickelt.

Herr Prof. Hecken (Vorsitzender): Danke schön. – Weitere Fragen? – Keine. – Doch, Herr Nell, bitte.

Herr Dr. Nell: Ich habe eine kurze Frage hinsichtlich Übertragbarkeit der Daten auf Patienten mit nicht klarzelligem Nierenzellkarzinom; sie waren in der Studie ja ausgeschlossen. Können Sie vielleicht aus anderer Forschung etwas dazu sagen – es gibt ja bei Kinase-Inhibitoren oder bei der antiangiogenetischen Therapie durchaus Unterschiede –, beispielsweise Remissionsraten?

Herr Prof. Hecken (Vorsitzender): Bitte schön, Frau Dr. Beckers.

Frau Dr. Beckers (BMS): Das Einschlusskriterium in diese Studie war, dass die Patienten eine klarzellige Komponente haben mussten. Das heißt aber nicht, dass es ausschließlich klarzellige Karzinome waren, sondern es konnten auch Mischanteile mit dabei sein. – Erster Punkt.

Zweiter Punkt: Alle anderen Substanzen, mit denen die Patienten entweder vortherapiert sind oder die heute zugelassen sind, haben die breite Zulassung sowohl für Patienten mit klarzelligem Nieren-

zellkarzinom als auch mit nicht klarzelligem Nierenzellkarzinom, weil die Gruppe dieser nicht klarzellig-nierenzellkarzinome sehr klein ist und eine Studie extra dafür zu etablieren sehr schwierig ist. Wir haben bereits auch Einzelberichte zur Wirksamkeit von Nivolumab bei Patienten mit nicht klarzelligem Nierenzellkarzinom, auch publiziert.

Herr Prof. Hecken (Vorsitzender): Das beantwortet die Frage, Herr Nell? – Ja. Weitere Fragen? – Das sehe ich nicht. – Ja, bitte, Herr Nell.

Herr Dr. Nell: Ich habe noch ein paar Fragen abschließend zu den Patientenzahlen, die das IQWiG ja auch aufgeworfen hat. Zum einen wurde Ihrerseits nur eine Einjahresprävalenz berücksichtigt. Da wollte ich fragen: Warum nur eine Einjahresprävalenz? Das wäre ja dann eigentlich eine Unterschätzung. Wenn man die Anteile des fortgeschrittenen Stadiums sieht, müsste der Anteil der Prävalenz niedriger sein. In einem weiteren Schritt kommen Sie dann zu einer Überschätzung, und am Schluss habe ich noch eine Frage zu der Obergrenze 20 Prozent bei den Patienten, die in das Stadium IV progredieren. Da haben Sie zwar auf Quellen verwiesen; diese Quellen geben aber 20 bis 30 Prozent an. Sie schreiben in der Stellungnahme dann, dass Sie sich dann nach neueren Werten auf 20 Prozent festgelegt haben, was für mich dann nicht nachvollziehbar ist.

Herr Prof. Hecken (Vorsitzender): Frau Herrmann, bitte.

Frau Dr. Herrmann (BMS): Ich kann nachvollziehen, dass Sie es als schwierig empfinden, diese Zahlen zu verfolgen. Wir hatten die gleiche Schwierigkeit wie Sie, sichere Quellen zu diesen Zahlen zu finden. Die Studien mit den Zahlenangaben, die wir gefunden haben, waren die beste Evidenz, die wir finden konnten. Vom IQWiG war ja auch entsprechende Kritik dazu geäußert worden.

Wir haben versucht, mit der Inzidenz zu den Zahlen eine Untergrenze festzustellen und eine Obergrenze zu finden, die möglichst konservativ geschätzt ist. Deshalb haben wir beschlossen, die Einjahresprävalenz zu nehmen, weil wir denken, sie trifft die Patientengruppe am besten. Wir haben ja auch entsprechend begründet, wie wir darauf gekommen sind. Wir haben einen Mischwert aus der Inzidenz plus Einjahresprävalenz gerechnet.

Wir sind dann in unserem zweiten Schritt, wie Sie richtig gesagt haben, auf diese 20 Prozent gekommen, die wiederum einen Mischwert aus den Zahlen, die wir aus den UICC-Daten und zu O'Reilly und Wolfert gefunden hatten, darstellen, und haben daraus dann diese 20 Prozent errechnet. Das schien oder scheint uns der Weg zu sein, wie man am besten eine konservative Schätzung für die Patientenanteile darstellen kann.

Herr Prof. Hecken (Vorsitzender): Herr Nell.

Herr Dr. Nell: Okay, den letzten Schritt schaue ich mir dann noch einmal an. Ich habe Sie jetzt so verstanden, dass Sie Patientenzahlen in Gänze praktisch konservativ – –, dass sie also unterschätzt sind oder potenziell unterschätzt sind.

Frau Dr. Herrmann (BMS): Es ist so, wie das IQWiG eigentlich auch beschrieben hat: Die Obergrenze ist eher eine Überschätzung, während wir hinsichtlich der Untergrenze denken, dass es natürlich richtig ist, wenn wir da die Inzidenz nehmen, um diese Grenze festzustellen. Das ist das, was wir festgestellt haben.

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

Herr Dr. Eyding: Ich habe noch eine ganz kurze Nachfrage, noch einmal zu der Altersrepräsentativität der Studie. Da würde ich gerne noch einmal um eine Stellungnahme der AkdÄ und auch der Fachgesellschaften bitten, inwiefern man wirklich so annehmen kann, dass es übertragbar ist, so wie

die EMA das eingeschätzt hat, wie das der pU einschätzt oder ob man da tatsächlich ein bisschen Probleme hat, dass das so junge Patienten sind. – Vielleicht könnten die jeweiligen Vertreter noch einmal etwas dazu sagen.

Herr Prof. Hecken (Vorsitzender): Herr Professor Staehler, bitte.

Herr Prof. Dr. Staehler (DGHO): Wir denken, dass die Studienpopulation die zu behandelnde Patientengruppe sehr gut abbildet. In der Regel ist es so, dass wir im Rahmen eines solchen Studienprotokolls natürlich Patienten haben, die wir ausschließen müssen. Insbesondere bei älteren Patienten sind Ausschlusskriterien sehr viel häufiger vorliegend als bei jüngeren Patienten, weswegen diese nicht Eingang in eine derartige Studie finden, was aber im klinischen Alltag nicht dagegen spricht, dass diese Patienten behandelbar sind, ganz im Gegenteil: Gerade ältere Patienten lassen sich mit dieser Substanz sehr gut behandeln.

Das andere ist auch: Die Patienteninzidenzen in einer metastasierten Situation sind etwas anders. Meist sehen wir hier die jüngeren Patienten und nicht die älteren, die in der Regel die Erkrankung dann schon einer Therapie zugeführt bekommen haben, die ja im überwiegenden Fall, sprich in 80 Prozent der Fälle, kurativ ist und somit nicht mehr in die metastasierte Situation übergeht. Das heißt, gerade bei den jüngeren Patienten finden wir die aggressiveren Tumorformen und die Wahrscheinlichkeit, dass sich diese Patienten in einem Studienprotokoll wiederfinden. Das heißt aber nicht, dass wir glauben würden, dass diese Patienten die Bevölkerung, so wie wir sie später behandeln würden, abdecken.

Herr Prof. Hecken (Vorsitzender): Herr Professor Miller.

Herr Prof. Dr. Miller (DGU): Vielleicht noch Folgendes: In der klinischen Entscheidungssituation würden Sie nicht das kalendarische Alter für irgendetwas heranziehen. Das ist einfach unreal. Da müssen Sie überlegen, welche Alternativen Sie haben, und sich fragen, ob das ein Patient ist, der, wie diskutiert worden ist, biologisch für eine Therapie geeignet ist. Dann würden Sie nicht sagen, so, und jetzt ist er 74 Jahre, jetzt bekommt er das nicht mehr. Das findet so einfach nicht statt.

Herr Prof. Hecken (Vorsitzender): Herr Fetscher.

Herr Dr. Fetscher (AkdÄ): Trotzdem besteht hier zwischen Glauben und Wissen ein Unterschied, der uns ja auch beschäftigen kann. Wenn diese Studien so sind, dass sie uns nur helfen zu glauben, dass diejenigen, die die Ausschlusskriterien der Studie negativ erfüllen, also nicht hineinkommen, in der klinischen Realität dann doch behandelbar sind, dann ist das Vertrauen in die beurteilenden Kliniker ja sehr groß und das Vertrauen in eine Studie, uns dabei zu helfen, dies zu definieren, sehr klein.

Das ist eigentlich bedauerlich; denn wenn diese Population so behandelbar ist, wie es hier gesagt wurde, dann wäre es ja sinnvoll gewesen, eine größere Gruppe davon einzuschließen, zu sagen, okay, Herr Miller, klar, wir müssen summarisch klinisch entscheiden, völlig klar, und wir wollen auf gar keinen Fall – die Kliniker nicht und die Industrie aus anderen Gründen auch nicht – durch zu rigide Kriterien gebunden werden.

Gleichzeitig ist es natürlich besser, wenn wir besser definierte Kriterien oder Grenzwerte kennen, die uns helfen, uns über das klinische Urteil hinaus zu entscheiden: Dieser Patient hat einen Nieren-Leber-Funktionsschaden, hat schon eine Pneumonitis hinter sich, hat ein anderes Problem. Wir haben eine Menge Immunprobleme, von denen wir wissen, dass sie kategorisch ausschließen – das ist ja auch gut so –; aber das Spektrum der Kriterien, die uns zu entscheiden helfen, diejenigen Patienten zu behandeln, die die hiesigen Ausschlusskriterien negativ erfüllen, ist für mich zu klein, und da ist auch die Industrie gefordert, uns Studien vorzulegen, die uns helfen, diese Generalisierung im nächsten Schritt, bei der Zulassung, etwas intelligenter zu machen.

Herr Prof. Hecken (Vorsitzender): Danke für diese Anmerkung. Diese Diskussion haben wir ja schon häufiger geführt, dass Studienpopulationen sehr häufig eben doch folgende Frage aufwerfen: Ist das jetzt das Klientel, das am Ende dann möglicherweise tatsächlich in der Masse hier behandelt wird? Aber, wie gesagt, das soll uns nicht von der ständigen Mahnung und Verpflichtung entbinden, möglichst nahe an die tatsächlichen Patientengruppen, die die Versorgungsrealität widerspiegeln, heranzukommen. Das hilft uns nur im Augenblick nicht. Jetzt müssen wir die Frage beantworten: Ist es übertragbar, oder ist es nicht übertragbar? – Weitere Fragen? – Bitte schön, Herr Dr. Bartmann.

Herr Dr. Bartmann: Im Anschluss an die gerade geführte Diskussion wollte ich noch einmal auf die Patienten zu sprechen kommen, die mit Temsirolimus vorbehandelt sind. Es ist gesagt worden, sie spielten in der klinischen Praxis keine so große Rolle oder hätten überhaupt keine Bedeutung. Aber die Frage, weil sie ja hier auch als Gruppe definiert sind, richtet sich noch einmal im Besonderen an die AkdÄ, ob man diese Patienten nicht auch als eine Patientenpopulation bewerten muss, die eben eine ungünstige Prognose hat. Mir ist noch nicht so ganz klar geworden, ob dies Patienten sind, die eben auch diesem anderen Kriterium, dem diskutierten Prognosekriterium, irgendwie zugerechnet werden können.

Herr Prof. Hecken (Vorsitzender): Herr Dr. Spehn, bitte.

Herr Dr. Spehn (AkdÄ): Ein klares Ja: Die mit Temsirolimus vorbehandelten Patienten müssen ja entsprechend der Zulassung einen erhöhten Motzer-Score aufweisen. Das sind ja Patienten mit Hochrisiko; ihnen aufgrund dieser Daten Nivolumab zu geben, halte ich für keinen Fehler, sondern für richtig. Nur geht es hier um Bewertung, und wir haben keine Daten, wie die Verläufe bei Temsirolimus-Vorbehandelten sind. Sie haben wir nicht; dass wir das in praxi machen, ist eine andere Frage.

Herr Prof. Hecken (Vorsitzender): Frau Dr. Beckers, bitte.

Frau Dr. Beckers (BMS): Es ist absolut richtig, dass wir aus einer randomisierten, klinischen, gegen eine aktive Vergleichssubstanz kontrollierten Situation keine Daten haben, wie Patienten ansprechen, die bereits mit Temsirolimus vorbehandelt wurden. Es gibt aber Daten aus Phase-I- und Phase-II-Studien zu Patienten, die mit einem mTOR-Inhibitor bzw. auch ausschließlich mit Temsirolimus vorbehandelt wurden. Das sind aber natürlich, weil hier die Kriterien anders waren und wir auch hier nur Daten aus einer randomisierten kontrollierten Phase-III-Studie in unserem Dossier vorlegen konnten, Daten, die wir jetzt hier nur indirekt heranzuführen konnten.

Herr Prof. Hecken (Vorsitzender): Danke schön. – Weitere Fragen? – Keine. Dann würde ich Ihnen erneut das Wort geben, um aus Ihrer Sicht die letzte Stunde kurz zusammenzufassen. Herr Dr. May, bitte schön.

Herr Dr. May (BMS): Herr Professor Hecken, meine Damen und Herren, vielen Dank. Ich hoffe, wir konnten Ihre Fragen weitgehend beantworten. Wie im Eingangsstatement bereits erwähnt, war es für uns wichtig, hier mit Ihnen nochmals in den Dialog eintreten zu können. Zusammenfassend möchte ich sagen, dass sich die heutige Diskussion sehr viel um die Gruppe in der Vorbehandlung mit Temsirolimus gedreht hat – wir hoffen zumindest von unserer Seite, diesen Aspekt auch eingehend beleuchten zu können –, des Weiteren natürlich um die Subgruppenbildung auf der Basis des MSKCC-Scores und dessen Korrelation mit dem Überlebensvorteil für Patienten. Aus unserer Sicht ist der MSKCC dafür nicht geeignet. Er dient aus unserer Sicht zur erfolgreichen Stratifizierung von Patientengruppen mit unterschiedlichen Prognosefaktoren in einer klinischen Studie.

Eine Frage ist offen geblieben. Herr Dr. Kaiser hatte die Rücklaufquoten der Patientenfragebögen angesprochen. Hier wollten Sie noch einmal Zahlen zu Nivolumab und Everolimus haben. Wir werden das im Nachgang nachliefern.

Ich möchte zusammenfassend sagen, dass sich unserer Auffassung nach in der Gesamtschau der vorhandenen Datenlage ein Beleg für einen erheblichen Zusatznutzen von Nivolumab bei der Behandlung von erwachsenen Patienten mit fortgeschrittenem Nierenzellkarzinom nach antiangiogener Vortherapie ergibt. Es handelt sich unserer Auffassung nach um eine nachhaltige und gegenüber der ZVT Everolimus um eine bisher nicht erreichte Verbesserung des therapielevanten Nutzens, und zwar in den Kategorien „Verlängertes Überleben“ und „Reduktion von Nebenwirkungen“, ein Thema, das in der heutigen Diskussion überhaupt nicht stattgefunden hat, „Verbesserte Morbidität“ und auch „Lebensqualität“.

Ich bin der Meinung, dass es heute um eine Substanz geht, die, wie vorhin schon angesprochen worden ist, für viele Patienten Grund zur Hoffnung auf ein längeres und besseres Leben gibt. Wenn wir Nivolumab als Ganzes betrachten, so sollte das therapeutische Potenzial dieser Substanz auch in seiner ganzen Breite betrachtet werden und unserer Auffassung nach auch in der abschließenden Bewertung entsprechend berücksichtigt werden. – Vielen Dank noch einmal von unserer Seite für die Diskussion.

Herr Prof. Hecken (Vorsitzender): Ja, ganz herzlichen Dank für diese Zusammenfassung. Herr Dr. May, wenn Sie jetzt gerade noch einmal ausdrücklich darauf abheben, dass bestimmte Dinge und bestimmte Daten, die die Überlebenszeitverlängerung, die das Nebenwirkungsspektrum und das UE-Spektrum ansprechen, hier nicht erörtert worden sind, dann ist das kein mangelndes Interesse; vielmehr können Sie davon ausgehen, dass wir a) die Studien gesehen haben, b) die Dossierbewertung und c) Ihre Stellungnahme. Deshalb muss man nicht unbedingt Eulen nach Athen tragen, zumal wir ja in der Stunde zuvor in ähnlichen Spektren auch etwas deutlicher und tiefer dann eben über die Vorteile, die sich bei den Nebenwirkungsprofilen zeigen, gesprochen haben.

Ich sage das nur deshalb, damit das auch im Wortprotokoll so festgehalten ist und der unbefangene Leser am Ende nicht den Eindruck gewinnen muss, dass hier bestimmte relevante Dinge bei der Entscheidungsfindung außer Betracht geblieben sind. Wenn wir hier jedes Mal alles herunterbeten wollen, dann singen wir uns gegenseitig die Stellungnahme vor, und dann dauert es halt noch ein bisschen länger. Das muss aber nicht unbedingt sein.

Danke, dass Sie da waren. Wir werden das jetzt auswerten und werden dann unsere Entscheidungen zu treffen haben.

Ende der Anhörung: 14.02 Uhr