

Mündliche Anhörung



**Gemeinsamer
Bundesausschuss**

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

hier: Wirkstoff Nintedanib

Sitzung im Hause des Gemeinsamen Bundesausschusses in Berlin
am 28. Juli 2015
von 10.07 Uhr bis 11.26 Uhr

– Stenografisches Wortprotokoll –

Angemeldete Teilnehmer der Firma **Boehringer Ingelheim Pharma GmbH & Co. KG:**

Frau Geier
Herr Kögler
Herr Dr. Pfannkuche
Herr Dr. Trampisch

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

Herr Prof. Dr. Ruof
Herr Prof. Dr. Leehmhuis

Angemeldete Teilnehmer für die **Deutsche Gesellschaft für Pneumologie und Beatmungsmedizin e. V. (DGP):**

Herr Prof. Dr. Behr
Herr Prof. Dr. Costabel

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

Frau Bungart
Herr Dr. Werner

Beginn der Anhörung: 10.07 Uhr

(Die angemeldeten Teilnehmer betreten den Raum)

Herr Dr. Vöcking (stellv. Vorsitzender): Ich begrüße ganz herzlich die Gäste für die erste heutige Anhörung. Herzlich willkommen! Viele Gesichter, die ich sehe, sind bekannt, aber nicht alle. Da ein Teil von Ihnen das Verfahren vielleicht nicht in extenso kennt, erlaube ich mir ein paar Hinweise.

Meine Bitte wäre, dass Sie, wenn Sie sich nachher äußern, Ihren Namen nennen und die Institution oder das Unternehmen, das Sie vertreten, damit es im Protokoll festgehalten werden kann. Es wird ein Wortprotokoll geführt; es findet auch eine Tonaufnahme statt. Das soll Sie nicht beunruhigen. Das dient alles der Wahrheitsfindung hier.

Ich werde erst einmal für das Protokoll feststellen, wer zur Anhörung erschienen ist, und fange so an, wie es auf der Agenda steht; da bitte ich wiederum um Nachsicht. Für den pharmazeutischen Unternehmer Boehringer Ingelheim sind anwesend Frau Geier, Herr Kögler, Herr Pfannkuche und Herr Trampisch. Herzlich willkommen! Dann sind für die Deutsche Gesellschaft für Pneumologie und Beatmungsmedizin (DGP) Herr Professor Behr und Herr Professor Costabel anwesend. Herzlich willkommen! Von Roche Pharma GmbH habe ich Herrn Professor Ruof und Herrn Professor Leehmhuis gesehen, vom vfa begrüße ich Herrn Werner und Frau Bungart. Herzlich willkommen!

Dann können wir in die mündliche Anhörung einsteigen. Zur Information fasse ich noch einmal ganz kurz zusammen: Es gibt Stellungnahmen des Unternehmers Boehringer, dann von Roche Pharma, von der Deutschen Gesellschaft für Pneumologie und Beatmungsmedizin (DGP), und eine Stellungnahme vom vfa. Es hat eine fünfte Stellungnahme vom Verein Lungenfibrose e. V. gegeben; ich weise darauf hin, dass dieser Verein eigentlich nicht zur Stellungnahme berechtigt war und dementsprechend auch nicht zur mündlichen Anhörung geladen ist.

Es geht um den Wirkstoff Nintedanib, Handelsname Ofev[®], ein Orphan Drug. Insofern gibt es keine Nutzenbewertung vom IQWiG. Es gibt auch keine Stellungnahme der AkdÄ, sondern nur die von mir eben schon erwähnte Stellungnahme der DGP.

Aus den Unterlagen ergibt sich aus meiner Sicht, dass wir heute zunächst einmal zwei Endpunkte diskutieren sollten: einmal die jährliche FVC-Abnahme und dann die Zeit bis zur ersten akuten Exazerbation. Beim ersten Punkt ist mir aufgefallen der Hinweis von Boehringer insbesondere auf die Resonanz in der internationalen wissenschaftlichen Literatur, dann die Positionierung der FDA etc.; das sind Themen, um die es hier geht. Auch von Relevanz dürfte die Frage sein, wie weit das schon ein validierter Endpunkt und ein validiertes Surrogat für Mortalität ist. Eine weitere Frage wäre in meinen Augen die Relevanz der Abnahme um 110 ml im Vergleich zum kompletten Lungenvolumen. Diese Aspekte sollten vielleicht neben dem zweiten Endpunkt „Zeit bis zur akuten Exazerbation“ besonders hervorgehoben werden. Meine Anregung wäre, auch noch das Thema Lebensqualität zu betrachten, und zwar im Grunde die Sichtweise von Boehringer, wie weit sich diese Lebensqualität verbessert. Darüber hat es ja Diskussionen gegeben. Ich glaube, das sind die drei wesentlichen Punkte, die jedenfalls aus meiner ersten Sicht eine Rolle spielen. Ich will Ihnen aber insofern natürlich nicht das Wort führen.

Sie kennen das Ritual hier: Ich gebe Ihnen gerne, wenn Sie es wünschen, die Gelegenheit zu einer grundsätzlichen Stellungnahme. Dann steigen wir in die Anhörung ein. – Sie haben das Wort.

Herr Dr. Pfannkuche (Boehringer Ingelheim Pharma): Vielen Dank, Herr Vöcking, für die wie immer nette Begrüßung und für die Erteilung des Wortes. – Sehr geehrte Mitglieder des Gemeinsamen Bundesausschusses! Das Arzneimittel Ofev[®], um das es heute geht, mit dem Wirkstoff Nintedanib ist

ein Arzneimittel zur Behandlung einer seltenen Erkrankung. Die idiopathische Lungenfibrose ist eine progrediente und bislang nicht heilbare Erkrankung, die bei der Mehrheit der Patienten innerhalb weniger Jahre zum Tode führt. Die Therapieoptionen, die wir derzeit am Markt haben, sind sehr eingeschränkt. Ofev[®] ermöglicht eine Verlangsamung des Fortschreitens der Erkrankung, und der maßgebliche Nutzen für die Patienten besteht darin, dass ihnen über die Zeit letztendlich mehr beatembares Lungenvolumen zur Verfügung steht. Das wirkt sich positiv auf die hier relevanten Endpunkte Mortalität, Morbidität und Lebensqualität aus.

Bevor ich diese Einschätzung weiter konkretisiere bzw. näher begründe, wie wir den Sachverhalt sehen, möchte ich Ihnen kurz meine Kollegen vorstellen, die mich heute hier begleiten. Neben mir sitzen Silke Geier – sie ist bei uns Teamleiterin im Bereich Markzugang und Gesundheitsökonomie für Atemwege und verantwortlich für das Frühbewertungsdossier – und Privatdozent Harald Kögler; er ist bei uns Teamleiter Medizin und war dort auch verantwortlich für das Frühbewertungsdossier zur idiopathischen Lungenfibrose. Herr Dr. Matthias Trampisch ist Mathematiker und von statistischer Seite aus mitverantwortlich für das Frühbewertungsdossier.

Für Ofev[®] liegen mit den beiden Zulassungsstudien INPULSIS-1 und INPULSIS-2 zwei RCTs mit einem niedrigen Verzerrungspotenzial auf Studienebene vor, und a priori war es geplant, beide Studien auch gepoolt auszuwerten. In den Studien selbst zeigte Ofev[®] eine klinisch relevante Reduktion der jährlichen Abnahme der forcierten Vitalkapazität um circa 50 Prozent und eine signifikante Senkung des Risikos für adjudizierte akute Exazerbationen um 68 Prozent. Hieraus ergibt sich aus unserer Sicht ein Hinweis für einen beträchtlichen Zusatznutzen, der eben in der klinisch relevanten Reduktion des mortalitätsassoziierten Endpunktes FVC sowie in dem reduzierten Risiko für akute Exazerbationen bei einem insgesamt akzeptablen Verträglichkeitsprofil gründet. Aus den dargestellten Daten geht hervor, dass hier die Vorteile die Nachteile überwiegen.

Für die heutige Anhörung sind aus unserer Sicht zwei Punkte relevant: Der erste Punkt – den haben auch Sie, Herr Vöcking, schon angesprochen – ist der in den Zulassungsstudien berichtete Endpunkt FVC und ist ein Surrogatparameter für die Mortalität. Der zweite Punkt, der für uns sehr wichtig ist, ist, dass der in den Zulassungsstudien berichtete Endpunkt Exazerbationen patientenrelevant ist und mit einer validen Methodik in den Studien erhoben worden ist. Zu beiden Punkten würde ich jetzt, wenn Sie erlauben, meinen Kollegen die Möglichkeit geben, ein paar mehr Details zu sagen, da sie Ihnen das besser erklären können als ich.

Herr Dr. Vöcking (stellv. Vorsitzender): Okay. – Herr Kögler.

Herr Kögler (Boehringer Ingelheim Pharma): Vielen Dank und guten Morgen auch von meiner Seite! Ich möchte zu dem Aspekt der akuten IPF-Exazerbation einige einleitende Bemerkungen machen. Die akute IPF-Exazerbation ist im Krankheitsverlauf eines IPF-Patienten ein einschneidendes Ereignis. Charakterisiert ist sie durch eine erhebliche Symptomverschlechterung. Das ist in der Regel die Dyspnoe, die Luftnot, die entweder schon zuvor bestand und erheblich schlimmer wird oder auch neu in erheblichem Maße über einen Zeitraum von vier Wochen auftritt. In den Studien, die wir im Modul 4 vorgelegt haben, war die Exazerbation in Übereinstimmung mit den Vorgaben der aktuell gültigen deutschen IPF-Leitlinie definiert als erhebliche akute Verschlechterung der respiratorischen klinischen Situation des Patienten, für die keine andere Ursache als die zugrundeliegende IPF identifiziert werden konnte.

Ich sprach von der erheblichen Symptomverschlechterung, aber die akute symptomatische Beeinträchtigung des Patienten ist nur einer der wichtigen Aspekte bei der IPF-Exazerbation. Viel entscheidender ist die erhebliche Auswirkung, die die Exazerbation auf den weiteren Krankheitsverlauf nimmt.

Man stellt sich die Exazerbation als einen akuten Fibroseschub vor, das heißt, es kommt massiv zur Vermehrung von narbigem Gewebe in der Lunge, und das geht einher mit einem erheblichen und unumkehrbaren Verlust an Lungenfunktion. Das ist für die Prognose, auch für die Überlebensprognose des Patienten ganz entscheidend. In der Mehrheit der Fälle führt eine Exazerbation zur Hospitalisierung. In unseren Studien war das bei zwei Dritteln aller berichteten Exazerbationen der Fall. Hier wird auch ein Unterschied zur Exazerbation beispielsweise der COPD deutlich – in diesem Hause ist in den letzten Jahren ja wiederholt der Nutzen von Produkten in diesem Indikationsgebiet bewertet worden –: Bei der COPD-Exazerbation kommt es viel seltener zu einer Krankenhausaufnahme; da führt etwa jede sechste bis zehnte Exazerbation den Patienten ins Krankenhaus. Im Gegensatz dazu beobachten wir es hier bei zwei Dritteln. Bei der COPD ist nach meiner Kenntnis die Patientenrelevanz der Exazerbation an und für sich in diesem Hause nie infrage gestellt worden.

Wenn der Patient ins Krankenhaus kommt, muss er oft intubiert und beatmet werden. Wenn er einmal auf der Intensivstation an der Beatmung ist, ist es sehr schwer, ihn wieder zu entwöhnen. Es gibt Daten, die zeigen, dass das mediane Überleben bei 2,2 Monaten liegt. Die akute Krankenhausmortalität beträgt 50 Prozent.

In unseren Studien wurden 76 Prozent aller IPF-Exazerbationen als schwerwiegende unerwünschte Ereignisse dokumentiert. Diese schwerwiegenden unerwünschten Ereignisse sind vom G-BA zuletzt immer als patientenrelevant anerkannt worden. Zudem zeigt sich in unseren Daten ein erheblicher Einfluss auf die Überlebensprognose. Von den 63 Patienten, die eine Exazerbation erlitten, verstarben innerhalb der Studie 31; das entspricht 49 Prozent. Von diesen Todesfällen traten fast drei Viertel innerhalb der ersten zwölf Wochen nach Beginn der Exazerbation auf. Damit war die Mortalität bei Patienten, die eine Exazerbation erlitten, etwa um den Faktor 10 höher als bei Patienten, die während der Studie keine Exazerbation erlitten. Das unterstreicht aus unserer Sicht noch einmal die Patientenrelevanz dieser Ereignisse.

Wir haben – das hat Herr Pfannkuche eben schon kurz angedeutet – in unseren Studien unter der Behandlung mit Ofev[®] eine statistisch signifikante Reduktion des Risikos einer akuten adjudizierten IPF-Exazerbation in den gepoolten Daten um 68 Prozent dokumentieren können. Und selbst wenn man nur diejenigen Exazerbationen betrachtet, die als schwerwiegendes unerwünschtes Ereignis dokumentiert worden sind, erweist sich dieser Effekt als robust; denn auch hier sehen wir in den gepoolten Daten eine 70-prozentige Reduktion des Risikos, ein solches Ereignis zu erleiden.

Damit darf ich zusammenfassen: Die in den INPULSIS-Studien dokumentierten IPF-Exazerbationen waren in hohem Maß patientenrelevante Ereignisse, und daher stellt der beobachtete Benefit unter der Behandlung mit Ofev[®] verglichen mit Placebo einen beträchtlichen patientenrelevanten Zusatznutzen dar. – Vielen Dank.

Herr Dr. Vöcking (stellv. Vorsitzender): Schönen Dank, Herr Kögler und Herr Pfannkuche. Wenn Sie erlauben, würde ich die Ausführungen gerne zum Anlass nehmen, eine Frage zu stellen – es wurde in Vorgesprächen einmal diskutiert –: Wie sieht es eigentlich mit einer Studiendauer von 52 Wochen aus? Das wurde verbunden mit einer entsprechend hohen Aussagekraft. Wie würden Sie das beurteilen bzw. wie sehen Sie das?

Herr Kögler (Boehringer Ingelheim Pharma): Wir denken, dass 52 Wochen, also eine Beobachtung von einem Jahr, adäquat ist. Es ist eine chronische Erkrankung. Bei anderen chronischen Erkrankungen wird in der Regel eine Studiendauer von mindestens 24 Wochen oder von einem halben Jahr als adäquat angesehen. Natürlich hätten wir in einer idealen Welt auch gerne noch längere Nachbeobachtungsdaten gehabt. Um diese zu generieren, haben die Patienten, die die INPULSIS-Studie in

der Regel regulär beendet haben, die Möglichkeit erhalten, in einer einarmigen offenen Anschlussstudie, Extensionsstudie, weiter behandelt zu werden. Es werden weiterhin unerwünschte Ereignisse und der Verlauf der forcierten Vitalkapazität bei diesen Patienten dokumentiert, sodass wir über die Zeit auch noch weitere klinische Informationen über 52 Wochen hinaus erhalten, dann allerdings nicht mehr im Placebovergleich.

Herr Dr. Vöcking (stellv. Vorsitzender): Schönen Dank. – Gibt es von Ihnen weitere Ausführungen, Herr Pfannkuche? Ich würde sonst gerne das Thema Lebensqualität noch einmal belegen. Ich hatte die Frage schon gestellt. Vielleicht könnten Sie das Thema aus Ihrer Sicht noch einmal anreißen.

Frau Geier (Boehringer Ingelheim Pharma): Wir hatten ja verschiedene Lebensqualitätsinstrumente erhoben und konnten sowohl für die generischen, zum Beispiel den PGI-C, bzw. auch für die spezifischen, zum Beispiel den SGRQ oder den SGRQ-I, einen positiven Trend zugunsten von Ofev[®] aufzeigen. Bislang gibt es noch keine validierten Lebensqualitätsinstrumente für die IPF, aber, wie ich eben schon ausführte, zeigen wir bei allen Instrumenten einen positiven Trend.

Herr Dr. Vöcking (stellv. Vorsitzender): Okay, schönen Dank. – Jetzt schaue ich in die Runde: Gibt es aus dem Kreis hier Fragen, Anmerkungen? – Frau Müller.

Frau Dr. Müller: Ich habe eine konkrete Nachfrage zu dem, was Sie, Herr Kögler, eben vortragen hatten. Sie sind ja auf die Relevanz der akuten Exazerbationen eingegangen, haben sie beschrieben und haben auch gesagt, dass es eben im Unterschied zu COPD sehr schwere Exazerbationen sind, dass es in zwei Dritteln der Fälle zu Hospitalisierungen kommt, und Sie hatten auch gesagt, dass sie in ungefähr 76 Prozent der Fälle, wenn ich das richtig verstanden haben, dementsprechend eben auch als SUEs – das folgt ja daraus – dokumentiert wurden. Da haben wir ja eine gewisse Doppelerfassung. Man kann diskutieren, ob die Doppelerfassung wünschenswert ist und welche Auswirkungen das hat. Das wurde ja sozusagen auch vor einem Sicherheits- und Wirksamkeitsendpunkt diskutiert. Aber Sie hatten gleichzeitig gesagt, die Nebenwirkungen sind – ich weiß nicht genau, wie Ihre Formulierung war, aber es ging in die Richtung – beherrschbar. Nun zu meiner Frage: Wenn ich eine Doppelerfassung habe und beim Wirksamkeitsendpunkt sozusagen weniger SUEs generiert werden, würde ich eigentlich erwarten, dass man bei den SUEs bei Nintedanib einen Vorteil sieht. Wir sehen aber keinen. Wenn Sie darauf noch einmal eingehen könnten, da wir ja hier eine Abwägung zwischen Nutzen und Risiko machen. Könnte es vielleicht sein, dass sozusagen diese Wirksamkeitsvorteile durch wirkliche Nebenwirkungen ausgeglichen werden und dass in der Gesamtsumme dann dazu führt, dass sich keine Unterschiede zwischen Nintedanib und Placebo bei den SUEs zeigen?

Herr Kögler (Boehringer Ingelheim Pharma): Die IPF-Exazerbationen haben nur etwa ein Fünftel aller Hospitalisierungen ausgemacht. Das heißt, wir haben ein großes Hintergrundrauschen anderer Ereignisse. Und in dem Fall ist es schwierig, einen statistisch signifikanten Effekt auf alle Hospitalisierungen nachzuweisen, wenn man bei einem Fünftel dieser Ereignisse, isoliert betrachtet, einen signifikanten Effekt sieht. Aus diesem Grund sehen wir nach unserer Auffassung bei den Hospitalisierungen bzw. auch bei den schwerwiegenden unerwünschten Ereignissen keinen signifikanten Unterschied.

Wir haben noch eine Analyse vorgenommen, bei der wir einmal gezielt die Exazerbationen, die im Rahmen der Efficacy-Analyse erfasst worden sind, aus der Gesamtheit der schwerwiegenden unerwünschten Ereignissen herausgerechnet haben; es zeigt sich da aber immer noch kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen, was die SUEs angeht, sodass wir davon

ausgehen, dass sich auch unter Herausrechnen der Exazerbationen kein möglicher Schaden durch die Therapie verglichen mit Placebo andeutet.

Herr Dr. Vöcking (stellv. Vorsitzender): Frau Müller, Nachfrage?

Frau Dr. Müller: Eigentlich nicht, nur ein kleiner Kommentar. Ich denke, für die Patienten ist es insofern wahrscheinlich in letzter Konsequenz egal, woher ihr schwerwiegendes unerwünschtes Ereignis kommt.

Herr Dr. Vöcking (stellv. Vorsitzender): Herr Eyding, ich habe Ihre Wortmeldung gesehen. Herr Professor Behr, unmittelbar dazu? – Ja, bitte, Sie haben das Wort.

Herr Prof. Dr. Behr (DGP): Ich denke, es ist schon ein Unterschied, ob eine akute Exazerbation oder eine andere Nebenwirkung einen Krankenhausaufenthalt erforderlich macht, weil die akute Exazerbation bei der idiopathischen Lungenfibrose eben mit einem extrem hohen Mortalitätsrisiko verbunden ist. Und nicht nur das: Die Patienten, die es überleben, sind in der Regel auch in ihrer Funktionalität eine ganze Stufe niedriger. Das heißt, es ist nicht so, wie das bei anderen Erkrankungen ist: Man hat eine Exazerbation, und wenn man sie übersteht, dann ist man wieder so wie vorher. – Das ist bei der idiopathischen Lungenfibrose ganz anders. Man müsste das eigentlich mehr als Krankheitsschub sehen, der auf einem schlechteren Niveau oder mortal endet. Ich glaube, das ist noch einmal ein ganz wichtiger Punkt. Deshalb kann man, glaube ich, diese SUEs nicht gleichsetzen.

Herr Dr. Vöcking (stellv. Vorsitzender): Danke für die Ergänzung. – Herr Eyding.

Herr Dr. Eyding: Ich habe auch eine Nachfrage zu den SUEs. Es wäre interessant, die Zahlen zu den SUEs ohne die Exazerbationen zu haben, die Sie gerade angesprochen haben; diese haben wir nicht, wenn ich das richtig sehe. Also, Sie haben jetzt zwar gesagt, dass das nicht signifikant ist, aber das würde ich natürlich gerne sehen, weil ich auch das für eine Doppelerfassung halte. Dass man die Nutzenendpunkte als UEs codiert, halte ich generell für ein Problem bei der ganzen Studiensituation, die wir da vorliegen haben; so haben wir nämlich sehr viele Doppelerfassungen.

Die zweite Doppelerfassung bezieht sich auf die Progresse. Bei den Progressen wurde uns, glaube ich, in der Nutzenbewertung berichtet, dass sie das häufigste SUE in beiden Armen sind. Die müsste man auch noch einmal herausrechnen, weil das ja auch Nutzenendpunkte sind und sie so einen Nutzen-Claim in Ihrem Dossier nach sich ziehen. Das heißt, man müsste auch die SUEs ohne die Progresse noch einmal separat dargestellt bekommen, um zu wissen, wie man die Nutzen- und die Schadensaspekte der beiden Endpunkte trennt.

Meine Nachfrage wäre also, ob es möglich ist, beides zu bekommen.

Herr Kögler (Boehringer Ingelheim Pharma): Die Analyse zur Berechnung der Exazerbationen können wir gern im Nachgang in Schriftform nachreichen, damit Sie sie sorgfältig prüfen können.

Bei den Progressen ist es, denke ich, nicht ganz so trivial. Denn was die Wirksamkeit angeht, waren die Progresse über einen Cut-off-Wert beim Verlust an der forcierten Vitalkapazität definiert. Das ist nicht notwendigerweise identisch mit der Angabe des Patienten, dass es ihm schlecht gehe, die dann den Prüfarzt dazu bewogen hat, den Progress als unerwünschtes Ereignis anzugeben. Dies richtete sich eher nach der klinischen Einschätzung der Situation – der Patient erlebt eine deutliche Verschlechterung seines Krankheitsbildes – und nicht am Erreichen eines Zahlenwertes bei der Lungenfunktionsprüfung. Von daher handelt es sich unserer Auffassung nach nicht um identische Ereignisse.

Es mag sein, dass es eine gewisse Koinzidenz gibt. Das im Einzelfall herauszurechnen ist auf jeden Fall nicht trivial.

Herr Dr. Vöcking (stellv. Vorsitzender): Herr Eyding.

Herr Dr. Eyding: Zum Verständnis eine Nachfrage. War der Progress auch allein als FVC-Abnahme definiert? Wenn die FVC-Abnahme über einer bestimmten Schwelle lag, war das dann auch ein Progress?

(Herr Kögler (Boehringer Ingelheim Pharma) nickt)

– Genau. Das heißt, Teile der Progresse haben Sie, so sage ich einmal, als patientenberichtete UE, und einen Teil der Progresse haben Sie als FVC-Werte, sozusagen als Laborwerte.

Aber aus den UE-Daten insgesamt, inklusive der SUE-Daten, kann man natürlich die Teile, die zum Progress geführt haben und die über das, was man bei den Exazerbationen hat, hinausgehen, auch herausrechnen. Das ist möglicherweise nicht trivial. Man kann die Progresse nicht einfach von den UE abziehen. Aber man kann natürlich die Fälle, die aufgrund von UE als Progress definiert worden sind, von den UE-Daten abziehen. Das geht ja. Oder verstehe ich das falsch?

Herr Dr. Vöcking (stellv. Vorsitzender): Herr Kögler, antworten Sie? Sie haben ja auch einen Statistiker dabei. – Herr Kögler.

Herr Kögler (Boehringer Ingelheim Pharma): Ein Progress wurde, wie ich sagte, dann als unerwünschtes Ereignis angegeben, wenn er an der Schwere der Beeinträchtigung festgemacht war. Ich bin mir nicht sicher, dass die Daten gesondert vorliegen. Ereignisse bei Patienten, die im Rahmen der Efficacy-Analyse als 10-prozentiger Abfall definiert waren – wir hatten als zweiten Schwellenwert auch noch einen 5-prozentigen Abfall genommen; 10 Prozent ist der anerkannt robustere –, werden zu einem gewissen Anteil möglicherweise auch als UE gemeldet worden sein. Da müssten wir wirklich noch einmal sehr tief graben, um sagen zu können, ob man diese einzelnen Fälle genau auseinandersortieren kann.

Herr Dr. Vöcking (stellv. Vorsitzender): Herr Eyding hat noch eine weitere Nachfrage oder Bemerkung.

Herr Dr. Eyding: Sie müssen einfach alle Progresse durchgehen und schauen, aufgrund welches Kriteriums sie als Progress gewertet worden sind. Diejenigen, die als UE bewertet worden sind, muss man abziehen. Natürlich muss man sich die Daten noch einmal anschauen. Aber prinzipiell ist das möglich, und es ist auch keine hohe Kunst, es zu tun. Aber es ist ein gewisser Arbeitsaufwand, das ist wahr.

Herr Dr. Vöcking (stellv. Vorsitzender): Herr Pfannkuche, ich nehme an, Sie wollen sich dazu äußern, was man nachliefert oder nicht nachliefert.

Herr Dr. Pfannkuche (Boehringer Ingelheim Pharma): Die SUE-Analyse ohne Exazerbation werden wir auf jeden Fall zeitnah in den nächsten Tagen nachreichen. Was die zweite angeht, werden wir prüfen, was wir aus den Daten herausholen können. Wir werden es uns noch einmal genau anschauen, Herr Eyding, und uns dann bei Ihnen melden.

Herr Dr. Vöcking (stellv. Vorsitzender): Schönen Dank, Herr Pfannkuche. Meine Bitte wäre, es zügig zu machen. – Herr Eyding, Sie wollen noch einmal nachhaken?

Herr Dr. Eyding: Es ist jetzt ein anderer Punkt. Darf ich ihn jetzt bringen?

Herr Dr. Vöcking (stellv. Vorsitzender): Machen Sie. Ansonsten haben wir noch andere, die Fragen stellen wollen. Aber sprechen Sie den Punkt an.

Herr Dr. Eyding: Ich habe noch einige Fragen. – Der Mortalitätsunterschied, den Sie schon angesprochen haben, ist nicht signifikant. Sie machen jetzt einen großen Claim auf, dass die Exazerbationen dazu führen, dass es da möglicherweise einen Mortalitätsvorteil geben könnte.

Ich habe die generelle Frage, warum Sie diese Studie nur auf 52 Wochen angelegt haben. Aus meiner Sicht ist das nicht adäquat, weil die Patienten keine sehr lange Lebenserwartung mehr haben. Eine solche Studiensituation ist der Idealfall, weil Sie so auf Overall Survival powern können. Wenn Sie die Studie etwas länger angelegt hätten, hätten Sie erheblich mehr Ereignisse gehabt als die 5 bis 7 Prozent, die Sie jetzt in dieser Studie hatten, und hätten ganz zweifelsfrei zeigen können, ob es einen Überlebensvorteil gibt oder nicht. Ich frage mich wirklich, warum Sie das nicht getan haben, und finde es völlig inadäquat. Andere Hersteller haben das gemacht. Pirfenidon-Studien sind immerhin über 72 Wochen gelaufen. Das hätte uns auch die Diskussion über die Surrogatvalidierung der FVC-Abnahme erspart, und wir hätten deutlich bessere Daten gehabt. Aus meiner Sicht war es, wie gesagt, eine ideale Situation, um Overall Survival als Endpunkt zu wählen, ganz anders als in einer onkologischen Situation im frühen Stadium mit Fünf-Jahres-Überlebensraten von 90 Prozent, wo das wirklich schwierig ist und wo man sich darum kümmern muss, Surrogatparameter festzulegen, oder das zumindest ein sinnvolles Unterfangen sein könnte. Hier sehe ich das überhaupt nicht ein. Das bitte ich Sie zu kommentieren.

Herr Dr. Vöcking (stellv. Vorsitzender): Herr Kögler erhält das Wort für den Versuch einer Erklärung. Ich hatte das Thema 52 Wochen ja eben schon angesprochen.

Herr Kögler (Boehringer Ingelheim Pharma): Natürlich wäre es prinzipiell auch möglich gewesen, die Studie mit einer längeren Dauer durchzuführen. Das steht außer Frage. Aber wenn man die Studie als confirmatorische Studie tatsächlich auf den Endpunkt Mortalität hätte ausrichten wollen, wäre man auch mit 72 Wochen nicht hingekommen. Die Studien des anderen Herstellers über 72 Wochen, die Sie ansprechen, haben für sich genommen auch nicht den Beleg eines Effekts auf die Mortalität erbringen können. Hierfür hätten wir die Studien entweder von der Patientenzahl her erheblich vergrößern müssen – ich bezweifle, dass das bei einer seltenen Erkrankung wie der idiopathischen Lungenfibrose ohne Weiteres möglich ist –, oder man hätte die Studie Event-getrieben deutlich verlängern müssen. Das hätte aber bedeutet, Patienten, für die es wenig Behandlungsalternativen gibt, eine potenziell relevante Innovation, die einen bedeutsamen Patientennutzen hat, noch eine unangemessen lange Zeit vorzuenthalten. Deswegen war die Entscheidung, die letztlich auch in Abstimmung mit den Zulassungsbehörden getroffen worden ist, die Studiendauer auf 52 Wochen zu beschränken.

Herr Dr. Vöcking (stellv. Vorsitzender): Herr Eyding.

Herr Dr. Eyding: Das verstehe ich ja alles. So wissen wir aber leider sehr wenig über den Effekt von Nintedanib, weil die Datenlage nicht wirklich optimal ist. Gerade das, was am relevantesten wäre, wissen wir eben überhaupt nicht. Das ist eben so.

Ich würde das Argument eher umgekehrt sehen: Wenn man eine definitive Situation erzeugen kann, nach deren Durchführung man weiß, ob eine Substanz wirkt oder nicht, dann ist das deutlich angemessener, als hier auf Grundlage einer relativ unsicheren Datenlage eine Entscheidung fällen zu müssen. Dabei geht es nicht darum, jemandem eine Innovation vorzuenthalten, bei der wir nur speku-

lieren können, ob sie so oder so aussieht. Wir haben hier ein paar Effekte, aber man hätte deutlich definitivere Daten erzeugen können, wenn man die Studie nur ein Jahr länger durchgeführt hätte. Zumindest nach der Epidemiologie hätte es erheblich mehr Ereignisse nach sich ziehen müssen, wenn die mediane Lebenserwartung 2 bis 4 Jahre beträgt und Patienten 1,6 Jahre nach Diagnose in die Studie eingeschlossen worden sind. Mir ist es völlig unverständlich, dass Sie in der Studie 5 bis 7 Prozent an Ereignissen haben. Es ist auch die Frage, wie das Patientengut selektiert worden ist und ob es besondere Patienten sind. Denn eigentlich würde man auch schon nach einem Jahr mehr Ereignisse erwarten. Das ist eine Frage nach der externen Validität der Studie.

Insgesamt ist die Situation aus meiner Sicht wirklich unbefriedigend.

Herr Dr. Vöcking (stellv. Vorsitzender): Herr Eyding, das ist eine andere Ebene. Wir haben jetzt gehört, welche Position der Unternehmer einnimmt. Die Frage der Bewertung kann vielleicht für die weitere Sitzung zurückgestellt werden. – Herr Professor Costabel hat sich gemeldet.

Herr Prof. Dr. Costabel (DGP): Ich möchte zu Ihrem Argument direkt Stellung beziehen und noch einmal klarstellen, dass die Mortalität natürlich der ideale Endpunkt wäre. Mortalitätsstudien sind aber nach Auffassung der Experten auf diesem Gebiet nicht wirklich durchführbar.

Sie haben erwähnt, dass man die Studie länger hätte laufen lassen können, über zwei Jahre hinweg. Das Problem ist, dass dann Patienten im Placeboarm zwei Jahre lang keine wirksame Therapie bekommen hätten. Der Drop-out wäre groß gewesen. Die Patienten wären wahrscheinlich auch früher aus der Studie ausgeschieden, oft auf eigenen Wunsch, weil sie gemerkt hätten, dass es immer schlechter wird. Dann hätten sie wieder verzweifelt nach anderen Therapien gesucht. Das ist der eine Punkt.

Die niedrige Mortalität erklärt sich aus dem eingeschlossenen Patientengut. Das sind Patienten mit früher bis moderater IPF. Wir wissen nicht nur aus diesen Studien, sondern auch aus Studien, die früher durchgeführt worden sind, etwa die Studien mit Pirfenidon, aber auch Studien, die negativ waren, dass die Mortalität bei diesen frühen Patienten in der Tat niedriger ist, als wir alle gedacht haben. Sie liegt eben nur bei 6 bis 8 Prozent pro Jahr. So müsste man, wenn man über 52 Wochen behandeln würde, bei dieser Mortalitätslage über 2.000 Patienten einschließen. Das würde enorme Kosten verursachen, sodass wir nach einem Surrogatmarker für die Mortalität gesucht und ihn in der FVC-Abnahme gefunden haben. Das hat sich ja auch durch die neue FDA-Publikation und die weitere Analyse der Firma Boehringer Ingelheim bestätigt.

Auf den Punkt der FVC werden wir vielleicht noch zu sprechen kommen. Ich möchte nur erwähnen, dass auch aus Expertensicht für alle IPF-Studien aktuell anerkannt ist, dass die FVC-Abnahme der geeignetste Surrogatmarker für die Mortalität ist, dass Mortalitätsstudien mit angemessenem Aufwand nicht durchführbar sind, sodass wir das nur dann erreichen könnten, wenn wir zum Beispiel Patienten in schweren Krankheitsstadien einschließen würden. Dann wären aber auch die Drop-outs höher. Man hätte mehr Komorbiditäten. Das wäre auch ein Problem.

Herr Dr. Vöcking (stellv. Vorsitzender): Schönen Dank, Herr Professor Costabel. – Ich habe eine Frage. Wenn man von vornherein versucht hätte, die Studiendauer länger auszulegen, würde das gerade wegen des Placeboarms überhaupt von der Ethikkommission gebilligt?

Herr Prof. Dr. Costabel (DGP): Ich könnte mir vorstellen, dass die Ethikkommission große Einwände hätte. In Zukunft sind solche Studien auch überhaupt nicht mehr durchführbar, weil wir jetzt ja zugelassene Medikamente haben. Es wird also für die Zukunft deutlich schwieriger, den Wirkungsnach-

weis zu führen. Bei zukünftigen Studien müssen wir wahrscheinlich auf schwerere Krankheitskollektive ausweichen, damit wir einfach höhere Event Rates haben.

Herr Dr. Vöcking (stellv. Vorsitzender): Schönen Dank. – Frau Geier, Sie hatten sich gemeldet.

Frau Geier (Boehringer Ingelheim Pharma): Schon mehrfach ist das Wort FVC gefallen. Sofern Sie es mir gestatten, würde ich gern noch ein paar einleitende Worte dazu fallen lassen, bevor wir in die Diskussion zu diesem Endpunkt starten.

Herr Dr. Vöcking (stellv. Vorsitzender): Frau Geier, können wir das Thema vielleicht noch etwas nach hinten schieben? Denn ich habe verschiedene Wortmeldungen. Als erste würde ich meine dazwischenschieben.

Herr Eyding hat eben das Thema UE erwähnt. Professor Ruof, ich habe an Sie eine Frage. Sie haben, wenn ich es richtig gelesen habe, in Ihrer Stellungnahme Tabelle 21 angesprochen – dabei geht es auch um die UE – und darauf hingewiesen, das sollte um eine Darstellung der Patienten und Behandlungsrelevanz ergänzt werden. Wie soll ich das verstehen? Oder habe ich das falsch gelesen?

Herr Prof. Dr. Ruof (Roche Pharma): Ich schaue es noch einmal kurz nach und gebe Ihnen dann gleich eine Rückmeldung.

Herr Dr. Vöcking (stellv. Vorsitzender): Dann nutze ich die Zeit. Es haben sich gemeldet Herr Kulig und dann Frau Müller. Das Thema FVC schieben wir etwas nach hinten. – Herr Kulig.

Herr Dr. Kulig: Das finde ich sehr gut. Denn ich habe noch zu den zwei vorgenannten Punkten Fragen. Zunächst will ich auch noch einmal zum Thema Mortalitätsendpunkt nachfragen. Einschlusskriterium war eine FVC von über 50 Prozent. Aber im Durchschnitt hatten sie noch 80 Prozent. Ich habe mich auch gefragt, warum relativ Gesunde mit einer etwas besseren Prognose in die Studie eingeschlossen wurden. Denn natürlich hätte auch ich gern belastbare Daten gesehen. Noch einmal: Warum nicht Kränkere mit unter 50 Prozent? Das eben gefallene Argument von Komorbiditäten ist mir nicht so einsichtig. Natürlich haben Patienten in anderen Studien auch Komorbiditäten. Aber man hätte einfach ein realistischeres Krankengut gehabt. Das ist die erste Frage.

Herr Kögler (Boehringer Ingelheim Pharma): Theoretisch wäre es vorstellbar gewesen, auch Patienten mit einer weiter fortgeschrittenen IPF, beispielsweise mit einem FVC-Schweregrad unter 50 Prozent, einzuschließen. Diese Patienten sind aber häufig nicht wirklich klinisch stabil.

Bei einer seltenen Erkrankung wie der IPF unternehmen Patienten zum Teil große Anstrengungen, nehmen auch Strecken auf sich, um sich an einem anerkannten Zentrum mit entsprechender Expertise bei interstitiellen Lungenerkrankungen in Ermangelung anderer Alternativen in einer Studie behandelt wiederzufinden. Dabei ist es nicht unüblich, dass 100 bis 150 km gefahren werden. Für Patienten, die durch ihre Grunderkrankung schon eine deutliche Einschränkung haben – Lungenfunktion 50 Prozent unterhalb vom Soll –, ist es unter Umständen schwierig, das Prüfzentrum zu den vorgegebenen Visiten zu erreichen, besonders dann, wenn etwas Unvorhergesehenes passiert. Solche Aspekte sind mit den Zulassungsbehörden diskutiert worden. Letztlich hat man sich darauf verständigt, eine untere Grenze von 50 Prozent für den Einschluss der Patienten zu setzen.

Dass am Ende die FVC im Mittel bei etwa 80 Prozent vom Soll liegt, ist eine Art Selbstselektion gewesen. Das ist nicht beeinflusst worden, indem man zum Beispiel stratifiziert in bestimmte Kategorien des FVC-Zustands bei Baseline eingeschlossen hätte. Das hat sich spontan ergeben. Als Hersteller

hat man, wenn nicht entsprechend präspezifiziert worden ist, dass zu stratifizieren ist, keine Möglichkeit mehr, da einzugreifen.

Herr Dr. Vöcking (stellv. Vorsitzender): Schönen Dank. – Herr Kulig, eine Nachfrage?

Herr Dr. Kulig: Eine Nachfrage und auch eine zweite Frage bzw. ein Kurzkomentar. Ich finde es fast ein bisschen schade, dass man die kränkeren Patienten nicht einschließen konnte, weil es gerade dann, wenn, wie Sie vorhin ja sagten, ein Unmet Medical Need besteht, ja ganz nett gewesen wäre es, wenn man für diese Patienten etwas hätte sagen können.

Meine Frage bezieht sich aber noch einmal auf die Exazerbationen. Wir fanden es sehr schwierig, deren Definition nachzuvollziehen bzw. deren Operationalisierung zu beurteilen. Man sieht das beispielsweise an den Häufigkeiten. Zunächst, als sie vom Prüfarzt festgestellt wurden, lag deren Zahl noch wesentlich höher als später, als sie von dem Expertenkomitee überprüft worden sind. Um welche Experten handelte es sich? Wie und nach welchen Kriterien war das Komitee zusammengesetzt? Was wurde ihm vorgelegt? Es hat ja nicht mehr alle Exazerbationen überprüft, sondern nach unserem Verständnis nur in die Unterlagen geschaut und versucht, die Fälle herauszufiltern, bei denen es sich nicht um Exazerbationen handelte. Von daher halten wir nach diesen ungenauen Angaben auch die Validität für schwer nachvollziehbar, obwohl die Idee, es von einem Expertenkomitee überprüfen zu lassen, vielleicht gut ist. Das fanden wir jedenfalls nicht ganz nachvollziehbar. Könnten Sie das bitte noch einmal erklären?

Herr Dr. Vöcking (stellv. Vorsitzender): Die Frage steht im Raum. – Herr Kögler.

Herr Kögler (Boehringer Ingelheim Pharma): Ich werde schon langsam heiser.

Herr Dr. Vöcking (stellv. Vorsitzender): Es gibt garantiert von irgendeinem Konkurrenten ein Präparat dagegen.

(Heiterkeit)

Herr Kögler (Boehringer Ingelheim Pharma): Leider liegt das aktuell nicht vor. – Sie sprechen zwei Aspekte an. Sie sprechen zum einen den Unterschied in den Zahlen der Events zwischen den prüfarztberichteten Exazerbationen einerseits und den adjudizierten andererseits an. Zum anderen ist der Adjudizierungsprozess als solcher hinterfragt worden.

Den Adjudizierungsprozess kann ich gern kurz erläutern. Die drei Experten, die ausgewählt worden sind, unterscheiden sich an Expertise wahrscheinlich nicht nennenswert von den erfahrenen Prüfern, die ohnehin in der Studie waren. Das Entscheidende war also nicht eine größere Expertise des Adjudizierungskomitees, sondern die Vollständigkeit der Unterlagen, die diesem zur Verfügung stand, um die Ereignisse retrospektiv umfassend zu bewerten.

Ich hatte es angesprochen: Die Patienten werden häufig in einem Prüfzentrum betreut, das nicht nahe bei ihrem Wohnort liegt. Wenn ein akutes Ereignis auftritt, setzen sie sich nicht ins Auto und fahren 150 km zum Prüfzentrum, sondern sie werden im nächsten akutversorgenden Krankenhaus an ihrem Wohnort vorstellig. Dort erfolgt dann eine Akutversorgung, in der Regel nicht durch Spezialisten in interstitiellen Lungenerkrankungen, sodass hier unter Umständen nicht alle Untersuchungen durchgeführt werden, die man bräuchte, um den Ausschluss anderer vorstellbarer Ursachen wirklich umfassend vorzunehmen. Die Exazerbation ist ja – ich habe es eingangs in meinem Eröffnungsstatement gesagt – definiert als eine erhebliche Verschlechterung, für die keine andere Ursache als die zugrundeliegende IPF identifiziert werden kann.

Konkret lief die Adjudizierung so ab, dass alle verfügbaren Unterlagen zu allen Exazerbationen, die gemeldet wurden – das möchte ich klarstellen, es wurden wirklich alle Exazerbationen adjudiziert –, dem Adjudication Committee vorgelegt wurden. Zunächst erhielt jeder der Experten einzeln und unabhängig voneinander diese Unterlagen und hat sich individuell seine eigene Meinung gebildet. Dazu gehörten die Entlassberichte der behandelnden Krankenhäuser, die dokumentierten Befunde, beispielsweise hochauflösende Computertomografie, Laborbefunde, Befunde zur Infektiologie und aus der bronchoalveolären Lavage, soweit vorhanden. Dann trafen sich die Kollegen zum Adjudication Meeting und haben ihre Einschätzungen jedes einzelnen Ereignisses zusammengetragen.

Es gab zu einem hohen Anteil einen Konsens darüber, ob es sich um eine definitive oder vermutete Exazerbation oder um keine Exazerbation handelte. Falls abweichende Meinungen auftraten, wurde in der Diskussion versucht, einen Konsens herbeizuführen. Nur wenn das nicht gelang, wurde die Adjudizierung nach dem einfachen Mehrheitsprinzip – bei drei adjudizierenden Spezialisten also mit zwei zu eins – vorgenommen.

Eine Bewertung, dass es sich um keine Exazerbation handelte, erfolgte dann, wenn aus den Unterlagen klar ersichtlich war, dass es eine greifbare andere Ursache gab. Das kann eine Pneumonie mit einem entsprechenden infektiologischen Befund gewesen sein oder vielleicht auch ein Ansprechen auf eine Antibiotikagabe; das kann ein Nachweis für eine Lungenembolie gewesen sein oder ein Linksherzversagen mit einer Lungenstauung. Alle diese Ereignisse können klinisch eine IPF-Exazerbation imitieren. Bei Vorliegen eines solchen beweisenden Befundes für eine andere Ursache wurde festgelegt, dass es sich um keine Exazerbation gehandelt hat.

Bei all den Exazerbationen, bei denen das nicht der Fall war, wurde zumindest eine wahrscheinliche Exazerbation angenommen. In den Fällen, in denen vollständige Unterlagen vorlagen, aus denen man alles ersehen konnte, hat man gesagt, es ist eine definitive Exazerbation. In den anderen Fällen war es eine wahrscheinliche Exazerbation.

Das führte in der Tat zu einer Reduktion der Event-Raten. Wir haben in beiden Studienarmen – so möchte ich es einmal nennen – ein vergleichbares Grundrauschen an Ereignissen, die klinisch bedeutsam waren – ohne Frage, sonst wären sie ja nicht auch als schwerwiegende unerwünschte Ereignisse berichtet worden –, die aber letztlich nicht als Exazerbationen bestätigt wurden. Die primäre Analyse war die geplante auf der prüfarztberichteten Ebene. Das entspricht dem Real-Life-Aspekt; denn in der klinischen Routine wird eine Adjudizierung im Normalfall nicht zur Verfügung stehen. Als Sensitivitätsanalyse war dann die Adjudizierung vorgesehen, um gezielt nur diejenigen Ereignisse zu betrachten, von denen man ausgehen konnte, dass sie durch eine spezifische antifibrotische Therapie, wie das mit Ofev[®] der Fall ist, auch beeinflussbar sind, während das bei den anderen Ereignissen, für die sich greifbare andere Ursachen fanden, einfach nicht anzunehmen ist. Und so kamen die beiden Analysen zustande.

Herr Dr. Vöcking (stellv. Vorsitzender): Schönen Dank, Herr Kögler, für die Erläuterung. – Herr Eyding hat eine Frage direkt dazu. Frau Müller, da muss ich Sie um Nachsicht bitten, dass Herr Eyding dazwischen kommt. – Herr Eyding.

Herr Dr. Eyding: Meine Nachfrage bezieht sich noch einmal darauf, welche Daten letztlich diesem Adjudizierungscommittee zur Verfügung standen. Ich fand es auffällig, dass gerade in beiden Studien jeweils im Behandlungsarm deutlich mehr Fälle durch das Adjudizierungscommittee verschwunden sind als im Placeboarm. Ich gehe einmal davon aus, dass das Adjudizierungscommittee alle Adjudizierungen vorgenommen hat, bevor die Studie entblindet worden ist. Das ist ja wahrscheinlich lege artis so gemacht worden. Entscheidend ist dann noch die Frage: Kann an anderer Stelle ein Bias dadurch ein-

geführt worden sein, dass zum Beispiel UE-Daten in großem Umfang dem Adjudizierungskomitee zur Verfügung standen, weil wir ja durchaus typische Nebenwirkungen der Behandlung haben? Es wäre also einfach noch einmal die Frage zu klären, welche Unterlagen genau dem Adjudizierungskomitee während der Studie zur Verfügung gestellt worden sind.

Herr Dr. Vöcking (stellv. Vorsitzender): Jetzt wird Herr Kögler durch Herrn Trampisch, Mathematiker und Statistiker, abgelöst. – Herr Trampisch, Sie haben das Wort.

Herr Dr. Trampisch (Boehringer Ingelheim Pharma): Dem Adjudizierungskomitee lagen alle Unterlagen bezüglich UEs und Vitalparameter vor. Die Ausführung erfolgte natürlich verblindet vor Datenbankabschluss. Zudem wurde von Suissa letztes Jahr eine Analyse bezüglich der Verbesserung des Exazerbationssignals durchgeführt. Dort wurde festgestellt, dass in beiden Populationen, Ofev[®] und Placebo, etwa 5 Prozent der Gesamtpopulation exazerbationsähnliche Ereignisse erlebt haben. Die höhere Anzahl von den daraus resultierenden, nicht als adjudizierte Exazerbationen identifizierten Ereignissen bezieht sich natürlich auch darauf, dass wir eine 3 : 2-Randomisierung durchgeführt hatten. Insofern ist die absolute Anzahl von entfernten Ereignissen auch höher. Und das ist der Grund, warum es augenscheinlich aussieht, dass mehr Exazerbationen aus dem aktiven Arm entfernt worden sind. Dies ist aber, berechnet auf die Gesamtpopulation, nicht der Fall.

Herr Dr. Vöcking (stellv. Vorsitzender): Vielen Dank für die Antwort. – Frau Müller, Sie haben das Wort. Sie hatten sich ja schon vor einiger Zeit gemeldet. Ich bitte um Nachsicht.

Frau Dr. Müller: Eine Frage wurde schon beantwortet, die sich eben in der Diskussion gestellt hat, nämlich warum diese konkurrierenden Ereignisse in der Vergleichsgruppe scheinbar häufiger aufgetreten sind. Wir haben ja tatsächlich den Fall, dass wir bei den nicht adjudizierten Exazerbationen nur in einer Studie einen Vorteil zeigen konnten. Nachdem dann aber die Experten ihre Beurteilung gemacht haben, haben wir auch in der gepoolten Analyse einen Vorteil. Insofern ist das natürlich ein sehr relevanter Punkt.

Ich würde gerne noch zwei Punkte diskutieren. Bisher noch gar nicht angesprochen wurde die Frage der Divergenzen bei den Ergebnisse in den beiden Studien. Das betrifft ja nicht nur die Zeit bis zur akuten Exazerbation, sondern auch die Lebensqualitätsdaten, die jeweils nur in der INPULSIS-2-Studie einen Vorteil zeigten. Das wurde ja auch von der EMA diskutiert. Sie haben einen Teil der Begründung der EMA auch in Ihrer Stellungnahme zitiert und gesagt, das gehe sozusagen in die gleiche Richtung, auch wenn es einmal signifikant wurde und einmal nicht. Aber die EMA hat ja auch darüber hinaus, soweit ich das gesehen habe, diskutiert, ob sich möglicherweise dieser positive Effekt auf die Lungenfunktion, der ja auch zu einer Verbesserung der Lebensqualität führen sollte und scheinbar auch führt, im Laufe der Behandlung abschwächt und dann möglicherweise bezüglich der Lebensqualität von den Nebenwirkungen überlagert wird, wo wir ja nun klare Nachteile sehen. Wenn Sie vielleicht noch etwas sagen könnten, wie Sie das sehen oder wie Sie sich diese Divergenzen erklären. Ich denke, das ist ein wichtiger Punkt.

Danach habe ich noch eine Frage zur Validierung der forcierten Vitalkapazität für Overall Survival, was ich auch ganz wichtig finde.

Herr Dr. Trampisch (Boehringer Ingelheim Pharma): Ich werde jetzt erst einmal auf Ihre Fragen bezüglich der von Ihnen als divergent dargestellten Ergebnisse in den Einzelstudien eingehen. Bei den Phase-III-Studien handelt es sich um sogenannte Zwillingsstudien. Insofern waren die Studienprotokolle identisch. Die Durchführung war doppelblind, randomisiert, multizentrisch, multinational. In-

sofern kann man davon ausgehen, dass es sich bei den eingeschlossenen Patientenkollektiven um ein mehr oder weniger identisches Kollektiv mit ähnlichen Erkrankungsschweren handelt. Dies ist auch aus den Baseline-Daten, den demografischen Tabellen und auch den Daten zu Vorerkrankungen ersichtlich. Insofern sind die von Ihnen angesprochenen, post hoc identifizierten Divergenzen in den Studienergebnissen, welche weitestgehend allesamt einen positiven Trend für Ofev[®] aufweisen, der Variabilität des beobachteten Effektparameters zuzuschreiben und nicht darauf zurückzuführen, dass die Studien inhomogene Ergebnisse liefern.

Herr Dr. Vöcking (stellv. Vorsitzender): Frau Müller.

Frau Dr. Müller: Sie sehen das also als zufällige Verteilung, zufällige Imbalance, die sich dann so auf die Studie verteilt, was ja sozusagen dann auch die Effektparameter doch etwas relativieren würde? – Ich will nur noch einmal darauf hinweisen, dass es auch bezüglich der Therapieabbrüche wegen UEs in der gleichen Richtung eine Verzerrung gab.

Herr Dr. Trampisch (Boehringer Ingelheim Pharma): Die Variabilität des beobachteten Parameters ist die in der Realität vorhandene Variabilität. Von daher geben wir Punktschätzer in den Studien ab mit Konfidenzintervallen. Die Konfidenzintervalle waren in den beiden Studien allesamt überlappend. Insofern wird dort die vorhandene Variabilität des Parameters abgebildet. Beispielhaft kann man es, falls ich noch kurz ein Beispiel erwähnen darf, bei Wachstumskurven von kleinen Kindern sehen. Dort wird typischerweise nicht der Mittelwert mit der vorhandenen Größe des Kindes verglichen, sondern es werden Quantile verglichen, weil man weiß, dass die Wachstumsrate bei Kindern unterschiedlich ausfällt, dass es unterschiedliche Schübe gibt. Insofern gibt es auch andere Beispiele, wo die Punktschätzer gegebenenfalls nicht alle identisch sind. Man wird aber nicht davon ausgehen, dass bei Zwillingstudien diese Unterschiede aufgrund von inhomogenen Patientenkollektiven zustande gekommen sind.

Herr Dr. Vöcking (stellv. Vorsitzender): Ist die zweite Frage auch schon beantwortet worden?

Frau Dr. Müller: Vielleicht noch eine kleine Nachbemerkung. Also die EMA stellt an irgendeiner Stelle die Frage, welche dieser beiden Studien eigentlich die Realität wiedergibt. Das ist ganz interessant. Möchten Sie zu diesem Kommentar etwas sagen? Sonst würde ich zu meinem zweiten Punkt kommen.

Herr Dr. Trampisch (Boehringer Ingelheim Pharma): Dazu ist die Position von Boehringer Ingelheim: Die gepoolten Ergebnisse geben den besten Schätzer und den präzisesten Schätzer ab.

Herr Dr. Vöcking (stellv. Vorsitzender): Schönen Dank. – Weitere Fragen, Frau Müller?

Frau Müller: Ja. – Die zweite Frage ist – ich denke, das ist ein ganz wichtiger Punkt, der auch von allen Stellungnehmern hier thematisiert wurde –, inwiefern die forcierte Vitalkapazität ein valides Surrogat für Overall Survival darstellt. Da hatten Sie ja auf die FDA-Bewertung von Karimi-Shah verwiesen, die Sie jetzt nachgereicht haben, was auch nachvollziehbar ist, weil sie so ungefähr zehn Tage nach Ihrem Dossier oder noch etwas später veröffentlicht wurde. Da ist meine Frage, wie Sie diese Analyse aus methodischer Sicht sehen. Sie haben sich ja Einzelstudien sowohl zu Pirfenidon als auch zu Nintedanib angesehen, haben gesagt, Sie können erst jetzt, weil erst jetzt genug Studien vorliegen, eine Validierung vornehmen. Sie haben ja auch auf Kritik reagiert bezüglich dieses Punktes. Von diesen sechs Studien zeigen ja fünf einen Vorteil bezüglich der forcierten Vitalkapazität. Aber bezüglich Overall Survival gibt es in diesen fünf Studien jeweils nur einen Trend. Es konnte ja in keiner einzigen

Studie ein signifikanter Vorteil für Overall Survival gezeigt werden. Meine Frage – ich bin keine Methodikerin – ist jetzt: Wie kann man sozusagen aus fünf Studien, die alle keinen signifikanten Overall-Survival-Vorteil zeigen, eine Validierung herleiten? Das ist wirklich einfach eine Verständnisfrage.

Herr Dr. Trampisch (Boehringer Ingelheim Pharma): Ich möchte zuerst einmal erwähnen, dass wir nicht den Anspruch erheben, diesen Surrogatendpunkt validiert zu haben. Die von uns grafisch aufbereitete Darstellung, die bezüglich der zulassungsrelevanten Studien für die Behandlung der IPF-Erkrankung der FDA vorgelegt worden ist, liefert weitere Evidenz. Hier sollte visuell in einer grafischen Darstellung der Zusammenhang noch einmal präsent gemacht werden.

Sie haben auch angesprochen, dass die Behandlungsdifferenz in einer Studie im Ergebnis keinen Einfluss auf FVC zeigt. Diese Studie – es handelt sich hierbei um eine Pirfenidon-Studie, die CAPACITY-1-Studie – hatte auch keinen Trend bezüglich Overall Survival. Das Schöne an dieser Darstellung, die Sie angesprochen haben, ist, dass wir keine weiteren Extrapolationen durchführen müssen, sondern dass wir hier Evidenz dafür haben, dass ein Medikament, wenn es nicht auf die forcierte Vitalkapazität wirkt, offensichtlich oder zumindest innerhalb von diesem Setting in dem eingeschlossenen Kollektiv auch kein Overall-Survival-Benefit für dieses Kollektiv gab. In den weiteren Studien, in denen ein FVC-Unterschied in den Behandlungsgruppen beobachtet worden ist, hat man überall einen Trend gesehen, der zugunsten des Präparates war. Diesen Zusammenhang wollten wir mit dieser Grafik noch einmal untermauern.

Herr Dr. Vöcking (stellv. Vorsitzender): Herr Professor Costabel, Sie hatten sich gemeldet. Vielleicht könnten Sie bitte gleich eine Frage von mir – ich bin von Hause aus kein Mediziner, sondern Jurist – mit beantworten, und zwar: Wie muss ich die Relevanz einer Verbesserung von 110 ml verstehen? Ist das viel? Ist das wenig?

Herr Prof. Dr. Costabel (DGP): Zunächst zu Ihrem Argument, dass alle fünf Studien nur einen Trend zeigen würden und keine dieser fünf Studien signifikant den Mortalitätseffekt zeigte. Das liegt daran – das haben wir vorher diskutiert –, dass alle diese Studien von der Fallzahl her nicht genügend gepowert waren, um einen signifikanten Effekt zu zeigen. Wenn die Patientenzahl größer gewesen wäre, hätte man vielleicht einen signifikanten Effekt zeigen können. Das ist irgendwo systemimmanent. Bei der sich durch die jetzige Studiengröße ergebenden Patientenzahl können wir nicht erwarten, dass die Mortalität signifikant ausfällt, warum auch von den Zulassungsbehörden FDA und EMA immer ausdrücklich gesagt wird: Wenn ihr einen Effekt auf die FVC zeigt, auf den primären Endpunkt, der parallel einen Trend in der Mortalität zugunsten der Substanz zeigt, reicht das völlig aus. – Das dazu.

Nun zu Ihrer Frage, Herr Vöcking, was 100 ml ausmachen. Ich würde das eher so interpretieren, dass 100 ml 50 Prozent des zu erwartenden Abfalls bei der IPF über ein Jahr sind. Also, 200 ml würde man erwarten. Das zeigt sich auch konsistent in vielen Studien, und es sind 100 ml gewesen, also 50 Prozent weniger. Das spürt der Patient in diesen frühen Stadien nicht. Das ist genau das, was immer so die Paradoxie ist: 100 ml bedeuten für den Patienten in diesen frühen Stadien noch nichts. Da sieht man auch wieder eine Analogie zu dem Lebensqualitätssignal im SGRQ. Man sieht dort, wenn man Patienten anhand einer Vitalkapazität von über oder von unter 70 Prozent aufteilt und eine Subgruppenanalyse macht, einen deutlich stärkeren Effekt auf die Lebensqualität in der schwerer erkrankten Gruppe. Das bedeutet, ein Patient, der schon in dem Bereich von 50 bis 60 Prozent Vitalkapazität ist und sich dann um 100 ml oder noch mehr – beispielsweise um 200 ml – verschlechtert, spürt das dann auch in seiner Lebensqualität.

Herr Dr. Vöcking (stellv. Vorsitzender): Herr Professor Behr.

Herr Prof. Dr. Behr (DGP): Ich möchte es noch dahin gehend ergänzen: 100 ml erscheinen zunächst vielleicht einmal wenig. Man muss wissen, dass der gesunde Mensch ab dem 25. Lebensjahr auch etwa 20 bis 30 ml pro Jahr an Vitalkapazität verliert. Dann ist es immerhin schon so das Drei- bis Fünffache von dem, was normal ist, je nachdem, wie man es rechnet. Das ist natürlich mit weiteren Konsequenzen für das Organ verbunden. Es ist ja nicht so, dass das Organ nur die 100 ml verliert und der Rest gleich bleibt. Vielmehr ist das ein sich in dem Organ quasi diffus ausbreitender Prozess, der dazu führt, dass es nicht nur sozusagen schrumpft und damit auch weniger Volumen hineingeht, sondern es auch steifer wird und auch der Gasaustausch natürlich schlechter wird. Dementsprechend wird die Funktion des Gesamtorgans dadurch wesentlich stärker beeinträchtigt.

Und es ist so, wie Herr Costabel es schon gerade ausgeführt hat: In den frühen Stadien ist es für die Patienten oftmals nicht so gut erkennbar, wobei auch da unterschiedliche Sensitivitäten da sind, aber sobald eine gewisse Schwelle überschritten ist – und die liegt wahrscheinlich so bei 70 Prozent –, dann merken die Patienten tatsächlich auch solche scheinbar geringen Veränderungen.

Herr Dr. Vöcking (stellv. Vorsitzender): Schönen Dank. – Herr Professor Ruof.

Herr Prof. Dr. Ruof (Roche Pharma): Die FVC ist natürlich auch eine Thematik, die uns mit am Herzen liegt. Generell wird bei der Beurteilung der Endpunkte aus Ihrem Hause, also in der G-BA-Stellungnahme, aus unserer Sicht immer zu wenig Rücksicht auf die Hierarchie der Endpunkte genommen. Die FVC ist natürlich der primäre Endpunkt, und das ganze Studiendesign, das ganze Konzept ist auf diesen Endpunkt ausgerichtet. Man kann nicht sagen, dass die anderen Endpunkte wie Overall Survival oder was auch immer hierarchisch gleich zu nehmen sind. Nein, der primäre Endpunkt, der das Studienkonstrukt rechtfertigt, ist natürlich die FVC. Und von daher soll dieser prioritär berücksichtigt werden.

Dann nur ganz am Rand: Sie sprachen vorher die Tabelle 21 an. Uns geht es nur darum, –

Herr Dr. Vöcking (stellv. Vorsitzender): Ich hätte es nicht vergessen.

Herr Prof. Dr. Ruof (Roche Pharma): – letztendlich die Frage der Identifikation, Behandelbarkeit und Reversibilität der Nebenwirkungen immer darzustellen, weil ich denke, das gibt dann auch noch einmal ein bisschen Klarheit.

Herr Dr. Vöcking (stellv. Vorsitzender): Okay, danke. – Gibt es weitere Anmerkungen, Fragen? – Herr Kulig.

Herr Dr. Kulig: Sie hatten, Herr Trampisch, auf die Nachfrage von Frau Müller schon angesprochen, wie homogen bzw. heterogen diese beiden Studien sind, um sie poolen zu können. Haben Sie sich einmal überlegt, das mit Random-Effects-Modellen zu berechnen? Auffällig ist es ja bei den akuten Exazerbationen. Da ist ja die Effektrichtung in den beiden Studien sogar umgekehrt. Wären da andere Modelle nicht geeigneter gewesen? Und wenn ja, haben Sie das gemacht und wie sind da die Ergebnisse gewesen?

Herr Dr. Trampisch (Boehringer Ingelheim Pharma): Wie bereits angesprochen, waren die Studien als Zwillingstudien ausgelegt. Es war von vornherein präspezifiziert, eine gepoolte Analyse durchzuführen. In dieser gepoolten Analyse wurde auch präspezifiziert, wie der Studieneinfluss zu modellie-

ren ist. Hier wurde ein Fixeffekt vorgesehen. Dies wurde auch von den Zulassungsbehörden akzeptiert.

Die von Ihnen geäußerte Auffassung, dass Inhomogenität oder gegenläufige Effekte bezüglich der Exazerbationen herrschten, teilen wir nicht. Wir haben ein statistisch signifikantes Ergebnis bezüglich der Investigator-reported Exazerbationen, und wir haben ein statistisch nicht signifikantes Ergebnis mit einem Effektschätzer, der etwas über 1 liegt. Allerdings sind die Konfidenzintervalle sehr breit und überlappend auch mit den signifikanten Ergebnissen der INPULSIS-2-Studie. Insofern liegen hier keine divergenten oder inhomogenen Ergebnisse vor. Wir haben nur ein statistisch nicht signifikantes Ergebnis mit einem Punktschätzer, der etwas über 1 liegt.

Bezüglich der Random-Effects-Modellierung der gepoolten Studien: Bei einem Random-Effects-Modell liegt das Problem vor, dass es bei unseren Modellen nicht straight forward wäre, einen Random-Effekt einzubeziehen. Allerdings teilen wir nicht den Standpunkt, dass hier ein Random-Effects-Modell adäquat wäre. Die Studien sind auf einem identischen Studienkollektiv identisch durchgeführt worden. Es ist nicht davon auszugehen, dass der fixe Behandlungseffekt von Ofev[®] in diesem Studienkollektiv aus einer zufälligen Behandlungseffektkurve gezogen worden ist, sondern dass ein fixer Effekt in beiden Studien vorliegt und dass die Variabilität, die dieser fixe Effekt in der Natur hat, über unsere beiden Studien abgebildet wurde. Insofern halten wir das Fixed-Effects-Modell natürlich für adäquat.

Haben wir ein Random-Effects-Modell durchgeführt? – Wir haben bezüglich Random Effects nur Metaanalysen auf den aggregierten gerechnet. Allerdings sind diese Ergebnisse bezüglich der Exazerbation konsistent zu den von uns berichteten Fixed-Effekten. Insofern ändert sich bezüglich der Interpretation der Ergebnisse von unserer Seite nichts.

Herr Dr. Vöcking (stellv. Vorsitzender): Herr Kollege Kulig.

Herr Dr. Kulig: Eine Verständnisfrage. Ich habe den letzten Teil nicht verstanden. Sie haben eine Metaanalyse durchgeführt und sagen, es ändert sich nichts, weil das Ergebnis gleich ist. Aber war das auch ein Fixed-Effects-Modell? Das habe ich noch nicht verstanden.

Herr Dr. Trampisch (Boehringer Ingelheim Pharma): Ja, das war auch ein Fixed-Effects-Modell. Wir haben auch ein Random-Effects-Modell gerechnet, allerdings nur auf den aggregierten Daten. Diese Daten liegen vor. Wir können diese Daten nachliefern, falls gewünscht. Sie können diese Daten aber auch selber generieren. Denn wir haben nur auf den aggregierten Daten gerechnet, nicht auf den patientenindividuellen Daten. Allerdings halten wir die Berechnung auf den individuellen Patientendaten für adäquat und auch relevant.

Herr Dr. Vöcking (stellv. Vorsitzender): Schönen Dank, Herr Trampisch. – Herr Eyding, ich habe jetzt Sie auf der Agenda, danach Frau Wegmann und dann Frau Müller.

Herr Dr. Eyding: Eine ganz kurze Nachfrage zu der Feststellung, dass es ein zufälliger Effekt sein kann. Gab es nicht Unterschiede in den Patientenpopulationen beider Studien? Hatten Sie nicht in der einen Studie deutlich mehr Asiaten eingeschlossen als in der anderen Studie? Könnte das nicht potenziell als Effektmodifikator eine Rolle spielen? Dann wäre das Random-Effects-Modell doch das eher angemessene.

Herr Dr. Vöcking (stellv. Vorsitzender): Jetzt antwortet, nachdem er sich erholt hat, Herr Kögler.

Herr Kögler (Boehringer Ingelheim Pharma): Danke für die Pause, die Sie mir eingeräumt haben. – In der Tat gab es in zwei Baseline erhobenen Patientencharakteristikaparametern geringfügige Abweichungen. Das eine ist der Anteil an Asiaten in der Studienbevölkerung, den Sie schon angesprochen haben. Er lag in der INPULSIS-2-Studie etwas höher als in der INPULSIS-1-Studie. Es gab auch einen Unterschied beim Anteil der Nichtraucher, der allerdings nicht so ausgeprägt war.

Als wir post hoc die Unterschiede in den Outcomes der Studien beobachtet haben, haben wir natürlich noch einmal intensiv nachgeschaut und getestet, ob die Parameter Ethnizität beziehungsweise Raucherstatus einen Effektmodifikator darstellen. Das ist nicht der Fall gewesen. Das haben wir so interpretiert, dass die Unterschiede nicht durch diese Faktoren bedingt sind.

Herr Dr. Vöcking (stellv. Vorsitzender): Herr Eyding.

Herr Dr. Eyding: Die Tests, die die Heterogenität nachweisen, sind relativ schlecht gepowert. Insofern wäre es trotzdem interessant, die Random-Effects-Daten zu sehen. Wie groß war I^2 , also der Interaktionstest? Wie groß war der p -Wert? All diese Dinge haben wir jetzt nicht. Insofern wäre es schon ganz gut, das noch zu bekommen, auch um zu sehen, wie groß die Übereinstimmung zwischen dem Fixed-Effects-Modell und dem Random-Effects-Modell in dieser Hinsicht ist. Wir können Ihre Aussage zur Kenntnis nehmen, aber wir können selber nicht beurteilen, auf welche Datengrundlage wir uns stützen. Ich weiß nicht, ob Sie das noch nachliefern könnten.

Herr Dr. Pfannkuche (Boehringer Ingelheim Pharma): Ich möchte gern versuchen, das noch einmal in anderen Worten wiederzugeben. Dem TSAP, unserem statistischen Analyseplan, der verblindet erstellt worden ist, lag die Grundannahme zugrunde, dass wir mit der INPULSIS-1-Studie und der INPULSIS-2-Studie zwei Zwillingsstudien haben, die vergleichbar oder gleich aufgesetzt sind. Auch aus diesem Grunde haben wir uns für das Fixed-Effects-Modell entschieden.

Man kann nachträglich ein Random-Effects-Modell rechnen. Nach den Anmerkungen aus der G-BA-Bewertung haben wir uns das auch angeschaut. Hierbei muss man natürlich im Hinterkopf behalten, was ursprünglich geplant war. Dann sind wir wieder bei der Frage, wie die Studie geplant wurde und was die Aussagekraft der Studie ist. Ursprünglich geplant wurde etwas anderes, nämlich das Fixed-Effects-Modell.

Schließlich sind die Analysen – das hat Herr Trampisch schon gesagt und das möchte ich hier auch noch einmal sagen – nur auf aggregierten Daten möglich, weil alle anderen Analysen für die primären und sekundären Endpunkte nicht unbedingt die Studie als Variable in der Analyse zulassen. Deswegen müssen wir hier auf aggregierte Daten zurückgreifen und können nicht auf patientenindividuelle Daten zurückgreifen, ohne sämtliche Analysen, die wir bisher durchgeführt haben, infrage zu stellen. Das würde die Aussagekraft nicht unbedingt weiter erhöhen.

Nichtsdestoweniger nehmen wir dieses Thema noch einmal mit und werden prüfen, ob wir die Daten nachreichen können.

Herr Dr. Vöcking (stellv. Vorsitzender): Schönen Dank. – Herr Kögler, Sie hatten sich auch gemeldet; ich nehme aber an, dass sich nun Ihre Stellungnahme erübrigt hat. – Dann habe ich jetzt Frau Wegmann und danach Frau Dr. Müller auf der Rednerliste. Frau Wegmann.

Frau van Bömmel-Wegmann: Ich habe eine Frage zum Nebenwirkungsprofil. Sie sind eingangs ja bereits darauf eingegangen, dass das Präparat gut verträglich sei. Jetzt haben wir hier ein deutliches Ausmaß an gastrointestinalen Nebenwirkungen, die als mild und moderat beschrieben werden. Mir ist aufgefallen, dass sie auch den größten Teil der UEs darstellen, die zum Therapieabbruch führen. Ich

bitte Sie, darauf einzugehen, ob das Präparat bei einigen Patienten sehr viel stärkere gastrointestinale Nebenwirkungen verursacht oder woran dieser Unterschied liegt.

Im Anschluss noch eine sehr spekulative Frage an die Stellungnehmer: Wir haben ja gehört, dass es sich um noch nicht so stark betroffene Patienten handelt. Wie würde sich dies verhalten, wenn die Patienten stärker betroffen sind? Würde bei stärker betroffenen Patienten mit schlechterem Allgemeinzustand die Gefahr des Therapieabbruchs oder von gastrointestinalen Nebenwirkungen ein noch höheres Gewicht bekommen?

Herr Kögler (Boehringer Ingelheim Pharma): Sie sprechen die gastrointestinalen Nebenwirkungen an, die registriert wurden. Hier gibt es in der Tat deutliche Gruppenunterschiede. Das ist auch klar dargelegt worden.

Ich möchte zunächst das Thema Diarrhö aufgreifen. Wenn wir über Diarrhöen nachdenken, haben wir alle möglicherweise die eigenen Erfahrungen im Kopf, die man einmal auf Reisen gemacht hat, wenn man – in Anführungszeichen – Magen-Darm gehabt hat und sich massiv kreislaufbeeinträchtigt vom Bett zum WC und wieder zurück schleppt. Vom Impact her wäre das eine schwere Diarrhö, und solche Ereignisse würden in der Studie als schwere Diarrhö erfasst. Es gab daneben milde und moderate Ereignisse. Das heißt, wir haben eine große Spannweite von sehr geringen Beeinträchtigungen bis hin zu deutlichen; letztere gab es auch, das möchte ich gar nicht in Abrede stellen. Aber bei den schweren Diarrhöen sprechen wir nicht über 60 Prozent, sondern über 3 Prozent.

Im Allgemeinen ließ sich durch eine Dosierungsunterbrechung und Wiederinitiierung mit niedrigerer Dosierung oder auch durch eine symptomatische Therapie mit Loperamid die Diarrhö insoweit beherrschen, dass die Therapie fortgesetzt werden konnte. Zugegebenermaßen traten einige Fälle auf, in denen es weiterhin Probleme gab. Da musste die Therapie tatsächlich abgebrochen werden. Nach unserer Ansicht sollte aber deswegen nicht die Entscheidung getroffen werden, erst gar nicht den Versuch einer Behandlung mit einer Therapie zu unternehmen, die potenziell 50 Prozent des jährlichen Verlustes an Lungenkapazität verhindern kann. Im Einzelfall gibt es aber die Möglichkeit, die Behandlung tatsächlich zu beenden.

Herr Dr. Vöcking (stellv. Vorsitzender): Schönen Dank. – Herr Professor Behr, wollen Sie sich direkt dazu äußern?

Herr Prof. Dr. Behr (DGP): Ich möchte zum Thema der Nebenwirkungen aus klinischer Sicht und als jemand, der viele solcher Patienten behandelt hat und noch behandelt, Stellung nehmen.

Die Diarrhö ist im Regelfall sehr gut beherrschbar. Es gibt nur ganz vereinzelt Patienten, die extrem ansprechen. In meinem Kollektiv kann ich mich an niemanden erinnern, den ich etwa in diesem Jahr zum Beispiel stationär hätte behandeln müssen. Das heißt, diese Beobachtung findet sich zwar in der Studie, keine Frage, aber in der klinischen Praxis sind diese Fälle gut beherrschbar. Wir kennen dieses Problem auch aus Studien zu und Behandlungen mit Tyrosinkinase-Inhibitoren, die wir bei Tumorpatienten durchführen. Ich denke, es ist auch eine Frage der Patientenschulung und der Patientenführung. Dann ist diese Nebenwirkung in aller Regel sehr gut beherrschbar. Von daher muss ich sagen: Das kann natürlich ein Problem sein. Es ist aber auf eine kleine Zahl von Patienten, bei der es wirklich dramatisch werden kann, begrenzt.

Herr Dr. Vöcking (stellv. Vorsitzender): Danke, Herr Professor Behr. – Ich schaue jetzt auf die Rednerliste und auf die Uhr. Frau Müller, Sie wären jetzt an der Reihe, dann käme Herr Eyding. Frau Müller.

Frau Dr. Müller: Eine der beiden Fragen, die ich hatte, wurde von Ihnen bereits beantwortet. Ich wollte den klinischen Experten zur Diarrhö fragen, auch gerade vor dem Hintergrund, dass die EMA das als relativ gravierende Nebenwirkung einschätzt, und vor dem Hintergrund der eingeschränkten Mobilität. Das betrifft dann, so nehme ich an, eher die fortgeschrittenen Stadien, die in der Studie nicht so stark vertreten waren. Sie haben ja thematisiert, dass eine Diarrhö sowohl für die Patienten selber als auch für eventuelle betreuende Personen ein besonderes Problem darstellt. Aber Sie haben klar gesagt, nach Ihrer klinischen Erfahrung ist es beherrschbar und nicht so gravierend. So habe ich es verstanden. Das ist insofern beantwortet.

Die zweite Frage richtet sich auf das methodische Vorgehen. Mich hat etwas irritiert, dass Sie die ganzen Auswertungen im Treated Set dargestellt haben, also etwa, dass Sie nur Patienten berücksichtigt haben, die mindestens eine Dosis Studienmedikation bekommen haben. Im Dossier haben Sie dargelegt, dass der Unterschied zum Randomised Set nicht so groß wäre – das stimmt ja auch, es sind bloß zwei bzw. drei Patienten herausgefallen – und insofern das Intention-to-Treat-Prinzip gewahrt wäre. Meine Frage ist folgende: Warum dann nicht gleich so, wie randomisiert? Warum haben Sie das verwendet?

Herr Dr. Trampisch (Boehringer Ingelheim Pharma): Es ist korrekt, wir haben insgesamt fünf Patienten, die randomisiert worden sind und keine Studienmedikation erhalten haben. Die Auswertung bezüglich Efficacy und Safety war präspezifiziert auf dem Treated Set, was insbesondere in Bezug auf die unerwünschten Ereignisse sinnvoll ist, um einen Effekt von Ofev[®] und Placebo nicht gegebenenfalls durch nicht behandelte Patienten zu verwässern.

Herr Dr. Vöcking (stellv. Vorsitzender): Frau Müller.

Frau Dr. Müller: Leider muss ich noch ganz kurz nachhaken. Ich habe schon gesehen, dass Sie das gleiche Set für Safety und Wirksamkeit verwendet haben. Aber normalerweise berücksichtigt man beim Safety Set ja, ob die Patienten eine Studienmedikation bekommen haben, und sie werden berücksichtigt „as treated“. Wenn sie anders behandelt wurden, als sie randomisiert wurden, wird das ja im Unterschied zur Wirksamkeitsbeurteilung berücksichtigt. Insofern kann ich dem jetzt, ehrlich gesagt, nicht folgen. Denn wenn Sie so vorgegangen wären, wie es üblich ist, wäre das Problem gar nicht erst aufgetreten.

Noch eine Frage. Sie haben gesagt, das Treated Set berücksichtigt nur die Patienten, die mindestens eine Dosis bekommen haben, allerdings wie randomisiert. Das ist also kein Safety Set, sondern es berücksichtigt die Patienten, wie sie randomisiert wurden, und nicht, wie sie wirklich behandelt wurden. Ist das richtig?

Herr Dr. Trampisch (Boehringer Ingelheim Pharma): Das ist korrekt.

Frau Dr. Müller: Okay. Auch für die Auswirkungen der Sicherheitsparameter?

(Herr Dr. Trampisch (Boehringer Ingelheim Pharma) nickt)

– Aha, das erstaunt mich. – Okay. Danke.

Herr Dr. Trampisch (Boehringer Ingelheim Pharma): Das ist korrekt. Das war so in den Protokollen definiert.

Herr Dr. Vöcking (stellv. Vorsitzender): Danke. – Herr Eyding.

Herr Dr. Eyding: Es ist ja schon spät, aber ich habe noch eine aus meiner Sicht klärende Nachfrage zu den Exazerbationen. Wir haben noch nicht über die definitiven und die vermuteten Exazerbationen gesprochen. Operationalisiert war ja die Vermutung, dass eines der Kriterien aus Dyspnoe und Bildgebung erfüllt sein muss. Ich habe die Diskussion so wahrgenommen, dass Exazerbationen eigentlich immer mit einer für den Patienten durchaus spürbaren und den Anlass gebenden Verschlechterung einhergehen. Aber der Form halber, weil es ja formal operationalisiert ist, die Frage: Gibt es vermutete Exazerbationen – diese sind ja bei den Exazerbationen mit in die Analyse eingegangen –, die ausschließlich auf Bilddaten beruhen?

Herr Kögler (Boehringer Ingelheim Pharma): Jede dieser Exazerbationen hatte die erhebliche Symptomverschlechterung, die die Voraussetzung dafür war, dass die primäre Erfassung auf dem Formular für unerwünschte Ereignisse erfolgte. Das war der Trigger. Erst sekundär wurde das Exazerbationsformular ausgefüllt. Für die Wahrnehmung als UE oder als SUE war die Symptomverschlechterung zwingend erforderlich. Das heißt, es gibt keine Ereignisse, die ausschließlich radiologisch charakterisiert worden sind.

Herr Dr. Vöcking (stellv. Vorsitzender): Eine Ergänzung von Herrn Professor Costabel.

Herr Prof. Dr. Costabel (DGP): Es musste beides, Dyspnoeverschlechterung, also Atemnot, und die Bildgebung, vorhanden sein, damit es als akute Investigator-reported Exazerbation berichtet wurde. Die Differenz zwischen suspected und definitiv bei Exazerbationen lag meistens darin, dass keine bakteriologischen Daten vorlagen, dass keine BAL durchgeführt werden konnte, weil es dem Patienten zu schlecht ging, dass keine Blutkulturen positiv waren, sodass man eine Pneumonie nicht mit Sicherheit ausschließen konnte. Das war der Hauptgrund, warum sie nur als suspected gewertet wurden.

Ganz interessant ist eine Arbeit von Amerikanern, in der gezeigt wird, dass, wenn man die Exazerbationen nach akut und suspected getrennt analysiert, im Hinblick auf das Überleben gar kein Unterschied existiert. Bei den Exazerbationen sind in dieser Analyse die suspected genauso schlecht verlaufen wie die definitiven

Herr Dr. Vöcking (stellv. Vorsitzender): Professor Costabel, danke für die Ergänzung. – Herr Eyding, eine Nachfrage?

Herr Dr. Eyding: Eine ganz kurze Nachfrage zu dem generellen Problem, dass die Wirksamkeits- oder Nutzenendpunkte hier auch auf den UE-Daten beruhen, was immer eine gewisse Unsicherheit mit sich bringt, weil UEs spontan berichtet werden oder eben auch nicht, zum Beispiel wenn ein Patient es nicht wahrnimmt. Haben Sie während der Studie irgendeine Art von prospektiver Abfrage für die nutzenrelevanten UE vorgesehen, etwa dass die Patienten bei jeder Visite routinemäßig in einer bestimmten Art abgefragt worden sind, ob es irgendwelchen Ereignisse hinsichtlich der Nutzenendpunkte, also etwa der Verschlechterung der IPF oder was auch immer, gegeben habe? Oder haben wir wirklich nur Daten über spontan berichtete UEs?

Herr Kögler (Boehringer Ingelheim Pharma): Es sind die spontan berichteten UEs gewesen.

Herr Dr. Vöcking (stellv. Vorsitzender): Danke für die kurze Antwort. – War sie zufriedenstellend?

(Herr Dr. Eyding nickt)

– Okay. – Ich schaue jetzt in die Runde und frage auch mit Blick auf die Uhr: Gibt es noch weitere Fragen, Anmerkungen oder Kommentare dazu? – Ich sehe, das ist nicht der Fall.

Herr Dr. Pfannkuche, Sie kennen das übliche Verfahren. Ich gebe Ihnen gern Gelegenheit, noch einmal abschließend zu resümieren. Meine Bitte wäre, nicht alles zu wiederholen, sondern es kurz und knackig zu machen. Ich denke, wir haben im Grunde alle wichtigen Punkte erwähnt. Ich erinnere daran, dass drei Themen noch – in Anführungszeichen – offen sind: Wir haben eine Zusage von Ihnen, Daten nachzuliefern, und zwei Zusagen, dass Sie prüfen, ob Sie nachliefern können, wollen oder sollten.

Herr Dr. Pfannkuche (Boehringer Ingelheim Pharma): Sehr geehrter Herr Vöcking! Sehr geehrte Mitglieder des Gemeinsamen Bundesausschusses! Vielen Dank für die gemeinsame Diskussion und die Möglichkeit, noch ein paar abschließende Worte zu sagen. Ich will die wichtigsten Punkte kurz zusammenfassen.

Erstens zum Punkt, dass der in den Zulassungsstudien berichtete Endpunkt Exazerbation patientenrelevant ist und mittels einer validen Methodik erhoben worden ist. Die Definition und die Dokumentation der in den Studien aufgetretenen Exazerbationen erfolgten in Übereinstimmung mit den Empfehlungen der Zulassungsbehörden und der entsprechenden Fachgesellschaften. Circa 85 Prozent der definitiven oder wahrscheinlichen Exazerbationen waren als SUE klassifiziert. Circa 55 Prozent der Patienten, die solch ein Ereignis erlitten hatten, verstarben am Ende.

Zweitens. Der in den Studien erfasste FVC-Abfall ist ein valider Surrogatparameter für die Mortalität. Diese Sichtweise wird weltweit auch von den Zulassungsbehörden und den Fachgesellschaften anerkannt. Auch die von den FDA-Mitarbeitern durchgeführte Analyse zeigt einen konsistenten und gleichgerichteten Zusammenhang zwischen der positiven Beeinflussung des Surrogats und des harten klinischen Endpunkts der Mortalität. Hier möchte ich kurz erwähnen, dass wir die Punkte, die in der G-BA-Bewertung erwähnt wurden – der erste Punkt war, die Intervention als unabhängige Größe zu betrachten; der zweite Punkt war, den Zusammenhang zwischen dem Effekt auf das Surrogat und dem Effekt auf den patientenrelevanten Endpunkt einander gegenüberzustellen –, mit der von uns durchgeführten Analyse eingereicht haben.

Zusammenfassend sehen wir für Ofev[®], ein Arzneimittel zur Behandlung einer seltenen Erkrankung, einen beträchtlichen Zusatznutzen, der sich aus der klinisch bedeutsamen Verbesserung des mortalitätsassoziierten Endpunktes FVC sowie der Reduktion der adjudizierten akuten Exazerbationen bei einem insgesamt guten Verträglichkeitsprofil, wie wir es hier auch diskutiert haben, begründet.

Damit möchte ich schließen und wünsche Ihnen einen schönen Tag. Bis bald!

Herr Dr. Vöcking (stellv. Vorsitzender): Herr Pfannkuche, schönen Dank. – Ich bedanke mich ganz herzlich für die in meinen Augen interessante Diskussion. Ich fand die Fragen und Antworten durchaus spannend. Es hat auch mehr Zeit in Anspruch genommen, als ich ursprünglich gedacht hatte. Das lag sicherlich an der Thematik und den interessanten Fragestellungen und Antworten. Wir werden dann später in die Beratungen einsteigen.

Damit beende ich die Anhörung zu diesem Punkt und wünsche allen Teilnehmern der Anhörung einen guten Tag und vor allem eine sichere Heimreise. Bis zum nächsten Mal!

Schluss der Anhörung: 11.26 Uhr