

Mündliche Anhörung



**Gemeinsamer
Bundesausschuss**

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

hier: Wirkstoff Decitabin

Sitzung im Hause des Gemeinsamen Bundesausschusses in Berlin
am 12. März 2013
von 10.10 Uhr bis 11.36 Uhr

– Stenografisches Wortprotokoll –

Angemeldete Teilnehmer der Firma **Janssen-Cilag GmbH:**

Herr Dr. Fleischmann

Herr Dr. Angermund, anstelle von Frau Dr. Rudolph

Herr Dr. Tapprich

Herr Dr. Tomeczkowski

Angemeldete Teilnehmer der Firma **Celgene GmbH:**

Frau Prof. Dr. Böhler

Herr Völkl

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

Herr Prof. Dr. Wörmann

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

Frau Dabisch

Herr Dr. Dintsios

Beginn der Anhörung: 10.10 Uhr

(Die angemeldeten Teilnehmer betreten den Raum)

Herr Hecken (Vorsitzender): Meine Damen! Meine sehr geehrten Herren! Herzlich willkommen zur heutigen mündlichen Anhörung zu Decitabin. Wir sind in dieser Zusammensetzung ja mittlerweile schon weitestgehend ein eingespieltes Team.

Wir unterhalten uns heute nach einem schriftlichen Stellungnahmeverfahren über ein Orphan Drug, ein Präparat bzw. einen Wirkstoff zur Behandlung der AML. Stellungnahmen haben im schriftlichen Stellungnahmeverfahren abgegeben der vfa, die DGHO, Celgene GmbH und Janssen-Cilag GmbH. Wir haben hier eine Reihe von Vertretern der genannten Gesellschaften und Organisationen. Ich begrüße ganz herzlich Herrn Dr. Fleischmann, Herrn Dr. Angermund, der anstelle von Frau Dr. Rudolph hier ist, Herrn Dr. Tapprich und Herrn Dr. Tomeczkowski von Janssen-Cilag; ich begrüße Frau Professor Böhler und Herrn Völkl von Celgene GmbH; Herr Professor Wörmann ist ständiger Gast als Vertreter der Deutschen Gesellschaft für Hämatologie und Onkologie; schließlich begrüße ich Frau Dabisch und Herrn Dr. Dintsios vom vfa.

Wir haben für die heutige Anhörung einen Zeitrahmen von anderthalb Stunden angesetzt. So viel Zeit ist angesichts der verschiedenen Punkte, die im schriftlichen Stellungnahmeverfahren vorgetragen worden sind, wohl auch notwendig.

Zum einen geht es – ich will das jetzt nur kurz anreißen – um rechtliche Fragen, hier konkret um die bei Orphan Drugs immer im Raum stehende Frage der Bindungswirkung der Zulassung unter besonderer Berücksichtigung der vom Gesetzgeber gewollten Privilegierung von Orphan Drugs. Hierzu ist auch eine rechtliche Stellungnahme vorgelegt worden.

Dann geht es um inhaltliche Fragen, insbesondere natürlich um Mortalität als primären Endpunkt. Hier stellt sich die Frage, wie mit den beiden im Raum stehenden Studien umzugehen ist. Die erste dokumentiert eine etwas längere Überlebensrate, während bei der zweiten, die aus Sicht der Vorbewertung ein Stück weit valider ist, ein etwas geringerer Überlebensvorteil am Ende herauskommt.

Dann geht es um weitere Fragen sekundärer Endpunkte. Es geht um Fragestellungen hinsichtlich der Morbidität. Es geht um die Frage, welchen Wert Transfusionsunabhängigkeit hat. Es geht um die Frage, in welchem Verhältnis positive Effekte, insbesondere der Effekt „Primärer Endpunkt Überleben“, zu Nebenwirkungen stehen, die mit dem Präparat verbunden sind, ob diese Nebenwirkungen im Vergleich zu anderen Nebenwirkungen signifikant gravierender oder nicht gravierender sind.

Neben der Frage der Zielpopulation, wo man sich noch einmal über zwei oder drei Annahmen, die das IQWiG gesetzt hat, unterhalten muss, geht es um die letztlich entscheidende Frage nach dem Ausmaß des Zusatznutzens, also ob 2,7 Monate oder 2,2 Monate oder 2,3 Monate ein beträchtlicher Überlebensvorteil – ja oder nein? – sind.

Damit habe ich grob umrissen, was Gegenstand der unterschiedlichen Stellungnahmen war und welche Fragestellungen heute erörtert werden sollen. Wir wollen so verfahren, dass Sie die Möglichkeit haben, zunächst in einleitenden Statements Stellung zu nehmen. Meine Bitte wäre wie immer, dass nicht in Gänze vorgelesen wird, was Sie uns hier zu Protokoll gegeben haben. Ich habe für die zwei Anhörungen, die heute Morgen stattfinden, einen Aktenordner. Wenn all das vorgelesen würde, dann hätten wir ein kleines Zeitproblem, und es würde letztlich auch dem Erkenntnisgewinn nicht förderlich sein, weil die hier Anwesenden die Stellungnahmen gelesen haben. Sie sind auch entsprechend ausgewertet und dann diskutiert worden. Meine Bitte wäre, dass Sie sich auf die aus Ihrer Sicht wesentli-

chen, relevanten bzw. streitigen Punkte beschränken, sodass wir auch seitens der Bänke Gelegenheit haben, in einen Diskurs mit Ihnen einzutreten.

Meine letzte Bitte wäre, dass Sie jeweils, bevor Sie das Wort hier ergreifen, Ihren Namen und die Organisation bzw. das Unternehmen, das Sie vertreten, nennen, damit das entsprechend protokolliert werden kann. Wir führen heute – wie immer bei diesen Veranstaltungen – ein Wortprotokoll. Es ist für die Protokollanten einfach besser, wenn man am Ende auch weiß, wer was gesagt hat; wobei es manchmal besser wäre, man wüsste nicht, was man gesagt hätte; aber das sei an dieser Stelle einfach so in den Raum gestellt. So viel zu den Präliminarien.

Ich frage, wer von Ihnen möchte beginnen? – Bitte schön, Herr Dr. Fleischmann.

Herr Dr. Fleischmann (Janssen-Cilag): Guten Morgen an Sie alle und vielen Dank für die Möglichkeit, hier Stellung zu nehmen. Ich erlaube mir, das aus Sicht von Janssen-Cilag zu tun. Eingangs möchte ich ganz kurz meine Kollegen vorstellen. Es gab auch eine kleine Änderung: Frau Dr. Rudolph darf ich erkrankungsbedingt entschuldigen. Wir hatten angekündigt, dass wir Herrn Dr. Angermund dafür mitbringen, und er ist auch hier zu meiner Linken vertreten. Zu meinen Kollegen: Dabei sind neben mir – ich bin Leiter der Abteilung Marktzugang bei Janssen-Cilag Deutschland – Dr. Ralf Angermund; er ist Internist, Hämatologe und internistischer Onkologe und leitet den medizinischen Fachbereich für Hämatologie und Onkologie bei Janssen-Cilag. Dr. Christoph Tapprich, auch Internist, Hämatologe und internistischer Onkologe, ist der zuständige Medical Manager für Hämatologie bei Janssen-Cilag. Dr. Jörg Tomeczkowski schließlich ist Biologe und in der Abteilung Marktzugang für das Produkt Dacogen® bzw. den Wirkstoff Decitabin zuständig.

Wir möchten an dieser Stelle – Herr Hecken, Sie hatten ja darum gebeten – nicht im Detail wiederholen, was im Dossier und in der Stellungnahme schon ausreichend vorgebracht wurde. Aber bitte lassen Sie uns nichtsdestotrotz einige wesentliche Aspekte hervorheben, die aus unserer Sicht für die Bewertung von Dacogen® von entscheidender Bedeutung sind und in jedem Falle in die Entscheidung über das Ausmaß des Zusatznutzens dieses Produktes einfließen sollten.

Kurz zum Hintergrund: Dacogen®, Wirkstoff Decitabin, ist indiziert zur Behandlung von erwachsenen Patienten ab einem Alter von 65 Jahren mit neu diagnostizierter *de novo* oder sekundärer akuter myeloischer Leukämie gemäß der WHO-Klassifizierung, und zwar Patienten, für die eine Standard-Induktionschemotherapie nicht infrage kommt. Dacogen® deckt einen erheblichen medizinischen Bedarf ab, denn einem großen Teil der relevanten Patientengruppe, nämlich älteren Patienten, für die eine intensive Chemotherapie nicht infrage kommt, stehen bisher keine oder nur sehr eingeschränkt zufriedenstellende Therapiealternativen zur Verfügung.

Bei der AML handelt es sich – das ist hier auch zu bedenken – um eine hoch aggressive, bösartige Erkrankung. Unbehandelt führt sie innerhalb von fünf Monaten nach den ersten Symptomen bei der Hälfte der Patienten und innerhalb eines Jahres bei fast allen Patienten zum Tode. Die ungünstige Prognose ist von einer hohen Morbidität begleitet. Das sind vor allem allgemeines Schwächegefühl, Müdigkeit, Blutungsneigungen und Infektionen. Die geringe Anzahl der therapeutischen Optionen mit unzureichender Wirksamkeit und Verträglichkeit definieren hier den therapeutischen Bedarf.

Dacogen® stellt, wie von uns im Dossier und in der Stellungnahme umfassend ausgeführt, eine im Sinne der Verfahrensordnung bisher nicht erreichte deutliche Verbesserung des therapielevanten Nutzens für die Patienten mit AML dar. Dies rechtfertigt aus unserer Sicht die Bewertung mit der Kategorie „beträchtlicher Zusatznutzen“. Unter Berücksichtigung der besonderen Schwere und der Besonderheit dieser speziellen Erkrankung sollten die primären Therapieziele sowie das Gesamtbild der verfügbaren Evidenz im Vordergrund der Entscheidung über den Zusatznutzen stehen und angemessen

sen gewürdigt werden. Lassen Sie mich hier die wesentlichen Elemente kurz zusammenfassen. Es sind vor allem zwei Punkte, auf die es ankommt: der Punkt Remission und – das hatten Sie ja auch schon angedeutet – der Punkt Gesamtüberleben.

Zunächst zum Punkt Remission. Eine fundamentale Bedeutung in der AML kommt der Remissionszielung zu. Die Remission ist entsprechend der Leitlinie das grundlegende Therapieziel. Die Patientenrelevanz der Remission steht aus unserer Sicht außer Frage. Remission heißt für den Patienten die Freiheit von den mit einer AML verbundenen Symptomen, also Schwäche, Müdigkeit, Blutungsneigung, Infektionen, und damit verbunden auch die Chance auf einen Überlebensvorteil, der sich in der hier relevanten Studie bei Patienten mit Remission durchaus auf 19 bis 20 Monate belaufen hat. Für die Remission wurde der Zusatznutzen von Dacogen® eindeutig festgestellt, und zwar in einer Bedeutung und Signifikanz, die nach unserer Ansicht schon alleine rechtfertigen, von einem beträchtlichen Zusatznutzen zu sprechen.

Dieser beträchtliche Zusatznutzen bestätigt sich aus unserer Sicht auch beim Punkt Überleben. Es wurde von der EMA festgestellt, dass es sich hier um einen klinisch hochrelevanten Nutzen handelt, wenn man die Verlängerung des Gesamtüberlebens anschaut. Diese beträgt 2,7 Monate oder, wenn man den Ausgangswert im Vergleichsarm anschaut, 54 Prozent. Diese Verlängerung des Überlebens ist relevant und entspricht aus unserer Sicht definitiv einem beträchtlichen Zusatznutzen.

Man muss nun Folgendes bedenken: In dem am 1. Februar veröffentlichten Berichtsdokument zur Bewertung von Dacogen® werden bezüglich der Interpretierbarkeit der Ergebnisse zum Gesamtüberleben sicher nachvollziehbare Einwände vorgebracht. Nichtsdestotrotz sollten wir aber gerade bei einer solchen Erkrankung immer einen angemessenen Fokus auf die Gesamtschau der Evidenz legen und eine angemessene und ausreichende Diskussion um potenziell verzerrende Effekte führen, um nicht in die Gefahr einer Fehleinschätzung des Zusatznutzenpotenzials zu laufen. Wichtig ist auch, dass man klar sieht, welche Analysen der Zulassungsstudie vorgeplant waren und welche Analysen auf Anforderung oder in Absprache mit den Zulassungsbehörden zustande kamen.

Bitte erlauben Sie uns, dass wir vor diesem Hintergrund und aufgrund von potenziellen Unklarheiten zum Punkt Gesamtüberleben hier noch etwas genauer ins Detail gehen. Die relevante Zulassungsstudie von Dacogen®, die DACO-016, zeigt eine Überlebensverlängerung zugunsten von Dacogen®: absolut die schon erwähnten 2,7 Monate und relativ 54 Prozent. Allerdings – und das wird auch im verfügbaren Berichtsdokument zur Nutzenbewertung von Dacogen® aufgeführt – ist diese Verlängerung im ursprünglichen Auswertungszeitpunkt statistisch nicht signifikant. Der dazugehörige p-Wert ist 0,107, das heißt, man sieht einen starken Trend hin zur Signifikanz und zum Vorteil von Dacogen®, aber es kann noch nicht definitiv von einem Vorteil von Dacogen® gesprochen werden.

Es wäre aus unserer Sicht nun aber nicht richtig und würde wesentliche Teile der vorliegenden Evidenz ignorieren, wenn der G-BA nur auf dieser Basis zu dem Schluss kommen würde, dass kein Überlebensvorteil von Dacogen® vorliegt. Der G-BA sollte vielmehr das Gesamtbild der Analysen in Bezug auf Überlebensvorteile anschauen; dann müsste er zu dem Schluss eines relevanten Vorteils kommen. Dabei sind einige Punkte zu beachten. Zu deren Erläuterung würde ich gerne an Herrn Dr. Tomeczkowski übergeben.

Herr Dr. Tomeczkowski (Janssen-Cilag): Guten Morgen! Ich möchte jetzt gerne ein paar Worte über die Phase-III-Studie verlieren. Es war eine randomisierte Phase-III-Studie. Zunächst ist dabei zu berücksichtigen, dass diese Studie zwei Auswertungszeitpunkte hatte: eine vorgeplante Analyse im Oktober 2009 und eine weitere Analyse ein Jahr später im Oktober 2010. Dabei möchten wir betonen, dass die weitere Analyse 2010 auf Anforderung des CHMP zustande kam und als Weiterführung der

Studie geplant war. Der nicht signifikante p-Wert von 0,107 – dies haben wir eben schon gehört – bezieht sich auf diese Analyse 2009. In der Analyse 2010 wurde die Signifikanz erreicht. Allerdings ergibt sich auch schon für die Analyse 2009 ein statistisch signifikanter Vorteil, und zwar dann, wenn man den eigentlich angebrachten Wilcoxon-Test anwendet. Bei einem stratifizierten Wilcoxon-Test bekommt man ein Ergebnis von $p=0,0458$; das liegt also unter 0,05. Dies gilt auch, wenn man eine vorgeplante Sensitivitätsanalyse betrachtet. Es gab also zwei Auswertungen zum Zeitpunkt 2009, die ein statistisch signifikantes Ergebnis liefern.

Zunächst möchte ich etwas zur Aussagekraft des statistischen Tests sagen. Die vorgeplante Analyse 2009 wurde mit dem sogenannten Log-Rank-Test gemacht; das Ergebnis war nicht signifikant. Es gibt allerdings gute Gründe dafür, dass dieser Test im Fall von Dacogen® gar nicht angemessen war und eigentlich durch einen anderen Test hätte ersetzt werden müssen. Wir beziehen uns hier auf einen Artikel der Autoren Andreas Ziegler, Stefan Lange und Ralf Bender – zwei der Autoren arbeiten beim IQWiG –, die darauf verweisen, dass der Log-Rank-Test bei sich überkreuzenden Überlebenszeitkurven nicht aussagekräftig ist; in diesem Fall sollte der Log-Rank-Test durch einen Wilcoxon-Test ersetzt werden. Bei Dacogen® liegt bei der Überlebenszeitanalyse genau diese Fallkonstellation vor: Die Überlebenszeitkurven überkreuzen sich zum Ende hin. Dies kann man auf Seite 97 des Berichtsdokuments des G-BA sehen.

Ersetzt man nun den Log-Rank-Test durch den Wilcoxon-Test, so ergibt sich für den Auswertungszeitraum 2009, also den primären Endpunkt, ein statistisch signifikantes Ergebnis. Weitere ergänzend durchgeführte Tests lieferten ebenfalls statistisch signifikante Ergebnisse. Das heißt, mit sehr hoher Wahrscheinlichkeit kam die Nichtsignifikanz des primären Endpunktes aufgrund eines nicht angemessenen und eigentlich gar nicht aussagekräftigen Testverfahrens zustande.

Der zweite Punkt wäre dann noch eine Sensitivitätsanalyse. Auch darauf möchte ich kurz eingehen. Ein Grund, warum sich die Überlebenszeitkurven überkreuzt haben, könnte der Einsatz weiterer nicht vorgesehener Therapien gewesen sein. Das heißt, ein Teil der Patienten erhielt, nachdem die Behandlung mit Dacogen® oder im Vergleichsarm beendet war, weitere Behandlungen mit aktiven Substanzen. Diese hatten natürlich einen Einfluss auf das Ergebnis. Insbesondere im Vergleichsarm wurden die Patienten unter anderem auch mit Dacogen® weiterbehandelt. Wenn diese weiteren Therapien aus der Analyse entfernt werden oder die Patienten in einer anderen Analyse mit vermindertem Gewicht dargestellt werden, ergibt sich zum vorgesehenen Auswertungszeitpunkt 2009 ebenfalls ein statistisch signifikanter Überlebensvorteil für Dacogen®. Diese Sensitivitätsanalyse war vorgeplant und findet sich auch im statistischen Analyseplan. Insofern ist sie bei der Interpretation der Daten zwingend zu berücksichtigen.

Jetzt noch einmal kurz zum Auswertungszeitpunkt 2010. Wie schon bereits erwähnt, wurde auf Anforderung der EMA das Gesamtüberleben zusätzlich noch einmal ein Jahr später ausgewertet. Das ist die Analyse 2010. Diese Auswertung wurde zwar erst nach dem Schließen der Datenbank geplant, allerdings handelte es sich lediglich um die Fortführung der Studie mit weiteren Todesfällen. Dadurch, dass wir weitere Todesfälle hatten, also mehr Events, hatten wir zu diesem Auswertungszeitpunkt auch eine höhere Power. Dies ist deshalb als zusätzliche Evidenz bei der Beurteilung des Endpunktes zu berücksichtigen und nicht als eine klassische datengetriebene Post-hoc-Analyse infrage zu stellen, bei der man im Nachhinein in die Datenbank geht und dann analysiert, ob es signifikante Effekte gab.

Abschließend möchte ich noch sagen, dass die EMA als zuständige Zulassungsbehörde diese Analysen anerkannt hat, insbesondere die Auswertung 2010, und in ihr positives Gesamturteil über Dacogen® einfließen ließ.

Herr Dr. Fleischmann (Janssen-Cilag): Zusammenfassend und abschließend lässt sich also sagen: In Bezug auf das Gesamtüberleben liegt lediglich eine statistisch signifikante Analyse vor, die auf Basis eines nicht angemessenen Testverfahrens zustande kam. Alle anderen Analysen, die zum Teil davor geplant waren, zeigen einen statistisch signifikanten Überlebensvorteil. Aus der Zusammenschau der Analysen zum Zeitpunkt 2009 und 2010 ergibt sich also insgesamt ein sehr deutliches Bild für einen Überlebensvorteil von Dacogen®. Dieses Gesamtbild zu ignorieren wäre nicht angemessen, würde zu verzerrten Entscheidungen führen und auch dem Bedarf in der Praxis nicht gerecht werden. Die Gesamtschau rechtfertigt aus unserer Sicht in jedem Fall die Feststellung eines beträchtlichen Zusatznutzens.

Aus unserer Sicht wirft die Bewertung von Dacogen® vor allem zwei Fragen auf, über die wir gerne mit dem Gemeinsamen Bundesausschuss bzw. dem Unterausschuss Arzneimittel diskutieren.

Erstens geht es um die Rolle der Remission. Remission ist ein zentraler und für die Patienten relevanter Parameter, insbesondere aufgrund der dadurch gegebenen Chance auf ein verlängertes Überleben. Wir sehen schon alleine auf Basis dieses Parameters einen beträchtlichen Zusatznutzen und würden natürlich gerne Ihre Sichtweise dazu erfahren.

Zweite Frage: Wie geht man mit einer Situation um, wie hier im Falle der Analyse des Overall Survival zum Zeitpunkt 2009, in der die vorgeplante Analyse aufgrund der erklärbaren Probleme nur einen Trend auf Vorteil liefert, während alle anderen Analysen ein klar positives Ergebnis im Sinne eines statistisch signifikanten Überlebensvorteils liefern?

Wir freuen uns auf die Diskussion. – Vielen Dank.

Herr Hecken (Vorsitzender): Ganz herzlichen Dank, Herr Dr. Fleischmann. – Ich glaube, wir sollten uns zunächst auf folgende Diskussion konzentrieren: Wie gehen wir mit den unterschiedlichen Studien, mit der vorgeplanten 2009er- und – so wie Sie es jetzt vorgetragen haben – der von Anfang an als zweite Planung vorgesehenen 2010er-Studie, und den hier generierten Daten um? Wie sind die Daten, die dort gewonnen worden sind, zu bewerten?

Ich frage jetzt die Bänke und das IQWiG: Gibt es hierzu Fragen, Erörterungsbedarfe, Nachfragen? Wer möchte? Wer fängt an? – Bitte schön, Herr Kulig.

Herr Dr. Kulig: Sie sprachen eben vom Endpunkt Remission. Wenn ich mich richtig erinnere, sind in die Definition nur Labor- und ähnliche Parameter, aber nicht die von Ihnen angesprochenen patientenrelevanten Symptome eingegangen. Ist das so korrekt?

Herr Dr. Angermund (Janssen-Cilag): Die Remissionskriterien sind im Protokoll festgelegt. Bei der akuten myeloischen Leukämie handelt es sich also um morphologische Kriterien, das heißt im Wesentlichen um die Blastenzahl. Die patientenrelevanten Outcomes, zum Beispiel Fatigue und Nebenwirkungen, kommen in der AML-Definition der Remission nicht vor.

Herr Dr. Tomeczkowski (Janssen-Cilag): Darf ich das noch ergänzen?

Herr Hecken (Vorsitzender): Ja klar, Sie dürfen.

Herr Dr. Tomeczkowski (Janssen-Cilag): Zur Definition der kompletten Remission gehörte auch die Transfusionsfreiheit. Das heißt, die Patienten mussten mindestens eine Woche transfusionsfrei sein, um überhaupt in die Gruppe „komplette Remission“ zu kommen. Wir haben festgestellt, dass die Patienten, die mit Dacogen® behandelt wurden, signifikant häufiger ein transfusionsfreies Intervall von

acht Wochen hatten. Deswegen sehen wir das Erreichen der Remission als patientenrelevant an. Zusätzlich gibt es neue Arbeiten, die zeigen, dass das Erreichen einer Remission positiv mit einer Lebensverlängerung korreliert; das ist in der neueren Arbeit auch validiert worden.

Herr Hecken (Vorsitzender): Frau Olberg, Sie hatten sich gemeldet.

Frau Olberg: Ich habe eine Nachfrage zu Ihren Ausführungen zur Auswertung 2010. Sie sagten anfangs, dass diese Analyse auf Aufforderung durchgeführt worden ist. Sie haben ja als Technik eine Alpha Spending Funktion gewählt, was eine sehr flexible Methodik ist. Sie erlaubt, dass man nicht alle Analysen vorab erklären muss. Können Sie vielleicht noch einmal ganz kurz erläutern, inwiefern die Auswertung 2010 sozusagen abgesprochen war, auf Aufforderung erfolgt ist oder wirklich a priori, sage ich jetzt einmal, bei der Erstellung der Studienpläne geplant worden ist?

Herr Dr. Tomczkowski (Janssen-Cilag): Ja, das ist richtig. Es gab den Auswertungszeitpunkt Oktober 2009. Die Datenbank wurde im Mai 2010 geschlossen, und man hat im Mai 2010 eine Analyse zum Endpunkt 2009 durchgeführt. Man hat aber die zusätzlichen Todesfälle, die in der Zeit bis zum Schließen der Datenbank aufgetreten sind, auch ausgewertet. Man hat gesehen, dass man ein signifikantes Ergebnis bekommt. Darüber hat man mit dem CHMP diskutiert; man hat das dem CHMP vorgelegt und gesagt: Es sieht so aus, dass wir ein signifikantes Ergebnis bekommen, wenn wir weitere Todesfälle haben. Daraufhin hat das CHMP angefordert, dass man ein Jahr später, also im Oktober 2010, noch einmal eine Analyse macht, und man hat diese Analyse dann ab dem Zeitpunkt geplant. Es gab also einen neuen CRF, man hat die Daten kontrolliert erhoben, und es gab auch für den Auswertungszeitpunkt 2010 ein zusätzliches Amendment, das auch im statistischen Analyseplan vorhanden ist. Die zweite Auswertung ist also praktisch nach dem Öffnen der Datenbank geplant worden.

Zu Ihrer Frage zum Signifikanzniveau. Ursprünglich waren zwei Interimsanalysen geplant. Man hat dann die p-Werte nach der Alpha Spending Function of the O'Brien-Fleming Type, die Sie eben erwähnt hatten, adjustiert. Man hat dann für den eigentlich vorgesehenen Auswertungszeitpunkt ein Signifikanzlevel von 0,046 ermittelt. Jetzt kam ja eine weitere Analyse dazu. Man hatte kein Signifikanzniveau und hat gesagt: Okay, dann betrachten wir den ursprünglich geplanten Auswertungszeitpunkt als Interimsanalyse und berechnen nach der gleichen Methode dann, welches Signifikanzlevel wir zum Zeitpunkt 2010 hätten. Dieses Signifikanzniveau, errechnet über diese Methode, würde 0,039 betragen, und unser p-Wert für 2010 lag bei 0,037, also darunter. Ich hatte ja eingangs schon erwähnt, dass das ein Ergebnis ist, das mit dem Log-Rank-Test berechnet wurde, der aber möglicherweise nicht aussagekräftig ist, weil die Ereignisse nicht proportional zueinander sind. Wenn man in dieser Situation dann noch einmal den Wilcoxon-Test macht, dann bekommt man sogar einen p-Wert von 0,03. – Beantwortet das Ihre Frage?

Herr Hecken (Vorsitzender): Frau Olberg? – Okay. Dann haben wir Frau Müller.

Frau Dr. Müller: Die Fragen, die ich danach hatte, wie das genau gelaufen ist, wurden jetzt eigentlich beantwortet. Sie haben also ein offizielles Amendment gemacht haben und auch einen statistischen Analyseplan, den Sie für multiples Testen korrigiert haben. Ist das richtig?

Herr Dr. Tomczkowski (Janssen-Cilag): Also, ich bin kein Statistiker. Das ist, glaube ich, nicht korrigiert für multiples Testen. Man hatte ja vorher festgelegt, dass man zwei Interimsanalysen machen möchte. Nachdem die Behörde eine weitere Analyse angefordert hatte, hat man gesagt: Okay, dann haben wir jetzt drei Interimsanalysen. Bei multiplem Testen würde man gucken, wie häufig man ei-

gentlich testen will, und dann eben entsprechend das Signifikanzniveau nach unten bringen. Aber hier war es eben nicht multiples Testen, sondern es gab eine zusätzliche Interimsanalyse, und man hat die gleiche Methode, die man für die Adjustierung der Interimsanalysen genutzt hat, ein weiteres Mal genutzt, um an dieses Signifikanzniveau zu kommen.

Herr Hecken (Vorsitzender): Ist das für Sie okay, Frau Müller, oder möchten Sie noch einmal nachfragen?

Frau Dr. Müller: Ich kann es im Moment nicht ganz nachvollziehen.

Herr Hecken (Vorsitzender): Ich habe es auch nicht so richtig verstanden.

Frau Dr. Müller: Es handelt sich also sozusagen um eine zusätzliche Interimsanalyse, die nicht geplant war. Die ursprünglich geplante Endanalyse wurde sozusagen – wie soll man sagen – umbewertet. Eine zusätzliche Analyse kam also auf jeden Fall dazu?

Herr Dr. Tomeczkowski (Janssen-Cilag): Richtig, ja. Es ist eine zusätzliche Analyse und es ist letztendlich auch bei der EMA der Diskussionspunkt, welche Aussagekraft sie hat. Sie ist zwar zusätzlich dazugekommen, war aber, nachdem man entschieden hatte, sie kommt dazu, vorab wiederum geplant mit einem entsprechenden Amendment. Aber man hatte natürlich schon einmal in die Datenbank hineingeguckt.

Festzuhalten ist aber – das möchte ich an dieser Stelle noch einmal betonen –, dass auch das Ergebnis aus 2009 signifikant sein kann, wenn man da den Wilcoxon-Test anwendet. Das ist zusätzliche Evidenz zum Zeitpunkt 2010. Wir können aber auch zum Zeitpunkt der Auswertung von 2009 schon Signifikanz sehen, wenn man diesen Wilcoxon-Test anwendet oder auch weitere Tests.

Herr Dr. Angermund (Janssen-Cilag): Vielleicht ganz kurz: Dieses Amendment wurde, wie besprochen, mit der EMA/CHMP abgestimmt. Damit ist auch der statistische Analyseplan und auch das Amendment einschließlich des statistischen Teiles und auch die zugrundeliegende Alpha-Wert-Signifikanz zu erreichen. Das Signifikanzniveau ist mit der EMA/CHMP besprochen und von diesen genehmigt worden als akzeptabel in diesem Falle und wurde dann entsprechend in einem statistischen Analysisreport mit den entsprechenden formalen Bestandteilen umgesetzt. Das wurde dann den Behörden vorgelegt und von diesen dann auch, weil es im Rahmen auch mit den Behörden abgestimmt war – das ist von denen entsprechend auch eingefordert worden –, umgesetzt und auch von der EMA in deren Bewertung akzeptiert.

Herr Hecken (Vorsitzender): Herzlichen Dank. – Ich habe jetzt Herrn Kulig, Herrn Carl und Herrn Wörmann. Herr Kulig, bitte.

Herr Dr. Kulig: Wir sammeln jetzt so langsam die Informationen. Trotzdem muss ich zum gleichen Punkt noch einmal nachhaken, weil es mir noch nicht ganz klar ist. Sie haben in Ihrer schriftlichen Stellungnahme auch ein Schaubild eingebracht, in dem der zeitliche Verlauf zu sehen ist. Sie haben eben gesagt, die primäre Studienplanung, worauf auch die Fallzahlplanung basierte, war der Auswertungszeitpunkt 2009, weil dann die Zahl der Todesfälle erreicht wurde, für die die primäre Fallzahlplanung gemacht wurde. Wurden die dann auch 2009 ausgewertet? Sie sagten ja, die Datenbank war offen und wurde erst im Mai 2010 geschlossen. Hat man die dann länger laufen lassen, weil man da dachte, ja gut, vielleicht gibt es doch noch Ereignisse, die eventuell eine Signifikanz ermöglichen? Oder wurde wirklich 2009 geschlossen und dann noch einmal im Mai 2010? Das ist mir noch nicht

klar. Und wann ist dann wirklich die 2010er-Analyse gelaufen? Wann ist da die Datenbank geschlossen worden? Noch einmal: Der zeitliche Ablauf – also Oktober 2009, Schaubild, Mai 2010, dann später noch einmal die Planung des Analyseplans – ist mir leider noch nicht so klar.

Herr Dr. Tomeczkowski (Janssen-Cilag): Oktober 2009 war der geplante Auswertungszeitpunkt. Das heißt aber nicht, dass ich dann in die Datenbank reingehe und sie öffne, sondern ich will zu diesem Zeitpunkt nur meine Ereignisse, also Todesfälle, auswerten. Geschlossen wurde die Datenbank erst im Mai 2010. Vorher hatte man also gar nicht die Möglichkeit, da hineinzugucken. Man hat das erst im Mai 2010 gemacht. Dann hat man ausgewertet zum Zeitpunkt 2009, so wie man sich das vorgenommen hatte. Aber jetzt war ja zwischen Oktober und Mai noch einmal ein halbes Jahr vergangen. Das heißt, wir hatten ja noch mehrere Ereignisse in dieser Datenbank. Man konnte jetzt sagen: Okay, wir gucken, was denn zum Zeitpunkt, wie es geplant war – 2009 – passiert ist. Und man hat gesehen: Es ist nicht signifikant. Weil aber in diesem halben Jahr noch weitere Patienten verstorben waren, hat man dann geguckt: Was würden wir denn kriegen, wenn Mai 2010 unser Auswertungszeitpunkt wäre? Das hat man sich auch angeguckt und hat gesehen, dass man einen p-Wert von 0,02 erreicht hatte. Das hat man der Behörde vorgelegt und hat gefragt, wie man damit umgehen soll. Und dann hat die Behörde gesagt: Okay, wenn das so ist, dann lassen wir die Studie jetzt weiterlaufen, dann haben wir noch mehr Todesfälle. Wir erhöhen die Power und werten dann noch einmal zum Zeitpunkt 2010 aus.

Man ist also nicht 2009 in die Datenbank reingegangen. Deswegen ist auch die eine Sensitivitätsanalyse, die auch positiv ist – das kann man nur schwer verstehen –, erst nach diesem ursprünglichen Auswertungszeitpunkt festgelegt worden. Aber geöffnet hat man die Datenbank erst im Mai 2010. Als man das gesehen hat, hat man sie komplett wieder geschlossen und bis Oktober 2010 gewartet. Die Datenbank wurde dann im Dezember 2010 erneut geöffnet; und da hat man alle Analysen gemacht, und zwar jetzt für zwei Auswertungszeitpunkte.

Herr Dr. Angermund (Janssen-Cilag): Noch einmal vielleicht zur Verdeutlichung: Wir haben den Clinical Cut-off – so nennen wir das – im Oktober 2009; das heißt, wir haben die Ereignisse bis zum Oktober 2009 bewertet. Das bedeutet aber nicht, dass im Oktober 2009 in der Datenbank schon alle Seiten des Dokumentationsbogens vorliegen. Da sind noch Fragen – „queries“ – offen, die zu klären sind. Das heißt, es dauert immer eine ganze Weile, bis dann die Datensätze zum Clinical Cut-off im Oktober 2009 in der Datenbank komplett sind, sodass man die Datenbank schließen kann. Diese Datenbank wurde dann im Mai 2010 geschlossen, und in dieser Zwischenzeit sind natürlich noch weitere Todesfälle gemeldet worden. So kam das zustande.

Im ersten Clinical Cut-off mit der entsprechenden Auswertung sind also alle Daten von Patienten enthalten, die das Ereignis, das Event, bis zum Oktober 2009 hatten. Wegen Rücklauf der Dokumentationsbögen, Eintragungen in die Datenbank, Klärung der offenen Fragen ist diese Datenbank für diese Analyse im Mai 2010 geschlossen worden und die Auswertung für Oktober 2009 gemacht worden. Des Weiteren wurde dann ein neuer Clinical Cut-off gesetzt – für die zweite Auswertung –, und zwar auf Basis Mai 2010; und da hat es bis Oktober gebraucht, bis die CAF-Seiten da waren und die Datenbank wieder geschlossen worden ist.

Herr Hecken (Vorsitzender): Direkt dazu Herr Eyding, dann Herr Carl, Herr Wörmann, Frau Müller.

Herr Dr. Eyding: Jetzt bin ich auch ein bisschen verwirrt. Sie haben jetzt ja eigentlich zwei Interimsanalysen. Sie haben die vom Oktober 2009 und die vom Mai 2010. Diese müsste man ja eigent-

lich auch als Interimsanalyse begreifen. Ist dafür auch adjustiert worden, oder ist das nicht der Fall gewesen? Hat man nur eine genommen?

Herr Dr. Tomeczkowski (Janssen-Cilag): Mai 2010 war keine offizielle Interimsanalyse. Das wüsste ich.

(Herrn Dr. Eyding: Sie haben die Daten!)

– Gut.

Herr Hecken (Vorsitzender): Gut oder nicht gut.

Herr Dr. Tomeczkowski (Janssen-Cilag): Ich weiß nicht, ob dafür adjustiert wurde. Ich kenne eben nur die beiden vorgesehenen Interimsanalysen und den Zeitpunkt 2009 als weitere Analyse, sodass man drei Interimsanalysen hat.

Herr Hecken (Vorsitzender): Okay. – Herr Carl, bitte.

Herr Dr. Carl: Ich würde gerne von Ihnen noch einmal wissen, warum Sie uns heute die Änderungen eines statistischen Signifikanztestes mitteilen und auf welcher mathematischen Grundlage das passiert. Also: Warum ist der Wilcoxon-Test in diesem Fall eher in der Lage, die Null-Hypothese abzulehnen? Warum war der Log-Rank-Test in dem Fall der falsche Test? Und wie sind Sie darauf gekommen?

Herr Dr. Tomeczkowski (Janssen-Cilag): Wir sind eigentlich über den Artikel von Professor Bender zum Log-Rank-Test darauf gekommen. Den Wilcoxon-Test haben wir ja schon im Dossier erwähnt. Die Rationale dafür, den zu machen, bzw. was wir aus dem Artikel gelernt haben, ist, dass der Log-Rank-Test die späten Ereignisse höher gewichtet, also genau dort, wo es das Cross-over, die Überkreuzung der Lebenszeitkurven, gegeben hat, was möglicherweise auch durch weitere Therapien verursacht worden ist, die hauptsächlich im Vergleichsarm gegeben wurden. Wenn jetzt der Log-Rank-Test gerade zum Ende der Ereignisse die Signifikanz zu ermitteln versucht, dann versucht er, zwei Aussagen zu treffen: einmal eben, wo Dacogen® besser ist, und einmal, wo Dacogen® zum Schluss hin schlechter ist. Und das kann er eigentlich nicht, also hat er gar keine Aussagekraft. So habe ich das zumindest gelernt. Der Wilcoxon-Test gewichtet dagegen mehr in der Mitte oder weiter vorne. Da sind die Kurven deutlich auseinander. Deswegen nimmt der Wilcoxon-Test, wenn Sie so wollen, eigentlich eine Zensierung der Patienten bzw. eine Zensierung der weiteren Therapien, die gegeben wurden, vor.

Das Gleiche haben wir auch mit einer Modellierung gemacht. Da kommen wir auf ähnliche Ergebnisse. Hier haben wir praktisch Patienten in der Analyse ein anderes Gewicht gegeben, wenn die Wahrscheinlichkeit für eine weitere Therapie steigt. Wenn die ein niedrigeres Gewicht bekommen und andere Patienten ein höheres Gewicht, dann kriegen Sie genau das gleiche positive Ergebnis. Diese IPCW-Methode, mit der man eben solche Cross-over-Ereignisse modellieren kann, macht also nichts anderes als der Wilcoxon-Test. Das haben wir mit Statistikern diskutiert und eben auch in dem Artikel von Professor Bender nachgelesen, der auch schreibt, dass der Log-Rank-Test in dieser Situation eigentlich keine Aussagekraft hat. Die Aussagekraft oder das Ergebnis zum 2009er-Zeitpunkt hängt eigentlich nur davon ab, wie der statistische Test gewichtet. Wir haben mit anderen Statistikern gesprochen, und die haben gesagt: Ja, dann guckt doch einmal, wenn ihr weitere Tests macht, ob die auch alle positiv sind. – Wir haben dann für die Statistiker auch noch Tests nach Tarone, Peto, modified-

Peto und Fleming gemacht; all diese sind zum Zeitpunkt 2009 positiv. Das haben wir zwar im Dossier nicht erwähnt, aber im Anhang zum Modul 5 sind all diese Tests zu finden.

Herr Hecken (Vorsitzender): Herr Carl, sofort dazu.

Herr Dr. Carl: Ich habe dazu eine Zusatzfrage an das IQWiG. Können Ihre Spezialisten für die Statistik diese Erklärungen fachlich unterstützen?

Herr Dr. Gerber-Grote: Wir haben beim IQWiG nur den Kostenteil und den Epidemiologieteil gemacht. Wir können das natürlich als Frage noch einmal mitnehmen. Aber wir sind ja heute nicht offiziell zum Nutzteile hier. Den haben wir nicht bewertet und nicht gelesen. Deswegen werden wir uns da heute eines Kommentars enthalten müssen.

Herr Hecken (Vorsitzender): Das ist auch korrekt so, weil das Fragestellungen sind, die wir intern im Unterausschuss diskutieren müssen, sollen und werden. – Wir gehen jetzt weiter. Herr Professor Wörmann, Frau Müller, Herr Bartmann. Herr Wörmann, bitte schön.

Herr Prof. Dr. Wörmann: Ich habe ein bisschen Sorge, dass wir hier in einem patientenfreien Raum diskutieren. Ich bin von der DGHO, aber ich habe über mehrere Jahre eine der großen deutschen AML-Studien betreut mit über 3.000 Patienten und war dort für die älteren Patienten zuständig. Das mittlere Alter für AML in Deutschland ist zurzeit 70 Jahre. Das heißt, mehr als die Hälfte der Patienten ist grundsätzlich für diese Art von Therapie qualifiziert. Die Diskussion, die wir haben, ist im Moment eigentlich nicht hauptsächlich, ob Decitabin ein gutes Präparat ist, sondern ob es einen kurativen Anspruch bei den Patienten gibt. Für diejenigen, bei denen das nicht so ist, ist Decitabin eine der Optionen.

Ich glaube – Sie dürfen das auch zitieren –, wenn nur die Überlebenszeit verlängert würde, wäre ich nicht sicher, ob ich dem Präparat einen Nutzen zusprechen würde. Zwei Monate Verlängerung würden mir nicht reichen. Es muss bei diesen Patienten um mehr und andere Dinge wie Lebensqualität gehen. Es handelt sich um akute Leukämie. Da kommen keine Patienten zum Arzt mit einem Knoten oder so. Vielmehr haben sie eine Sepsis oder eine schwere Entzündung oder eine Anämie und brauchen Bluttransfusionen oder bluten. Zur Diskussion kommt also hinzu: Kann ich noch eine lebensqualitätsbewertete Zeit herausholen, oder ist das bei den Patienten nicht mehr drin? Decitabin würden wir deswegen dort positionieren, weil die Möglichkeit besteht, eine Remission zu haben. Ich sehe auch, dass das, was wir bei einer Remission erleben, wenn einer also keine Bluttransfusion mehr braucht, keine Sepsis mehr hat und nicht blutet, in der Studie nicht gut abgebildet ist. Das ist manchmal auch nicht einfach, einen 80- oder 85-Jährigen, wenn Sie ihm solch ein Präparat geben, am Anfang einen Lebensqualitätsbogen ausfüllen zu lassen, um dann sehen zu können, ob es ihm wirklich besser geht.

Klinisch ist Remission ein hoher Wert für diese Leute. Ich glaube, die Position von Decitabin ist, dass das Medikament es schafft, bei einem Teil der Patienten Remissionen zu erzielen, also ein progressionsfreies Überleben, bei dem dann die Symptome nicht mehr auftreten. Wenn dann auch noch mit irgendeiner statistischen Methode eine signifikant höhere Überlebenszeit gemessen wird, ist das ein Plus. Das würde mir aber als einzelner Zusatznutzen nicht reichen. Ich würde mich dagegen wehren, hier nur über Überlebenszeit und nur über Statistik zu reden.

Herr Hecken (Vorsitzender): Herzlichen Dank, Herr Wörmann. Damit keine Missverständnisse aufkommen: Wir werden nicht nur über die Frage der Überlebenszeit reden. Diese 2,7 oder 2,3 oder 2,2 Monate sind nicht das alleinige Kriterium. Aber der pharmazeutische Unternehmer hat angeführt,

dass allein hieraus – insofern sehe ich da einen kleinen Widerspruch – ein beträchtlicher Vorteil und ein beträchtlicher Zusatznutzen zu generieren sind. Hierbei spielen noch eine Reihe von anderen Komponenten, die Sie ansprechen, eine Rolle: Wie ist die Compliance des Patienten? Welche Symptome hat er in Zukunft?

Wir sind jetzt zunächst beim Abschichten des ersten Teils. Wir diskutieren deshalb auch nicht im patientenfreien Raum, sondern wir sind jetzt bei Statistiken und Wahrscheinlichkeitsrechnung, um dann im zweiten Schritt zu folgenden Fragen zu kommen: Welche Auswirkungen hat die Gabe dieses Wirkstoffs auf den Zustand des Patienten? Mit welchen Nebenwirkungen ist dies verbunden? Welche Vorteile treten auf? Beides muss getrennt betrachtet werden, weil wir es auch nach der Arzneimittel-Nutzenbewertungsverordnung getrennt klassifizieren und bewerten. Ich nehme aber mit Interesse zur Kenntnis, dass Sie als doch sehr profunder Kenner der Materie sagen, dass allein die Verlängerung des Überlebens, egal ob es um 2 oder 2,7 Monate geht, für sich genommen kein glücklich machendes Faktum ist – ich formuliere es jetzt einfach einmal um –, sondern dass andere wesentliche Parameter, die die Situation des Patienten und seine Allgemeinbefindlichkeit während des Überlebens verbessern, hinzutreten müssen.

Wir gehen jetzt in der Rednerliste weiter. Frau Müller, Herr Bartmann, Herr Eyding, Frau Teupen.

Frau Dr. Müller: Herzlichen Dank für Ihre Ausführungen. Sie haben recht. Allein ein Überlebensvorteil, auch wenn es Overall Survival ist, ohne Betrachtung des Gesamtbildes kann noch nicht zu einem Urteil führen. Trotzdem ist es für uns wichtig, zu wissen, wo wir dabei stehen. Darüber zu diskutieren, wie wir das einordnen, ist ebenfalls ganz wichtig.

Dazu, wo wir beim Survival stehen, hätte ich noch eine letzte Frage an den pharmazeutischen Unternehmer – ich fasse mich kurz –, um das noch einmal klarzumachen; denn es ging aus meiner Sicht etwas durcheinander. Zu welchen Zeitpunkten gab es einen Data Cut-off? Welcher davon war geplant? Stichwort Database Lock: Wann war der erste Database Lock? Das hatten Sie, glaube ich, schon gesagt: im Mai 2010. Sie haben dann – so habe ich es verstanden – zwei Analysen gemacht. Die erste Analyse war eine geplante, die später in eine nicht geplante Interimsanalyse umgedeutet wurde, und die zweite war eine sowieso nicht geplante Interimsanalyse. Das war sozusagen ein Database Lock, aber zu zwei unterschiedlichen Zeitpunkten des Data Cut-off. Ist das richtig? Könnten Sie das noch einmal darstellen? Wie ging es dann weiter? Wie oft wurde die Datenbank entlocked, und zu welchen Zeitpunkten gab es einen Data Cut-off? Die Frage ist also: Bis zu welchem Zeitpunkt gucke ich mir die Daten an? Welche Zeitpunkte haben Sie für die verschiedenen Analysen gewählt? Das ist mir immer noch nicht ganz klar. Vielleicht könnten Sie ganz kurz etwas dazu sagen.

Herr Dr. Tomczkowski (Janssen-Cilag): Der erste Database Lock – das haben wir auch im Zeitstrahl in der Stellungnahme abgebildet – war am 5. Mai 2010. Vorgesehener Auswertungszeitpunkt war der 28. Oktober 2009. Für diesen Zeitpunkt wurde analysiert. Weil es zu dem Zeitpunkt weitere Todesfälle gab, wurde dann noch einmal auch zu dem Zeitpunkt analysiert. Dann wurde wieder verschlossen; der nächste Database Lock war im Dezember 2010. Man hat dann noch einmal für Oktober 2010 analysiert. Es gab also zwei Database Locks, im Mai 2010 und Dezember 2010, und zwei Auswertungszeitpunkte.

(Frau Dr. Müller: Zwei oder drei?)

- Es müssten drei sein.

Herr Hecken (Vorsitzender): Streiche zwei, setze drei. – Herr Bartmann.

Herr Dr. Bartmann: Ich ziehe meine Frage zurück.

Herr Hecken (Vorsitzender): Herr Eyding, bitte.

Herr Dr. Eyding: Ich habe noch eine Frage zum Überleben. Es tut mir leid, aber ich muss noch einmal darauf zurückkommen. Bei meiner Frage geht es um die regionalen Differenzen, die man gesehen hat. Man könnte jetzt ketzerisch sein und sagen: Wir haben mehr Todesfälle unter Behandlung mit Decitabin, wir haben mehr SUE, wir haben mehr Infektionen. Wir haben in Westeuropa, wo wir eine bessere Begleitbehandlung haben und wo mehr Patienten mit hypomethylierenden Substanzen im Vergleichsarm behandelt worden sind, keinen Überlebensvorteil. Wäre es nicht eigentlich die angemessene Strategie, nicht Decitabin zu geben, sondern nur dann, wenn man den Anlass sieht, auf Decitabin oder eine andere hypomethylierende Substanz zu wechseln, wenn also die Response verloren geht oder andere Gründe dafür sprechen, das Therapieregime zu wechseln? Denn eigentlich sieht man keinen Vorteil, sondern eher Nachteile bei den Aspekten, die Herr Wörmann zu Recht angesprochen hat, nämlich mehr SUE, mehr Todesfälle unter der Behandlung und mehr Infektionen bei Patienten unter Decitabin. Diese regionalen Differenzen finde ich interessant, weil es ja auch Erklärungen dafür gibt, wie die regionalen Differenzen zustande gekommen sind.

Herr Dr. Tomczkowski (Janssen-Cilag): Wenn Sie mehr Todesfälle, mehr SUE ansprechen, dann muss man sagen, dass das nicht signifikant ist, wenn man das wirklich direkt mit Cytarabin vergleicht und wenn man die längere Behandlungszeit berücksichtigt; für Dacogen[®] waren es ja 4 Zyklen vs. 2 Zyklen im Vergleichsarm. Wenn man das mitberücksichtigt, dann ist die Todesrate im Dacogen[®]-Arm niedriger; das gilt auch für die SUE. Im Vergleich zum Cytarabin gibt es keinen signifikanten Unterschied, und auch expositionszeitadjustiert gibt es keinen Nachteil. Die EMA hat ja im EPAR geschrieben, dass das Adverse Event Profile ausgeglichen ist.

Wenn Sie jetzt die Vorteile im TC-Arm für Westeuropa ansprechen, dann muss man sagen, dass dort überproportional viele Patienten im TC-Arm weitere Behandlung bekommen haben. Es waren 35 Patienten, und davon haben 50 Prozent noch eine Folgebehandlung bekommen, von der sie anscheinend profitiert haben. Wir haben auch überproportional viele Patienten mit einem sehr guten ECOG-Status, ECOG 0, gehabt. Auch das ist eher ein Vorteil für Cytarabin gewesen; denn bei diesen Patienten wirkte Cytarabin noch sehr gut. Dadurch sind die Ergebnisse für Westeuropa zustande gekommen.

Herr Dr. Angermund (Janssen-Cilag): Ich darf dazu ergänzen. Die Kaplan-Meier-Überlebenskurve für Westeuropa im TC-Arm ist deutlich besser im Verlauf als in Osteuropa, in Nordamerika und in Asien. Wenn Sie zustimmen, dass dieser Effekt – wir spekulieren nur; wir wissen natürlich nicht, was diesen Effekt bewirkt hat; wir nehmen es hypothetisch an – durch die Sekundärtherapie mit hypomethylierenden Substanzen zustande gekommen ist, dann hat dies zu einem bestimmten Zeitpunkt, wie Sie sagten, zum Beispiel im Progress, stattgefunden. In den ersten drei Monaten haben wir hier einen Abfall von 5, 10, 20 Prozent der Patienten; diese sind in den ersten zwei, drei Monaten gestorben. Für diese Patienten würde die Sekundärtherapie natürlich nicht in Betracht kommen, sodass man dann eigentlich schließen müsste: Wenn das Decitabin verantwortlich ist, wenn die sekundären hypomethylierenden Substanzen verantwortlich sind für die Überlebenszeitverlängerung, dann würden die Patienten, die es am Anfang bekommen haben, davon nicht profitieren.

Herr Hecken (Vorsitzender): Frau Teupen, bitte.

Frau Teupen: Die erste Frage hat sich schon fast erledigt; Sie sprachen ja von den Kaplan-Meier-Kurven. Diese werden Sie ja auch zur Proportionalitätsannahme, zum Aussuchen des Testes gemacht haben.

Ich habe noch eine andere Anmerkung, weil der Patientenvertretung Lebensqualität immer sehr wichtig ist. Sie haben ja auch den QLQ-C30 auf die Lebensqualität erhoben. Er ist ja nicht signifikant, auch wenn es natürlich eine offene Studie ist. Ich habe nun eine Frage zu der Fatigue-Subskala. Ich weiß es nicht genau, ich kenne sie nicht so detailliert, aber eigentlich gibt es ja drei Domänen in dem Fragebogen, die validiert sind. Wie gehen Sie mit den einzelnen Fatigue-Subskalen um? Vielleicht könnten Sie kurz etwas dazu sagen.

Herr Dr. Angermund (Janssen-Cilag): Die Lebensqualität mit dem EORTC QLQ-C30-Fragebogen zu messen ist eine recht komplexe Angelegenheit, weil es dafür keinen EORTC-Subscore der -Summenscore gibt, er vielmehr aus einzelnen verschiedenen Domänen innerhalb dieser 30 Fragen besteht. Die verschiedenen Domänen können Sie unterschiedlich zählen. Einige sind symptom-basiert, zum Beispiel Schmerz, Übelkeit, Erbrechen, Appetit, Magen-Darm-Problematik, andere sind nicht symptomorientiert, zum Beispiel Fatigue. Diese Domänen an sich sind auch validiert. Für die Fatigue sind die Fragen 10, 12 und 18 des EORTC-Fragebogens relevant. Es sind also insgesamt drei von diesen 30 Fragen. Es gibt andere Domänen. Zum Beispiel ist die globale Quality of Life unterschiedlich davon zu sehen. Das sind zwei andere Fragen, nämlich die Fragen 29 und 30 des EORTC-Fragebogens. Darüber hinaus gibt es noch weitere Domänen, zum Beispiel die soziale Situation, die finanzielle Situation des Patienten und, wie auch schon erwähnt, die Symptomatik.

Die Analyse des Ganzen ist sehr komplex. Nehmen wir einmal als Beispiel einen 75-jährigen AML-Patienten; diese Altersgruppe wurde von Herrn Wörmann auch schon angesprochen. Er ist wahrscheinlich Rentner. Nehmen wir einmal an, dass er verheiratet ist, ein Kind oder ein Enkelkind hat. Nun bekommt er eine akute myeloische Leukämie und leidet an den Symptomen Infektion, Anämie mit dem Fatigue-Syndrom. Die AML-Patienten, die zur stationären Aufnahme kommen, sind in der Regel schwerkranke Patienten. Betrachten wir jetzt einmal nur die Lebensqualität dieses Patienten. Er hat einerseits eine Fatigue. Ist das die gesamte Einschränkung der Lebensqualität, die dieser Patient hat? Mit Sicherheit nicht. Vielleicht hat er darüber hinaus Sorgen, dass seine Ehefrau, wenn er diese Erkrankung nicht überlebt, eventuell mit der Rente in finanzielle Schwierigkeiten kommt; das kann ihm schon Sorge machen. Das heißt, die finanzielle Dimension ist wichtig. Es könnte sein, dass er im sozialen Umfeld etwas isoliert wird. Das heißt, es gibt eine soziale Dimension, die bei der Gesamtlebensqualität ebenfalls eine Rolle spielt. Es gibt zum Beispiel die Rollenfunktion. Welche Rolle im Leben spielt dieser individuelle Patient? Nehmen wir auch hier ein Beispiel: Vor seiner AML kam er mit seinen 75 Jahren relativ gut zurecht und ist vielleicht mit seinem Enkelkind zum Fußballplatz gegangen. Das war für ihn die Rolle des Großvaters, die er jetzt mit der Erkrankung nicht mehr wahrnehmen kann. Führt das bei diesem Patienten auch zu einer Einschränkung der Lebensqualität? Dies wird dann in dem Rollenfunktionsscore, also einem Subscore des Lebensqualitätsfragebogens, abgebildet.

Kommt dieser Patient nun in eine komplette Remission, dann kann sich zum Beispiel durch die Transfusionseinheit die Fatigue-Symptomatik deutlich bessern. Viele andere Aspekte könnten dann ebenfalls besser werden, zum Beispiel die Rollenfunktion. Andere Dimensionen könnten sich dadurch auch verschlechtern, zum Beispiel könnte er mehr darüber nachdenken, wann er ein Rezidiv bekommt, und darüber, dass er sterben könnte. Auch könnte dann die finanzielle Dimension für diesen Patienten höhere Relevanz haben.

Das Problem bei unseren Auswertungen ist folgendes: Wir haben uns in der Auswertung auf zwei Kernfragen fokussiert. Das ist zum einen die Overall Quality of Life, also die Frage, wie die Lebensqualität im Ganzen ist, und das ist zum anderen die Fatigue, weil sie beispielsweise mit Anämie klinisch sehr eng korreliert. Das Problem, das wir bei dieser Auswertung haben – hier müssen wir in unserer Studiendurchführung besser werden –, liegt bei der Compliance. Wir haben Missing Values in einer Größenordnung von über 30 Prozent. Das bedeutet: Wir hätten es gerne ausgewertet, aber leider ist die Auswertung zu unsicher. In diesen Punkten hat unsere Auswertung keine signifikanten Ergebnisse gebracht. Aber, wie gesagt, wir haben Missing Values von über 30 Prozent. Wir sagen: Normalerweise ist unser Cut-off bei der Lebensqualitätsanalyse. Da müssen wir eine Compliance von mindestens 80 Prozent haben. Wir gehen teilweise mit unter 60 Prozent in diese Analysen, und das ist schwierig.

Herr Hecken (Vorsitzender): Herzlichen Dank, Herr Angermund. – Frau Olberg.

Frau Olberg: Ich habe jetzt noch eine Rückfrage, und zwar zur Auswahl der Domänen. Sie haben ja eben erläutert, dass Sie zwei Domänen sozusagen herausgepickt haben: Overall Quality of Life und Fatigue. Könnten Sie noch einmal kurz erläutern, wieso Sie keine anderen Domänen als Endpunkt definiert haben. Beispielsweise – das haben Sie ja eben erwähnt – wird dort auch der Schmerz abgefragt, oder es geht um Rollenfunktionen. Es geht also quasi darum, wie die Auswahl zustande gekommen ist und welche Fragen das genau betrifft. Das wäre die erste Frage.

Die zweite Frage betrifft nicht die Lebensqualität, sondern ist ganz allgemein gemeint. Eben habe ich aufgeschnappt, dass die Patientenpopulation in der Studie einen recht guten allgemeinen Zustand hat. Sie haben circa 75 Prozent der Patienten in der Studie, die den ECOG 0 und 1 abdecken. Gleichzeitig wären 6 Prozent der Patienten in der Studie noch für eine Standard-Induktionstherapie infrage gekommen. Jetzt lautet Ihr Indikationsgebiet aber: für die keine Standard-Induktionstherapie infrage kommt. Wie schätzen Sie das ein? Haben diese Patienten, die wahrscheinlich intensiver therapiert worden wären, weil sie die Therapie noch bekommen hätten, einen Einfluss auf die Ergebnisse? Warum haben Sie sozusagen den Ausschluss einer Standard-Induktionstherapie nicht als Kriterium definiert? - Dazu würde ich gerne noch einmal um Erläuterung bitten.

Herr Dr. Tomeczkowski (Janssen-Cilag): Bei Quality of Life sind alle zehn Domänen ausgewertet worden. Die Ergebnisse finden sich über alle Zyklen. Die Ergebnisse finden sich auch im Clinical Study Report. Es sind aber nur die beiden, Fatigue und Global Quality of Life, wie Herr Angermund schon ausführte – Fatigue mit drei Fragen und Global Quality mit zwei Fragen – vorher im statistischen Analyseplan festgelegt worden; da wollte man die Veränderung von Baseline zu Zyklus 3 ermitteln. Aber ausgewertet ist alles.

Herr Hecken (Vorsitzender): Herzlichen Dank. – Frau Teupen.

Frau Teupen: Ich danke Frau Olberg noch einmal für die Konkretisierung meiner Frage. Das hatte ich eigentlich auch so gemeint.

Noch ein Hinweis: Die Patientenvertretung bevorzugt eigentlich den Begriff „Adherence“ anstelle von Compliance. Darauf wollte ich noch einmal hinweisen.

Herr Hecken (Vorsitzender): Frau Olberg erinnert an die Beantwortung der zweiten Frage.

Herr Dr. Tomeczkowski (Janssen-Cilag): Ich hatte eben gesagt, dass es gute Risiken gab, aber damit meinte ich nicht „in der Studie“, sondern „für Westeuropa“. Dort gab es überproportional gute

Risiken. In der Studie selber war eine Hochrisikogruppe. Dort hatten 36 Prozent eine Hochrisiko-Zytogenetik, 26 Prozent einen ECOG-2-Status und 35 Prozent eine schwer zu behandelnde sekundäre AML als Folge einer Vorerkrankung wie MDS oder ausgelöst durch Chemotherapie etc. Die Population selber war schon eine Hochrisikogruppe. Nur in Westeuropa gab es überproportional viele gute Risiken, wovon dann überproportional der Vergleichsarm profitiert hat. Deswegen sehen wir dort ja keine eindeutigen Ergebnisse.

Dann zu Ihrer Frage der intensiven Chemotherapie: Das ist richtig. Wir haben im Nachhinein die ECOG-Kriterien angewendet und sind auf 6 Prozent Patienten gekommen, die eine intensive Chemotherapie hätten erhalten können. Wir haben diese dann von der Auswertung ausgeschlossen, haben aber die gleichen Effektstärken bekommen. Die Effektstärken waren 5 Monate Überleben im Vergleichsarm und 7,7 im Dacogen[®]-Arm, also zusätzlich 2,7; wenn man noch die für intensive Chemotherapie geeigneten Patienten herausrechnet, dann erhält man 4,8 zu 7,5. Auf der anderen Seite haben die Patienten, die für eine intensive Chemotherapie geeignet waren, in der Tat länger gelebt, aber hier haben die Patienten im Dacogen[®]-Arm genauso davon profitiert wie die nicht für eine intensive Chemotherapie geeigneten Patienten. Dort war der Unterschied 6,8 Monate zu 12 Monate. Auf der anderen Seite: Wenn die Patienten wirklich noch eine intensive Chemotherapie bekommen haben – das waren immerhin 10 Prozent pro Arm –, dann gab es eine sehr hohe Anfangsmortalität von 18,9 Prozent. Normalerweise liegt die Anfangsmortalität bei hypomethylierenden Substanzen niedriger; bei uns lag sie nur bei 9 Prozent. Das heißt, wenn Sie wirklich eine intensive Chemotherapie geben, dann erkaufen Sie das auch durch eine höhere Anfangsmortalität. In der Folge haben die Patienten, die eine bekommen haben, dann aber nicht gelebt. Das waren dann im Vergleichsarm 3 Monate und im Dacogen[®]-Arm 3,4 Monate.

Herr Hecken (Vorsitzender): Frau Olberg, Sie haben eine Nachfrage.

Frau Olberg: Sie hatten eben erwähnt, Sie haben die Patienten herausgerechnet. Jetzt die Frage: Sind die Patienten zu den vorliegenden Auswertungen 2009/2010 herausgerechnet gewesen?

Herr Dr. Tomczkowski (Janssen-Cilag): Nein, da sind sie mit drin. Es gibt diese zensierten Analysen; das ist richtig. Es gibt zum Zeitpunkt 2009 und 2010 zensierte Analysen, aber nicht dort, wo die für intensive Chemotherapie Geeigneten herausgerechnet wurden, sondern es wurden diejenigen herausgerechnet, die noch weitere Therapien bekommen haben, oder dann noch einmal speziell diejenigen, die nur hypomethylierende Substanzen erhalten haben.

Es gibt also zu zwei Auswertungszeitpunkten jeweils drei Analysen: einmal ohne Zensierung und zweimal mit Zensierung. Und von diesen sechs Analysen ist eigentlich nur die erste nicht signifikant gewesen mit Log-Rank. Alle anderen Analysen sind signifikant; und die erste ist auch mit Wilcoxon signifikant.

Herr Hecken (Vorsitzender): Herzlichen Dank. – Jetzt habe ich noch eine Wortmeldung von Frau Günther und eine von Herrn Eyding. Dann würde ich vorschlagen, dass Sie als pharmazeutischer Unternehmer noch kurz bei einer der Beantwortungen anreißen, was es aus Ihrer Sicht für Probleme bei der Zielpopulation – also IQWiG: Zahlen zu niedrig angenommen – gibt. Dann würde ich gerne noch kurz auf die Dinge, die Herr Wörmann vorgetragen hat, auf die Dinge, die der Mitbewerber vorgetragen hat, und auf das, was vfa vorgetragen hat, eingehen. Ich glaube nämlich, wir drehen uns jetzt so ein bisschen im Kreis. Wir können noch bis heute Nachmittag um 15 Uhr Exegese betreiben, ohne dass am Ende eine weitergehende Klarheit kommt. – Also, Frau Günther und dann Herr Eyding.

Frau Günther: Ich habe in der Tat eine Frage zur Zielpopulation. Sie haben ja in der Stellungnahme die Zielpopulation deutlich nach oben korrigiert, auch unter der Annahme, dass jetzt nicht mehr 73 Prozent, sondern nur noch 49,5 Prozent für eine Standard-Induktionstherapie infrage kommen. Deswegen kommen jetzt deutlich mehr Patienten für Decitabin infrage. Wie lässt sich das erklären, dass der Anteil der Patienten auf einmal viel niedriger geworden ist? Wie definiert sich denn „für eine Standardtherapie nicht mehr infrage kommen“, und warum sind das jetzt offenbar andere Kriterien als vorher, da das ja um ein Drittel gesunken zu sein scheint?

Herr Dr. Fleischmann (Janssen-Cilag): Ich habe noch einen Kommentar zu den Patientenzahlen, die wir in der Stellungnahme angegeben haben. Wir hatten in unserem Dossier eine Zielpopulation von jährlich 598 Patienten geschlussfolgert auf Basis dieser Quote von 73 Prozent. Wenn ich den IQWiG-Bericht richtig verstehe, war das so weit in Ordnung bis auf die Unsicherheit, die bei diesen 73 Prozent herrschte. Deshalb wurde sozusagen ein Range um diesen Wert herumgelegt, sodass sich die Zahl dann auf 306 bis 781 Patienten beläuft.

Wir hatten jetzt in unserer Stellungnahme lediglich noch einmal auf einige frühe Auswertungen aus dem Register hingewiesen, die darauf hindeuten – wobei da natürlich sehr viele verschiedenste Einflüsse reinspielen können; das Medikament ist ja erst seit kurzem im Markt, und es war auch die Weihnachtszeit dazwischen –, dass sich möglicherweise diese 73 Prozent verändern. Damit deutet sich eine Tendenz an, dass diese 781, der obere Wert, der im IQWiG-Bericht genannt wird, noch einmal überschritten werden könnten. Ob das wirklich so ist, das wissen wir zum jetzigen Zeitpunkt nicht. Das muss sich, glaube ich, durch weitere Analysen bestätigen. Aber das war einfach für uns aus Sicht dieser Analyse aus dem Register wichtig, dass wir hier darauf hinweisen: Möglicherweise wird dieser obere Wert von 781 noch einmal überschritten.

Herr Hecken (Vorsitzender): Herzlichen Dank. – Beantwortet das Ihre Frage, Frau Günther?

Frau Günther: Noch eine kurze Nachfrage: Mir ging es noch um „für eine Standard-Induktionstherapie nicht infrage kommen“. Wenn man dafür vorher infrage gekommen ist, warum ist jetzt, wo es Decitabin gibt, auf einmal die Situation, dass man dafür nicht mehr infrage kommt? Welche Kriterien sprechen also dafür, dass man für eine Standard-Induktionstherapie infrage kommt?

Herr Hecken (Vorsitzender): Herr Wörmann, wir müssen zuerst den pharmazeutischen Unternehmer lassen, und dann dürfen Sie.

(Herr Prof. Dr. Wörmann (DGHO): Selbstverständlich!)

Herr Dr. Tomeczkowski (Janssen-Cilag): In der Tat sehen wir, was die intensive Chemotherapie angeht, einen Rückgang von 10 Prozentpunkten, von 60 auf 49 Prozent. Wir können auch nur spekulieren, warum das so ist. Wir haben das mit dem Autor der Studie auch diskutiert, warum er glaubt, dass das jetzt zurückgegangen ist. Er hatte das vorher auch so ein bisschen angekündigt. Deswegen haben wir eben noch eine weitere Analyse gemacht und geguckt, ob das in der Tat so eintritt. Er hat gesagt, dass aufgrund zu wenig zugelassener Medikamente die Patienten vielleicht doch zu häufig intensive Chemotherapien bekommen und dass sich das möglicherweise jetzt verändert.

Ich hatte ja schon gesagt: Einen kurativen Ansatz über eine intensive Chemotherapie erkauft man sich gerade bei Älteren immer über eine hohe Mortalität. Es ist sicherlich immer eine Abwägungssache, auch des Arztes, zu entscheiden: Mache ich eine intensive Chemotherapie oder nicht? Das Risiko – das haben wir ja in der Studie gesehen, und ich habe es auch gerade gesagt – von Patienten, die eine intensive Chemotherapie bekommen haben, innerhalb der ersten 30 Tage zu versterben,

liegt bei 20 Prozent. Der Arzt muss ständig abwägen, ob er dieses Risiko eingeht. Vor dem Hintergrund weiterer zugelassener Medikamente nimmt er wahrscheinlich eine andere Risikoabwägung vor.

Herr Hecken (Vorsitzender): Herzlichen Dank. – Ergänzend, Herr Wörmann.

Herr Prof. Dr. Wörmann (DGHO): Da muss ich widersprechen. Entscheiden tut immer der Patient. Vielleicht sollte Herr Rambach die Frage beantworten, ob das so ist. Die Situation ist: Patienten werden aufgeklärt, welche Möglichkeiten sie haben. In der Vergangenheit war die Alternative: entweder intensive Chemotherapie oder Abwarten. Das ist für Patienten schwer aushaltbar. In dem Moment, wo ein ambulant applizierbares Medikament da ist, entscheiden sich mehr Patienten für eine ambulante Version, wenn die Endergebnisse einigermaßen identisch sind. Insofern hat eine ambulant zu gebende Therapie durchaus einen Vorteil für Patienten.

Dazu kommt als Zweites: Wir lernen zunehmend, welche älteren Patienten nicht von einer intensiven Therapie profitieren. Das ist ein Lerneffekt der letzten fünf bis sieben Jahre, der sich bei den Ärzten zunehmend durchsetzt. Deswegen auch mehr die Beratung, Patienten nicht mehr intensiv zu behandeln, wenn wir keine besseren Medikamente haben.

Herr Hecken (Vorsitzender): Herzlichen Dank, Herr Wörmann. – Jetzt haben wir noch Herrn Eyding.

Herr Dr. Eyding: Ich habe eine kurze Frage an Herrn Wörmann. Ich hadere immer noch ein bisschen mit den Westeuropa-Daten. ECOG 0 ist jetzt als eine Erklärung angeboten worden. Sind denn ECOG-0-Patienten diejenigen, die in der Praxis mit der Erkrankung am ehesten in Westeuropa auftauchen, das heißt, sind die Daten anwendbar, so wie sie sind, oder ist es eher nicht so? Kommen die eher in einem schlechteren Zustand?

Herr Prof. Dr. Wörmann (DGHO): Auch ich hadere mit den Westeuropa-Daten. Ich habe auch in unserer Stellungnahme gesagt: Wir sind nicht sicher, wie weit die Studie wirklich auf Deutschland komplett übertragbar ist. ECOG-0-Patienten mit einer AML habe ich lange nicht mehr gesehen.

Herr Hecken (Vorsitzender): Herzlichen Dank. – Gibt es jetzt noch weitere unmittelbare Fragen an den pharmazeutischen Unternehmer? Ansonsten würden wir das, was Herr Wörmann für die DGHO vorgetragen hat, und das, was der zweite pharmazeutische Unternehmer und der vfa in ihren Stellungnahmen vorgetragen haben, jetzt noch kurz durchgehen. Ich würde vorschlagen, wir fangen mit Herrn Wörmann an. – Herr Bartmann, Sie hatten sich noch gemeldet. Entschuldigung.

Herr Dr. Bartmann: Ich hätte noch eine Frage an Herrn Wörmann. Ich komme dabei noch einmal auf diese Studie zurück. Das Cytarabin wurde ja in einer bestimmten Dosierung gegeben. Ich wollte, weil das auch erwähnt wurde, nachfragen, inwieweit die Gabe von 20 mg der Praxis entspricht.

Herr Prof. Dr. Wörmann (DGHO): Es gibt 20 oder 40 mg. Es gibt eine randomisierte Studie zu Cytarabin gegen Hydroxyharnstoff, aber keine zu den Dosierungen. Viele Ärzte, die Patienten in einem etwas besseren Zustand haben, neigen dazu, 40 mg/m² zu geben. Das Risiko, wenn man in dieser Studie 40 statt 20 gibt, ist allerdings, dass die Rate unerwünschter Ereignisse unter 40 mg höher ist als unter 20 mg. Insofern ist es ein bisschen Kaffeesatzleserei, weil wir die Studie nicht haben.

Herr Hecken (Vorsitzender): Damit sind, glaube ich, auch schon zwei Punkte angesprochen, die Sie in Ihrer Stellungnahme für die DGHO erwähnt hatten. Zum einen war die ECOG-0-Studie angesprochen worden. Sie hatten sich zur Übertragbarkeit der Studiendaten auf die Versorgungssituation in Deutschland eingelassen. Sie hatten angesprochen, dass es hier Unterschiede bei der Prognose ge-

ben könnte. Dann – eben hatten wir es auch schon kurz angerissen – hatten Sie auch Anmerkungen zur Häufigkeit der Erkrankungen gemacht und die strenge Altersbegrenzung von 60 und älter aus ärztlicher Sicht infrage gestellt. Sie hatten sich auch zur Stellung von Decitabin im Versorgungsgeschehen insgesamt eingelassen. Sie haben eben wiederholt, dass es zur Behandlung von MDS- und AML-Patienten wirksam ist und dass die Ergebnisse bei AML und MDS vergleichbar sind, dass es zu einer Verbesserung der Remissionsrate und zu einer Verlängerung der Überlebenszeit kommt. Sie hatten sich zur Dosierung eingelassen und hatten eben auch gesagt, dass höhere Dosierungen höhere Remissionsraten, aber auch höhere Toxizität zur Folge haben. Dies ist relativ logisch; das ist selbst für mich als Jurist eindeutig nachvollziehbar, ohne vorher noch ein medizinisches Proseminar besucht zu haben.

Ich frage jetzt die Bänke: Gibt es zu dem, was Herr Professor Wörmann in seiner schriftlichen Stellungnahme vorgetragen hat, ergänzende Fragen über das hinaus, was hier schon beantwortet worden ist? Ich glaube, wir sind hier auf alle Punkte eingegangen. – Das ist nicht der Fall. Herr Wörmann, möchten Sie noch ergänzend etwas vortragen?

(Herr Prof. Dr. Wörmann verneint)

Dann haben wir die Stellungnahme von Celgene. Celgene übt Kritik an den Aussagen im Dossier des pharmazeutischen Unternehmers. Dabei geht es zum einen um die Wertung, Decitabin sei weitaus potenter als Azacitidin als erste zugelassene Therapieoption, zum anderen werden die Unterschiedlichkeiten bei den Kosten herausgearbeitet. Ich frage die Bänke – wir haben das ja alles gelesen –: Gibt es dazu Fragen? – Das ist nicht der Fall. Möchten Sie ergänzend noch etwas vortragen? Sie haben es ja schriftlich vorgetragen. Ich glaube, das, was Sie schreiben, ist nachvollziehbar, und es wird von uns auch überprüft werden.

Dann frage ich den vfa: Gibt es aus Ihrer Sicht noch etwas, das vorzutragen wäre? Sie hatten ja wie üblich gemeinhin kritisiert, dass wir im Verfahren der frühen Nutzenbewertung von Orphan Drugs im Dossier, das von der Geschäftsstelle als Entwurf gefertigt und vom Unterausschuss zum Gegenstand des Stellungnahmeverfahrens gemacht wird, keine eindeutige Klassifizierung mit einer möglichen „Endbenotung“ haben. Wir hatten schon mehrfach darüber diskutiert. Wir sagen: Es gibt hierzu zum einen keine Rechtspflicht, zum anderen ist es auch der besonderen Situation der Orphan Drugs geschuldet, bei denen ja kraft Gesetzes ein Zusatznutzen generell schon als belegt gilt und bei denen man dann gerade im Stellungnahmeverfahren wirklich in den Feinabschichtungen diskutieren muss.

Das ist anders als bei den klassischen Bewertungsverfahren, bei denen möglicherweise dann auch kein Zusatznutzen am Ende steht. Vor diesem Hintergrund muss für einen vernünftigen Diskurs und für eine vernünftige Argumentation zumindest einmal in eine bestimmte Richtung Klarheit bestehen, wenngleich – das möchte ich an dieser Stelle auch betonen; das habe ich Ihnen ja auch beim letzten Mal gesagt, weil wir es ja regelmäßig auf der Tagesordnung haben und die jeweiligen Antworten schon entsprechend einrücken können – dem, was das IQWiG in seiner Nutzenbewertung, in seinem Dossier schreibt, zwar kraft der Rechtsprechung eine gewisse Richtigkeitsvermutung zukommt, dies aber keineswegs bindend für den Gemeinsamen Bundesausschuss ist. In mehreren Fällen sind wir nach oben und nach unten davon abgewichen. Dies ist also sicherlich ein wertvoller Anhaltspunkt, aber keine verbindliche Festlegung, auf deren Basis rechtsicher von Ihnen argumentiert werden kann. Das war im Wesentlichen das, was Sie vorgetragen hatten. Ich gebe Ihnen gerne Gelegenheit, das hier noch vertieft darzustellen. Sie können darauf auch verzichten. Beides wäre uns gleichermaßen recht oder würde uns erfreuen.

Herr Dr. Dintsios (vfa): Ich mache es sehr kurz. Ich kann die rechtlichen Erläuterungen, die Sie hier vorgetragen haben, nachvollziehen. Ich beziehe mich in unserer Stellungnahme nur auf die Schwierigkeit, dann auch entsprechend Stellung zu nehmen. Das macht die Sache nicht einfacher.

Es gibt trotzdem einen Punkt, an dem ich festhalten möchte. Dabei geht es um die Annahme des IQWiG, dass eine Studie, die aus dem Jahr 1999 stammt und damit älter ist, nicht einbezogen wird. Dafür gibt es zwei Gründe. Das IQWiG nennt die Unsicherheit der Daten. Diesen Punkt kann ich akzeptieren. Das Alter reicht mir aber bei einer Studie mit epidemiologischen Auswertungen als Begründung nicht aus. Bei uns ist die Datenlage schlecht; in Deutschland ist die Epidemiologie aus bekannten Gründen nun einmal nicht wie in Skandinavien. Ich verstehe diese Vorgehensweise nicht so ganz. Ich würde ältere Studien nicht generell ausschließen. Die Epidemiologie ändert sich nicht so schnell, dass man innerhalb von zehn Jahren ein anderes Krankheitsbild hat oder dass sich die Bevölkerungsanteile innerhalb von zehn Jahren rapide oder massiv verändern. So ist der Genpool in Deutschland nicht veränderbar. Also das missfällt mir; das muss ich sagen.

Das Zweite ist: Das IQWiG hat einen Stresstest gemacht mit 10 Prozent nach oben und nach unten. Wie Sie sehen, entsteht so die Spanne von 73 auf 80 Prozent und von 73 runter auf 66 Prozent. Das muss man dann methodisch allerdings auch begründen. Man kann Stressanalysen auch mit der Hälfte und mit Verdopplung machen: mit 5 Prozent, mit 1 Prozent, mit 0,5 Prozent. Ich hätte gerne – das habe ich in der Stellungnahme auch ein bisschen moniert – die entsprechende Herleitung.

Übrigens habe ich heute zum ersten Mal erfahren, dass man sich im IQWiG im Rahmen von Orphan Drugs, wenn das IQWiG nur die Zielpopulation und die Kosten zu bewerten hat, die anderen Module nicht anschaut. Ich persönlich muss sagen, dass ich mir diese schon aus eigenem Interesse, aus Neugier anschauen würde. Das hat mich schon ein bisschen verwundert.

Herr Hecken (Vorsitzender): Sie sollten Ihre persönliche Neugier nicht an die Stelle der Regelungen in der Verfahrensordnung und Geschäftsordnung des Gemeinsamen Bundesausschusses setzen. Wir sind auch alle neugiergetriebene und neugiergenerierte Wesen. Faktum ist aber, dass wir ausdrücklich festgelegt haben, dass die Bewertung und die Dossierbewertung bis auf die zwei in Rede stehenden Aspekte eben beim Gemeinsamen Bundesausschuss im Unterausschuss Arzneimittel erfolgt. Es gibt jenseits der Frage der Neugier sicherlich auch das Faktum, dass a) Zeit begrenzt ist und dass b) Wissen nur dann, wenn nötig, gewonnen werden kann; es geht ja auch um Unterlagenschutz. Ob der eine oder andere heimlich etwas liest, ist etwas anderes, als ob er es offiziell bewertet. Zur Bewertung ist – diese Option lässt das Gesetz ausdrücklich offen – der Unterausschuss Arzneimittel berufen. Erste Feststellung.

Zweite Feststellung. Ich nehme mit Befriedigung zur Kenntnis, dass Sie meine rechtlichen Betrachtungen zumindest als nachvollziehbar bezeichnet haben. Das ist mehr, als ich gemeinhin an einem solchen Tag erwarten kann.

Drittens. Sie hatten eine methodische Frage an das IQWiG gestellt, die Herr Gerber-Grote Ihnen jetzt gerne beantworten wird. Bitte schön, Herr Gerber-Grote.

Herr Dr. Gerber-Grote: Danke, Herr Hecken. – Es geht um die Studie Hellenbrecht et al. (2003), die im Dossier auf Seite 44 zitiert wird. Da sagt ja selbst der pharmazeutische Hersteller, da es sich um eine Studie handelt, in der Daten aus den USA von 1995 bis 1999 auf die deutsche Bevölkerung hochgerechnet werden, dass dies die Limitation ist, die in dieser Studie herrscht, also dass regionale Faktoren etc. – all dies wird in dem entsprechenden Ansatz genannt – eben nicht bewertet werden können und die Studie deswegen nicht herangezogen werden kann. Ich kann Ihnen insofern recht geben, dass das Argument der Aktualität hier vielleicht etwas verkürzt dargestellt war. Wir haben aber

darauf verzichtet, diesen Absatz zu zitieren, in dem der pU ja selber die weiteren Argumente beibringt.

Das Zweite ist sicherlich eine längere methodische Diskussion. Wie geht man damit um, wenn man entsprechend Unsicherheit in der Studie hat und diese in irgendeiner Form als Sensitivitätsanalyse hier vorlegen muss? Es ist bekannt, dass man auch Meta-Analysen zu epidemiologischen Daten machen kann; dies wird aber derzeit noch sehr problematisch gesehen. Deswegen verzichten wir an den Stellen, wo vielleicht mehrere Studien zusammenzufassen sind, darauf, diese Meta-Analyse zu machen. Wenn natürlich nur eine Studie vorliegt, müssen wir in irgendeiner Form pragmatisch vorgehen. Ich denke, eine 10-Prozent-Regel ist ein sinnvolles pragmatisches Vorgehen.

Herr Hecken (Vorsitzender): Herzlichen Dank. Aus meiner Sicht sind alle Fragen gestellt. Ich sehe auch keine Wortmeldungen mehr. – Dann bedanke ich mich ganz herzlich bei Ihnen allen, die Sie an der heutigen Anhörung teilgenommen haben. Wir werden das, was im schriftlichen Verfahren vorgetragen wurde, und die neuen Erkenntnisse, die heute vorgetragen worden sind, bewerten – Sie werden auch im Wortprotokoll der heutigen Anhörung wiederfinden, was Sie hier gesagt haben – und dann zu einer Entscheidungsfindung kommen, die hoffentlich alle vorgetragenen Argumente in der angemessenen Art und Weise aus Sicht einer Optimierung der Patientenversorgung und unter Berücksichtigung der wissenschaftlichen Evidenz am Ende inkludieren wird.

Danke, dass Sie da waren. Guten Heimweg!

Schluss der Anhörung: 11.36 Uhr