

MDK

**Medizinischer Dienst der
Krankenversicherung Nordrhein**



**MDK
Nordrhein**

Medizinischer Dienst der Krankenversicherung Nordrhein
Geschäftsführung: W. Machnik, PD Dr. med. H.P. Buszello

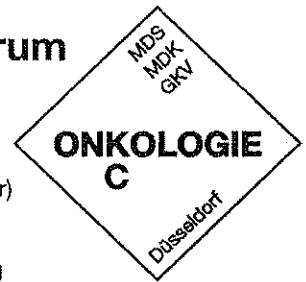
MDK Nordrhein • KC Onkologie • Postfach 10 37 44 • 40028 Düsseldorf

Kompetenz Centrum Onkologie

Prof. Dr. med. A. Heyll (Leiter)
Dr. med. K.-P. Thiele (stellv. Leiter)
Dr. med. B. Zimmer MPH
Dr. med. T. Wehkopf MSc
Dr. med. A. Niederste-Hollenberg
Dr. med. P. Schüller

Postanschrift: Hausanschrift:
Postfach 10 37 44 Bismarckstr. 43
40028 Düsseldorf 40210 Düsseldorf

Telefon: 02 11 13 82 - 451, - 452, - 459
Telefax: 02 11 13 82 - 461
E-Mail: info@kconkologie.de
internet: www.kconkologie.de



Ihr(e) Zeichen / Nachricht vom	Unsere E-Nummer	Unser Zeichen	Auskunft erteilt	Datum
	16147/2010	260.01 He / is /ss	Prof. Dr. Heyll	18.11.2010

Standpunktpapier des KC Onkologie zur frühen Nutzenbewertung onkologischer Arzneimittel

29.11.2010

Prof. Dr. med. Axel Heyll
Dr. med. Klaus-Peter Thiele
Kompetenz Centrum Onkologie der Medizinischen Dienste

Inhaltsverzeichnis

1	Beleg des Nutzens durch geeigneten Endpunkt.....	2
1.1	Auswahl der Endpunkte	2
1.2	Relative Bewertung von Überlebenszeitverlängerungen	6
2	Versorgungsrelevanz des Studienkonzepts.....	7
3	Sicherheit der Erkenntnisse	9
4	Zusammenfassung.....	12
5	Zusätzliche Anregungen des KC Onkologie	13
5	Literatur	14

1 Beleg des Nutzens durch geeigneten Endpunkt

1.1 Auswahl der Endpunkte

Behandlungsziel bei onkologischen Patienten ist in der Regel eine Überlebenszeitverlängerung, im Idealfall eine Heilung. Weitere mögliche Behandlungsziele sind – insbesondere bei palliativen Behandlungen mit ungünstiger Prognose – Verringerung der Toxizität der Therapie und/oder Linderung von Symptomen der Tumorerkrankung.

Aus diesen Therapiezielen ergeben sich zwangsläufig die für eine Nutzenbewertung geeigneten Endpunkte, nämlich Überlebenszeit, Toxizität und Lebensqualität.

Überlebenszeit (OS „overall survival“)

Die Vorteile der Überlebenszeit sind neben einem unmittelbaren Bezug zum Patientennutzen ein geringer Aufwand bei der Erfassung und eine absolut eindeutige Erhebung aufgrund der klaren Definition (nur Tod zählt als Ereignis). Zudem ist eine Verblindung in der Regel nicht notwendig.

Trotz dieser Vorteile wird Überlebenszeit als Endpunkt von vielen pharmazeutischen Unternehmen bei Zulassungsstudien mit folgenden Argumenten abgelehnt:

- *Die notwendige Patientenzahl ist zu hoch.*
Dieses Argument ist nicht nachvollziehbar. Um z. B. einen Anstieg von 10 % (von 50 % auf 60 %) im 1-Jahres-Überleben mit einer „Power“ von 0,8 erkennen zu können, sind ca. 840 Patienten notwendig (Di Leo A et al., 2003). Diese Patientenzahl kann bei häufigen onkologischen Erkrankungen und Indikationen (z. B. adjuvante oder palliative Therapie beim Mammakarzinom oder kolorektalem Karzinom) allein in Deutschland ohne Probleme in einer multizentrischen Studie in einem Jahr rekrutiert werden, da die Inzidenz vielfach höher ist. Aber auch bei seltenen onkologischen Erkrankungen kann durch internationale Kooperationen die notwendige Patientenzahl in einer angemessenen Rekrutierungszeit erreicht werden. Nachdem z. B. von der europäischen Arzneimittelbehörde EMA die Zulassungserweiterung von Lenalidomide für die seltene Unterform des myelodysplastischen Syndroms (MDS) mit 5q-Veränderung und günstiger Prognose (IPSS low oder intermediate¹) auf Grundlage einer Phase II-Studie abgelehnt worden war, hat der Hersteller (Celgene) durch internationale Kooperation eine Phase III-Studie initiieren können, die angemessen rekrutiert, wie auf dem US-amerikanischen Hämatologenkongress (ASH) 2009 berichtet wurde (Fenaux P et al., 2009).
- *Die notwendige Nachbeobachtungszeit ist zu lang.*
Bei palliativen Behandlungen von Patienten im fortgeschrittenen Tumorstadium (meist Karzinome mit Fernmetastasen) ist dieses Argument nicht nachvollziehbar. Die mediane Überlebenszeit der Patienten beträgt in der Regel nicht mehr als 1 bis 2 Jahre und Ereignisse (Todesfälle) sind häufig von Beginn der Beobachtung an. Stichhaltig ist dieses Argument nur bei Patienten mit einer günstigen Prognose und einer potenziell kurativen Therapie, z. B. bei Patienten mit lokalisiertem, radikal operiertem Kolonkarzinom, bei denen verschiedene Formen der adjuvanten Therapie erprobt werden. In diesem Falle ist es akzeptabel, als primären Endpunkt der Studie krankheitsfreies Überleben (DFS) auszuwählen. Dabei wird neben Tod auch ein Krankheitsrezidiv als Ereignis gewertet. In der Regel überträgt sich bei diesen Patientenkollektiven ein Vorteil im krankheitsfreien Überleben in einen Vorteil im Gesamtüberleben (André T et al., 2009). Bei diesen Indikationen besteht auch eine deutlich engere Korrelation zwischen DFS und Überlebenszeit als bei palliativen Indikationen zwischen RFS (progressionsfreiem Überleben) und Überlebenszeit (Sargent DJ et al., 2007). Letztendlich handelt es sich

um einen Kompromiss, um mit möglichst sicheren Aussagen einen Therapiefortschritt in möglichst überschaubarer Zeit zu ermöglichen.

- *Durch „cross over“ kann das Studienergebnis bei Auswahl des Endpunktes „Überlebenszeit“ verfälscht werden.*
 „Cross over“ bedeutet, dass Patienten in der Kontrollgruppe nach Eintreten eines Tumorrezidivs oder -progresses das Prüfmedikament erhalten. Eine solche Situation kann auch bei nicht zugelassenen Arzneimitteln eintreten, da bei den meisten Zulassungsstudien Zwischenauswertungen im Protokoll vorgesehen sind. Zeigt sich bei einer Zwischenauswertung eine Differenz für den primären Endpunkt in einem im Studienkonzept vorab festgelegten Ausmaß, wird die Randomisation abgebrochen und die Studie entblindet, da dann mit ausreichender Wahrscheinlichkeit von einer Überlegenheit des Prüfmedikaments ausgegangen wird. In diesen Fällen ist die Studienleitung aus ethischen Gründen verpflichtet, auch Patienten im Kontrollarm das Prüfmedikament anzubieten.

Dieses Problem kann dadurch vermieden werden, dass OS als primärer Endpunkt der Studie vorgesehen wird. In diesem Fall besteht bei Abbruch der Studie zum Zeitpunkt der Zwischenauswertung ein Vorteil im OS, der mit einer ausreichend geringen Irrtumswahrscheinlichkeit festgestellt wurde. Sollte sich dieser Überlebensvorteil im weiteren Verlauf der Studie durch „cross over“ verringern, ist dies für die Nutzenbewertung nur mit geringen Problemen verbunden, da ein Überlebensvorteil im Grundsatz bereits gesichert wurde. Außerdem erreicht die „cross over“-Rate niemals 100 % und viele Medikamente sind aufgrund einer zunehmenden Resistenzentwicklung der Tumorzellen bei Gabe in einer späteren Therapielinie weniger wirksam. Bei Medikamenten mit einer ausgeprägten Wirksamkeit konnte deshalb auch bei „cross over“-Raten von 31 % bzw. 66 % ein signifikanter Überlebensvorteil gesichert werden (Douillard JY et al., 2000; Slamon DJ et al., 2001). Fraglich ist deshalb lediglich, wie groß der Überlebensvorteil ohne Einfluss von „cross over“ bei der Endauswertung gewesen wäre. Dies ist z. B. für eine Nutzenbewertung nach dem QALY-Konzept (Ausmaß der Überlebenszeitverlängerung und in diesem Zeitraum gemessene Lebensqualität werden ins Verhältnis gesetzt) relevant. Häufig kann das erreichbare Ausmaß des Überlebensvorteils, besonders bei geringer „cross over“-Rate, mit ausreichender Sicherheit abgeschätzt werden. Letztendlich verbleibt bei hoher „cross over“-Rate eine Unsicherheit, die aus ethischen Gründen in Kauf genommen werden muss und damit unvermeidbar ist.

Nicht lösbare Probleme für die Nutzenbewertung durch „cross over“ nach Beendigung der Randomisation ergeben sich nur, wenn ein anderer Endpunkt als Überleben, z. B. Zeit bis zum Tumorprogress gewählt wird. Wird die Studie bereits abgebrochen, wenn die für diesen Endpunkt für einen Studienabbruch geforderte Differenz nachgewiesen wurde und überträgt sich das längere Intervall bis zum Tumorprogress nicht in einen signifikanten Vorteil im OS, ist es bei einer relevanten Rate von „cross over“ nicht möglich zu bewerten, ob ohne „cross over“ überhaupt ein signifikanter Überlebensvorteil durch das Prüfmedikament erreicht werden kann und wie groß dieser Überlebensvorteil ggf. sein könnte. Die Unsicherheit bei der Nutzenbewertung ist in jedem Fall erheblich größer als bei Auswahl von OS als primären Endpunkt.

Diese Betrachtung spricht eindeutig dafür, OS als primären Endpunkt, insbesondere bei palliativer Therapie auszuwählen. Auch ist es nicht unethisch, ein „cross over“ erst dann zuzulassen, wenn ein Überlebensvorteil mit ausreichender Sicherheit festgestellt wurde, da ein signifikanter Vorteil in der Zeit bis zum Tumorprogress keinen Patientenutzen belegt.

- *Wenn Medikamente in einem frühen Therapieabschnitt getestet werden (z. B. Erstlinientherapie) kann durch nachfolgende Behandlungen im Falle eines Tumorrezidivs*

bzw. –progresses (z. B. Zweitlinientherapie) der eigentlich signifikante Überlebensvorteils verwischt werden.

Auch dieses Argument ist nicht stichhaltig. Erstens ist es möglich durch entsprechende Vorgaben im Studienprotokoll oder den Verweis auf eine S3-Leitlinie die nachfolgenden Behandlungen, so weit festzulegen, wie dies medizinisch sinnvoll ist. Bei kurzer Rekrutierungsdauer dürften sich ohnehin keine relevanten Unterschiede in den nachfolgenden Therapielinien ergeben, wenn es sich um Therapiezentren handelt, die nach evidenzbasierten Leitlinien behandeln. Zweitens darf bezweifelt werden, dass ein Überlebensvorteil klinisch relevant ist, der bereits durch geringe Unterschiede in den nachfolgenden Therapielinien verwischt wird. In diesem Fall erscheint das Medikament in der Versorgung verzichtbar, da eine vergleichbare Überlebenszeit ohne das neue Medikament erreicht werden kann. Dann sollte Patienten die zusätzliche Toxizität und der GKV die zusätzlichen Kosten erspart werden.

- *Bei einigen Tumorerkrankungen, wie z. B. metastasiertes kolorektales Karzinom gibt es bereits 3 oder 4 evidenzbasierte Therapielinien bei der palliativen Behandlung, die jeweils bei Auftreten einer Therapieresistenz eingesetzt werden. In diesem Fall ist der Einfluss der einzelnen Therapielinie auf die Gesamtüberlebenszeit so gering, dass dieser mit einem angemessenen Aufwand nicht nachweisbar ist. Ein Nutzenbeleg ist nur dadurch zu erbringen, dass Tumorregister oder ähnliche Erhebungen eine Überlebenszeitverlängerung von Patienten im metastasierten Stadium im Vergleich zu historischen Kontrollen belegen.*

Auch hier ergibt sich die Frage, ob eine so marginale Wirkung des Prüfmedikaments auf den Endpunkt Überlebenszeit klinisch relevant sein kann, wenn er nicht mit einer akzeptablen Anzahl von Patienten darstellbar ist. Außerdem erfordert der Nachweis einer geringen Überlebenszeitverlängerung nicht regelhaft eine für häufige Tumorerkrankungen unrealistisch hohe Patientenzahl. So konnte bei der palliativen Therapie des Pankreaskarzinoms eine Überlebenszeitverlängerung von nur 2 Wochen im Median (HR für Tod 0,82, 95%KI 0,69 – 0,99) durch zusätzliche Gabe von Erlotinib bei Rekrutierung von 569 Patienten als signifikant detektiert werden (Moore MJ et al., 2007). Eine verbesserte Prognose in der Gesamtpopulation, die z. B. durch Krebsregister erkennbar wird, zeigt nicht einmal in jedem Fall einen Therapiefortschritt an, erst recht ist es – abgesehen von epochalen Therapiefortschritten – nicht möglich, die Prognoseverbesserung auf eine bestimmte Änderung der Therapie zurückzuführen. Durch Fortschritte in der bildgebenden Diagnostik werden zunehmend auch sehr kleine Metastasen erfasst, so dass dem Tumorstadium M1 (Nachweis von Fernmetastasen) heute Patienten zugerechnet werden, bei denen man früher von einem lokalisierten Stadium ausging und die eine vergleichsweise günstige Prognose aufweisen. Außerdem hat sich nicht nur die medikamentöse Tumorthherapie fortentwickelt, auch in der chirurgischen Therapie wurden Fortschritte erzielt. So werden auch in Deutschland zunehmend geeignete Patienten mit Kolonkarzinom und Lebermetastasen sowohl lokal als auch in der Leber radikal operiert, was erheblich zur Verbesserung der Prognose von Patienten mit Kolonkarzinom im Stadium M1 beigetragen haben dürfte (Neumann UP et al., 2010). Geeignete klinische Studien der Phase III mit dem Endpunkt OS sind deshalb für die Nutzenbewertung unverzichtbar, um den Einfluss eines einzelnen Medikamentes auf die Überlebenszeit bewerten zu können.

Toxizität der Behandlung als Endpunkt

Besonders bei palliativen Therapien und sehr begrenzter Prognose ist ein entscheidendes Behandlungsziel, die Restlebenszeit des Patienten nicht durch ausgeprägte unerwünschte Wirkungen der medikamentösen Tumorthherapie zu belasten. Angestrebt werden deshalb nebenwirkungsarme Therapieprotokolle, die nach Möglichkeit auch ambulant eingesetzt werden können.

Für eine aussagekräftige Auswertung ist die Erhebung nach einer anerkannten Skala entscheidend, z. B. nach WHO oder NCI. Kann bei vergleichbarer Überlebenszeit eine signifikante Reduktion unerwünschter Wirkungen erreicht werden, insbesondere solche die tödlich sind oder die mit erheblichen Einschränkungen verbunden sind, kann ein Patientennutzen vorausgesetzt werden.

Symptome bzw. gesundheitsbezogene Lebensqualität als Endpunkt

Der Vorteil dieses Parameters ist der unmittelbare Bezug zum Patientennutzen. Andererseits sind Erhebung und Auswertung dieses Parameters methodisch äußerst anspruchsvoll, so dass das Risiko von Verzerrungen aufgrund methodischer Probleme häufig sehr hoch ist. Zur Erfassung müssen vom Patienten Fragebögen ausgefüllt werden. Dabei sind validierte, diagnosebezogene Fragebögen zu verwenden, wie sie z. B. von der EORTC zur Verfügung gestellt werden. Da die Erhebung das subjektive Erleben des Patienten widerspiegelt, ist eine doppelte Verblindung immer notwendig; das heißt, dass weder Arzt noch Patient die Zuordnung zum Studien- bzw. Kontrollarm erkennen können. Die Fragebögen müssen exakt zu den vorab festgelegten Zeitpunkten ausgefüllt werden. Außerdem ist eine fast vollständige Rücklaufquote notwendig. Gibt ein relevanter Anteil von Patienten die Fragebögen nicht ab oder bestehen erhebliche Abweichungen beim Zeitpunkt der Ausfüllung der Fragebögen in beiden Armen, besteht ein hohes Risiko für systematische Verzerrungen. Zudem können im Gegensatz zu den Endpunkten OS oder unerwünschte Wirkungen, die durch regelmäßige klinische oder Laboruntersuchungen erhoben und dokumentiert werden, verpasste Erhebungen zur Messung der Lebensqualität nicht nachgeholt werden. Nach Erfahrung des KC Onkologie gibt es nur sehr wenige Studien, in denen die Methodik der Erhebung und Auswertung der Lebensqualität bei onkologischen Patienten für eine sichere Nutzenbewertung ausreichend war.

Surrogatendpunkte

Häufig werden bei Zulassungsstudien Surrogatendpunkte verwandt, um mit möglichst wenigen Patienten bei einer möglichst kurzen Nachbeobachtungsdauer zu einem signifikanten, „positiven“ Ergebnis zu gelangen. Weit verbreitet ist z. B. der Endpunkt „Zeit bis zum Tumorprogress“ (TTP) bzw. progressionsfreies Überleben (PFS) (Mathoulin-Pelissier S et al., 2008). Entscheidendes Ereignis ist in beiden Fällen ein radiologisch nachgewiesener Tumorprogress. Bei PFS wird zwar auch Tod, teilweise auch nur Tod durch die Tumorerkrankung als Ereignis gewertet, allerdings tritt bei den meisten Patienten zunächst der Tumorprogress ein, so dass dieser auch bei PFS die Masse der Ereignisse ausmacht. Die radiologischen Kontrolluntersuchungen müssen in beiden Armen zum gleichen Zeitpunkt durchgeführt werden. Die Auswertung, üblicherweise bei soliden Tumoren nach der RECIST-Methodik durchgeführt, ist aufwändig und mit erheblichen Unsicherheiten behaftet (Therasse P et al., 2000). Selbst bei erfahrenen Radiologen kann es aufgrund der Unschärfe der Methodik, z. B. unterschiedliche Auswahl von Markerläsionen, zu relevanten Abweichungen bei der Bewertung kommen (Geyer CE et al., 2006). Somit besteht schon aus methodischen Gründen eine erheblich größere Unsicherheit bei der Nutzenbewertung als beim Endpunkt Überlebenszeit.

Auch wenn es verschiedene Hinweise darauf gibt, dass bei bestimmten Indikationen eine Korrelation zwischen TTP bzw. PFS und Gesamtüberleben besteht, führt die Auswahl dieser Endpunkte regelhaft zu einer deutlichen Zunahme der Irrtumswahrscheinlichkeit, da die Korrelation niemals 100 % bzw. 1 erreicht. Teilweise wird bei Zulassungsstudien auch ein signifikanter Vorteil im PFS nachgewiesen, obwohl die Überlebensraten absolut identisch sind, was erhebliche Zweifel an der Eignung dieses Endpunktes für die Nutzenbewertung begründet (Geyer CE et al., 2006). Noch unsicherer ist es vom Ausmaß der TTP- oder PFS-Differenz auf das Ausmaß der Überlebenszeitverlängerung zu schließen (Tang PA et al., 2007).

Für eine ausreichend sichere Nutzenbewertung sind diese Surrogatendpunkte deshalb ungeeignet (Anonym, 2009).

1.2 Relative Bewertung von Überlebenszeitverlängerungen

Der Nachweis einer signifikanten Überlebenszeitverlängerung in einem begrenzten Ausmaß belegt nicht in jedem Fall einen Nutzen für Patienten. Wenn die Prüfmedikation mit zusätzlicher Toxizität verbunden ist, muss die Häufigkeit und der Schweregrad der Toxizität gegen das Ausmaß der Überlebenszeitverlängerung abgewogen werden. Diese Abwägung kann nicht nach zuvor festgelegten biometrischen Kriterien erfolgen, sondern muss von klinisch erfahrenen Onkologen vorgenommen werden. Als Beispiel wird auf die Zulassungsstudie von Erlotinib beim metastasierten Pankreaskarzinom verwiesen. Standardmedikation in dieser Indikation ist Gemcitabine. Geprüft wurde Erlotinib in Kombination mit Gemcitabine gegen Placebo und Gemcitabine im Kontrollarm. Es konnte ein signifikanter Überlebensvorteil von im Median 0,45 Monaten (6,24 versus 5,91 Monate, $p=0,038$) nachgewiesen werden. Gleichzeitig traten unter Erlotinib mehr unerwünschte Wirkungen auf, insbesondere Hautrötung, Diarrhö, Infektionen und Stomatitis. 8 Patienten entwickelten eine interstitielle Lungenerkrankung (ILD), davon 7 unter Erlotinib. Bei Auswertung der Lebensqualität zeigte sich - bei sonst vergleichbaren Parametern in beiden Armen - eine signifikante Zunahme von Diarrhö unter Erlotinib ($p<0,001$) (Moore MJ et al., 2007). Patienten mit Pankreaskarzinom weisen regelhaft schwere Ernährungsstörungen auf, so dass die Zunahme von Diarrhö klinisch relevant und bedenklich ist. Die Zunahme der Toxizität und die beschriebene Einschränkung der Lebensqualität muss gegen einen Überlebensvorteil von ca. 2 Wochen abgewogen werden. Industrieunabhängige Onkologen gelangen deshalb zu einer ungünstigen Nutzen-Risiko-Abwägung für Erlotinib beim Pankreaskarzinom und raten von der Anwendung in dieser Indikation ab (http://www.arznei-telegramm.de/html/2008_02/0802022_02.html).

2 Versorgungsrelevanz des Studienkonzepts

Ein- und Ausschlusskriterien

Durch Wahl entsprechender Ein- bzw. Ausschlusskriterien werden für Zulassungsstudien häufig Patientenkollektive rekrutiert, die in wesentlichen Merkmalen von dem für diese Erkrankung typischen Patientenkollektiv abweichen. Selbst bei Studien, die Patienten in einem weit fortgeschrittenes Tumorstadium mit umfangreicher Vorbehandlung einschließen, wird häufig ein nicht oder nicht wesentlich beeinflusster Allgemeinzustand (nach ECOG 0 oder 1) gefordert. Wesentliche Einschränkungen der Organfunktionen (Leber, Herz, Niere, Knochenmark) werden als Ausschlusskriterien festgelegt. Das mediane Alter des eingeschlossenen Patientenkollektivs liegt häufig 10 oder mehr Jahre unterhalb des medianen Erkrankungsalters. Bei so stark selektionierten Patienten ergibt sich die Frage der Versorgungsrelevanz der Studie. Selbst wenn zahlreiche Ein- und Ausschlusskriterien in die Fachinformation übernommen werden, zeigt die klinische Praxis, dass – insbesondere wenn therapeutische Alternativen fehlen – in der Versorgung davon regelhaft abgewichen wird.

Wahl des Kontrollarms bzw. der Standardtherapie

Wesentlich für die Versorgungsrelevanz der Studie ist die angemessene Auswahl des Kontrollarms. Die Behandlung des Kontrollarms sollte nach den Prinzipien der evidenzbasierten Medizin festgelegt werden, so wie in der Regel in S3-Leitlinien beschrieben. Eine Testung gegen alleinige Placebotherapie ist immer kritisch zu hinterfragen, da dies meistens nicht der Versorgungsrealität entspricht. Häufig stehen als Standardtherapie Alternativen zur Verfügung, entweder weil sich diese als gleichwertig erwiesen haben oder weil vergleichende Studien fehlen und deshalb Unterschiede nicht bekannt sind. In diesen Fällen erscheint ein Vergleich mit der preisgünstigsten Therapiealternative attraktiv, um festzustellen, ob mit geringem finanziellem Aufwand ein vergleichbares Behandlungsergebnis erzielt werden kann. Bei Medikamenten, die einer neuen Substanzklasse angehören, ist immer dann ein randomisierter Vergleich gegen ein anderes Medikament dieser Substanzklasse zu fordern, wenn dieses bereits für die gleiche Indikation zugelassen wurde. So existieren z. B. für die Behandlung des metastasierten Nierenzellkarzinoms zwei Multikinaseinhibitoren (Sorafenib und Sunitinib), die beide arzneimittelrechtlich zugelassen, aber nie gegeneinander randomisiert verglichen wurden.

Ein Beispiel für eine unzweckmäßige Auswahl des Kontrollarms ist die neue Phase-III-Prüfung der Firmen Bayer und Onyx von Sorafenib beim metastasierten bzw. fortgeschrittenen Schilddrüsenkarzinom (<http://www.onyx-pharm.com/view.cfm/629/Bayer-and-Onyx-Initiate-Phase-3-Trial-of-Nexavar-in-Patients-with-Non-Responsive-Thyroid-Cancer>). Obwohl Doxorubicin für diese Indikation seit mehr als 20 Jahren arzneimittelrechtlich zugelassen ist und in vielen Onkologielehrbüchern und Leitlinien als Standard empfohlen wird, wird Sorafenib gegen alleinige Gabe von Placebo getestet. Entweder hätte man Sorafenib primär gegen Doxorubicin oder erst nach Versagen von Doxorubicin gegen Placebo vergleichen müssen, um zu versorgungsrelevanten Ergebnissen zu gelangen.

Noch problematischer ist nach Bewertung des KC Onkologie die Phase-III-Studie der französischen Myelomstudiengruppe IFM (2005-02), die diese mit dem Ziel der arzneimittelrechtlichen Zulassungserweiterung von Lenalidomide in Kooperation mit dem Hersteller Celgene durchgeführt hat. Lenalidomide wurde als Erhaltungstherapie bei Myelompatienten nach Hochdosis-Melphalan und autologer Stammzelltransplantation in der Erstlinientherapie bis zum Tumorprogress eingesetzt und gegen die alleinige Gabe von Placebo im Kontrollarm verglichen (Attal M et al., 2010). Dieses Vorgehen ist absolut unverständlich, da von der gleichen Myelomstudiengruppe in einer vorherigen klinischen Studie der Phase III in gleicher Indikation eine überlebensverlängernde Wirksamkeit des Strukturanalogons Thalidomid gesi-

chert werden konnte (Attal M et al., 2006). Die Frage, ob mit Thalidomid eine vergleichbare Überlebenszeitverlängerung wie mit Lenalidomide erreicht werden kann, ist für die Wirtschaftlichkeit von essentieller Bedeutung. Die Erhaltungstherapie mit Thalidomid (100 mg/Tag) für ein Jahr kostet ca. 12.500 Euro, während die Erhaltungstherapie mit Lenalidomide (10 mg/Tag) für 1 Jahr ca. 119.200 Euro kostet (Preise nach Rote Liste 2010). Dabei wird Thalidomid nur für etwa ein Jahr eingesetzt, während Lenalidomide für einen deutlich längeren Zeitraum, nämlich bis zum Tumorprogress eingesetzt werden soll.

3 Sicherheit der Erkenntnisse

Neben der Auswahl geeigneter Endpunkte und der Versorgungsrelevanz des Studiendesigns ist die Sicherheit der Erkenntnisse für die Nutzenbewertung entscheidend. Es handelt sich um eine klinische Einschätzung des KC Onkologie, die durch Experten in medizinischer Statistik ergänzt werden muss.

Adäquate Gruppenzuordnung

Notwendig für eine Vergleichbarkeit der Gruppen bzw. Patientencharakteristika in den Therapiearmen als Voraussetzung für eine ausreichende Sicherheit der Nutzenbewertung ist das „Concealment“ bei Randomisation, d. h. weder Patient noch Arzt haben die Möglichkeit das Ergebnis der Randomisation zu beeinflussen. Außerdem sollte, sofern besonders bedeutsame Prognosefaktoren bekannt sind, für diese Prognosefaktoren eine Stratifikation bei der Randomisation vorgesehen werden. Bei multizentrischen Studien, insbesondere wenn diese in verschiedenen Ländern durchgeführt werden, sollte eine Randomisation stratifiziert nach Ländern oder Zentren vorgesehen werden. Aufgrund der unterschiedlichen Versorgungsniveaus wäre es problematisch, wenn beispielsweise der überwiegende Teil der Patienten eines Armes in Deutschland und der überwiegende Teil der Patienten im anderen Arm in einem Land mit einem deutlich weniger entwickelten Gesundheitssystem behandelt worden wäre. Auch eine unterschiedliche Verteilung der Ethnien kann das Ergebnis beeinflussen, was ebenfalls für eine zentrumsbezogene Stratifikation spricht.

Protokollverstöße in einem akzeptablen Rahmen

Kritisch ist ein Verstoß gegen die Ein- und Ausschlusskriterien bei prognostisch besonders relevanten Patientenmerkmalen wie histologische Diagnose oder Tumorstadium. Im Zweifelsfall ist zu prüfen, ob bei schwierigen histologischen Diagnosen, wie z. B. malignem Lymphom, eine Referenzpathologie durchgeführt wurde und ob die „Staging“-Untersuchungen adäquat und vollständig durchgeführt wurden. Kritisch für die Sicherheit der Nutzenbewertung sind erhebliche Abweichungen vom Behandlungsprotokoll bei einem relevanten Anteil von Patienten, die nicht medizinisch begründet sind. Die Durchführung der Behandlung entgegen dem Ergebnis der Randomisation stellt dabei den ungünstigsten Fall dar und entwertet, wenn dies bei einem relevanten Anteil von Patienten vorliegt, das Ergebnis der Studie vollständig.

Vollständigkeit der Datenerfassung

Im Idealfall dürfen keine oder nur sehr wenige Patienten im Rahmen der Nachbeobachtung verloren gehen („lost to follow-up“). Die vorgeschriebenen Untersuchungen vor Beginn der Therapie und zur Kontrolle nach Einleitung der Therapie müssen vollständig durchgeführt und dokumentiert worden sein. Die Dauer der Nachbeobachtung soll in Abhängigkeit von der Prognose ausreichend sein, um den Einfluss der Prüfmedikation auf Überlebenszeit und Toxizität sicher bewerten zu können.

Auswertung der Studie

Zu fordern ist immer eine Auswertung nach dem „intention-to-treat“-Prinzip. Dies bedeutet, dass unabhängig von der tatsächlich durchgeführten Behandlung die Patienten immer entsprechend Randomisation zuzuordnen und auszuwerten sind. Ausschlüsse von Patienten nach Randomisation, selbst bei stichhaltiger Begründung (z. B. Fehldiagnose), sind deshalb

problematisch. Allein die Auswertung nach dem „intention-to-treat“-Prinzip ist maßgeblich für die Nutzenbewertung, da dieses die Umsetzung der Therapie in der klinischen Praxis beschreibt. Wenn selbst unter Studienbedingungen eine konsequente Umsetzung der im Protokoll vorgegebenen Therapie nicht möglich war, erscheint es umso weniger wahrscheinlich, dass dies im klinischen Alltag besser funktioniert. Eine „per protocol“-Auswertung, die nur Patienten berücksichtigt, die die protokollgemäße Behandlung vollständig erhalten haben, ist informativ, um das therapeutische Potenzial eines Protokolls bewerten zu können. Daraus können z. B. Konsequenzen für nachfolgende Studienkonzepte abgeleitet werden. Für die Nutzenbewertung kann eine solche Auswertung aber nicht maßgeblich sein.

Untergruppenanalysen

Untergruppenanalysen sind zur Verbesserung von Qualität und Wirtschaftlichkeit sinnvoll, um Merkmale von Patienten zu identifizieren, die auf die Behandlung ansprechen („Responder“). Häufig sind die Ansprechraten bei onkologischen Medikamenten niedrig, so dass die Mehrzahl der Patienten nicht auf die Behandlung anspricht und nicht von dieser Behandlung profitiert. Um den Patienten unnötige Belastungen zu ersparen und um die Wirtschaftlichkeit zu verbessern, sollten deshalb die Medikamente möglichst gezielt eingesetzt werden (günstiges Verhältnis von behandelten Patienten und Behandlungserfolgen, entspricht einer niedrigen „number needed to treat“).

Das Ergebnis einer Untergruppenanalyse ist ausreichend sicher, wenn diese im Studiendesign geplant und damit bei der Fallzahlschätzung berücksichtigt wurde oder wenn Untergruppenanalysen anderer Phase-III-Studien bzw. Metaanalysen bei gleicher Fragestellung zu einem vergleichbaren Ergebnis gekommen sind.

Ein Beispiel ist die Frage des Nutzens einer Wiederholung von Hochdosistherapie und autologer Stammzelltransplantation (Tandemtransplantation) bei Patienten mit multiplen Myelom. Dazu wurden mehrere klinische Studien der Phase III durchgeführt, in denen Tandemtransplantation mit einfacher autologer Stammzelltransplantation für den Endpunkt Überlebenszeit verglichen wurden. 2 große Phase-III-Studien wiesen übereinstimmend durch Untergruppenanalysen nach, dass nur bei Patienten, die unzureichend auf die erste Behandlung mit Hochdosistherapie und autologer Stammzelltransplantation angesprochen haben (partielle Remission oder schlechter) von einer Wiederholung der Behandlung durch signifikante Verlängerung der Überlebenszeit profitieren (Attal M et al., 2003; Cavo M et al., 2007). Entsprechende Therapieempfehlungen haben international Eingang in evidenzbasierte Leitlinien gefunden (Reece DE, ASH 2005), so dass diese kostspielige und toxische Therapie gezielter eingesetzt werden kann.

Andernfalls sind Untergruppenanalysen aufgrund des Problems des multiplen Testens eher kritisch zu bewerten (je häufiger getestet wird, umso wahrscheinlicher werden rein zufällig auftretende, signifikante Differenzen). Ausnahmen sind ggf. denkbar bei Untergruppenanalysen mit einer sehr großen Patientenzahl und/oder deutlich distanzierten Konfidenzintervallen. Außerdem sollten im Zweifelsfall nur Untergruppenanalysen akzeptiert werden, bei denen das Ergebnis nachvollziehbar medizinisch begründet werden kann.

Ein Beispiel für eine Untergruppenanalyse mit einem für die Nutzenbewertung irrelevanten Ergebnis ist in der bereits zitierten Studie von Geyer GE et al., (2006) enthalten. Beschrieben wurde in der Publikation, dass in der Kontrollgruppe 11 Patienten Hirnmetastasen entwickelt hatten im Vergleich zu nur 4 Patienten im Arm mit der Prüfmedikation (Lapatinib). Dieser Unterschied war statistisch nicht signifikant ($p=0,1$). Neben der Tatsache, dass die Differenz in der Untergruppenanalyse nicht signifikant war, ist zu kritisieren, dass ZNS-Metastasen – laut Publikation – kein vorab definierter Endpunkt der Studie waren, die Untergruppenanalyse nicht geplant war und andere Untergruppenanalysen nicht dargestellt wurden, so dass das Problem des multiplen Testens nicht transparent gemacht wurde. Trotz dieser Mängel zeigt

die Begutachtungspraxis des KC Onkologie, dass mit Hinweis auf diese Untergruppenanalyse einige behandelnde Ärzte Lapatinib bevorzugt bei Patientinnen mit Hirnmetastasen bei HER2-positivem Brustkrebs einsetzen.

4 Zusammenfassung

Zusammenfassend schlägt das KC Onkologie eine „Checkliste“ für Studienauswertungen zur Nutzenbewertung vor. Wir schlagen vor diese Checkliste nach den Hauptpunkten:

- 1) Beleg des Nutzen durch geeignete Endpunkte
- 2) Versorgungsrelevanz des Studienkonzepts
- 3) Sicherheit der Erkenntnisse

zu gliedern und durch entsprechende Unterpunkte aufzufüllen. Dabei könnte sich beispielsweise folgende Vorlage für eine Checkliste ergeben:

1. Beleg des Nutzens durch geeignete(n) Endpunkt(e)

- In der Regel Überlebenszeit und Toxizität
 - Überlebenszeitverlängerung gesichert?
 - Angemessenes Verhältnis zu ggf. zusätzlicher Toxizität?
- Alternativ: gesundheitsbezogene Lebensqualität
- Ggf. abweichende Endpunktauswahl adäquat begründet?

2. Versorgungsrelevanz des Studienkonzepts

- Für die jeweilige Erkrankung repräsentatives Patientenkollektiv?
 - Ein- und Ausschlusskriterien
- Standardarm akzeptabel?
 - Best mögliche Therapie nach EBM-Kriterien?
 - Berücksichtigung preiswerterer Behandlungen?
 - Berücksichtigung neuerer Medikamente der gleichen Substanzklasse?

3. Sicherheit der Erkenntnisse

- Vergleichbarkeit der Gruppen gewährleistet?
- Protokollverstöße in akzeptablem Rahmen?
- Daten ausreichend vollständig?
 - Rate von „loss to follow up“?
 - Untersuchungen termingerecht und vollständig?
- Nachbeobachtung ausreichend lange?
- Auswertung nach dem „Intention to treat“ – Prinzip?
- Untergruppenanalysen ausreichend abgesichert?

5 Zusätzliche Anregungen des KC Onkologie

Das KC Onkologie regt an, für die Nutzenbewertung onkologischer Arzneimittel klinisch-onkologischen Sachverstand einzubeziehen, z. B. durch Beteiligung der Arzneimittelkommission der Deutschen Ärzteschaft (AkdÄ) und klinischen Onkologen der für die jeweilige Indikation „zuständigen“ Studiengruppen und Fachgesellschaften. Durch eine Kooperation von G-BA und IQWiG mit diesen Experten könnte nach Erstzulassung oder Zulassungserweiterung eines onkologischen Arzneimittels in der notwendigen Kürze der Zeit ein kompetentes Expertenteam gebildet werden, welches umgehend anhand der verfügbaren Daten zu einer Entscheidung gelangen könnte. Vorbild sind u. a. die Expertenkomitees der Arzneimittelbehörden (z. B. CHMP der EMA). Auch in Hinblick auf die öffentliche Akzeptanz der Bewertungen erscheint eine solche Kooperation vorteilhaft.

Das KC Onkologie empfiehlt zusätzlich zur Nutzenbewertung, die häufig auf Grundlage der Daten der Zulassungsstudie nicht möglich sein wird, die noch offenen, versorgungsrelevanten Fragestellungen durch das Expertenteam beschrieben und geeignete Konzepte für klinische Studien nach Zulassung des Medikamentes entwickeln zu lassen. Das KC Onkologie empfiehlt, die industrieunabhängige klinische Forschung zu unterstützen. Dazu sollte es der GKV bzw. dem G-BA oder dem IQWiG ermöglicht werden, finanzielle Mittel zur Studienförderung zur Verfügung zu stellen. Um dieses Ziel zu erreichen würde sich eine Kooperation von G-BA, IQWiG und GKV-Spitzenverband mit den onkologischen Studiengruppen anbieten. Mit einem solchen Konzept könnte die finanzielle Abhängigkeit der Studiengruppen von der pharmazeutischen Industrie beendet werden. Damit wären auch Studienkonzepte möglich, die zu Einsparungen von Medikamentenkosten führen und somit den Interessen der Hersteller zuwider laufen. Wir sind sicher, dass die zu erzielenden Einsparungen für die GKV durch eine Verbesserung der Wirtschaftlichkeit der Versorgung die für die Förderung industrieunabhängiger Forschung eingesetzten Mittel um ein Vielfaches übersteigen würden.

Durch geeignete Beschlüsse des G-BA oder ggf. notwendige Gesetzesänderungen sollte zudem erreicht werden, dass – falls die Zulassungsstudie nicht alle für eine Nutzenbewertung und einen gezielten Medikamenteneinsatz notwendige Daten liefert – die breite Verordnungsmöglichkeit des Medikaments zu Lasten der GKV verhindert wird. Ausschließlich Ärzte an den Prüfzentren der kooperierenden Studiengruppen (Kliniken mit onkologischer Fachabteilung oder onkologische Schwerpunktpraxen) sollten diese Medikamente zu Lasten der GKV verordnen können. Die Behandlungen sollten vorzugsweise in klinischen Studien durchgeführt werden. Nur bei Patienten, bei denen entsprechend der Rechtsprechung des Bundesverfassungsgerichts, die Voraussetzungen für einen Heilversuch zu Lasten der GKV erfüllt sind (lebensbedrohliche Erkrankung, Fehlen von anerkannten Behandlungsmöglichkeiten), sollten auch Behandlungen außerhalb klinischer Studien möglich sein. Auch diese Behandlungen sollten nur in den Prüfzentren durchgeführt werden, damit diese genau wie Studienbehandlungen umfassend dokumentiert und ausgewertet werden können. Erst wenn durch die Auswertung der Studiendaten ein Nutzen gesichert werden konnte und ein ausreichender gezielter Einsatz der Medikamente durch Beschreibung der für ein Ansprechen prognostisch relevanten Merkmale möglich ist, sollten die neuen medikamentösen Behandlungen in die bestehenden S3-Leitlinien der Deutschen Krebsgesellschaft und die Arzneimittelrichtlinie des G-BA implementiert werden und auf dieser Grundlage eine breite Verordnung ermöglicht werden. Diese Daten würden es dem IQWiG außerdem ermöglichen, eine Nutzen-Kosten-Bewertung auf einer wissenschaftlich gesicherten Grundlage durchzuführen, z. B. nach dem QUALY-Konzept.

5 Literatur

André T et al.: Improved overall survival with oxaliplatin, fluorouracil and leucovorin as adjuvant treatment in stage II or III colon cancer in the MOSAIC trial. *J Clin Oncol* 27:3109-3116, 2009

Anonym.: Effects of cancer drug on survival: often poorly evaluated. *Rev Prescrire* 29: 218-221, 2009

Attal M et al.: Single versus double autologous stem-cell transplantation for multiple myeloma. *N Engl J Med* 349: 2495-2502, 2003

Attal M et al.: Lenalidomide maintenance after transplantation for myeloma. Abstract Nr. 8018, ASCO Kongress 2010
(http://www.asco.org/ASCOv2/Meetings/Abstracts?&vmview=abst_detail_view&confID=74&abstractID=41503)

Attal M et al.: Maintenance therapy with thalidomide improves survival in patients with multiple myeloma. *Blood* 108: 3289-3294, 2006

Cavo M et al.: Prospective, randomized study of single compared with double autologous stem-cell transplantation for multiple myeloma: Bologna 96 Clinical Study. *J Clin Oncol* 25 No. 17 June 10 2007 (Epub ahead of print)

Di Leo A et al.: Overall survival is not a realistic end point for clinical trials of new drugs in advanced solid tumors: A critical assessment based on recently reported phase III trials in colorectal and breast cancer. *Letter. J Clin Oncol* 21:2045-5047, 2003

Douillard JY et al.: Irinotecan combined with fluorouracil compared with fluorouracil alone as first-line treatment for metastatic colorectal cancer: a multicentre randomised trial. *Lancet* 355:1041-1047, 2000

Fenaux P et al.: RBC Transfusion Independence and Safety Profile of Lenalidomide 5 or 10 mg in Pts with Low- or Int-1-Risk MDS with Del5q: Results From a Randomized Phase III Trial (MDS-004). *Blood (ASH Annual Meeting Abstracts)* 114: Abstract No 944, 2009

Geyer CE et al.: Lapatinib plus Capecitabine for HER2-positive advanced breast cancer. *N Engl J Med* 355:2733-2743, 2006

Moore MJ et al.: Erlotinib plus gemcitabine compared with gemcitabine alone in patients with advanced pancreatic cancer: A phase III trial of the National Cancer Institute of Canada Clinical Trials Group. *J Clin Oncol* 25:1960-1966, 2007

Mathoulin-Pelissier S et al.: Survival end point reporting in randomized cancer clinical trials: a review of major journals. *J Clin Oncol* 26:3721-3726, 2008

Neumann UP et al.: The surgical treatment of hepatic metastases in colorectal carcinoma. *Dt Arztebl Int* 107: 335-342, 2010

Reece DE: An update of the management of multiple myeloma: The changing landscape. *Hematology* 2005: 353-359 (<http://asheducationbook.hematologylibrary.org/cgi/reprint/2005/1/353>)

Sargent DJ et al.: End points for colon cancer adjuvant trials: Observations and recommendations based on individual patient data from 20,898 patients enrolled onto 18 randomized trials from the AC-CENT group. *J Clin Oncol* 25: 4569-4574, 2007

Slamon DJ et al.: Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med* 344:783-792, 2001

Tang PA et al.: Surrogate end points for median overall survival in metastatic colorectal cancer: Literature-based analysis from 39 randomized controlled trials of first-line chemotherapy. *J Clin Oncol* 25: 4562-4568, 2007

Therasse P et al.: New guidelines to evaluate the response to treatment in solid tumors. J Nat Cancer Inst 92:205-216, 2000