

IQWiG-Berichte – Nr. 171

**Addendum zum Auftrag A13-06
(Abirateronacetat [neues
Anwendungsgebiet])**

Addendum

Auftrag: A13-22
Version: 1.0
Stand: 13.06.2013

Impressum

Herausgeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema:

Addendum zum Auftrag A13-06 (Abirateronacetat [neues Anwendungsgebiet])

Auftraggeber:

Gemeinsamer Bundesausschuss

Datum des Auftrags:

28.05.2013

Interne Auftragsnummer:

A13-22

Anschrift des Herausgebers:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Im Mediapark 8 (KölnTurm)
50670 Köln

Tel.: +49 (0)221 – 35685-0

Fax: +49 (0)221 – 35685-1

E-Mail: berichte@iqwig.de

Internet: www.iqwig.de

ISSN: 1864-2500

An dem Addendum beteiligte Mitarbeiter des IQWiG¹:

- Volker Vervölgyi
- Corinna Kiefer
- Beate Wieseler

Schlagwörter: Abirateron, Prostatatumoren, Nutzenbewertung

Keywords: Abiraterone, Prostatic Neoplasms, Benefit Assessment

¹ Aufgrund gesetzlicher Datenschutzbestimmungen haben Mitarbeiter das Recht, ihrer Namensnennung nicht zuzustimmen.

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	iv
Abkürzungsverzeichnis.....	v
1 Hintergrund.....	1
2 Bewertung.....	2
2.1 Auswahl der Analysen für die Nutzenbewertung	2
2.2 Verzerrungspotenzial.....	4
2.3 Ergebnisse	5
2.4 Ausmaß und Wahrscheinlichkeit des Zusatznutzens	9
2.4.1 Beurteilung des Zusatznutzens auf Endpunktebene	10
2.4.2 Gesamtaussage zum Zusatznutzen	13
3 Literatur	14

Tabellenverzeichnis

	Seite
Tabelle 1: Gegenüberstellung der Angaben zu aufgetretenen Ereignissen in der Stellungnahme und dem Dossier bzw. Studienbericht der Studie COU-AA-302.....	3
Tabelle 2: Verzerrungspotenzial auf Studien- und Endpunktebene – RCT, direkter Vergleich: Abirateron vs. abwartendes Vorgehen	5
Tabelle 3: Ergebnisse (stetige Endpunkte) zu Morbidität und Lebensqualität – RCT, direkter Vergleich: Abirateron vs. abwartendes Vorgehen	6
Tabelle 4: Ausmaß des Zusatznutzens auf Endpunktebene (Nutzenendpunkte): Abirateron vs. abwartendes Vorgehen.....	11
Tabelle 5: Ausmaß des Zusatznutzens auf Endpunktebene (Schadensendpunkte): Abirateron vs. abwartendes Vorgehen	12
Tabelle 6: Positive und negative Effekte aus der Bewertung von Abirateron im Vergleich zum abwartenden Vorgehen.....	13

Abkürzungsverzeichnis

Abkürzung	Bedeutung
ADT	Androgenentzugstherapie
BPI-SF	Brief Pain Inventory – Short Form
FACT-P	Functional Assessment Cancer Therapy – Prostate
G-BA	Gemeinsamer Bundesausschuss
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
KI	Konfidenzintervall
MedDRA	Medical Dictionary of Medical Activities
MMRM	Mixed-Effects-Model-Repeated-Measures
pU	pharmazeutischer Unternehmer
PT	Preferred Term
RCT	randomized controlled trial (randomisierte kontrollierte Studie)
SMD	Standardisierte Mittelwertsdifferenz
UE	unerwünschtes Ereignis
vs.	versus

1 Hintergrund

Der Gemeinsame Bundesausschuss (G-BA) hat das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) am 28.05.2013 mit einer ergänzenden Bewertung zum Auftrag A13-06 (Nutzenbewertung von Abirateronacetat [neues Anwendungsgebiet] [1]) beauftragt.

Im Rahmen des Stellungnahmeverfahrens zur Bewertung von Abirateronacetat (im Folgenden als „Abirateron“ bezeichnet) wurden vom pharmazeutischen Unternehmer (pU) am 06.05.2013 weitere Daten an den G-BA übermittelt, die über die Angaben im Dossier hinausgehen. Dabei handelt es sich um Daten zur Studie COU-AA-302 (Vergleich Abirateron + Prednison + konventionelle Androgenentzugstherapie [ADT] bzw. Placebo + Prednison + ADT). Diese Studie war bereits im Dossier des pU enthalten und wurde in der Bewertung A13-06 vom IQWiG als relevant eingeschlossen. Auf Basis der im Dossier vom pU vorgelegten Daten lagen jedoch insbesondere für die Endpunkte gesundheitsbezogene Lebensqualität und Nebenwirkungen entweder keine oder größtenteils keine verwertbaren Daten vor. Die nachgereichten Daten umfassen im Wesentlichen neue Auswertungen zu diesen Endpunkten aus der Studie COU-AA-302.

Der G-BA hat das IQWiG mit der Bewertung der im Stellungnahmeverfahren für die Studie COU-AA-302 nachgereichten Auswertungen beauftragt. Dabei sollen die Daten unter der Fragestellung bewertet werden, ob unter Berücksichtigung der vom pharmazeutischen Unternehmer eingereichten Auswertungen zur gesundheitsbezogenen Lebensqualität (Responsekriterium) sowie zu Nebenwirkungen ein Zusatznutzen für Abirateron hinsichtlich der Lebensqualität belegt ist und Unsicherheiten auf der Schadensseite behoben werden.

Im nachfolgenden Kapitel 2 werden die zusätzlichen Ergebnisse für die Studie COU-AA-302 auftragsgemäß dargestellt und bewertet. Anschließend werden das Ausmaß und die Wahrscheinlichkeit des Zusatznutzens von Abirateron unter Berücksichtigung der nachgereichten Auswertungen beschrieben.

Die Verantwortung für die vorliegende Bewertung und das Bewertungsergebnis liegt ausschließlich beim IQWiG. Die Bewertung wird an den G-BA übermittelt. Über den Zusatznutzen beschließt der G-BA.

2 Bewertung

2.1 Auswahl der Analysen für die Nutzenbewertung

In seiner Stellungnahme [2] legte der pU zum einen, wie in der Bewertung A13-06 vorgeschlagen [1], Analysen zu unerwünschten Ereignissen (UE) auf Basis der Zeit bis zum ersten Ereignis vor. Zum anderen reichte er Mixed-Effects-Model-Repeated-Measures (MMRM)-Auswertungen zur gesundheitsbezogenen Lebensqualität, gemessen mit dem Functional Assessment Cancer Therapy – Prostate (FACT-P), vor. Während für die Bewertung A13-06 nur Ergebnisse pro Behandlungszyklus vorgelegen hatten, wurden in den nachgereichten Auswertungen, wie ebenfalls in der Bewertung A13-06 vorgeschlagen, die Mittelwerte über den gesamten Studienverlauf zwischen den Behandlungsgruppen der Studie COU-AA-302 verglichen, sodass nun eine Bewertung des Gesamteffekts über den gesamten Studienverlauf möglich war.

Neben den im Auftrag des G-BA beschriebenen Auswertungen zu UE und zur gesundheitsbezogenen Lebensqualität lagen in den vom des pU in der Stellungnahme vorgelegten Auswertungen auch Angaben zur Änderung des stärksten Schmerzes und Beeinträchtigung des Patienten durch den Schmerz im Studienverlauf, jeweils gemessen mit dem Brief Pain Inventory – Short Form (BPI-SF) vor. Wie für den FACT-P vergleicht der pU auch für die BPI-SF die Mittelwerte über den gesamten Studienverlauf zwischen den Behandlungsgruppen. Diese Endpunkte wurden grundsätzlich als patientenrelevant und der verwendete Fragebogen als hinreichend valide angesehen. Aus diesem Grund wurden die Auswertungen auf Basis des BPI-SF im vorliegenden Addendum zusätzlich bewertet.

Für die in der Stellungnahme vorgelegten Auswertungen zu den Endpunkten zu UE ergaben sich Diskrepanzen im Vergleich zum Dossier [3] und zum Studienbericht der Studie COU-AA-302 [4,5] hinsichtlich der in der Überlebenszeitanalyse berücksichtigten Anzahl der Patienten mit Ereignissen. Aus den verfügbaren Informationen in der Stellungnahme und in den Anlagen zur Stellungnahme war nicht nachvollziehbar, worauf diese Diskrepanzen zurückzuführen waren. Neben dem Hazard Ratio mit zugehörigem Konfidenzintervall fanden sich in der Stellungnahme Angaben zur medianen Zeit bis zum Ereignis pro Gruppe sowie die „eingetretenen Ereignisse“ pro Gruppe. Aus der Stellungnahme selbst ging nicht hervor, ob mit der Angabe „eingetretene Ereignisse“ die Anzahl der Patienten mit Ereignissen oder die Gesamtzahl der Ereignisse gemeint war. Lediglich anhand der in der Anlage zur Stellungnahme mitgelieferten Kaplan-Meier-Kurven ließ sich ableiten, dass es sich um die Anzahl der Patienten mit Ereignis handelte. Es ist allerdings nicht nachvollziehbar, warum sich diese Angaben in der Stellungnahme von denjenigen zur Anzahl der Patienten mit Ereignis im Dossier und dem Studienbericht der Studie COU-AA-302 unterscheidet. Teilweise haben laut den Angaben in der Stellungnahme mehr Patienten ein Ereignis als im Studienbericht, teilweise aber auch weniger. Dies ist auch deshalb nicht plausibel, als aus den Angaben hervorgeht, dass die Analysen auf dem gleichen Datenschnitt beruhen. Beim Endpunkt Gesamtrate UEs beruht die Überlebenszeitanalyse auf einer deutlich geringeren Patientenzahl als die Safety-Population der Studie COU-AA-302. Auch dieses Vorgehen ist

weder durch den pU begründet noch vor dem Hintergrund des gleichen zugrundeliegenden Datenschnitts nachvollziehbar. Zur Verdeutlichung sind die Diskrepanzen in der folgenden Tabelle 1 dargestellt.

Tabelle 1: Gegenüberstellung der Angaben zu aufgetretenen Ereignissen in der Stellungnahme und dem Dossier bzw. Studienbericht der Studie COU-AA-302

Endpunkt	Angabe laut Stellungnahme	Angabe laut Dossier bzw. Studienbericht
	eingetretene Ereignisse Abirateron / Prednison / ADT vs. Placebo / Prednison / ADT ^a n/N	Anzahl Patienten mit mindestens einem Ereignis Abirateron / Prednison / ADT vs. Placebo / Prednison / ADT ^a n/N
UEs	522/527 vs. 507/522	538/542 vs. 524/540
Schwere UEs (CTCAE -Grade 3 und 4)	268/542 vs. 234/538	267/542 vs. 235/540
SUEs	199/542 vs. 163/539	188/542 vs. 146/540
Therapieabbrüche wegen UEs	57/542 vs. 50/540	58/542 vs. 53/540
Frakturen	_ ^b	_ ^b
Flüssigkeitsretention / Ödeme	158/541 vs. 122/540	159/542 vs. 130/540
Ischämische Herzerkrankungen	26/542 vs. 20/540	25/542 vs. 20/540
Herzinsuffizienz	12/542 vs. 2/540	12/542 vs. 2/540
<p>a: Operationalisierung des abwartenden Vorgehens b: Für das UE Frakturen war nicht klar, auf Basis welcher Operationalisierung die Auswertung durchgeführt wurde (siehe Text), deshalb ist kein Referenzwert aus den Dossier bzw. den Studienbericht der Studie COU-AA-302 verfügbar. Aus diesem Grund erfolgt hier keine Darstellung. ADT: Androgendeprivationstherapie; CTCAE: Common Terminology Criteria for Adverse Events; n: Anzahl der Patienten mit Ereignis; N: Anzahl der ausgewerteten Patienten; SUE: schwerwiegendes unerwünschtes Ereignis; UE: unerwünschtes Ereignis</p>		

Die Zahlen unterscheiden sich insbesondere bei den Endpunkten Gesamtrate UEs und SUEs deutlich. Für den Endpunkt SUEs zeigte sich in der vom pU nachgelieferten Überlebenszeitanalyse zwar kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen, die untere Grenze des Konfidenzintervalls liegt jedoch nur geringfügig unterhalb des Nulleffekts (Hazard Ratio 1,21; 95 %-Konfidenzintervall [0,98; 1,49]). Für diesen Endpunkt ist nicht auszuschließen, dass sich das Ergebnis unter Berücksichtigung der Diskrepanz der berichteten Zahlen zwischen Stellungnahme und Dossier bzw. Studienbericht hinsichtlich der statistischen Signifikanz ändern würde. Eine Überprüfung des Einflusses der Diskrepanzen auf das Ergebnis ist aufgrund der Überlebenszeitanalysen allerdings nicht möglich.

Des Weiteren sind die in den vom pU in der Stellungnahme nachgereichten Kaplan-Meier-Kurven dargestellten hohen Anzahlen der Patienten, die zu einem bestimmten Zeitpunkt noch unter Risiko stehen, unter Berücksichtigung des im Dossier beschriebenen Abbruchs der

Studienmedikation und des damit verbundenen Abbruchs der Beobachtung nicht vereinbar. Der pU äußert sich in seiner Stellungnahme nicht dazu, sodass auch keine Informationen vorliegen, die diese Unklarheiten auflösen könnten.

Für den Endpunkt Frakturen ging aus den Angaben in der Stellungnahme zudem nicht hervor, auf Basis welcher Operationalisierung die Auswertung durchgeführt wurde. Im Studienbericht der Studie COU-AA-302 wurden Frakturen als UEs nicht separat ausgewertet. Daher war unklar, ob die vorgelegten Ergebnisse auf einem einzelnen Preferred Term (PT) nach MedDRA oder auf einer post hoc definierten Zusammenstellung von mehreren PTs beruhten.

Zusammenfassend sind die in der Stellungnahme nachgereichten Überlebenszeitanalysen aufgrund der unbegründeten diskrepanten Angaben im Vergleich zum Dossier und zum Studienbericht der Studie COU-AA-302 sowie aufgrund der fehlenden Angaben zur Operationalisierung bei Frakturen nicht verwertbar und wurden daher in der vorliegenden Bewertung nicht berücksichtigt.

Auch für die nachgereichten Ergebnisse zu den Endpunkten Schmerz und gesundheitsbezogene Lebensqualität ergaben sich aus der Stellungnahme sowie der Anlage zur Stellungnahme Unklarheiten. Zum einen war unklar, auf welchem Datenschnitt die Analysen beruhten. Zudem ging aus den Informationen nicht hervor, wie viele Patienten tatsächlich in der Auswertung berücksichtigt wurden. In der entsprechenden vorgelegten Ergebnistabelle fand sich lediglich die Angabe zur Anzahl der randomisierten Patienten. Dass alle randomisierten Patienten in die Auswertung eingingen, ist jedoch nicht wahrscheinlich, da bereits zu Studienbeginn nicht für jeden Patienten Daten vorlagen und die Zahl der Patienten mit jedem Erhebungszeitpunkt niedriger wurde. Für die vorliegende Bewertung wurde daher angenommen, dass diejenige Zahl der Patienten in die Auswertung einging, für die zum ersten Beobachtungszeitpunkt nach Studienbeginn Werte vorlagen. Insgesamt wurde für die Ergebnisse zum Schmerz und zur gesundheitsbezogenen Lebensqualität nicht davon ausgegangen, dass die beschriebenen Unklarheiten einen relevanten Einfluss auf das Ergebnis haben, sodass die vorgelegten Daten in der vorliegenden Bewertung berücksichtigt wurden.

2.2 Verzerrungspotenzial

Tabelle 2 zeigt das Verzerrungspotenzial auf Studienebene (zur Begründung siehe Dossierbewertung A13-06 [1]) sowie das Verzerrungspotenzial der Ergebnisse zu den Endpunkten stärkster Schmerz, Beeinträchtigung durch den Schmerz und gesundheitsbezogene Lebensqualität.

Tabelle 2: Verzerrungspotenzial auf Studien- und Endpunktebene – RCT, direkter Vergleich: Abirateron vs. abwartendes Vorgehen

Studie	Studienebene	Endpunkte		
		Stärkster Schmerz (gemessen mit dem BPI-SF)	Beeinträchtigung durch den Schmerz (gemessen mit dem BPI-SF)	Gesundheitsbezogene Lebensqualität (gemessen mit dem FACT-P)
COU-AA-302	n	h	h	h

h: hoch; n: niedrig; BPI-SF: Brief Pain Inventory – Short Form; FACT-P: Functional Assessment Cancer Therapy – Prostata; RCT: randomisierte kontrollierte Studie

Das Verzerrungspotenzial ist auf Studienebene niedrig. Endpunktbezogen wurde das Verzerrungspotenzial für alle im vorliegenden Addendum betrachteten Endpunkte als hoch bewertet.

Die Endpunkte zum BPI-SF und zum FACT-P wurden nur bis zum Ende der Behandlungsphase mit der Studienmedikation erhoben. Die Studienbehandlung wurde bis zum Auftreten einer Progression fortgeführt. Bei einer Progression der Erkrankung beendeten die Patienten die randomisierte Behandlung. Dabei ist der Anteil der Patienten, die die Studienbehandlung beendet haben oder die Studie aus anderen Gründen abgebrochen haben, in der Placebo-Gruppe deutlich höher als in der Abiraterongruppe, sodass auch die medianen Behandlungsdauern (und damit auch die Beobachtungsdauer für diese Endpunkte) sich deutlich unterscheiden (13,8 vs. 8,3 Monate). Aufgrund des möglichen Zusammenhangs zwischen der Progression und der gesundheitsbezogenen Lebensqualität bzw. dem Schmerz ist fraglich, ob die Annahme für eine MMRM-Analyse des „missing at random“ erfüllt ist. Somit kann für die Endpunkte zum BPI-SF und zum FACT-P nicht von einer unverzerrten Schätzung des Behandlungseffekts ausgegangen werden. Ein weiterer möglicherweise verzerrender Aspekt sind die in Abschnitt 2.1 beschriebenen Unklarheiten hinsichtlich der in der Analyse berücksichtigten Patientenzahlen.

2.3 Ergebnisse

Tabelle 3 stellt die stetigen Daten zu den Auswertungen zum Schmerz, gemessen mit dem BPI-SF, und zur gesundheitsbezogenen Lebensqualität, gemessen mit dem FACT-P, dar.

Tabelle 3: Ergebnisse (stetige Endpunkte) zu Morbidität und Lebensqualität – RCT, direkter Vergleich: Abirateron vs. abwartendes Vorgehen

Studie Endpunktkategorie Endpunkt	Abirateron / Prednison / ADT			Placebo / Prednison / ADT ^a			Abirateron / Prednison / ADT vs. Placebo / Prednison / ADT ^a	
	N ^b	Werte Studien- beginn MW (SE)	Mittelwerte über Beobachtungs- zeitraum MW ^c (SE)	N ^b	Werte Studien- beginn MW (SE)	Mittelwerte über Beobachtungs- zeitraum MW ^c (SE)	Mittelwert- differenz ^d [95 %-KI]; p-Wert	Hedges' g ^e [95 %-KI]; p-Wert
COU-AA-302^f								
Morbidität								
BPI-SF								
Stärkster Schmerz	511	1,11 (0,08)	1,31 (0,06)	502	1,13 (0,08)	1,67(0,06)	-0,36 [-0,53; -0,19]; p <0,001	-0,26 [-0,38; -0,14]; p < 0,001
Beeinträchtigung durch den Schmerz	497	0,68 (0,06)	0,96 (0,05)	493	0,68 (0,06)	1,13 (0,05)	-0,17 [-0,31; -0,02]; p = 0,025	-0,15 [-0,27; -0,02]; p = 0,022
Gesundheitsbezogene Lebensqualität								
FACT-P								
Gesamtscore	513	123,10 (0,63)	123,64 (0,55)	507	123,23 (0,63)	120,83 (0,61)	2,81 [1,19; 4,42]; p = 0,001	0,21 [0,09; 0,34]; p < 0,001
Physisches Wohlbefinden	512	25,27 (0,13)	24,75 (0,12)	506	25,31 (0,13)	24,17 (0,13)	0,58 [0,24; 0,92]; p = 0,001	0,21 [0,09; 0,33]; p < 0,001
Soziales Wohlbefinden	512	22,80 (0,17)	22,73 (0,15)	504	22,73 (0,17)	22,78 (0,16)	-0,05 [-0,48; 0,37]; p = 0,802	-0,01 [-0,14; 0,11]; p = 0,816
Emotionales Wohlbefinden	509	18,73 (0,12)	19,74 (0,11)	499	18,83 (0,12)	19,19 (0,12)	0,56 [0,24; 0,87]; p = 0,001	0,22 [0,09; 0,34]; p < 0,001
Funktionales Wohlbefinden	509	21,52 (0,17)	21,61 (0,16)	499	21,59 (0,17)	20,90 (0,17)	0,71 [0,26; 1,17]; p = 0,002	0,19 [0,07; 0,32]; p = 0,002
FACT-G	513	88,11 (0,45)	88,34 (0,40)	506	88,17 (0,45)	86,42 (0,44)	1,91 [0,75; 3,08]; p = 0,001	0,20 [0,08; 0,33]; p = 0,001
Prostatakarzinom spezifischer Score	501	35,42 (0,21)	35,83 (0,19)	498	35,50 (0,21)	34,82 (0,21)	1,01 [0,45; 1,56]; p < 0,001	0,23 [0,10; 0,35]; p < 0,001

(Fortsetzung)

Tabelle 3: Ergebnisse (stetige Endpunkte) zu Morbidität und Lebensqualität – RCT, direkter Vergleich: Abirateron vs. abwartendes Vorgehen (Fortsetzung)

Studie Endpunktkategorie Endpunkt	Abirateron / Prednison / ADT				Placebo / Prednison / ADT ^a		Abirateron / Prednison / ADT vs. Placebo / Prednison / ADT ^a	
	N ^b	Werte Studien- beginn MW (SE)	Mittelwerte über Beobachtungs- zeitraum MW ^c (SE)	N ^b	Werte Studien- beginn MW (SE)	Mittelwerte über Beobachtungs- zeitraum MW ^c (SE)	Mittelwert- differenz ^d [95 %-KI]; p-Wert	Hedges' g ^e [95 %-KI]; p-Wert
Trial Outcome Index	513	81,62 (0,48)	81,42 (0,42)	507	81,79 (0,48)	79,29 (0,46)	2,13 [0,91; 3,35]; p = 0,001	0,21 [0,09; 0,34]; p < 0,001

a: Operationalisierung des abwartenden Vorgehens
b: Anzahl der Patienten zum ersten Beobachtungszeitpunkt (unter der Annahme, dass dies die Zahl der Patienten ist, die in die Schätzung der Mittelwerte über den gesamten Beobachtungszeitraum eingehen), die Werte zu Studienbeginn können auf anderen Patientenzahlen basieren.
c: Mittelwerte (least square means) über den gesamten Beobachtungszeitraum aus einem Modell mit wiederholten Messungen (MMRM, Modell mit gemischten Effekten) mit folgenden erklärenden Variablen: Baseline-Wert, Behandlungszyklus, Behandlung, Behandlung-Behandlungszyklus-Interaktion (feste Effekte) und Patient (zufälliger Effekt).
d: Differenz der Mittelwerte über den gesamten Beobachtungszeitraum aus der MMRM-Auswertung.
e: Eigene Berechnung aus den Mittelwerten über den gesamten Beobachtungszeitraum, den zugehörigen SEs und der Anzahl der Patienten in den beiden Gruppen.
f: Die Angaben der Werte zu Studienbeginn beruhen auf dem 2. Datenschnitt (20.12.2011). Für die Mittelwerte über den Studienverlauf sowie die Effektschätzer waren keine Angaben über den verwendeten Datenschnitt verfügbar.
ADT: Androgendeprivationstherapie; BPI-SF: Brief Pain Inventory – Short Form; FACT-G: Functional Assessment Cancer Therapy - General; FACT-P: Functional Assessment Cancer Therapy - Prostate; KI: Konfidenzintervall; MMRM: Mixed-Effects-Model-Repeated-Measures; MW: Mittelwert; RCT: randomisierte kontrollierte Studie; SE: Standardfehler.

Morbidität

Schmerz, gemessen mit dem BPI-SF

Der BPI-SF besteht aus 4 Fragen zur Schmerzintensität, 2 Fragen zum Schmerzmittelverbrauch und 7 Fragen zur Beeinträchtigung durch den Schmerz. Der pU hat in seiner Stellungnahme Analysen zum einen zum stärksten Schmerz in den letzten 24 Stunden (Frage 3 des BPI-SF) und zum anderen zur Beeinträchtigung des Patienten durch den Schmerz in den letzten 24 Stunden (Fragen 9 A bis G des BPI-SF) nachgereicht. Diese Analysen waren jeweils im statistischen Analyseplan der Studie COU-AA-302 prädefiniert. Jede Frage wird vom Patienten auf einer Skala von 0 (kein Schmerz bzw. keine Beeinträchtigung) bis 10 (stärkste vorstellbare Schmerzen bzw. vollständige Beeinträchtigung) eingeschätzt. Für die Beeinträchtigung durch den Schmerz wird jeweils der Mittelwert aus den 7 Fragen gebildet. Wie in Tabelle 3 dargestellt, werden in den vorliegenden Auswertungen die Mittelwerte über den gesamten Studienverlauf zwischen den Behandlungsgruppen verglichen.

Sowohl hinsichtlich des stärksten Schmerzes in den letzten 24 Stunden als auch hinsichtlich der Beeinträchtigung durch den Schmerz zeigte sich in den nachgereichten Analysen der Studie COU-AA-302 jeweils ein statistisch signifikanter Unterschied zugunsten von Abirateron im Vergleich zum abwartenden Vorgehen. Da keine adäquaten Responderanalysen und auch skalenspezifisch validierten bzw. etablierten Relevanzkriterien für den Gruppenunterschied vorliegen, wurde auf die Betrachtung der standardisierten Mittelwertdifferenz (SMD in Form von Hedges' g) für die Relevanzbewertung zurückgegriffen. Das 95 %-KI der SMD lag jeweils nicht vollständig unterhalb der Irrelevanzschwelle von $-0,2$. Somit konnte ein irrelevanter Effekt nicht sicher ausgeschlossen werden.

Insgesamt ergibt sich aus den in der Stellungnahme nachgereichten Analysen kein Beleg für einen Zusatznutzen von Abirateron im Vergleich zum abwartenden Vorgehen unter Beibehaltung der bestehenden konventionellen Androgendeprivation für die Endpunkte stärkster Schmerz und Beeinträchtigung durch den Schmerz, jeweils gemessen mit dem BPI-SF. Da Schmerz mit dem BPI-SF nur bis zur Progression der Erkrankung erhoben, beschreiben die Ergebnisse ausschließlich diesen Zeitraum.

Gesundheitsbezogene Lebensqualität gemessen mit dem FACT-P

Der FACT-P ist ein krankheitsspezifisches Instrument zur Messung der gesundheitsbezogenen Lebensqualität. Er besteht aus dem FACT-G sowie einer für Prostatakrebs spezifischen Subskala. Jedes der insgesamt 39 Items wird vom Patienten auf einer 5-Punkte-Likert-Skala beantwortet. Für den FACT-P existiert ein Gesamtscore, der für die Bewertung maßgeblich herangezogen wird. Wie in Tabelle 3 dargestellt, werden in den vorliegenden Auswertungen die Mittelwerte über den gesamten Studienverlauf zwischen den Behandlungsgruppen verglichen.

Sowohl hinsichtlich des Gesamtscores des FACT-P als auch für die meisten der Subskalen zeigte sich in den nachgereichten Analysen ein statistisch signifikanter Unterschied zugunsten von Abirateron im Vergleich zum abwartenden Vorgehen. Da keine adäquaten Responderanalysen und auch skalenspezifisch validierten bzw. etablierten Relevanzkriterien für den Gruppenunterschied vorliegen, wurde auf die Betrachtung der standardisierten Mittelwertdifferenz (SMD in Form von Hedges' g) für die Relevanzbewertung zurückgegriffen. Das 95 %-KI der SMD lag für den Gesamtscore des FACT-P, aber auch für alle Subskalen, jeweils nicht vollständig oberhalb der Irrelevanzschwelle von 0,2. Somit konnte ein irrelevanter Effekt nicht sicher ausgeschlossen werden.

Insgesamt ergibt sich aus den in der Stellungnahme nachgereichten Analysen kein Beleg für einen Zusatznutzen von Abirateron im Vergleich zum abwartenden Vorgehen unter Beibehaltung der bestehenden konventionellen Androgendeprivation für die gesundheitsbezogene Lebensqualität, gemessen mit dem FACT-P. Da die gesundheitsbezogene Lebensqualität mit dem FACT-P nur bis zur Progression der Erkrankung erhoben, beschreiben die Ergebnisse ausschließlich diesen Zeitraum.

Nebenwirkungen

Wie in Abschnitt 2.1 beschrieben, waren die vom pU nachgereichten Überlebenszeitanalysen zu den Endpunkten zu unerwünschten Ereignissen aufgrund nicht nachvollziehbarer Diskrepanzen in den Angaben nicht verwertbar und wurden in der vorliegenden Bewertung deshalb nicht berücksichtigt. Für die Bewertung der Nebenwirkungen stehen damit nach wie vor nur die Ergebnisse, die bereits in der Dossierbewertung A13-06 beschrieben wurden, zur Verfügung.

Subgruppenanalysen

Der pU hat in seiner Stellungnahme keine Subgruppenanalysen nachgereicht. Demzufolge liegen insgesamt für die vorliegende Nutzenbewertung von Abirateron Subgruppenanalysen zu den in der Bewertung A13-06 beschriebenen relevanten Subgruppenmerkmalen nach wie vor nur für den Endpunkt Gesamtüberleben vor.

2.4 Ausmaß und Wahrscheinlichkeit des Zusatznutzens

Im Folgenden wird die Herleitung von Ausmaß und Wahrscheinlichkeit des Zusatznutzens auf Endpunktebene dargestellt. Dabei werden die verschiedenen Endpunktkategorien und die Effektgrößen berücksichtigt. Die hierzu verwendete Methodik ist in Anhang A der Nutzenbewertung A11-02 erläutert [6].

Das Vorgehen zur Ableitung einer Gesamtaussage zum Zusatznutzen anhand der Aggregation der auf Endpunktebene hergeleiteten Aussagen stellt einen Vorschlag des IQWiG dar. Über den Zusatznutzen beschließt der G-BA.

2.4.1 Beurteilung des Zusatznutzens auf Endpunktebene

Aus der Gesamtschau der in Abschnitt 2.4 der Bewertung A13-06 [1] und Abschnitt 2.3 des vorliegenden Addendums präsentierten Datenlage ergab jeweils einen Hinweis auf einen Zusatznutzen (Endpunkte Gesamtüberleben und schwerer Schmerz gemessen anhand des Beginns einer Opiattherapie) von Abirateron gegenüber der zweckmäßigen Vergleichstherapie (abwartendes Vorgehen unter Beibehaltung der bestehenden konventionellen Androgendeprivation). Ausgehend von diesen Ergebnissen wurde das Ausmaß des jeweiligen Zusatznutzens auf Endpunktniveau eingeschätzt. Die folgenden Tabellen (Tabelle 4 und Tabelle 5) stellen eine Aktualisierung der Tabellen 12 und 13 der Bewertung A13-06 dar, die um die im vorliegenden Addendum berücksichtigten Ergebnisse ergänzt wurden.

Tabelle 4: Ausmaß des Zusatznutzens auf Endpunktebene (Nutzenendpunkte): Abirateron vs. abwartendes Vorgehen

Endpunkt	Effektschätzer [95 %-KI] p-Wert Zeit bis zum Ereignis^a Abirateron / Prednison / ADT vs. Placebo / Prednison / ADT^b (Monate) Wahrscheinlichkeit^c	Ableitung des Ausmaßes^d
Mortalität^e		
Gesamtüberleben	HR: 0,79 [0,66; 0,96] p = 0,015 Median: 35,3 vs. 30,1 Wahrscheinlichkeit: Hinweis	Endpunktkategorie: Überlebensdauer $0,95 \leq KI_o < 1$ Zusatznutzen, Ausmaß: gering
Morbidität		
Schwerer Schmerz gemessen anhand des Beginns einer Opiattherapie ^e	HR: 0,71 [0,59; 0,85] p < 0,001 25 %-Quantil ^f : 14,8 vs. 12,0 Wahrscheinlichkeit: Hinweis	Endpunktkategorie: schwerwiegende / schwere Symptome / Folgekomplikationen $0,75 \leq KI_o < 0,90$ Zusatznutzen, Ausmaß: beträchtlich
Stärkster Schmerz (gemessen mit BPI-SF) ^g	MD: -0,36 [-0,53; -0,19] p < 0,001 SMD ^h : -0,26 [-0,38; -0,14] p < 0,001	geringerer Nutzen / Zusatznutzen nicht belegt
Beeinträchtigung durch den Schmerz (gemessen mit BPI-SF) ^g	MD: -0,17 [-0,31; -0,02] p = 0,025 SMD ^h : -0,15 [-0,27; -0,02] p = 0,022	geringerer Nutzen / Zusatznutzen nicht belegt
Gesundheitsbezogene Lebensqualität		
FACT-P Gesamtscore ^g	MD: 2,81 [1,19; 4,42] p = 0,001 SMD ^h : 0,21 [0,09; 0,34] p < 0,001	geringerer Nutzen / Zusatznutzen nicht belegt
<p>a: sofern zutreffend b: Operationalisierung des abwartenden Vorgehens c: Angabe der Wahrscheinlichkeit, sofern statistisch signifikante Unterschiede bzw. relevante Effekte vorliegen. d: Einschätzungen zur Effektgröße erfolgen je nach Endpunktkategorie mit unterschiedlichen Grenzen anhand der oberen Grenze des Konfidenzintervalls (KI_o). e: 3. Datenschnitt (22.05.2012) f: Die mediane Zeit bis zum Ereignis konnte in mindestens einem Behandlungsarm aufgrund des hohen Anteils zensierter Daten nicht geschätzt werden. Das 25 %-Quantil gibt die Zeit an, zu der der Kaplan-Meier-Schätzer der Survivalfunktion erstmalig 75 % unterschreitet. g: keine Angaben zum zugrundeliegenden Datenschnitt. h: SMD in Form von Hedges' g zur Bewertung der Relevanz des statistisch signifikanten Unterschieds. Liegt das 95 %-KI nicht vollständig unterhalb der Irrelevanzschwelle von -0,2 bzw. oberhalb der Irrelevanzschwelle von 0,2, kann ein irrelevanter Effekt nicht sicher ausgeschlossen werden. ADT: Androgendeprivationstherapie; BPI-SF: Brief Pain Inventory – Short Form; FACT-P: Functional Assessment of Cancer Therapy-Prostate; HR: Hazard Ratio; KI: Konfidenzintervall; KI_o: obere Grenze Konfidenzintervall; MD: Mittelwertdifferenz; pU: pharmazeutischer Unternehmer; SMD: standardisierte Mittelwertdifferenz</p>		

Tabelle 5: Ausmaß des Zusatznutzens auf Endpunktebene (Schadensendpunkte): Abirateron vs. abwartendes Vorgehen

Endpunkt	Effektschätzer [95 %-KI] p-Wert Ereignisanteil^a Abirateron / Prednison / ADT vs. Placebo / Prednison / ADT^b Wahrscheinlichkeit^c	Ableitung des Ausmaßes^d
Nebenwirkungen^e		
Gesamtrate UEs	Keine verwertbaren Daten verfügbar	Größerer / geringerer Schaden nicht belegt
Schwere UEs (CTCAE-Grade 3 und 4) ^f	RR: 1,06 [0,82; 1,37] p = 0,652 18,1 % vs. 17,0 %	Größerer / geringerer Schaden nicht belegt
SUEs	Keine verwertbaren Daten verfügbar	Größerer / geringerer Schaden nicht belegt
Therapieabbrüche wegen UEs	IDR: 0,77 [0,53; 1,11] p = 0,160 58 (8,2) vs. 53 (10,7) ^g	Größerer / geringerer Schaden nicht belegt
Frakturen	Keine verwertbaren Daten verfügbar	Größerer / geringerer Schaden nicht belegt
Flüssigkeitsretention / Ödeme	Keine verwertbaren Daten verfügbar	Größerer / geringerer Schaden nicht belegt
Ischämische Herzerkrankungen	IDR: 0,87 [0,49; 1,57] p = 0,655 25 (3,5) vs. 20 (4,0) ^g	Größerer / geringerer Schaden nicht belegt
Herzinsuffizienz	IDR: 4,20 [0,94; 18,76] p = 0,060 12 (1,7) vs. 2 (0,4) ^g	Größerer / geringerer Schaden nicht belegt
<p>a: sofern nicht anders angegeben b: Operationalisierung des abwartenden Vorgehens c: Angabe der Wahrscheinlichkeit, sofern statistisch signifikante Unterschiede vorliegen. d: Einschätzungen zur Effektgröße erfolgen je nach Endpunktkategorie mit unterschiedlichen Grenzen anhand der oberen Grenze des Konfidenzintervalls. e: 3. Datenschnitt (22.05.2012) f: schwere UEs (CTCAE-Grade 3 und 4) innerhalb von 3 Monaten nach Beginn der Studienbehandlung g: Angabe als Anzahl Patienten mit Ereignis (n/100 Patientenjahre) ADT: Androgendeprivationstherapie; CTCAE: Common Terminology Criteria for Adverse Events; HR: Hazard Ratio; IDR: Inzidenzdichteverhältnis (Incidence Density Ratio); KI: Konfidenzintervall; RR: relatives Risiko; SUE: schwerwiegendes unerwünschtes Ereignis; UE: unerwünschtes Ereignis</p>		

2.4.2 Gesamtaussage zum Zusatznutzen

Tabelle 6 fasst die Resultate zusammen, die in die Gesamtaussage zum Ausmaß des Zusatznutzens einfließen.

Tabelle 6: Positive und negative Effekte aus der Bewertung von Abirateron im Vergleich zum abwartenden Vorgehen

Positive Effekte	Negative Effekte
Hinweis für einen Zusatznutzen – Ausmaß: gering (Mortalität: Gesamtüberleben)	
Hinweis für einen Zusatznutzen – Ausmaß: beträchtlich (Morbidität, schwerwiegende / schwere Symptome / Folgekomplikationen: schwerer Schmerz gemessen anhand des Beginns einer Opiattherapie)	

Auf der Grundlage der verfügbaren bzw. verwertbaren Ergebnisse verbleiben in der Gesamtschau auf Endpunktebene ausschließlich positive Effekte. Diese bestehen in einem Hinweis auf einen geringen Zusatznutzen in der Endpunktkategorie Mortalität (Gesamtüberleben) und einem Hinweis auf einen beträchtlichen Zusatznutzen für einen Endpunkt in der Kategorie schwerwiegende / schwere Symptome / Folgekomplikationen (schwerer Schmerz gemessen anhand des Beginns einer Opiattherapie). In der Aggregation dieser positiven Effekte ergibt sich zunächst ein Hinweis auf einen beträchtlichen Zusatznutzen von Abirateron im Vergleich zum abwartenden Vorgehen unter Beibehaltung der konventionellen Androgendeprivation. Für einen großen Teil der Endpunkte zum Schaden, insbesondere für SUE, liegen aufgrund der in Abschnitt 2.1 beschriebenen Unklarheiten in den mit der Stellungnahme nachgereichten Überlebenszeitanalysen nach wie vor keine adäquaten Auswertungen vor. Daher lässt sich zum Schaden noch immer keine abschließende Aussage treffen. Aufgrund dessen bleibt die in der Bewertung A13-06 vorgenommene Herabstufung der Wahrscheinlichkeit des Zusatznutzens von Abirateron auf einen Anhaltspunkt bestehen.

Zusammenfassend ergibt sich ein Anhaltspunkt für einen beträchtlichen Zusatznutzen von Abirateron gegenüber der zweckmäßigen Vergleichstherapie (abwartendes Vorgehen unter Beibehaltung der bestehenden konventionellen Androgendeprivation) zur Behandlung des metastasierten kastrationsresistenten Prostatakarzinoms bei erwachsenen Männern mit asymptomatischem oder mild symptomatischem Verlauf der Erkrankung nach Versagen der ADT, bei denen eine Chemotherapie noch nicht klinisch indiziert ist.

3 Literatur

1. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Abirateronacetat (neues Anwendungsgebiet): Nutzenbewertung gemäß § 35a SGB V; Dossierbewertung; Auftrag A13-06 [online]. 11.04.2013 [Zugriff: 15.04.2013]. (IQWiG-Berichte; Band 160). URL: https://www.iqwig.de/download/A13-06_Abirateronacetat_neues_Anwendungsgebiet_Nutzenbewertung_35a_SGB_V.pdf.
2. Janssen-Cilag. Stellungnahme zum IQWiG-Bericht Nr. 160: Abirateronacetat (neues Anwendungsgebiet); Nutzenbewertung gemäß § 35a SGB V; Dossierbewertung; Auftrag A13-06. [Demnächst verfügbar unter: <http://www.g-ba.de/informationen/nutzenbewertung/60/#tab/beschluesse> im Dokument "Zusammenfassende Dokumentation"].
3. Janssen-Cilag. Nutzenbewertungsverfahren zum Wirkstoff Abirateronacetat (neues Anwendungsgebiet): Dossier [online]. [Zugriff: 07.06.2013]. URL: <http://www.g-ba.de/informationen/nutzenbewertung/60/#tab/dossier>.
4. Janssen Research & Development. A phase 3, randomized, double-blind, placebo-controlled study of abiraterone acetate plus prednisone in asymptomatic or mildly symptomatic subjects with metastatic castration-resistant prostate cancer: study COU-AA-302; clinical study report [unveröffentlicht]. 2012.
5. Janssen Research & Development. A phase 3, randomized, double-blind, placebo-controlled study of abiraterone acetate plus prednisone in asymptomatic or mildly symptomatic subjects with metastatic castration-resistant prostate cancer: study COU-AA-302; report of updated data [unveröffentlicht]. 2012.
6. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Ticagrelor: Nutzenbewertung gemäß § 35a SGB V; Dossierbewertung; Auftrag A11-02 [online]. 29.09.2011 [Zugriff: 11.03.2013]. (IQWiG-Berichte; Band 96). URL: https://www.iqwig.de/download/A11-02_Ticagrelor_Nutzenbewertung_%C2%A735a_SGB_V_.pdf.