

IQWiG-Berichte – Nr. 202

**Teriflunomid –
Nutzenbewertung
gemäß § 35a SGB V**

Dossierbewertung

Auftrag: A13-38
Version: 1.0
Stand: 20.12.2013

Impressum

Herausgeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema:

Teriflunomid – Nutzenbewertung gemäß § 35a SGB V

Auftraggeber:

Gemeinsamer Bundesausschuss

Datum des Auftrags:

01.10.2013

Interne Auftragsnummer:

A13-38

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

Medizinisch-fachliche Beratung:

- Bernd C. Kieseier, Universitätsklinikum der Heinrich-Heine-Universität, Düsseldorf

Das IQWiG dankt dem medizinisch-fachlichen Berater für seinen Beitrag zur Dossierbewertung. Der Berater war jedoch nicht in die Erstellung der Dossierbewertung eingebunden. Für die Inhalte der Dossierbewertung ist allein das IQWiG verantwortlich.

An der Dossierbewertung beteiligte Mitarbeiter des IQWiG¹:

- Sebastian Werner
- Gertrud Egger
- Andreas Gerber-Grote
- Ulrich Grouven
- Thomas Kaiser
- Sarah Mostardt
- Stefanie Reken
- Anke Schulz
- Siw Waffenschmidt
- Min Zhou

Schlagwörter: Teriflunomid, Multiple Sklerose, Nutzenbewertung

Keywords: Teriflunomide, Multiple Sclerosis, Benefit Assessment

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

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	vi
Abbildungsverzeichnis	vii
Abkürzungsverzeichnis	viii
1 Hintergrund	1
1.1 Verlauf des Projekts	1
1.2 Verfahren der frühen Nutzenbewertung	1
1.3 Erläuterungen zum Aufbau des Dokuments	2
2 Nutzenbewertung	3
2.1 Kurzfassung der Nutzenbewertung	3
2.2 Fragestellung	8
2.3 Informationsbeschaffung und Studienpool	8
2.3.1 Eingeschlossene Studien	9
2.3.2 Studiencharakteristika	9
2.4 Ergebnisse zum Zusatznutzen	15
2.4.1 Relevante Endpunkte.....	15
2.4.2 Datenverfügbarkeit und Verzerrungspotenzial.....	16
2.4.3 Ergebnisse	18
2.4.4 Subgruppenanalysen.....	25
2.5 Ausmaß und Wahrscheinlichkeit des Zusatznutzens	26
2.5.1 Beurteilung des Zusatznutzens auf Endpunktebene	26
2.5.2 Gesamtaussage zum Zusatznutzen	29
2.6 Liste der eingeschlossenen Studien	31
2.7 Kommentare zum Dossier des pharmazeutischen Unternehmers	32
2.7.1 Kommentar zur zweckmäßigen Vergleichstherapie (Modul 3, Abschnitt 3.1)....	32
2.7.2 Kommentar zur Darstellung von Nutzen und Zusatznutzen (Modul 4).....	32
2.7.2.1 Fragestellung / Einschlusskriterien.....	32
2.7.2.2 Methodik zur Bewertung der Aussagekraft der Nachweise sowie zur Informationssynthese und -analyse.....	33
2.7.2.3 Methodik und Ergebnisse der Informationsbeschaffung	37
2.7.2.3.1 Informationsbeschaffung.....	37
2.7.2.3.2 Studienpool	38
2.7.2.4 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel	41

2.7.2.4.1	Studiendesign und Population.....	41
2.7.2.4.2	Verzerrungspotenzial	41
2.7.2.4.3	Ergebnisse	42
2.7.2.5	Kommentar zu Ergebnissen aus weiteren Unterlagen – indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	47
2.7.2.6	Kommentar zu Ergebnissen aus weiteren Unterlagen – nicht randomisierte vergleichende Studien.....	47
2.7.2.7	Kommentar zu Ergebnissen aus weiteren Unterlagen – weitere Untersuchungen	47
2.7.2.8	Kommentar zur abschließenden Bewertung der Unterlagen zum Nachweis des Zusatznutzens	47
2.7.2.8.1	Beurteilung der Aussagekraft der Nachweise	47
2.7.2.8.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß / Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht.....	48
2.7.2.9	Kommentar zur Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	49
2.7.2.9.1	Begründung für die Vorlage indirekter Vergleiche	49
2.7.2.9.2	Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen	49
2.7.2.9.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen.....	49
2.7.2.9.4	Verwendung von Surrogatendpunkten.....	49
3	Kosten der Therapie	50
3.1	Kommentar zur Anzahl der Patienten mit therapeutisch bedeutsamem Zusatznutzen (Modul 3, Abschnitt 3.2).....	50
3.1.1	Beschreibung der Erkrankung und Charakterisierung der Zielpopulation.....	50
3.1.2	Therapeutischer Bedarf	50
3.1.3	Prävalenz und Inzidenz	50
3.1.4	Anzahl der Patienten mit therapeutisch bedeutsamen Zusatznutzen.....	51
3.2	Kommentar zu den Kosten der Therapie für die gesetzliche Krankenversicherung (Modul 3, Abschnitt 3.3)	52
3.2.1	Behandlungsdauer	52
3.2.2	Verbrauch	52
3.2.3	Kosten.....	52
3.2.4	Kosten für zusätzlich notwendige Leistungen.....	52
3.2.5	Jahrestherapiekosten.....	52
3.2.6	Versorgungsanteile	52
3.3	Konsequenzen für die Bewertung.....	53

4	Kommentare zu sonstigen Angaben im Dossier des pharmazeutischen Unternehmers.....	54
4.1	Kommentar zur Zusammenfassung der Aussagen im Dossier (Modul 1).....	54
4.2	Kommentar zu allgemeinen Angaben zum Arzneimittel und zu den zugelassenen Anwendungsgebieten (Modul 2)	54
4.3	Kommentar zur qualitätsgesicherten Anwendung (Modul 3, Abschnitt 3.4)	54
5	Zusammenfassung der Dossierbewertung.....	55
5.1	Zugelassene Anwendungsgebiete.....	55
5.2	Medizinischer Nutzen und medizinischer Zusatznutzen im Verhältnis zur zweckmäßigen Vergleichstherapie	55
5.3	Anzahl der Patienten in den für die Behandlung infrage kommenden Patientengruppen	55
5.4	Kosten der Therapie für die gesetzliche Krankenversicherung	56
5.5	Anforderungen an eine qualitätsgesicherte Anwendung	57
6	Literatur	59
	Anhang A – Darlegung potenzieller Interessenkonflikte (externe Sachverständige sowie Patientinnen und Patienten bzw. Patientenorganisationen).....	61

Tabellenverzeichnis

	Seite
Tabelle 1: Erläuterungen zum Aufbau des Dokuments	2
Tabelle 2: Teriflunomid – Ausmaß und Wahrscheinlichkeit des Zusatznutzens.....	7
Tabelle 3: Studienpool – RCT, direkter Vergleich: Teriflunomid vs. IFN β -1a.....	9
Tabelle 4: Charakterisierung der eingeschlossenen Studien – RCT, direkter Vergleich: Teriflunomid vs. IFN β -1a	10
Tabelle 5: Charakterisierung der Interventionen – RCT, direkter Vergleich: Teriflunomid vs. IFN β -1a.....	11
Tabelle 6: Charakterisierung der Studienpopulationen – RCT, direkter Vergleich: Teriflunomid vs. IFN β -1a (Studie TENERE)	13
Tabelle 7: Verzerrungspotenzial auf Studienebene – RCT, direkter Vergleich: Teriflunomid vs. IFN β -1a	14
Tabelle 8: Matrix der Endpunkte – RCT, direkter Vergleich: Teriflunomid vs. IFN β -1a.....	16
Tabelle 9: Verzerrungspotenzial auf Studien- und Endpunktebene – RCT, direkter Vergleich: Teriflunomid vs. IFN β -1a.....	17
Tabelle 10: Ergebnisse (Mortalität und Morbidität) – RCT, direkter Vergleich: Teriflunomid vs. IFN β -1a (Studie TENERE)	19
Tabelle 11: Ergebnisse (gesundheitsbezogene Lebensqualität) – RCT, direkter Vergleich: Teriflunomid vs. IFN β -1a (Studie TENERE)	22
Tabelle 12: Ergebnisse (Nebenwirkungen) – RCT, direkter Vergleich: Teriflunomid vs. IFN β -1a (Studie TENERE)	24
Tabelle 13: Ausmaß des Zusatznutzens auf Endpunktebene: Teriflunomid vs. IFN β -1a.....	27
Tabelle 14: Positive und negative Effekte aus der Bewertung von Teriflunomid im Vergleich zu IFN β -1a.....	29
Tabelle 15: Teriflunomid – Ausmaß und Wahrscheinlichkeit des Zusatznutzens.....	30
Tabelle 16: Teriflunomid – Ausmaß und Wahrscheinlichkeit des Zusatznutzens.....	55
Tabelle 17: Anzahl der GKV-Patienten in der Zielpopulation	55
Tabelle 18: Jahrestherapiekosten für die GKV für das zu bewertende Arzneimittel und die zweckmäßige Vergleichstherapie pro Patient	56

Abbildungsverzeichnis

	Seite
Abbildung 1: Kaplan-Meier-Kurven der Zeit bis zum bestätigten Schub (ITT-Population)...	21
Abbildung 2: Kaplan-Meier-Kurven der Zeit bis zur (mindestens 12 Wochen anhaltenden) Behinderungsprogression (ITT-Population).....	21

Abkürzungsverzeichnis

Abkürzung	Bedeutung
ALT	Alanin-Aminotransferase
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
EDSS	Expanded Disability Status Scale
EMA	European Medicines Agency
EPAR	European Public Assessment Report
EU	European Union
FIS	Fatigue Impact Scale
FS	funktionelles System
G-BA	Gemeinsamer Bundesausschuss
GKV	gesetzliche Krankenversicherung
HLT	High Level Term
HMG	hierarchisierte Morbiditätsgruppe
i. m.	intramuskulär
ICTRP	International Clinical Trials Registry Platform (WHO)
IFN β -1a	Interferon beta-1a
IMP	investigational medicinal products
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	intention to treat
IVRS	Interactive Voice Response System
MedDRA	Medical Dictionary for Regulatory Activities
Morbi-RSA	morbiditätsorientierter Risikostrukturausgleich
MS	Multiple Sklerose
MSFC	Multiple Sclerosis Functional Composite
OR	Odds Ratio
OTC	over the counter
PP	per protocol
PT	Preferred Term
pU	pharmazeutischer Unternehmer
RCT	randomised controlled trial (randomisierte kontrollierte Studie)
RMS	relapsing multiple sclerosis (schubförmige Multiple Sklerose)
RR	relatives Risiko
RRMS	relapsing remitting multiple sclerosis (schubförmig-remittierende Multiple Sklerose)
s. c.	subkutan

SGB	Sozialgesetzbuch
SOC	System Organ Class
SUE	schwerwiegendes unerwünschtes Ereignis
TSQM	Treatment Satisfaction Questionnaire for Medication
UE	unerwünschtes Ereignis
WHO	World Health Organization

1 Hintergrund

1.1 Verlauf des Projekts

Der Gemeinsame Bundesausschuss (G-BA) hat das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Nutzenbewertung des Wirkstoffs Teriflunomid gemäß § 35a SGB V beauftragt. Die Bewertung erfolgte auf Basis eines Dossiers des pharmazeutischen Unternehmers (pU). Das Dossier wurde dem IQWiG am 01.10.2013 übermittelt.

Die Verantwortung für die vorliegende Bewertung und für das Bewertungsergebnis liegt ausschließlich beim IQWiG. Die Bewertung wird zur Veröffentlichung an den G-BA übermittelt, der die Nutzenbewertung zur Anhörung stellt. Die Beschlussfassung über den Zusatznutzen erfolgt durch den G-BA im Anschluss an die Anhörung.

Die vorliegende Bewertung wurde unter Einbindung externer Sachverständiger (einer Beraterin / eines Beraters zu medizinisch-fachlichen Fragen) erstellt. Diese Beratung beinhaltete die schriftliche Beantwortung von Fragen zu den Themenbereichen Krankheitsbild / Krankheitsfolgen, Therapieziele, Patienten im deutschen Versorgungsalltag, Therapieoptionen, therapeutischer Bedarf und Stand der medizinischen Praxis. Darüber hinaus konnte eine Einbindung im Projektverlauf zu weiteren spezifischen Fragen erfolgen.

Die Bewertung wurde zudem unter Einbindung von Patientinnen und Patienten bzw. Patientenorganisationen erstellt. Diese Einbindung beinhaltete die schriftliche Beantwortung von Fragen zu den Themenbereichen Erfahrungen mit der Erkrankung, Notwendigkeit der Betrachtung spezieller Patientengruppen, Erfahrungen mit den derzeit verfügbaren Therapien für das Anwendungsgebiet, Erwartungen an eine neue Therapie und ggf. zusätzliche Informationen.

Die Beteiligten außerhalb des IQWiG, die in das Projekt eingebunden wurden, erhielten keine Einsicht in das Dossier des pU.

Für die vorliegende Nutzenbewertung war ergänzend zu den Angaben in den Modulen 1 bis 4 die Verwendung von Informationen aus Modul 5 des Dossiers des pU notwendig. Es handelte sich dabei um Informationen zu Studienmethodik und Studienergebnissen. Die entsprechenden Angaben wurden in den vorliegenden Bericht zur Nutzenbewertung aufgenommen.

1.2 Verfahren der frühen Nutzenbewertung

Die vorliegende Dossierbewertung ist Teil des Gesamtverfahrens zur frühen Nutzenbewertung. Sie wird gemeinsam mit dem Dossier des pU (Module 1 bis 4) auf der Website des G-BA veröffentlicht. Im Anschluss daran führt der G-BA ein Stellungnahmeverfahren zu dieser Bewertung durch. Der G-BA trifft seinen Beschluss zur frühen Nutzenbewertung nach Abschluss des Stellungnahmeverfahrens. Durch den Beschluss des G-BA werden ggf. die in der Dossierbewertung dargestellten Informationen ergänzt.

Weitere Informationen zum Stellungnahmeverfahren und zur Beschlussfassung des G-BA sowie das Dossier des pharmazeutischen Unternehmers finden sich auf der Website des G-BA (www.g-ba.de).

1.3 Erläuterungen zum Aufbau des Dokuments

Die vorliegende Dossierbewertung gliedert sich in 6 Kapitel plus Anhänge. In Kapitel 2 bis 5 sind die wesentlichen Inhalte der Dossierbewertung dargestellt. Die nachfolgende Tabelle 1 zeigt den Aufbau des Dokuments im Detail.

Tabelle 1: Erläuterungen zum Aufbau des Dokuments

Kapitel 2 – Nutzenbewertung	
Abschnitt 2.1	<ul style="list-style-type: none"> ▪ Zusammenfassung der Ergebnisse der Nutzenbewertung
Abschnitte 2.2 bis 2.6	<ul style="list-style-type: none"> ▪ Darstellung des Ergebnisses der Nutzenbewertung im Detail ▪ Angabe, ob und inwieweit die vorliegende Bewertung von der Einschätzung des pU im Dossier abweicht
Abschnitt 2.7	Kommentare zu folgenden Modulen des Dossiers des pU: <ul style="list-style-type: none"> ▪ Modul 3, Abschnitt 3.1 (Bestimmung der zweckmäßigen Vergleichstherapie) ▪ Modul 4 (Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen)
Kapitel 3 – Kosten der Therapie	
Abschnitte 3.1 und 3.2	Kommentare zu folgenden Modulen des Dossiers des pU: <ul style="list-style-type: none"> ▪ Modul 3, Abschnitt 3.2 (Anzahl der Patienten mit therapeutisch bedeutsamem Zusatznutzen) ▪ Modul 3, Abschnitt 3.3 (Kosten der Therapie für die gesetzliche Krankenversicherung)
Abschnitt 3.3	<ul style="list-style-type: none"> ▪ Zusammenfassung der daraus entstehenden Konsequenzen für die Bewertung
Kapitel 4 – Kommentare zu sonstigen Angaben im Dossier	
Abschnitte 4.1 bis 4.3	Kommentare zu folgenden Modulen des Dossiers des pU: <ul style="list-style-type: none"> ▪ Modul 1 (Zusammenfassung der Aussagen im Dossier) ▪ Modul 2 (Allgemeine Angaben zum Arzneimittel, zugelassene Anwendungsgebiete) ▪ Modul 3, Abschnitt 3.4 (Anforderungen an eine qualitätsgesicherte Anwendung)
Kapitel 5 – Zusammenfassung der Dossierbewertung	
Abschnitte 5.1 bis 5.5	<ul style="list-style-type: none"> ▪ Zusammenfassung der wesentlichen Aussagen als Bewertung der Angaben im Dossier des pU nach § 4 Absatz 1 AM-NutzenV [1]
AM-NutzenV: Arzneimittel-Nutzenbewertungsverordnung; pU: pharmazeutischer Unternehmer	

Die Kommentierung der Angaben im Dossier des pU erfolgte unter Berücksichtigung der Anforderungen, die in den vom G-BA bereitgestellten Dossievorlagen beschrieben sind (siehe Verfahrensordnung des G-BA [2]).

Bei Abschnittsverweisen, die sich auf Abschnitte im Dossier des pU beziehen, ist zusätzlich das betroffene Modul des Dossiers angegeben. Abschnittsverweise ohne Angabe eines Moduls beziehen sich auf den vorliegenden Bericht zur Nutzenbewertung.

2 Nutzenbewertung

2.1 Kurzfassung der Nutzenbewertung

Hintergrund

Der G-BA hat das IQWiG mit der Nutzenbewertung des Wirkstoffs Teriflunomid gemäß § 35a SGB V beauftragt. Die Bewertung erfolgte auf Basis eines Dossiers des pharmazeutischen Unternehmers (pU). Das Dossier wurde dem IQWiG am 01.10.2013 übermittelt.

Fragestellung

Das Ziel des vorliegenden Berichts ist die Bewertung des Zusatznutzens von Teriflunomid im Vergleich zur zweckmäßigen Vergleichstherapie bei erwachsenen Patienten mit schubförmig-remittierender Multipler Sklerose (relapsing remitting multiple sclerosis [RRMS]).

Der G-BA hat für das Anwendungsgebiet der schubförmigen Multiplen Sklerose (relapsing multiple Sclerosis [RMS]) die zweckmäßige Vergleichstherapie wie folgt festgelegt: Beta-Interferone (1a oder 1b) oder Glatirameracetat unter Beachtung des jeweils zugelassenen Anwendungsgebietes.

Da die RRMS eine Teilmenge der RMS ist, gilt die Festlegung des G-BA auch für das zugelassene Anwendungsgebiet von Teriflunomid.

Der pU wählt Interferon beta-1a (IFN β -1a) unter den vom G-BA festgelegten Alternativen aus, beschränkt sich jedoch auf IFN β -1a 44 μ g s. c. 3-mal pro Woche (Rebif), eines der Präparate aus dieser Wirkstoffgruppe. Die zum Vergleich mit Rebif ausgestaltete Recherche würde Studien zum Vergleich anderer Präparate innerhalb der Wirkstoffgruppe nicht identifizieren. Gemäß der Festlegung des G-BA sind jedoch alle Darreichungsformen und damit alle IFN β -1a Präparate zu berücksichtigen. Insgesamt hat die Einschränkung der Vergleichstherapie des pU jedoch keine Konsequenz auf das vorliegende Bewertungsergebnis.

Die vorliegende Nutzenbewertung wurde im Vergleich zur zweckmäßigen Vergleichstherapie des G-BA Beta-Interferon 1a (IFN β -1a) durchgeführt.

Die Bewertung wurde bezüglich patientenrelevanter Endpunkte vorgenommen.

Ergebnisse

Studienpool

In die Bewertung wurde eine randomisierte kontrollierte Studie zum direkten Vergleich von Teriflunomid mit der zweckmäßigen Vergleichstherapie eingeschlossen (TENERE). Die Studie wurde offen durchgeführt und verglich Teriflunomid mit IFN β -1a 44 μ g s. c. (Rebif). Die Behandlungsdauer von Patienten variierte zwischen minimal 48 Wochen und (erwartet) maximal 118 Wochen.

Der pU führte unterstützend zum direkten Vergleich, einen indirekten Vergleich von Teriflunomid gegenüber IFN β -1a 44 μ g s. c. (Rebif) durch, für den er 3 Studien vorlegte. Dabei verglichen 2 Studien Teriflunomid mit Placebo (TEMPO, TOWER) und eine Studie IFN β -1a 44 μ g s. c. (Rebif) mit Placebo (PRISMS). Der pU führte einen indirekten Vergleich nach Bucher über den Brückenkomparator Placebo durch („indirekter Vergleich“) und fasste den resultierenden Effektschätzer mit dem Ergebnis der direkt vergleichenden TENERE-Studie zu einer Gesamtaussage zusammen („indirekter Vergleich + TENERE“). Dieser indirekte Vergleich ist jedoch nicht geeignet, Aussagen zum Zusatznutzen von Teriflunomid aus dem direkten Vergleich zu stützen. Zum einen ist der Studienpool des indirekten Vergleichs aufgrund der vom pU gewählten Einschränkung auf das Wirkstoffpräparat IFN β -1a 44 μ g s. c. (Rebif) unvollständig. Zum anderen bestehen bei mehreren Punkten Zweifel an der inhaltlichen Ähnlichkeit der Studien. Darüber hinaus wurde die vorliegende Heterogenität zwischen den Studien TEMPO und TOWER im Rahmen der indirekten Vergleiche nicht adäquat berücksichtigt.

Da der indirekte Vergleich nicht geeignet war, Aussagen zum Zusatznutzen von Teriflunomid zu stützen, erfolgte die Bewertung allein auf Basis der direkt vergleichenden Studie TENERE. Das Verzerrungspotenzial auf Studienebene wurde für diese Studie als niedrig eingestuft. Das Verzerrungspotenzial der relevanten Endpunkte (mit Ausnahme der Gesamtmortalität) wurde aufgrund der fehlenden Verblindung von Patient und behandelnder Person sowie der grundsätzlich vorhandenen subjektiven Komponente bei der Erfassung der Endpunkte als hoch eingestuft.

Mortalität

In den Behandlungsarmen der Studie TENERE sind keine Patienten verstorben. Die Studie war aufgrund ihrer Größe und auch ihrer Dauer aber nicht darauf ausgelegt, Unterschiede hinsichtlich der Mortalität aufzudecken. Ein geringerer Nutzen / Zusatznutzen von Teriflunomid ist für diesen Endpunkt nicht belegt.

Morbidität

Weder hinsichtlich der schubbezogenen Endpunkte noch hinsichtlich der Endpunkte zur Behinderungsprogression zeigten sich statistisch signifikante Unterschiede zwischen den Behandlungsgruppen.

Bei den schubbezogenen Endpunkten zeigten die Effektschätzer eher einen ungünstigen Effekt von Teriflunomid. Bei den Endpunkten zur Behinderungsprogression lagen die Effektschätzer um den Nulleffekt. Aufgrund der breiten Konfidenzintervalle kann bei beiden Endpunkten mehr als eine Verdoppelung des Risikos zuungunsten von Teriflunomid nicht ausgeschlossen werden. Die Studie war aufgrund der kurzen Studiendauer nicht geeignet, um relevante Effekte hinsichtlich der Behinderungsprogression nachzuweisen. Ein negativer Effekt von Teriflunomid gegenüber IFN β -1a kann hinsichtlich der Endpunkte zur Morbidität nicht sicher ausgeschlossen werden. Ein Zusatznutzen von Teriflunomid ist für diese Morbiditätsendpunkte nicht belegt.

Gesundheitsbezogene Lebensqualität

In der Studie TENERE wurde die symptomsspezifische gesundheitsbezogene Lebensqualität erhoben, die mit der Symptomatik der Erschöpfung (Fatigue) einhergeht. Sie wurde mittels Fatigue Impact Scale (FIS) gemessen. Beim FIS-Gesamtwert sowie den Werten der 3 Subskalen (kognitive, physische und psychosoziale Dimension) zeigte sich jeweils kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen. Bei der Interpretation der vorliegenden Ergebnisse ist zu beachten, dass ein großer Teil der Daten nach 48 Wochen nicht explizit beobachtet wurden (Teriflunomid: 23 [22 %], Rebif: 32 [33 %]). Es bleibt unklar, inwieweit die dem verwendeten MMRM-Modell (Mixed-effects Model with Repeated Measures) zugrunde liegenden Annahmen tatsächlich zutrafen.

Andere Instrumente zur Erfassung der gesundheitsbezogenen Lebensqualität (z. B. generische Instrumente) wurden in der Studie TENERE nicht angewendet. Zusammenfassend ist hinsichtlich der gesundheitsbezogenen Lebensqualität ein Zusatznutzen von Teriflunomid nicht belegt.

Nebenwirkungen

Der Anteil von Patienten mit schwerwiegenden unerwünschten Ereignissen unterschied sich zwischen den Behandlungsgruppen nicht maßgeblich und das Ergebnis war nicht statistisch signifikant. Für diesen Endpunkt ist ein geringerer / größerer Schaden von Teriflunomid nicht belegt.

Beim Endpunkt Abbruch wegen unerwünschter Ereignisse zeigte sich kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen, wobei Behandlungsabbrüche aufgrund von Schwangerschaft bzw. aufgrund von Untersuchungen (Laborwertveränderungen) nicht als relevante Ereignisse gewertet wurden. Eine Sensitivitätsanalyse, in der auch Behandlungsabbrüche aufgrund von Untersuchungen (Laborwertveränderungen) berücksichtigt wurden, zeigte ebenfalls keinen statistisch signifikanten Unterschied zwischen den Behandlungsgruppen. Für den Endpunkt Abbruch wegen unerwünschter Ereignisse ist ein geringerer / größerer Schaden von Teriflunomid nicht belegt.

Reaktionen an der Injektionsstelle und grippeähnliche Symptome traten bei der Behandlung mit IFN β -1a häufiger auf als mit Teriflunomid. Das Ergebnis war jeweils statistisch signifikant. Die Mehrheit dieser Ereignisse war von milder bzw. moderater Intensität. Unter Berücksichtigung des hohen Verzerrungspotenzials beider Endpunkte ergibt sich für Reaktionen an der Injektionsstelle und grippeähnliche Symptome jeweils ein Anhaltspunkt für einen geringeren Schaden von Teriflunomid.

Alopezie und Diarrhö traten bei der Behandlung mit Teriflunomid häufiger auf als mit IFN β -1a. Das Ergebnis war jeweils statistisch signifikant. Die Mehrheit dieser Ereignisse war von milder bzw. moderater Intensität. Unter Berücksichtigung des hohen Verzerrungspotenzials beider Endpunkte, ergibt sich für Alopezie und Diarrhö jeweils ein Anhaltspunkt für einen größeren Schaden von Teriflunomid.

Übelkeit und Erbrechen traten bei der Behandlung mit Teriflunomid häufiger auf als mit IFN β -1a. Das Ergebnis war statistisch signifikant bei allerdings nur geringfügiger Effektstärke. Für diesen Endpunkt ist, im Kontext der frühen Nutzenbewertung, ein größerer Schaden von Teriflunomid nicht belegt.

Subgruppenanalysen

Subgruppenauswertungen lagen im Dossier des pU nur für den relevanten Endpunkt jährliche Schubrate vor. Für diesen Endpunkt zeigten sich keine relevanten Subgruppeneffekte bzw. Effektmodifikationen für die untersuchten Merkmale (Geschlecht, Alter, Region, Expanded Disability Status Scale [EDSS]-Wert zu Studienbeginn, Anzahl der Schübe in den letzten 2 Jahren und Vorbehandlung mit krankheitsmodifizierenden MS-Wirkstoffen). Aufgrund der fehlenden Subgruppenauswertungen zu anderen relevanten Endpunkten können heterogene Behandlungseffekte insgesamt nicht beurteilt werden.

Wahrscheinlichkeit und Ausmaß des Zusatznutzens, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen

Auf Basis der dargestellten Ergebnisse werden das Ausmaß und die Wahrscheinlichkeit des Zusatznutzens von Teriflunomid im Vergleich zur zweckmäßigen Vergleichstherapie wie folgt bewertet:

Auf der Grundlage von besonderen Nebenwirkungen ergeben sich positive und negative Effekte mit gleicher Aussagesicherheit (Anhaltspunkt) und Endpunktkategorie (nicht schwerwiegende / schwere Nebenwirkungen). Auf beiden Seiten wird das Ausmaß „beträchtlich“ erreicht. Dabei liegen die positiven Effekte (Reaktionen an der Injektionsstelle, grippeähnliche Symptome) jeweils mit Ausmaß „beträchtlich“ vor, während auf der negativen Seite ein Effekt das Ausmaß „beträchtlich“ (Alopezie) und ein Effekt das Ausmaß „gering“ (Diarrhö) aufweist. Bezüglich der Morbiditätsendpunkte Schübe bzw. Behinderungsprogression kann ein negativer Effekt von Teriflunomid gegenüber IFN β -1a nicht sicher ausgeschlossen werden.

In der Abwägung der positiven und negativen Effekte ergibt sich kein Beleg für einen Zusatznutzen von Teriflunomid gegenüber der zweckmäßigen Vergleichstherapie.

Demnach gibt es auch keine Patientengruppen, für die sich ein therapeutisch bedeutsamer Zusatznutzen ableiten lässt.

Tabelle 2 fasst die Ergebnisse der Bewertung des Zusatznutzens von Teriflunomid gegenüber der zweckmäßigen Vergleichstherapie im zugelassenen Anwendungsgebiet zusammen.

Tabelle 2: Teriflunomid – Ausmaß und Wahrscheinlichkeit des Zusatznutzens

Anwendungsgebiet	Zweckmäßige Vergleichstherapie ^a	Ausmaß und Wahrscheinlichkeit des Zusatznutzens
Erwachsene Patienten mit schubförmig-remittierender Multipler Sklerose	Beta-Interferon (1a oder 1b) oder Glatirameracetat	Zusatznutzen nicht belegt
<p>a: Dargestellt ist jeweils die vom G-BA festgelegte zweckmäßige Vergleichstherapie. In den Fällen, in denen der pU aufgrund der Festlegung der zweckmäßigen Vergleichstherapie durch den G-BA aus mehreren Alternativen eine Vergleichstherapie auswählen kann, ist die entsprechende Auswahl des pU fett markiert. G-BA: Gemeinsamer Bundesausschuss</p>		

Das Vorgehen zur Ableitung einer Gesamtaussage zum Zusatznutzen stellt einen Vorschlag des IQWiG dar. Über den Zusatznutzen beschließt der G-BA.

2.2 Fragestellung

Das Ziel des vorliegenden Berichts ist die Bewertung des Zusatznutzens von Teriflunomid im Vergleich zur zweckmäßigen Vergleichstherapie bei erwachsenen Patienten mit schubförmig-remittierender Multipler Sklerose (relapsing remitting multiple sclerosis [RRMS]) [3].

Der G-BA hat für das Anwendungsgebiet der schubförmigen Multiplen Sklerose (relapsing multiple Sclerosis [RMS]) die zweckmäßige Vergleichstherapie wie folgt festgelegt:

- Beta-Interferone (1a oder 1b) oder Glatirameracetat unter Beachtung des jeweils zugelassenen Anwendungsgebietes.

Da die RRMS eine Teilmenge der RMS ist, gilt die Festlegung des G-BA auch für das zugelassene Anwendungsgebiet von Teriflunomid.

Der pU wählt Interferon beta-1a (IFN β -1a) unter den vom G-BA als zweckmäßige Vergleichstherapie festgelegten Alternativen aus, beschränkt sich jedoch auf IFN β -1a 44 μ g s. c. 3-mal pro Woche (Rebif [4]), eines der Präparate aus dieser Wirkstoffgruppe. Die zum Vergleich mit Rebif ausgestaltete Recherche würde Studien zum Vergleich anderer Präparate innerhalb der Wirkstoffgruppe nicht identifizieren (siehe Abschnitt 2.7.1). Gemäß der Festlegung des G-BA sind jedoch alle Darreichungsformen und damit alle IFN β -1a Präparate zu berücksichtigen. Insgesamt hat die Einschränkung der Vergleichstherapie des pU jedoch keine Konsequenz auf das vorliegende Bewertungsergebnis.

Die vorliegende Nutzenbewertung wurde gegenüber der zweckmäßigen Vergleichstherapie IFN β -1a durchgeführt.

Die Bewertung erfolgte bezüglich patientenrelevanter Endpunkte.

Weitere Informationen zur Fragestellung befinden sich in Modul 3, Abschnitt 3.1 sowie Modul 4, Abschnitt 4.2.1 des Dossiers und in Abschnitt 2.7.1 sowie 2.7.2.1 der vorliegenden Dossierbewertung.

2.3 Informationsbeschaffung und Studienpool

Der Studienpool der Bewertung wurde anhand der folgenden Angaben zusammengestellt:

Quellen des pU im Dossier:

- Studienliste zu Teriflunomid (bis zum 29.07.2013 abgeschlossene Studien)
- bibliografische Literaturrecherche zu Teriflunomid (letzte Suche am 19.08.2013)
- Suche in Studienregistern zu Teriflunomid (letzte Suche am 29.07.2013)

Eigene Recherchen zur Überprüfung der Vollständigkeit des Studienpools:

- bibliografische Literaturrecherche zu Teriflunomid (letzte Suche am 31.10.2013)

- Suche in Studienregistern zu Teriflunomid (letzte Suche am 15.10.2013)

Mit den benannten Schritten zur Informationsbeschaffung wurde eine relevante Studie zum direkten Vergleich von Teriflunomid mit der zweckmäßigen Vergleichstherapie identifiziert. Der resultierende Studienpool für direkt vergleichende Studien entsprach dem des pU.

Darüber hinaus legte der pU einen indirekten Vergleich vor. Dieser war jedoch für die vorliegende Fragestellung nicht geeignet (siehe Abschnitt 2.7.2.3.2).

Weitere Informationen zu den Einschlusskriterien für Studien in die vorliegende Nutzenbewertung und zur Methodik der Informationsbeschaffung befinden sich in Modul 4, Abschnitte 4.2.2 sowie 4.2.3 des Dossiers und in den Abschnitten 2.7.2.1 sowie 2.7.2.3 der vorliegenden Dossierbewertung.

2.3.1 Eingeschlossene Studien

Tabelle 3 zeigt die in die Nutzenbewertung eingeschlossenen Studien.

Tabelle 3: Studienpool – RCT, direkter Vergleich: Teriflunomid vs. IFN β -1a

Studie	Studienkategorie		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja / nein)	gesponserte Studie ^a (ja / nein)	Studie Dritter (ja / nein)
TENERE	ja	ja	nein

a: Studie, für die der Unternehmer Sponsor war oder an der der Unternehmer anderweitig finanziell beteiligt war.
IFN β -1a: Interferon beta-1a; RCT: randomisierte kontrollierte Studie

Der Studienpool für direkt vergleichende Studien mit Teriflunomid und der zweckmäßigen Vergleichstherapie entsprach dem des pU.

Abschnitt 2.6 enthält eine Liste der Quellen für die eingeschlossenen Studien.

Weitere Informationen zum Ergebnis der Informationsbeschaffung und zum daraus hervorgehenden Studienpool befinden sich in Modul 4, Abschnitte 4.3.1.1 sowie 4.3.2.1.1 des Dossiers und in den Abschnitten 2.7.2.3.1 sowie 2.7.2.3.2 der vorliegenden Dossierbewertung.

2.3.2 Studiencharakteristika

Tabelle 4 und Tabelle 5 beschreiben die in der Nutzenbewertung eingeschlossene Studie TENERE.

Tabelle 4: Charakterisierung der eingeschlossenen Studien – RCT, direkter Vergleich: Teriflunomid vs. IFN β -1a

Studie	Studien- design	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; sekundäre Endpunkte ^a
TENERE	RCT, offen ^b , parallel	Patienten (≥ 18 Jahre) mit schubförmigen Formen der Multiplen Sklerose (MS) mit gesicherter MS- Diagnose nach McDonald's Kriterien (2005 Revision [5]) und EDSS-Score $\leq 5,5$	Teriflunomid 7 mg (N = 109) Teriflunomid 14 mg (N = 111) IFN β -1a 44 μ g (Rebif) (N = 104) Davon relevante Teilpopulation mit RRMS ^c : Teriflunomid 14 mg (n = 108) IFN β -1a 44 μ g (Rebif) (n = 104)	Screening: 4 Wochen Behandlungsdauer (erwartet): zwischen minimal 48 und maximal 118 Wochen ^d (festes Ende für alle Patienten nachdem der letzte randomisierte Patient 48 Wochen behandelt wurde). Follow-up (inklusive Eliminationsphase): 16 Wochen	53 Zentren in 13 Ländern (Belgien, Deutschland, Frankreich, Griechenland, Großbritannien, Italien, Kanada, Polen, Schweiz, Spanien, Tschechien, Tunesien, Ungarn) 4/2009–9/2011	Primär: Zeit bis zum Therapieversagen (Schub oder Abbruch der Behandlung) Sekundär: Mortalität, schubbezogene Endpunkte, Behinderungsprogression, gesundheitsbezogene Lebensqualität (symptomspezifisch [FIS]), unerwünschte Ereignisse
<p>a: Primäre Endpunkte beinhalten Angaben ohne Berücksichtigung der Relevanz für diese Nutzenbewertung. Sekundäre Endpunkte beinhalten ausschließlich Angaben zu relevanten verfügbaren Endpunkten für diese Nutzenbewertung.</p> <p>b: bezogen auf die zu vergleichenden Wirkstoffe Teriflunomid und IFN β-1a 44 μg (Rebif); Die Teriflunomidarme (7 mg und 14 mg) waren doppelt verblindet.</p> <p>c: Dargestellt ist nur die zulassungskonforme Dosis von Teriflunomid 14 mg.</p> <p>d: Die tatsächliche maximale Behandlungsdauer (erster randomisierter Patient) betrug ca. 115 Wochen. Die tatsächlichen Behandlungsdauern mit den untersuchten Wirkstoffen Teriflunomid (14 mg) vs. IFN β-1a 44 μg (Rebif) betragen zum Ende der Studie im Median 449,5 vs. 421,0 Tage (mit Minimum 27 vs. 19 Tage bzw. Maximum 755 vs. 800 Tage).</p> <p>EDSS: Expanded Disability Status Scale; FIS: Fatigue Impact Scale; IFN β-1a: Interferon beta-1a; MS: Multiple Sklerose; N: Anzahl randomisierter Patienten; n: relevante Teilpopulation; RCT: randomisierte kontrollierte Studie; RRMS: schubförmig-remittierende Multiple Sklerose</p>						

Tabelle 5: Charakterisierung der Interventionen – RCT, direkter Vergleich: Teriflunomid vs. IFN β -1a

Studie	Intervention	Vergleich
TENERE	Teriflunomid 14 mg p. o. 1 x täglich <ul style="list-style-type: none"> ▪ Eliminationsphase (Wash-out): bei Patienten, die die Behandlung abbrachen oder die Studie beendeten und nicht an der optionalen Extensionsstudie teilnahmen, wurde ein Verfahren zur beschleunigten Elimination durchgeführt (Cholestyramin 8 g alle 8 h [24 g/Tag] oder 50 g Aktivkohlepulver alle 6 h [200 g/Tag] jeweils über einen Zeitraum von 11 Tagen). ▪ Behandlung von akuten Schüben mit Methylprednisolon Natriumsuccinat 1 g/Tag i. v. für 3–5 Tage 	IFN β -1a 44 μ g s. c. (Rebif) 3 x wöchentlich <ul style="list-style-type: none"> ▪ Auftrationsphase von 2 Wochen mit 8,8 μg, anschließenden 2 Wochen mit 22 μg und ab Woche 5 Gabe von 44 μg. ▪ Bei schlechter Verträglichkeit war eine Reduktion auf die niedrigere 22 μg Dosis erlaubt.^a ▪ Vor der Injektion war die Gabe eines nicht steroidalen antientzündlichen Präparates empfohlen ▪ Behandlung von akuten Schüben mit Methylprednisolon Natriumsuccinat 1 g/Tag i. v. für 3–5 Tage
a: Im Laufe der Studie wurde bei 12 Patienten eine entsprechende Dosisreduktion vorgenommen. IFN β -1a: Interferon beta-1a; i. v.: intravenös; p. o.: oral; RCT: randomisierte kontrollierte Studie; s. c.: subkutan		

Die Studie TENERE war eine multizentrische, randomisierte, kontrollierte Zulassungsstudie zum Vergleich von Teriflunomid (in 2 Dosierungen) mit IFN β -1a 44 μ g (Rebif). Die Studie wurde in Bezug auf den relevanten Vergleich offen durchgeführt.

In die Studie wurden Patienten mit schubförmiger Multipler Sklerose eingeschlossen. Die Diagnose der Multiplen Sklerose erfolgte über die revidierten McDonald Kriterien (2005) [5]. Es wurden nur Patienten eingeschlossen, die einen Expanded Disability Status Scale (EDSS)-Wert $\leq 5,5$ aufwiesen. Somit wurde nur ein Teil der Patientenpopulation untersucht, für die Teriflunomid zugelassen ist. Patienten mit höherer Krankheitsschwere, die laut EDSS bereits auf Gehhilfen angewiesen sind, wurden aus der Studie ausgeschlossen.

Die Patienten wurden im Verhältnis 1:1:1 zufällig den 3 Behandlungsarmen (Teriflunomid 7 mg, Teriflunomid 14 mg und IFN β -1a 44 μ g [Rebif]) zugeteilt. Die Randomisierung war stratifiziert nach Region und EDSS-Ausgangswert ($\leq 3,5$ / $> 3,5$). Der Anteil der für die Nutzenbewertung relevanten Teilpopulation mit RRMS betrug im Behandlungsarm mit der zulassungskonformen Dosierung von Teriflunomid (14 mg) 97,3 % und im Behandlungsarm mit IFN β -1a 44 μ g (Rebif) 100 %. Aufgrund des nur sehr kleinen Anteils der nicht zulassungskonformen Studienpopulation im Studienarm mit Teriflunomid (14 mg) werden die Gesamtergebnisse der Studie herangezogen. Der Behandlungsarm mit Teriflunomid 7 mg wird im Folgenden nicht weiter betrachtet.

Patienten im Interventionsarm (Teriflunomid) erhielten täglich eine Tablette mit Teriflunomid 14 mg. Patienten im Vergleichsarm erhielten IFN β -1a 44 μ g 3-mal wöchentlich (Rebif) als subkutane Injektion. Die Gabe von IFN β -1a 44 μ g (Rebif) erfolgte gemäß Fachinformation [4] mit einer Auftrationsphase von 2 Wochen mit 8,8 μ g IFN β -1a, anschließenden 2

Wochen mit 22 µg IFN β-1a und ab Woche 5 mit der Gabe der Erhaltungsdosis von 44 µg IFN β-1a jeweils 3-mal wöchentlich. Des Weiteren war bei schlechter Verträglichkeit, wie gemäß Zulassung vorgesehen, eine Reduktion auf die niedrigere 22 µg Dosis erlaubt. Vor der Injektion war die Gabe eines nicht steroidal entzündlichen Präparates empfohlen.

Die Behandlungsdauer von Patienten variierte zwischen minimal 48 Wochen und (erwartet) maximalen 118 Wochen (die tatsächliche maximale Behandlungsdauer für den ersten randomisierten Patienten betrug ca. 115 Wochen). Alle Patienten der Studie wurden solange behandelt, bis der letzte randomisierte Patient die 48 Wochen Behandlung beendet hatte. Die tatsächlichen Behandlungsdauern mit den untersuchten Wirkstoffen Teriflunomid vs. IFN β-1a 44 µg (Rebif) betragen zum Ende der Studie im Median 449,5 vs. 421,0 Tage. Die Behandlungsdauer war damit im Median kürzer als von den Zulassungsbehörden empfohlen, die eine Mindestbeobachtungsdauer von 2 Jahren fordern [6,7]. Dieser Aspekt ist bei der Interpretation der Ergebnisse bzw. der Beurteilung der Aussagekraft der Nachweise zu berücksichtigen, insbesondere bei in der Regel längerfristig auftretenden Ereignissen wie Mortalität, Behinderungsprogression und Malignitäten. Der geringfügige Unterschied zwischen den Behandlungsdauern der beiden Behandlungsgruppen (28,5 Tage) wird für die vorliegende Bewertung als vernachlässigbar hinsichtlich der Interpretierbarkeit der Studienergebnisse bewertet.

Im Anschluss an die Behandlungsphase erfolgte eine 16-wöchige Follow-up Phase. Diese beinhaltete im Behandlungsarm mit Teriflunomid eine 11-tägige Eliminationsphase („Wash-out“) für Patienten, die die Behandlung mit Teriflunomid abbrachen oder die Studie beendeten und nicht an der optional anschließenden Extensionsstudie teilnahmen. Um eine Verzerrung bezüglich des eingeleiteten Eliminationsverfahrens zu vermeiden, wird für die Nutzenbewertung die Beobachtungsperiode während der Behandlung (also vor der Eliminationsphase) betrachtet und entsprechende Ergebnisse dargestellt (siehe Abschnitt 2.7.2.4.3).

Tabelle 6 zeigt die Charakteristika der Patienten in der eingeschlossenen Studie.

Tabelle 6: Charakterisierung der Studienpopulationen – RCT, direkter Vergleich:
Teriflunomid vs. IFN β -1a (Studie TENERE)

Charakteristika Kategorie	Teriflunomid N = 111	IFN β-1a N = 104
Alter (Jahre)		
MW (SD)	36,8 (10,3)	37,0 (10,6)
< 38 Jahre (n [%])	62 (55,9)	60 (57,7)
\geq 38 Jahre (n [%])	49 (44,1)	44 (42,3)
Geschlecht (n [%])		
Weiblich	78 (70,3)	71 (68,3)
Männlich	33 (29,7)	33 (31,7)
Zeit seit erster MS-Diagnose (Jahre)		
MW (SD)	3,68 (6,2)	3,82 (5,7)
Median (Min–Max)	0,75 (0,1–36,5)	1,00 (0,1–30,3)
Anzahl Schübe in letzten 2 Jahren		
MW (SD)	1,7 (0,9)	1,7 (1,1)
0 (n [%])	7 (6,3)	11 (10,6)
1 (n [%])	41 (36,9)	39 (37,5)
2 (n [%])	41 (36,9)	30 (28,8)
3 (n [%])	20 (18,0)	18 (17,3)
\geq 4 (n [%])	2 (1,8)	6 (5,8)
MS-Verlaufsform (n [%])		
RRMS	108 (97,3)	104 (100)
SPMS	1 (0,9)	0 (0)
PRMS	2 (1,8)	0 (0)
Vorbehandlung mit MS-Wirkstoffen in den letzten 2 Jahren (n [%])		
Ja	13 (11,7)	25 (24,0)
Nein	98 (88,3)	79 (76,0)
EDSS-Wert zu Studienbeginn		
MW (SD)	2,33 (1,35)	2,04 (1,19)
\leq 3,5 (n [%])	95 (85,6)	94 (90,4)
$>$ 3,5 (n [%])	16 (14,4)	10 (9,6)
Behandlungsabbrecher (n [%])	22 (19,8)	30 (28,8)
EDSS: Expanded Disability Status Scale; IFN β -1a: Interferon beta-1a; MS: Multiple Sklerose; MW: Mittelwert; n: Anzahl Patienten in Kategorie; N: Anzahl der randomisierten Patienten; PRMS: progredient schubförmige MS; RCT: randomisierte kontrollierte Studie; RRMS: schubförmig- remittierende MS; SD: Standardabweichung; SPMS: sekundär progrediente MS		

Die Patienten in den Behandlungsgruppen waren im Mittelwert 37 Jahre alt bzw. seit ca. 4 Jahren an MS erkrankt und mehrheitlich weiblich (ca. 70 %). Die meisten Patienten waren therapienaiv (88 % der Patienten im Teriflunomidarm bzw. 76 % IFN β -1a-Arm der Studie). Die in die Studie eingeschlossenen Patienten wiesen eine eher leichte Symptomatik auf

(hinsichtlich des EDSS-Werts zu Studienbeginn oder der Anzahl der Schübe in den letzten 2 Jahren), was im Hinblick auf die Einschlusskriterien erwartungsgemäß ist.

Tabelle 7 zeigt das Verzerrungspotenzial auf Studienebene.

Tabelle 7: Verzerrungspotenzial auf Studienebene – RCT, direkter Vergleich: Teriflunomid vs. IFN β -1a

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung			Keine sonstigen Aspekte	Verzerrungspotenzial auf Studienebene
			Patient	Behandelnde Personen	Ergebnisunabhängige Berichterstattung		
TENERE	ja	ja	nein	nein	ja	ja	niedrig

IFN β -1a: Interferon beta-1a; RCT: randomisierte kontrollierte Studie

Das Verzerrungspotenzial auf Studienebene wurde für die Studie TENERE als niedrig eingestuft. Dies entspricht der Einschätzung des pU (siehe Abschnitt 2.7.2.4.2).

Weitere Informationen zum Studiendesign und zu den Studienpopulationen sowie zum Verzerrungspotenzial auf Studienebene befinden sich in Modul 4, Abschnitte 4.3.1.2.1, 4.3.1.2.2, 4.3.2.1.2 sowie Anhang 4-G des Dossiers und in den Abschnitten 2.7.2.4.1 sowie 2.7.2.4.2 der vorliegenden Dossierbewertung.

2.4 Ergebnisse zum Zusatznutzen

2.4.1 Relevante Endpunkte

In die vorliegende Bewertung gingen folgende patientenrelevante Endpunkte ein (zur Begründung siehe Abschnitt 2.7.2.4.3):

- Mortalität
 - Todesfälle
- Morbidität
 - schubbezogene Endpunkte
 - Patienten mit bestätigtem Schub
 - Zeit bis zum bestätigten Schub
 - jährliche Schubrate
 - Behinderungsprogression
 - Patienten mit (mindestens 12 Wochen anhaltender) Behinderungsprogression
 - Zeit bis zur (mindestens 12 Wochen anhaltenden) Behinderungsprogression
- gesundheitsbezogene Lebensqualität
 - Fatigue Impact Scale (FIS) zur Erfassung der symptomspezifischen gesundheitsbezogenen Lebensqualität (Fatigue)
- Nebenwirkungen
 - Gesamtrate SUE
 - Abbrüche wegen UE
 - Reaktionen an der Injektionsstelle
 - grippeähnliche Symptome
 - Alopezie
 - Diarrhö
 - Übelkeit und Erbrechen

Die Auswahl der patientenrelevanten Endpunkte weicht von der Auswahl des pU ab, der in Modul 4 des Dossiers weitere Endpunkte herangezogen hat. Dazu gehörten die folgenden Endpunkte: schubfreie Patienten, mittlere Änderung des EDSS-Werts, Zeit bis zum Therapieversagen, Patientenzufriedenheit, Gesamtrate unerwünschter Ereignisse, Leberfunktionsstörung, Überempfindlichkeit / Hautreaktionen, Neutropenie, Infektionen und parasitäre Erkrankungen, Bluthochdruck, periphere Neuropathie (inkl. Parästhesie) und Kopfschmerzen (siehe Abschnitt 2.7.2.4.3).

2.4.2 Datenverfügbarkeit und Verzerrungspotenzial

Tabelle 8 zeigt, für welche Endpunkte in der eingeschlossenen Studie Daten zur Verfügung standen. Tabelle 9 beschreibt das Verzerrungspotenzial für diese Endpunkte.

Tabelle 8: Matrix der Endpunkte – RCT, direkter Vergleich: Teriflunomid vs. IFN β -1a

Studie	Endpunkte											
	Gesamtmortalität	Schubbezogene Endpunkte	Behinderungsprogression	Gesundheitsbezogene Lebensqualität (generisch / krankheitsspezifisch)	Gesundheitsbezogene Lebensqualität (symptomspezifisch [Fatigue])	Schwerwiegende unerwünschte Ereignisse	Abbruch wegen unerwünschter Ereignisse	Reaktionen an der Injektionsstelle	Grippeähnliche Symptome	Alopezie	Diarrhö	Übelkeit und Erbrechen
TENERE	ja	ja	ja	nein	ja	ja	ja	ja	ja	ja	ja	ja
IFN β -1a: Interferon beta-1a; RCT: randomisierte kontrollierte Studie												

Tabelle 9: Verzerrungspotenzial auf Studien- und Endpunktebene – RCT, direkter Vergleich: Teriflunomid vs. IFN β-1a

Studie	Studienebene	Endpunkte											
		Gesamtmortalität	Schubbezogene Endpunkte	Behinderungsprogression	Gesundheitsbezogene Lebensqualität (generisch / krankheitsspezifisch)	Gesundheitsbezogene Lebensqualität (symptomspezifisch [Fatigue])	Schwerwiegende unerwünschte Ereignisse	Abbruch wegen unerwünschter Ereignisse	Reaktionen an der Injektionsstelle	Grippeähnliche Symptome	Alopezie	Diarrhö	Übelkeit und Erbrechen
TENERE	n	n	H	H	- ^a	H	H	H	H	H	H	H	H

a: Zielgröße wurde nicht erhoben
h: hoch; IFN β-1a: Interferon beta-1a; n: niedrig; RCT: randomisierte kontrollierte Studie

Das Verzerrungspotenzial der relevanten Endpunkte wird – mit Ausnahme der Gesamtmortalität – aufgrund der fehlenden Verblindung von Patient und behandelnder Person sowie der grundsätzlich vorhandenen subjektiven Komponente bei der Erfassung der Endpunkte als hoch eingestuft. Dies weicht von der Einschätzung des pU ab, der für schubbezogene Endpunkte, unerwünschte Ereignisse, schwerwiegende unerwünschte Ereignisse, Abbruch wegen unerwünschter Ereignisse, Reaktionen an der Injektionsstelle, grippeähnliche Symptome sowie Alopezie ein niedriges Verzerrungspotenzial angenommen hatte (siehe Abschnitt 2.7.2.4.2 zur umfassenden Kommentierung des Verzerrungspotenzials auf Endpunktebene).

Weitere Informationen zu Endpunktauswahl sowie Verzerrungspotenzial auf Endpunktebene befinden sich in Modul 4, Abschnitte 4.3.1.2.2, 4.3.1.3 sowie 4.3.2.1.3 des Dossiers und in den Abschnitten 2.7.2.4.2 und 2.7.2.4.3 der vorliegenden Dossierbewertung.

2.4.3 Ergebnisse

Tabelle 10, Tabelle 11 und Tabelle 12 fassen die Ergebnisse zum Vergleich von Teriflunomid mit IFN β -1a zusammen. Die Daten aus dem Dossier des pU wurden, wo notwendig, durch eigene Berechnungen ergänzt².

Aufgrund des hohen Verzerrungspotenzials fast aller Endpunkte der TENERE Studie ist bezüglich der Wahrscheinlichkeit des Zusatznutzens grundsätzlich nur die Ableitung von Anhaltspunkten möglich. Lediglich für Mortalität ist die Ableitung eines Hinweises möglich.

² Das Odds Ratio bietet bei geringen Ereigniszahlen eine gute Approximation des relativen Risikos. Daher wurde bei Ereignisraten von $\leq 1\%$ (in mindestens einer Zelle) das Peto-OR statt des relativen Risikos als Effektmaß berechnet und zur Bewertung verwendet.

Tabelle 10: Ergebnisse (Mortalität und Morbidität) – RCT, direkter Vergleich: Teriflunomid vs. IFN β -1a (Studie TENERE)

Kategorie Endpunkt	Teriflunomid		IFN β -1a		Teriflunomid vs. IFN β -1a	
	N	Patienten mit Ereignis n (%)	N	Patienten mit Ereignis n (%)	RR [95 %-KI]	p-Wert
Mortalität						
Todesfälle	110	0 (0)	101	0 (0)	-	-
Morbidität						
<i>Schubbezogene Endpunkte (EDSS-basiert)</i>						
Patienten mit bestätigtem Schub	111	26 (23,4)	104	16 (15,4)	1,52 [0,87; 2,67]	0,143
	N	KM-Schätzer ^a [95 %-KI]	N	KM-Schätzer ^a [95 %-KI]	HR [95 %-KI]	p-Wert
Zeit bis zum bestätigtem Schub	111	0,29 [0,19; 0,40]	104	0,19 [0,10; 0,27]	1,46 [0,78; 2,73]	0,229
	N	Jährliche Schubrate ^b [95 %-KI]	N	Jährliche Schubrate ^b [95 %-KI]	IDR [95 %-KI]	p-Wert
jährliche Schubrate ^c	111	0,26 [0,15; 0,44]	104	0,22 [0,11; 0,42]	1,20 [0,62; 2,30]	0,590 ^d
<i>Behinderungsprogression^e (EDSS-basiert)</i>						
	N	Patienten mit Ereignis n (%)	N	Patienten mit Ereignis n (%)	RR [95 %-KI]	p-Wert
Patienten mit Behinderungsprogression	111	10 (9,0)	104	9 (8,7)	1,04 [0,44; 2,46]	0,927
	N	KM-Schätzer ^a [95 %-KI]	N	KM-Schätzer ^a [95 %-KI]	HR [95 %-KI]	p-Wert
Zeit bis zur Behinderungsprogression	111	0,12 [0,05; 0,20]	104	0,10 [0,04; 0,16]	0,98 [0,40; 2,42]	0,969
<p>a: Wahrscheinlichkeit eines Ereignisses zur Woche 96 (Zeit-zum-Ereignis-Analyse [definiert als Zeit vom Tag der Randomisierung bis zum Tag des ersten Auftretens des Ereignisses])</p> <p>b: Adjustierte jährliche Schubrate erhalten aus einem Poisson-Modell mit der Gesamtzahl bestätigter Schübe zwischen Randomisierung und letzter Dosierung als abhängiger Variable, adjustiert nach Baseline EDSS-Wert und Region, log-transformierte Behandlungsdauer als Offset-Variable.</p> <p>c: Anzahl bestätigter Schübe (während der Behandlungsphase) dividiert durch Anzahl an Patientenjahren</p> <p>d: Chi-Quadrat-Test aus der Schätzung des Rate Ratios</p> <p>e: mindestens 12 Wochen anhaltende Behinderungsprogression</p> <p>EDSS: Expanded Disability Status Scale; HR: Hazard Ratio; IDR: Inzidenzdichteverhältnis (Rate Ratio); IFN β-1a: Interferon beta-1a; ITT: intention to treat; KI: Konfidenzintervall; KM: Kaplan-Meier; N: Anzahl ausgewerteter Patienten bezogen auf die ITT Population, außer bei Mortalität (Safety-Population); RCT: randomisierte kontrollierte Studie; RR: relatives Risiko</p>						

Mortalität

In den beiden relevanten Behandlungsarmen der Studie TENERE sind keine Patienten verstorben. Die Studie war aufgrund ihrer Größe und auch ihrer Dauer aber nicht darauf ausgelegt, Unterschiede hinsichtlich der Mortalität aufzudecken. Ein Zusatznutzen von Teriflunomid ist für diesen Endpunkt nicht belegt.

Morbidität

Weder hinsichtlich der schubbezogenen Endpunkte noch hinsichtlich der Endpunkte zur Behinderungsprogression zeigten sich statistisch signifikante Unterschiede zwischen den Behandlungsgruppen.

Bei den schubbezogenen Endpunkten zeigten die Effektschätzer eher einen ungünstigen Effekt von Teriflunomid. Bei den Endpunkten zur Behinderungsprogression lagen die Effektschätzer um den Nulleffekt. Aufgrund der breiten Konfidenzintervalle kann bei beiden Endpunkten mehr als eine Verdoppelung des Risikos zuungunsten von Teriflunomid nicht ausgeschlossen werden (Tabelle 10). Die Studie war aufgrund der kurzen Studiendauer nicht geeignet, um relevante Effekte hinsichtlich der Behinderungsprogression nachzuweisen. In Abbildung 1 und Abbildung 2 sind zur Illustration die Kaplan-Meier-Kurven der Zeit bis zum bestätigten Schub und der Zeit bis zur (mindestens 12 Wochen anhaltenden) Behinderungsprogression dargestellt. Wie für andere Langzeitereignisse sind insbesondere für die Behinderungsprogression Beobachtungsperioden von mindestens 2 Jahren für eine angemessene Interpretierbarkeit zu fordern [6].

Insgesamt kann hinsichtlich der Endpunkte zur Morbidität ein negativer Effekt von Teriflunomid gegenüber IFN β -1a nicht sicher ausgeschlossen werden. Ein Zusatznutzen von Teriflunomid ist für diese Morbiditätsendpunkte nicht belegt.

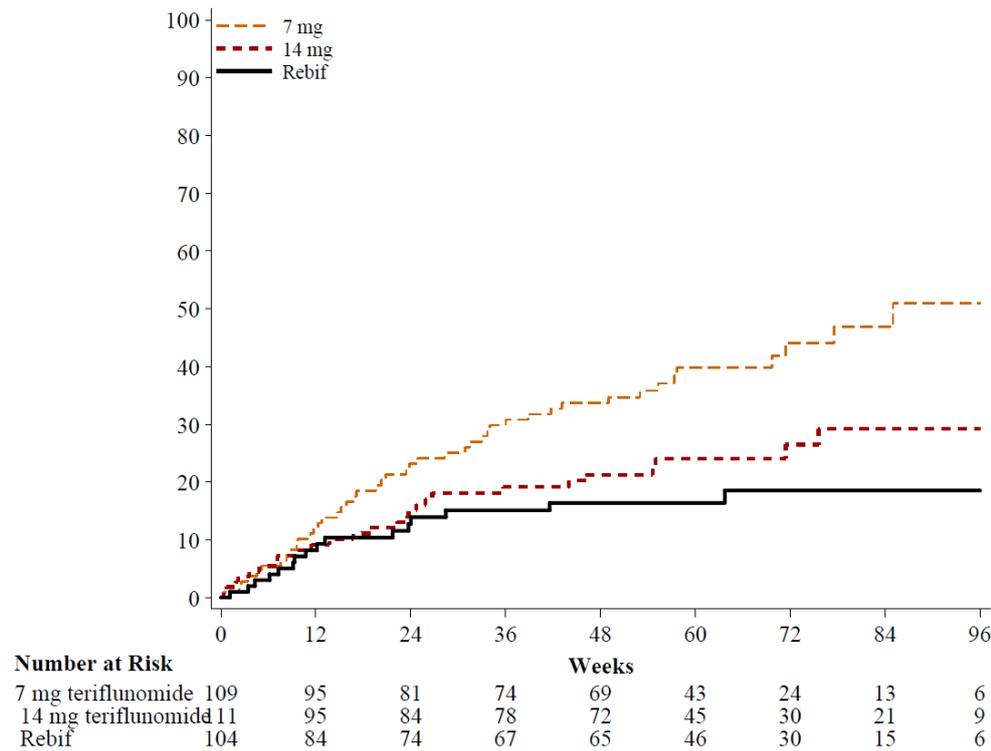


Abbildung 1: Kaplan-Meier-Kurven der Zeit bis zum bestätigten Schub (ITT-Population)

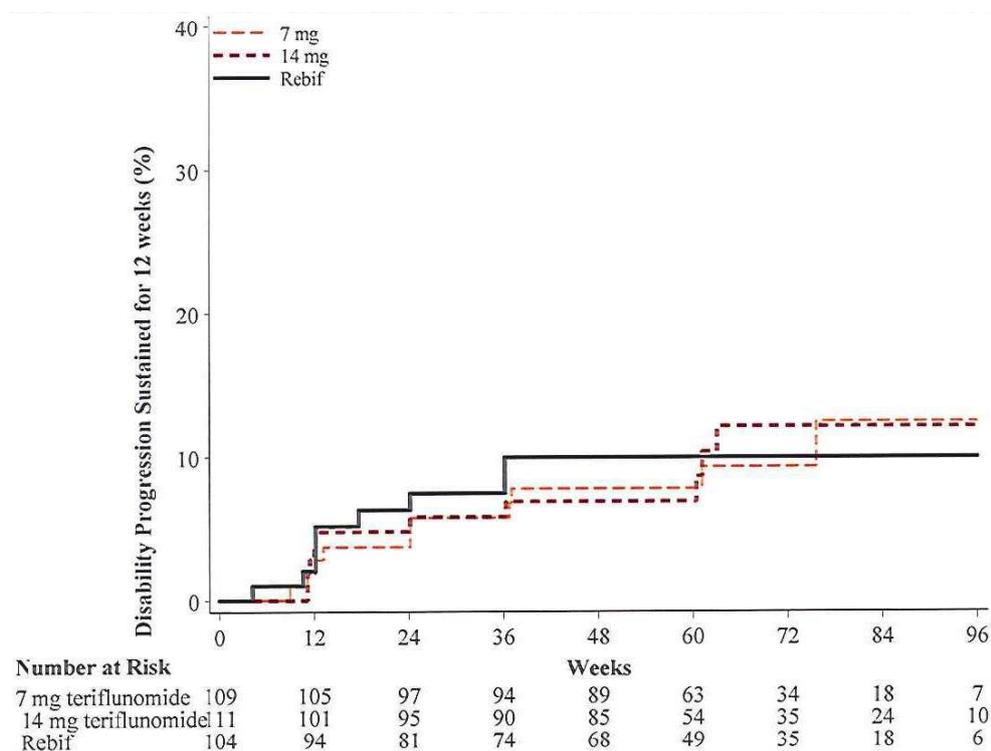


Abbildung 2: Kaplan-Meier-Kurven der Zeit bis zur (mindestens 12 Wochen anhaltenden) Behinderungsprogression (ITT-Population)

Gesundheitsbezogene Lebensqualität

Die Ergebnisse zur gesundheitsbezogenen Lebensqualität sind in Tabelle 11 dargestellt.

Tabelle 11: Ergebnisse (gesundheitsbezogene Lebensqualität) – RCT, direkter Vergleich: Teriflunomid vs. IFN β -1a (Studie TENERE)

Endpunktkategorie Endpunkt	Teriflunomid			IFN β -1a			Teriflunomid vs. IFN β -1a MD ^b [95 %-KI]; p-Wert
	N	Werte Studien- beginn MW (SD)	Änderung Studien- ende MW ^a (SD)	N	Werte Studien- beginn MW (SD)	Änderung Studien- ende MW ^a (SD)	
Gesundheitsbezogene Lebensqualität (generisch/krankheitsspezifisch)							
keine Daten verfügbar							
Gesundheitsbezogene Lebensqualität (symptomspezifisch [Fatigue])							
FIS ^c -Gesamtwert	106	42,5 (37,8)	4,10 (3,03)	97	34,2 (32,7)	9,10 (3,21)	-5,00 [-12,31; 2,31] p = 0,179
kognitive Dimension	106	10,2 (10,2)	0,87 (0,84)	97	7,8 (8,5)	2,34 (0,89)	-1,47 [-3,51; 0,58] p = 0,160
physische Dimension	106	12,6 (10,2)	1,19 (0,87)	97	11,1 (9,7)	1,51 (0,92)	-0,32 [-2,37; 1,73] p = 0,762
psychosoziale Dimension	106	19,7 (19,5)	2,70 (1,53)	97	15,4 (16,0)	5,52 (1,62)	-2,81 [-6,51; 0,89] p = 0,136
a: adjustierte Schätzer aus einem MMRM-Modell zum Erhebungszeitpunkt Woche 48 b: MMRM, adjustiert nach Baseline EDSS-Wert, Region, Visite, Behandlung x Visite-Wechselwirkung, Baseline-Wert und Behandlung x Visite-Wechselwirkung. c: hohe FIS-Werte bedeuten hohen negativen Einfluss von Fatigue auf die Lebensqualität (hohe FIS-Werte = niedrige symptomspezifische gesundheitsbezogene Lebensqualität) FIS: Fatigue Impact Scale; IFN β -1a: Interferon beta-1a; KI: Konfidenzintervall; MD: Mittelwertdifferenz; MMRM: Mixed-effects Model with Repeated Measures; MW: Mittelwert; N: Anzahl ausgewerteter Patienten bezogen auf Anzahl zu Studienbeginn (Anzahl real vorhandener Datensätze nach 48 Wochen im MMRM-Modell waren 83 [Teriflunomid] und 65 [IFN β -1a]); RCT: randomisierte kontrollierte Studie; SD: Standardabweichung							

In der Studie TENERE wurde die symptomspezifische gesundheitsbezogene Lebensqualität erhoben, die mit der Symptomatik der Erschöpfung (Fatigue) einhergeht. Sie wurde mittels Fatigue Impact Scale (FIS) gemessen. Beim FIS-Gesamtwert sowie den Werten der 3 Subskalen (kognitive, physische und psychosoziale Dimension) zeigte sich jeweils kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen. Bei der Interpretation der vorliegenden Ergebnisse ist zu beachten, dass ein großer Teil der Daten nach 48 Wochen nicht explizit beobachtet wurde (Teriflunomid: 23 [22 %], Rebif: 32 [33 %]). Es bleibt unklar, inwieweit die dem verwendeten MMRM-Modell zugrunde liegende Annahmen tatsächlich zuträfen. Darüber hinaus zeigte sich ein auffälliger Unterschied der Ausgangswerte zwischen den Behandlungsgruppen. Patienten mit Teriflunomid hatten dabei konsistent über die 3 Subskalen hinweg höhere Werte (im Mittel jeweils ca. 0,2 Standardabweichungen).

Andere Instrumente zur Erfassung der gesundheitsbezogenen Lebensqualität (z. B. generische Instrumente) wurden in der Studie TENERE nicht erhoben.

Zusammenfassend ist hinsichtlich der gesundheitsbezogenen Lebensqualität ein Zusatznutzen von Teriflunomid nicht belegt.

Nebenwirkungen

Die Ergebnisse zu Nebenwirkungen sind in Tabelle 12 dargestellt.

Tabelle 12: Ergebnisse (Nebenwirkungen) – RCT, direkter Vergleich: Teriflunomid vs. IFN β -1a (Studie TENERE)

Endpunktkategorie Endpunkt	Teriflunomid		IFN β -1a		Teriflunomid vs. IFN β -1a RR [95 %-KI]; p-Wert
	N	Patienten mit Ereignis n (%)	N	Patienten mit Ereignis n (%)	
Nebenwirkungen					
UE	110	102 (92,7)	101	97 (96,0)	-
SUE	110	4 (3,6)	101	6 (5,9)	0,61 [0,18; 2,11]; 0,525 ^a
Abbruch wegen UE ^b	110	7 (6,4)	101	9 (8,9)	0,71 [0,28; 1,85]; 0,529 ^a
Abbruch wegen UE ^c (mit Laborwerten)	110	12 (10,9)	101	20 (19,8)	0,55 [0,28; 1,07]; 0,081 ^a
Besondere Nebenwirkungen^d					
Reaktionen an der Injektionsstelle ^e	110	0	101	22 (21,8)	0,10 [0,04; 0,24] ^f ; < 0,001
grippeähnliche Symptome ^g	110	3 (2,7)	101	54 (53,5)	0,05 [0,02; 0,16]; < 0,001 ^a
Alopezie ^h	110	22 (20,0)	101	1 (1,0)	7,01 [2,95; 16,65] ^f ; < 0,001 ^a
Diarrhö ⁱ	110	23 (20,9)	101	8 (7,9)	2,64 [1,24; 5,63]; 0,008 ^a
Übelkeit und Erbrechen ^j	110	15 (13,6)	101	5 (5,0)	2,75 [1,04; 7,31]; 0,033 ^a
<p>a: eigene Berechnung b: Auswertung, in die keine Ereignisse eingingen, die aufgrund von Schwangerschaft (SOC Pregnancy, puerperium and perinatal conditions) oder aufgrund von Untersuchungen (SOC Investigations) zum Behandlungsabbruch führten. c: Auswertung, in die keine Ereignisse eingingen, die aufgrund von Schwangerschaft (SOC Pregnancy, puerperium and perinatal conditions) zum Behandlungsabbruch führten. d: die Mehrheit der Ereignisse waren von milder/moderater Intensität e: erfasst über MedDRA HLT Injection site reactions f: Peto OR g: erfasst über MedDRA PT Influenza like illness h: erfasst über MedDRA HLT Alopecia i: erfasst über MedDRA HLT Diarrhoea (excl. infective) j: erfasst über MedDRA HLT Nausea and vomiting symptoms HLT: High Level Term; IFN β-1a: Interferon beta-1a; KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; N: Anzahl ausgewerteter Patienten bezogen auf Safety Population; n: Anzahl Patienten mit Ereignis während der Behandlungsphase (during treatment period); OR: Odds Ratio; PT: Preferred Term; RCT: randomisierte kontrollierte Studie; RR: relatives Risiko; SOC: System Organ Class; SUE: schwerwiegendes unerwünschtes Ereignis; UE: unerwünschtes Ereignis</p>					

Die Auswertung der unerwünschten Ereignisse betrifft die Beobachtungsperiode während der Behandlung („during treatment period“). Dies weicht vom Vorgehen des pU ab, der in Modul 4 des Dossiers die Beobachtungsperiode während der Behandlung („during treatment

period“) plus Follow-up Periode (inkl. „Wash-out“ [beschleunigte Elimination] Phase) betrachtete (siehe Abschnitt 2.7.2.4.3).

Der Anteil von Patienten mit schwerwiegenden unerwünschten Ereignissen unterschied sich zwischen den Behandlungsgruppen nicht maßgeblich und das Ergebnis war nicht statistisch signifikant. Für diesen Endpunkt ist ein geringerer / größerer Schaden von Teriflunomid nicht belegt.

Beim Endpunkt Abbruch wegen unerwünschter Ereignisse zeigte sich kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen, wobei Behandlungsabbrüche aufgrund von Schwangerschaft bzw. aufgrund von Untersuchungen (Laborwertveränderungen) nicht als relevante Ereignisse gewertet wurden (siehe Abschnitt 2.7.2.4.3). Dies unterscheidet sich maßgeblich von der in Modul 4 des Dossiers vorgelegten Auswertung, in der diese Ereignisse erfasst wurden. Im Dossier zeigte das Ergebnis einen statistisch signifikanten Unterschied zugunsten von Teriflunomid. Eine Sensitivitätsanalyse, in der auch Behandlungsabbrüche aufgrund von Untersuchungen (Laborwertveränderungen) berücksichtigt wurden (jedoch keine aufgrund von Schwangerschaft), zeigte ebenfalls keinen statistisch signifikanten Unterschied zwischen den Behandlungsgruppen. Dies stützt das vorliegende Ergebnis. Für den Endpunkt Abbruch wegen UE ist ein geringerer / größerer Schaden von Teriflunomid nicht belegt.

Reaktionen an der Injektionsstelle und grippeähnliche Symptome traten bei der Behandlung mit IFN β -1a häufiger auf als mit Teriflunomid. Das Ergebnis war jeweils statistisch signifikant. Die Mehrheit dieser Ereignisse war von milder bzw. moderater Intensität. Für die beiden Endpunkte Reaktionen an der Injektionsstelle und grippeähnliche Symptome ergibt sich jeweils ein Anhaltspunkt für einen geringeren Schaden von Teriflunomid.

Alopezie und Diarrhö traten bei der Behandlung mit Teriflunomid häufiger auf als mit IFN β -1a. Das Ergebnis war jeweils statistisch signifikant. Die Mehrheit dieser Ereignisse war von milder bzw. moderater Intensität. Für die beiden Endpunkte Alopezie und Diarrhö ergibt sich jeweils ein Anhaltspunkt für einen größeren Schaden von Teriflunomid.

Übelkeit und Erbrechen traten bei der Behandlung mit Teriflunomid häufiger auf als mit IFN β -1a. Das Ergebnis war statistisch signifikant, bei allerdings nur geringfügiger Effektstärke (siehe Abschnitt 2.5.1 Tabelle 13). Für diesen Endpunkt ist, im Kontext der frühen Nutzenbewertung, ein größerer Schaden von Teriflunomid nicht belegt.

2.4.4 Subgruppenanalysen

Subgruppenauswertungen lagen im Dossier des pU nur für den relevanten Endpunkt jährliche Schubrate vor. Mögliche heterogene Behandlungseffekte wurden für die Merkmale Geschlecht, Alter, Region, EDSS-Wert zu Studienbeginn, Anzahl der Schübe in den letzten 2 Jahren und Vorbehandlung mit krankheitsmodifizierendem MS-Wirkstoffen untersucht. Die

Ergebnisse der Interaktionstests lieferten dabei keine Hinweise oder Belege für Subgruppeneffekte bzw. Effektmodifikationen für diesen Endpunkt.

Aufgrund der fehlenden Subgruppenauswertungen zu anderen relevanten Endpunkten (insbesondere zu unerwünschten Ereignissen) können potenziell heterogene Behandlungseffekte nicht umfassend beurteilt werden.

Weitere Informationen zu Endpunktauswahl, Verzerrungspotenzial auf Endpunktebene sowie zu Endpunktergebnissen befinden sich in Modul 4, Abschnitte 4.3.1.2.2, 4.3.1.3 sowie 4.3.2.1.3 des Dossiers und in den Abschnitten 2.7.2.4.2 und 2.7.2.4.3 der vorliegenden Dossierbewertung.

2.5 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 im allgemeinen Methodenpapier des IQWiG [8] erläutert.

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.

2.5.1 Beurteilung des Zusatznutzens auf Endpunktebene

Die in Abschnitt 2.4 präsentierte Datenlage ergab 2 Anhaltspunkte für einen geringeren Schaden sowie 2 Anhaltspunkte für einen größeren Schaden von Teriflunomid im Vergleich zu IFN β -1a.

Ausgehend von diesen Ergebnissen wurde das Ausmaß des jeweiligen Zusatznutzens (bzw. geringeren / größeren Schadens) auf Endpunktniveau eingeschätzt (siehe Tabelle 13).

Tabelle 13: Ausmaß des Zusatznutzens auf Endpunktebene: Teriflunomid vs. IFN β -1a

Endpunktkategorie Endpunkt	Teriflunomid vs. IFN β-1a Ereignisanteil / KM-Schätzer / Schubrate / MW Effektschätzer [95 %-KI] p-Wert Wahrscheinlichkeit^a	Ableitung des Ausmaßes^b
Mortalität		
Todesfälle	0 % vs. 0 %	geringerer Nutzen / Zusatznutzen nicht belegt
Morbidität		
Patienten mit bestätigtem Schub	23,4 % vs. 15,4 % RR 1,52 [0,87; 2,67] p = 0,143	Zusatznutzen nicht belegt
Zeit bis zum bestätigtem Schub	0,29 ^c vs. 0,19 ^c HR 1,46 [0,78; 2,73] p = 0,229	Zusatznutzen nicht belegt
jährliche Schubrate	0,26 ^d vs. 0,22 ^d IDR 1,20 [0,62; 2,30] p = 0,590	Zusatznutzen nicht belegt
Patienten mit Behinderungs- progression	9,0 % vs. 8,7 % RR 1,04 [0,44; 2,46] p = 0,927	Zusatznutzen nicht belegt
Zeit bis zur Behinderungs- progression	0,12 ^c vs. 0,10 ^c HR 0,98 [0,40; 2,42] p = 0,969	Zusatznutzen nicht belegt
Gesundheitsbezogene Lebensqualität		
Symptomspezifisch (Fatigue Impact Scale) Gesamtwert (FIS)	4,10 ^e vs. 9,10 ^e MD -5,00 [-12,31; 2,31] p = 0,179	Zusatznutzen nicht belegt
kognitive Dimension (FIS)	0,87 ^e vs. 2,34 ^e MD -1,47 [-3,51; 0,58] p = 0,160	Zusatznutzen nicht belegt
physische Dimension (FIS)	1,19 ^e vs. 1,51 ^e MD -0,32 [-2,37; 1,73] p = 0,762	Zusatznutzen nicht belegt
psychosoziale Dimension (FIS)	2,70 ^e vs. 5,52 ^e MD -2,81 [-6,51; 0,89] p = 0,136	Zusatznutzen nicht belegt

(Fortsetzung)

Tabelle 13: Ausmaß des Zusatznutzens auf Endpunktebene: Teriflunomid vs. IFN β -1a
(Fortsetzung)

Endpunktkategorie Endpunkt	Teriflunomid vs. IFN β-1a Ereignisanteil / KM-Schätzer / Schubrate / MW Effektschätzer [95 %-KI] p-Wert Wahrscheinlichkeit^a	Ableitung des Ausmaßes^b
Nebenwirkungen		
Gesamtrate SUE	3,6 % vs. 5,9 % RR 0,61 [0,18; 2,11] p = 0,525	geringerer / größerer Schaden nicht belegt
Abbruch wegen UE	6,4 % vs. 8,9 % RR 0,71 [0,28; 1,85] p = 0,529 ^f	geringerer / größerer Schaden nicht belegt
Reaktionen an der Injektionsstelle	0 % vs. 21,8 % Peto OR 0,10 [0,04; 0,24] p < 0,001 Wahrscheinlichkeit: Anhaltspunkt	Endpunktkategorie: nicht schwerwiegende / schwere Nebenwirkungen KI _o < 0,8 geringerer Schaden, Ausmaß: beträchtlich
grippeähnliche Symptome	2,7 % vs. 53,5 % RR 0,05 [0,02; 0,16] p < 0,001 Wahrscheinlichkeit: Anhaltspunkt	Endpunktkategorie: nicht schwerwiegende / schwere Nebenwirkungen KI _o < 0,8 geringerer Schaden, Ausmaß: beträchtlich
Alopezie	20,0 % vs. 1,0 % Peto OR 7,01 [2,95; 16,65] Peto OR ^g 0,14 [0,06; 0,34] p < 0,001 Wahrscheinlichkeit: Anhaltspunkt	Endpunktkategorie: nicht schwerwiegende / schwere Nebenwirkungen KI _o < 0,8 größerer Schaden, Ausmaß: beträchtlich
Diarrhö	20,9 % vs. 7,9 % RR 2,64 [1,24; 5,63] RR ^g 0,38 [0,18; 0,81] p = 0,008 Wahrscheinlichkeit: Anhaltspunkt	Endpunktkategorie: nicht schwerwiegende / schwere Nebenwirkungen 0,8 ≤ KI _o < 0,9 größerer Schaden, Ausmaß: gering
Übelkeit und Erbrechen	13,6 % vs. 5,0 % RR 2,75 [1,04; 7,31] RR ^g 0,36 [0,14; 0,96] p = 0,033 Wahrscheinlichkeit: Anhaltspunkt	Endpunktkategorie: nicht schwerwiegende / schwere Nebenwirkungen KI _o ≥ 0,90 geringerer / größerer Schaden nicht belegt

(Fortsetzung)

Tabelle 13: Ausmaß des Zusatznutzens auf Endpunktebene: Teriflunomid vs. IFN β -1a
(Fortsetzung)

<p>a: Angabe der Wahrscheinlichkeit, sofern statistisch signifikante Unterschiede vorliegen</p> <p>b: Einschätzungen zur Effektgröße erfolgen je nach Endpunktkategorie mit unterschiedlichen Grenzen anhand der oberen Grenze des Konfidenzintervalls (KI_o)</p> <p>c: KM-Schätzer der Wahrscheinlichkeit eines Ereignisses zur Woche 96</p> <p>d: Anzahl bestätigter Schübe während der Behandlungsphase dividiert durch Anzahl an Patientenjahren</p> <p>e: mittlere Änderung nach Woche 48 im Vergleich zu Studienbeginn (hohe FIS-Werte = niedrige symptomsspezifische gesundheitsbezogene Lebensqualität)</p> <p>f: Abbrüche wegen UE, die keine Ereignisse aufgrund von Schwangerschaft oder aufgrund von Untersuchungen (Laborwerten) enthalten. Der Effekt bleibt weiterhin nicht statistisch signifikant, auch wenn Laborwerte in die Analyse eingehen (RR 0,55 [0,28; 1,07]; p = 0,081).</p> <p>g: Ereignisanteil IFN β-1a vs. Teriflunomid (andere Effektrichtung zur Ableitung des Ausmaßes des Zusatznutzens)</p> <p>FIS: Fatigue Impact Scale; HR: Hazard Ratio; IDR: Inzidenzdichteverhältnis (Rate Ratio); IFN β-1a: Interferon beta-1a; KI: Konfidenzintervall, KI_o: obere Grenze Konfidenzintervall; KM: Kaplan-Meier; MD: Mittelwertdifferenz; MW: Mittelwert; OR: Odds Ratio; RR: relatives Risiko; SUE: schwerwiegendes unerwünschtes Ereignis; UE: unerwünschtes Ereignis; vs.: versus</p>
--

2.5.2 Gesamtaussage zum Zusatznutzen

Tabelle 14 fasst die Resultate zusammen, die in die Gesamtaussage zum Ausmaß des Zusatznutzens einfließen.

Tabelle 14: Positive und negative Effekte aus der Bewertung von Teriflunomid im Vergleich zu IFN β -1a

Positive Effekte	Negative Effekte
Anhaltspunkt für einen geringeren Schaden – Ausmaß: beträchtlich (nicht schwerwiegende / schwere Nebenwirkungen: Reaktionen an der Injektionsstelle)	Anhaltspunkt für einen größeren Schaden – Ausmaß: beträchtlich (nicht schwerwiegende / schwere Nebenwirkungen: Alopezie)
Anhaltspunkt für einen geringeren Schaden – Ausmaß: beträchtlich (nicht schwerwiegende / schwere Nebenwirkungen: grippeähnliche Symptome)	Anhaltspunkt für einen größeren Schaden – Ausmaß: gering (nicht schwerwiegende / schwere Nebenwirkungen: Diarrhö)
Bei den Morbiditätsendpunkten Schübe / Behinderungsprogression kann ein negativer Effekt von Teriflunomid gegenüber IFN β -1a nicht sicher ausgeschlossen werden.	

Auf der Grundlage von besonderen Nebenwirkungen ergeben sich positive und negative Effekte mit gleicher Ergebnissicherheit (Anhaltspunkt) und Endpunktkategorie (nicht schwerwiegende / schwere Nebenwirkungen). Auf beiden (positiven und negativen) Seiten wird das Ausmaß „beträchtlich“ erreicht. Dabei liegen die positiven Effekte (Reaktionen an der Injektionsstelle, grippeähnliche Symptome) jeweils mit Ausmaß „beträchtlich“ vor, während auf der negativen Seite ein Effekt das Ausmaß „beträchtlich“ (Alopezie) und ein Effekt das Ausmaß „gering“ (Diarrhö) aufweist. Bezüglich der Morbiditätsendpunkte Schübe bzw. Behinderungsprogression kann ein negativer Effekt von Teriflunomid gegenüber IFN β -1a nicht sicher ausgeschlossen werden.

In der Abwägung der positiven und negativen Effekte ergibt sich kein Beleg für einen Zusatznutzen von Teriflunomid gegenüber der zweckmäßigen Vergleichstherapie IFN β -1a.

Diese Bewertung weicht maßgeblich von der des pU ab, der einen Beleg für einen beträchtlichen Zusatznutzen beanspruchte.

Tabelle 15 fasst die Ergebnisse der Bewertung zu Ausmaß und Wahrscheinlichkeit des Zusatznutzens von Teriflunomid gegenüber der zweckmäßigen Vergleichstherapie im zugelassenen Anwendungsgebiet zusammen.

Tabelle 15: Teriflunomid – Ausmaß und Wahrscheinlichkeit des Zusatznutzens

Anwendungsgebiet	Zweckmäßige Vergleichstherapie ^a	Ausmaß und Wahrscheinlichkeit des Zusatznutzens
Erwachsene Patienten mit schubförmig-remittierender Multipler Sklerose	Beta-Interferon (1a oder 1b) oder Glatirameracetat	Zusatznutzen nicht belegt
<p>a: Dargestellt ist jeweils die vom G-BA festgelegte zweckmäßige Vergleichstherapie. In den Fällen, in denen der pU aufgrund der Festlegung der zweckmäßigen Vergleichstherapie durch den G-BA aus mehreren Alternativen eine Vergleichstherapie auswählen kann, ist die entsprechende Auswahl des pU fett markiert. G-BA: Gemeinsamer Bundesausschuss</p>		

Über den Zusatznutzen beschließt der G-BA.

Weitere Informationen zum Ausmaß und zur Wahrscheinlichkeit des Zusatznutzens befinden sich in Modul 4, Abschnitt 4.4 des Dossiers und in Abschnitt 2.7.2.8 der vorliegenden Dossierbewertung.

2.6 Liste der eingeschlossenen Studien

TENERE

Sanofi Aventis. A multi-center, randomized, parallel-group, rater-blinded study comparing the effectiveness and safety of teriflunomide and interferon beta-1a in patients with relapsing multiple sclerosis: study EFC10891; clinical study report [unveröffentlicht]. 2011.

Sanofi. A study comparing the effectiveness and safety of teriflunomide and interferon beta-1a in patients with relapsing multiple sclerosis (TENERE): full text view [online]. In: Clinicaltrials.gov. 17.07.2013 [Zugriff: 10.12.2013]. URL: <http://www.clinicaltrials.gov/ct2/show/NCT00883337>.

Vermersch P, Czlonkowska A, Grimaldi LM, Confavreux C, Comi G, Kappos L et al. Teriflunomide versus subcutaneous interferon beta-1a in patients with relapsing multiple sclerosis: a randomised, controlled phase 3 trial. *Mult Scler* 21.11.2013 [Epub ahead of print].

2.7 Kommentare zum Dossier des pharmazeutischen Unternehmers

2.7.1 Kommentar zur zweckmäßigen Vergleichstherapie (Modul 3, Abschnitt 3.1)

Die Angaben des pU zur zweckmäßigen Vergleichstherapie befinden sich in Modul 3 (Abschnitt 3.1) des Dossiers.

Der G-BA hat für das Anwendungsgebiet der schubförmigen Multiplen Sklerose (relapsing multiple Sclerosis [RMS]) folgende zweckmäßige Vergleichstherapie auf Stoffebene festgelegt:

- Beta-Interferone (1a oder 1b) oder Glatirameracetat unter Beachtung des jeweils zugelassenen Anwendungsgebietes (Glatirameracetat ist nicht zugelassen für Patienten mit primärer oder sekundär progredienter Multipler Sklerose).

Der pU wählt Interferon beta-1a (IFN β -1a) als eine der vom G-BA als zweckmäßige Vergleichstherapie festgelegten Alternativen aus, beschränkt seine Wahl jedoch auf IFN β -1a 44 μ g s. c. 3-mal pro Woche (Rebif [4]), eines der Präparate aus dieser Wirkstoffgruppe. Neben Rebif, das subkutan angewendet wird, gibt es in der Gruppe ein weiteres, intramuskulär verabreichtes Präparat (Avonex [9]), welches aufgrund der Beschränkung nicht in der Fragestellung des pU berücksichtigt wird. Die zum Vergleich mit Rebif ausgestaltete Recherche würde Studien zum direkten Vergleich von Teriflunomid mit IFN β -1a (Avonex) nicht identifizieren. Gemäß der Festlegung des G-BA sind jedoch alle Darreichungsformen und damit alle IFN β -1a Präparate zu berücksichtigen. Die Begründung des pU (Rebif sei wirksamer) widerspricht seinem eigenen Vorgehen, sich im Grundsatz der Festlegung des G-BA anzuschließen. Die Nutzenbewertung wird gegenüber der zweckmäßigen Vergleichstherapie IFN β -1a durchgeführt.

Insgesamt hat die Einschränkung der Vergleichstherapie des pU jedoch keine Konsequenz auf das vorliegende Bewertungsergebnis, da sich der Studienpool für direkt vergleichende Studien mit dem zu bewertenden Arzneimittel durch eine Beschränkung auf Rebif nicht ändert und der indirekte Vergleich auch aus anderen Gründen nicht geeignet ist, die Ergebnisse des direkten Vergleichs zu stützen (siehe Abschnitt 2.7.2.3.2).

2.7.2 Kommentar zur Darstellung von Nutzen und Zusatznutzen (Modul 4)

2.7.2.1 Fragestellung / Einschlusskriterien

Die Angaben des pU zur Fragestellung und zu Einschlusskriterien für die Studienauswahl befinden sich in Modul 4 (Abschnitte 4.2.1, 4.2.2) des Dossiers.

Die vom pU gewählte Fragestellung und die gewählten Einschlusskriterien sind nur teilweise nachvollziehbar. Die nachfolgenden Kommentare beziehen sich auf die Angaben des pU zum direkten und indirekten Vergleich, welche inhaltlich übereinstimmen.

Die Fragestellung im Dossier richtet sich auf die Patientenpopulation, für die gemäß Zulassung eine Behandlung mit Teriflunomid infrage kommt (erwachsene Patienten mit RRMS [3]). Der pU schränkt jedoch die Population auf Patienten mit einem EDSS-Wert $\leq 5,5$ ein, wodurch Studien zur zugelassenen Teilpopulation (Patienten mit EDSS $> 5,5$) ausgeschlossen werden. Der pU begründet diese Einschränkung mit der Einschätzung, dass die untersuchte Patientenpopulation hinreichend homogen sein sollte. Die Begründung des pU ist nicht nachvollziehbar, da heterogene Patientengruppen innerhalb der relevanten Population getrennt betrachtet werden können. Für die Nutzenbewertung wird keine Einschränkung der zugelassenen Patientenpopulation (z. B. nach Krankheitsschwere oder zur Gehfähigkeit, gemessen über den EDSS) vorgenommen.

Das Einschlusskriterium zur Vergleichstherapie wird innerhalb der IFN β -1a Wirkstoffgruppe auf IFN β -1a 44 μ g s. c. (Rebif [4]) beschränkt. Wie bereits unter Abschnitt 2.7.1 beschrieben wird diesem Vorgehen nicht gefolgt und die zweckmäßige Vergleichstherapie IFN β -1a für die Dossierbewertung herangezogen.

Der pU schließt Studien mit einer Studiendauer von mindestens 1 Jahr bzw. einer geplanten Behandlungsdauer von mindestens 48 Wochen ein, da er die Aussagekraft von Studien von kürzeren Dauern als nicht ausreichend hoch einstuft. Zur Begründung zieht er die Leitlinie für klinische Untersuchungen zur Multiplen Sklerose der europäischen Zulassungsbehörde heran [6]. Diese empfiehlt für den Wirksamkeitsnachweis bei RRMS eine Mindestbeobachtungsdauer von 2 Jahren, da gerade eine valide Erfassung der Behinderungsprogression in kürzeren Studien mit Unsicherheit verbunden sei. Für den Nachweis von Effekten auf die Schubaktivität würde auch ein Zeitraum von einem Jahr als prinzipiell ausreichend angesehen, auch wenn längere Beobachtungszeiträume zu bevorzugen seien [6,7]. Für eine Nutzenbewertung im vorliegenden Anwendungsgebiet wird ebenfalls eine Beobachtungsdauer von mindestens 1 bis 2 Jahren als sinnvoll betrachtet.

Ergänzend ist anzumerken, dass der pU neben aktiv kontrollierten Studien auch placebo-kontrollierte Studien in seinem Dossier darstellt. Diese werden für die Bewertung des Zusatznutzens gemäß der vorliegenden Fragestellung und dem hier beschriebenen Kriterium zur Vergleichstherapie ausschließlich im Kontext indirekter Vergleiche betrachtet. Der indirekte Vergleich wird in Abschnitt 2.7.2.3.2 kommentiert; in anderen Abschnitten werden Angaben, die ausschließlich placebokontrollierte Studien betreffen, nicht kommentiert.

2.7.2.2 Methodik zur Bewertung der Aussagekraft der Nachweise sowie zur Informationssynthese und -analyse

Die Angaben des pU zur Methodik für die Bewertung der Aussagekraft der Nachweise sowie für die Informationssynthese und -analyse befinden sich in Modul 4 (Abschnitte 4.2.4, 4.2.5) des Dokuments.

Verzerrungsaspekte

Die Angaben des pU zum Vorgehen bei der Bewertung von Verzerrungsaspekten befinden sich in Modul 4 (Abschnitt 4.2.4) des Dossiers.

Der pU schließt sich der vorgegebenen Methodik der Dossievorlage zur endpunktübergreifenden und endpunktspezifischen Bewertung der Verzerrungsaspekte der eingeschlossenen Studien an.

Studiendesign / Patientencharakteristika / Endpunkte

Die Angaben des pU zur Darstellung von Studiendesign und Patientencharakteristika sowie zur Auswahl der Endpunkte für die Bewertung befinden sich in Modul 4 (Abschnitte 4.2.5.1, 4.2.5.2) des Dossiers.

Studiendesign

Der pU gibt in diesem Abschnitt an, welchen Kriterien die Beschreibung des Designs und der Methodik der eingeschlossenen Studien folgt. Diesen Angaben wird gefolgt.

Patientencharakteristika

Der pU gibt an, welche Patientencharakteristika im Dossier dargestellt werden und nennt Alter, EDSS-Wert zu Studienbeginn, Verlaufsform der MS, Geschlecht, Anzahl der Schübe im vorangegangenen Jahr und Zeit seit der ersten MS-Diagnose.

In der Dossierbewertung wird zudem der Anteil der Patienten, die mit krankheitsmodifizierenden Therapien vorbehandelt waren (Vorbehandlung mit anderen MS-Präparaten), dargestellt. Dieses Merkmal wird auch vom pU als relevantes Merkmal für Subgruppenuntersuchungen identifiziert.

Endpunkte

Der pU benennt zur Bewertung des Zusatznutzens patientenrelevante Endpunkte der Kategorien Mortalität, Morbidität, gesundheitsbezogene Lebensqualität, unerwünschte Ereignisse und zusätzlich Patientenzufriedenheit. Es werden nachfolgend nur Endpunkte kommentiert, die im Zusammenhang mit den für den direkten Vergleich relevanten Studien stehen.

Im Dossier wurden zur Morbidität die Endpunkte Behinderungsprogression und Schübe über EDSS-Messungen dargestellt. Dabei wurden Schübe anhand eines a priori definierten Responsekriteriums ausgewertet, die Kriterien zur Auswertung der Behinderungsprogression wurden laut pU jedoch post hoc und unverblindet angewendet (siehe Abschnitt 2.7.2.4.2 zum Verzerrungspotenzial). Die EDSS ist als Skala im vorliegenden Anwendungsgebiet etabliert, hat aber Schwächen [6,10]. Eine bekannte Schwäche der EDSS ist die begrenzte Intra- und Interobserver-Reliabilität. Aus diesem Grund empfiehlt die EMA, dass alle möglichen Maßnahmen getroffen werden sollten, um dieses Problem zu minimieren, wie z. B. eine spezifische Schulung der Endpunkterheber oder die Durchführungen aller Messungen durch

dieselbe Person am Patienten [6]. Diese Maßnahmen sind gemäß den Ausführungen im Studienprotokoll für die Studie TENERE nicht gegeben. Es wurde über die Einschätzung ausreichender Kenntnisse hinaus keine Schulung der behandelnden Personen bzw. Endpunkterheber durchgeführt. Weiterhin ist bei den verwendeten EDSS-Responsekriterien nicht bei allen Änderungen sicher, ob es sich um spürbare Änderungen handelt. Bei Änderungen von einem EDSS-Ausgangswert 0 auf 1 handelt es sich in fast allen funktionellen Systemen (FS) um eine der Änderung zugrunde liegende Verschlechterung klinischer Zeichen, nicht jedoch einer spürbaren Symptomatik [11,12]. Es ist unklar, inwiefern durch die vorliegende Schubdefinition im konkreten Patientenpool der TENERE Studie klinische und patientenrelevante Änderungen zur Erfassung der Schübe beigetragen haben. Die EDSS-basierten Endpunkte Behinderungsprogression und Schübe sind jedoch für die Nutzenbewertung relevant. Abschnitt 2.7.2.4.3 beschreibt den Ein- bzw. Ausschluss der im Dossier erhältlichen Operationalisierungen zu den einzelnen EDSS-basierten Endpunkten.

Der Endpunkt Zeit bis zum Therapieversagen ist ein kombinierter Endpunkt, der in der TENERE-Studie als primärer Endpunkt definiert wurde als die Zeit bis zum Auftreten eines Schubes oder bis zum Abbruch der Studienmedikation aus jedwedem Grund. Aus Sicht des pU bildet dieser Endpunkt neben der klinischen Wirksamkeit die Bedeutung einer dauerhaften Therapie bei der Behandlung einer chronischen Erkrankung ab und sei damit von Bedeutung für den klinischen Alltag. Der Endpunkt Zeit bis zum Therapieversagen wird in der Nutzenbewertung nicht berücksichtigt, da die 2 Komponenten unterschiedliche Konstrukte mit unterschiedlicher klinischer Bedeutung darstellen und eine Zusammenfassung nicht sinnvoll ist. Die relevanten Konstrukte Schübe und Abbrüche der Behandlung (aufgrund unerwünschter Ereignisse) werden über separate Endpunkte erfasst.

Der Endpunkt Fatigue wurde in der TENERE Studie als belastendes Symptom mittels Fatigue Impact Scale (FIS) erhoben. Es handelt sich um ein validiertes, patientenbezogenes Messinstrument, das entwickelt wurde, um den Einfluss der wahrgenommenen Erschöpfung auf das Leben von Menschen mit MS zu untersuchen. Demnach handelt es sich um ein Instrument zur Erfassung der symptomspezifischen gesundheitsbezogenen Lebensqualität. Der Endpunkt wird in der Nutzenbewertung herangezogen.

Es finden sich Angaben zur Patientenzufriedenheit (erhoben mittels Treatment Satisfaction Questionnaire for Medication). Nach Einschätzung des pU ist die Zufriedenheit der Patienten mit der Behandlung für die Beurteilung des Zusatznutzens relevant, kann jedoch nicht allein als Basis für die Bewertung des Nutzens oder Zusatznutzens herangezogen werden. Die Patientenzufriedenheit ist kein patientenrelevanter Endpunkt gemäß AM-NutzenV [1] und wird deshalb nicht in der Nutzenbewertung berücksichtigt.

Statistische Methodik (Meta-Analysen, Sensitivitätsanalysen, Subgruppen bzw. Effektmodifikatoren, indirekte Vergleiche)

Die Angaben des pU zur statistischen Methodik (Meta-Analysen, Sensitivitätsanalysen, Subgruppen / Effektmodifikatoren, indirekte Vergleiche) befinden sich in Modul 4 (Abschnitte 4.2.5.3 bis 4.2.5.6) des Dossiers.

Meta-Analysen

Der pU beschreibt in Abschnitt 4.2.5.3 des Dossiers die statistische Methodik zur Durchführung von Meta-Analysen und zum Umgang mit Heterogenität. Dieser Methodik wird weitestgehend gefolgt. Für die Bewertung gleichgerichteter Effekte bei heterogener Studienlage verwendet der pU vom Methodenpapier des Instituts [8] abweichende Kriterien. Auf die Auswertungen zum Zusatznutzen hat dies jedoch keinen Einfluss, da dort nur eine Studie vorliegt und keine Meta-Analysen durchgeführt werden.

Sensitivitätsanalysen

Das Dossier enthält an dieser Stelle zum einen eine kurze Ergebnisbeschreibung der in den eingeschlossenen Studien durchgeführten Sensitivitätsanalysen (Vergleich der Ergebnisse von ITT-Analysen mit denen von PP-Analysen) mit Verweis auf die Studienberichte. Im Ergebnisteil des Dossiers fehlt eine numerische Darstellung der Ergebnisse dieser Sensitivitätsanalysen. Darüber hinaus führt der pU eine Sensitivitätsanalyse für den Endpunkt Schübe an, in der er das Gegenereignis „Anteil der Patienten mit mindestens einem Schub“ betrachtet. Diese Auswertung betrifft jedoch nur die Meta-Analyse der placebokontrollierten Studien, die für den stützenden indirekten Vergleich verwendet wurden. Dieser indirekte Vergleich wird jedoch nicht in der vorliegenden Dossierbewertung herangezogen (siehe Abschnitt 2.7.2.3.2). Somit wird diese Sensitivitätsanalyse hier nicht weiter kommentiert.

Subgruppen / Effektmodifikatoren

Das Dossier enthält in den Abschnitten 4.2.5.5 und 4.3.1.3.12.1 Angaben dazu, welche Methodik zur Evaluation von Effektmodifikatoren eingesetzt wurde. Dieser Methodik wird gefolgt. Der pU legt neben Subgruppenanalysen zu in den Studienunterlagen a priori festgelegten Merkmalen eine post hoc definierte Subgruppenanalyse zu dem potenziellen Effektmodifikator Vorbehandlung mit einem krankheitsmodifizierenden (MS) Wirkstoff vor. Insgesamt wurden folgende Merkmale betrachtet:

- Geschlecht (männlich/weiblich)
- Alter (< 38 / ≥ 38 Jahre)
- Region (Osteuropa / Westeuropa und Afrika / Amerika)
- EDSS zu Studienbeginn ($\leq 3,5$ / $> 3,5$)
- Anzahl der Schübe in letzten 2 Jahren (0 / 1 / 2 / 3 / ≥ 4)
- Vorbehandlung mit krankheitsmodifizierendem (MS) Wirkstoff

Die Wahl der Trennpunkte zur Kategorisierung quantitativer Merkmale (Alter, EDSS zu Studienbeginn) war prädefiniert.

Die Untersuchung auf Effektmodifikation beschränkt der pU bei der TENERE-Studie auf den primären Endpunkt (Zeit bis zum Therapieversagen) sowie die jährliche Schubrate. Die Untersuchung von möglichen Effektmodifikatoren im Rahmen systematischer Übersichten sollte jedoch alle patientenrelevanten Endpunkte einschließen, da sonst nicht überprüft werden kann, ob eventuelle Effektunterschiede über mehrere Endpunkte konsistent sind. Zudem ist bei relevanten Effektmodifikationen eine Abwägung des Zusatznutzens innerhalb der entsprechenden Subgruppen erforderlich.

Indirekter Vergleich

Der unterstützende indirekte Vergleich wird in Abschnitt 2.7.2.3.2 kommentiert. Dort wird auch begründet, warum der indirekte Vergleich den direkten Vergleich nicht stützen kann.

2.7.2.3 Methodik und Ergebnisse der Informationsbeschaffung

Die Kommentierung der Methodik und der Ergebnisse der Informationsbeschaffung wird nachfolgend in 2 Themenbereiche aufgeteilt:

- Informationsbeschaffung
- Studienpool

2.7.2.3.1 Informationsbeschaffung

Die Angaben des pU zur Informationsbeschaffung befinden sich in Modul 4 (Abschnitte 4.2.3, 4.3.1.1, 4.3.2.1.1, Anhang 4-A bis 4-C) des Dossiers.

Direkter Vergleich

Studienliste des pU

Bezüglich der Studienliste des pU gibt es keinen Hinweis auf Unvollständigkeit.

Bibliografische Literaturrecherche

Um die relevanten Studien zu Teriflunomid zu identifizieren, war eine bibliografische Literaturrecherche gefordert, da Zulassungen vorliegen, die mehr als 12 Monate vor Einreichung des Dossiers erteilt wurden. Diese führte der pU auch durch.

Die Recherche des pU ist geeignet, die Vollständigkeit des Suchergebnisses sicherzustellen.

Suche in Studienregistern

Die laut Dossierunterlagen geforderte Suche in Studienregistern wurde vom pU durchgeführt. Die Suche des pU ist geeignet, die Vollständigkeit des Suchergebnisses sicherzustellen.

Zusammenfassung

Die Recherchen des pU zum direkten Vergleich sind geeignet, die Vollständigkeit der Suche sicherzustellen. Die Fragestellung und damit auch die Studienselektion des pU war aufgrund der Beschränkung der Fragestellung auf die Vergleichstherapie IFN β -1a 44 μ g s. c. (Rebif) nicht geeignet, um (auf Basis der eigentlich geeigneten Suche) die Vollständigkeit des Studienpools sicherzustellen (siehe Abschnitt 2.7.2.1). Deshalb erfolgte eine Überprüfung der Vollständigkeit des angegebenen Studienpools mittels einer Suche in den Studienregistern ClinicalTrials.gov, ICTRP Search Portal und EU Clinical Trials Register. Weiterhin wurde mittels Kombination verschiedener Suchtechniken (einfache Boolesche Suche sowie „related citations“ Funktion in PubMed) die Vollständigkeit der bibliografischen Literaturrecherche überprüft.

Durch die Vollständigkeitsprüfung wurden keine zusätzlichen relevanten Studien zum direkten Vergleich von Teriflunomid mit der zweckmäßigen Vergleichstherapie identifiziert.

Indirekter Vergleich

Zur Identifizierung relevanter Studien für einen unterstützenden indirekten Vergleich hat der pU separate Recherchen durchgeführt. Die Recherchen zum indirekten Vergleich werden im nachfolgenden Abschnitt 2.7.2.3.2 kommentiert. Dort wird auch begründet, warum der indirekte Vergleich für die vorliegende Nutzenbewertung nicht herangezogen werden kann.

2.7.2.3.2 Studienpool

Die Angaben des pU zum Studienpool befinden sich in Modul 4 (Abschnitte 4.3.1.1, 4.3.2.1.1, Anhang 4-D, 4-E) des Dossiers.

Direkter Vergleich

Nach Überprüfung der Suchergebnisse mit den Einschlusskriterien für die Dossierbewertung liegt eine direkt vergleichende Studie von Teriflunomid mit der zweckmäßigen Vergleichstherapie vor (TENERE). Dieser Studienpool stimmt mit dem Studienpool für den Vergleich mit der zweckmäßigen Vergleichstherapie im Dossier überein.

Indirekter Vergleich

Der pU legt unterstützend zum direkten Vergleich einen indirekten Vergleich vor. Er begründet dies wie folgt: das Studienziel der TENERE zum Nachweis der besseren Wirksamkeit im primären Endpunkt wurde nicht erbracht. Ein Nachweis einer Nichtunterlegenheit war in der Studie nicht geplant, weder bezogen auf Therapieabbruch noch auf andere Endpunkte. Das geht u. a. aus den Abschnitten 4.3.1.3.13.2 und 4.5.1 des Dossiers hervor. Zwar sei ein Hinweis darauf erkennbar, dass die Effektivität in sekundären Endpunkten zwischen Teriflunomid und IFN β -1a 44 μ g (Rebif) vergleichbar sei. Die Studie sei jedoch für eine ausreichende Präzision in der Erhebung der sekundären Endpunkte zu klein angelegt worden.

Gemäß der AM-NutzenV können Studien in die Bewertung einbezogen werden, die sich für einen indirekten Vergleich eignen, wenn keine direkt vergleichenden Studien vorliegen oder diese keine Aussagen über den Zusatznutzen zulassen [1]. Der Begründung zur Vorlage der indirekten Vergleiche im Dossier zur Erhöhung der Präzision ist nachvollziehbar. Die dabei herangezogenen Studien müssen jedoch zur Beantwortung der Fragestellung grundsätzlich geeignet sein.

Im Ergebnis ist der vorgelegte indirekte Vergleich nach Prüfung der Informationsbeschaffung, der identifizierten Studien und der verwendeten statistischen Methodik für die vorliegende Nutzenbewertung nicht verwertbar. Dies ist wie folgt begründet.

Studienpool unvollständig

Für den unterstützenden indirekten Vergleich soll Teriflunomid mit IFN β -1a 44 μ g s. c. (Rebif) über den Brückenkomparator Placebo verglichen werden. Der pU beschränkt die Einschlusskriterien zur Identifikation der relevanten Studien zur zweckmäßigen Vergleichstherapie auf IFN β -1a 44 μ g s. c. (Rebif) (siehe Abschnitt 2.7.2.1). Aus Modul 3 des Dossiers geht hervor, dass randomisierte kontrollierte Studien zum direkten Vergleich von IFN β -1a 30 μ g i. m. (Avonex) mit dem Brückenkomparator Placebo vorliegen [13,14]. Der vorliegende Studienpool für den indirekten Vergleich ist deshalb unvollständig, da aufgrund der Beschränkung der Einschlusskriterien auf Rebif diese Studien zu Avonex nicht in den Studienpool des indirekten Vergleichs eingingen.

Studien inhaltlich unterschiedlich

Für den stützenden indirekten Vergleich geht der pU in 2 Schritten vor. Der pU legte 3 Studien vor: 2 Studien verglichen Teriflunomid mit Placebo (TEMPO, TOWER) und eine Studie IFN β -1a 44 μ g s. c. (Rebif) mit Placebo (PRISMS). Mit den Ergebnissen dieser 3 Studien führt der pU einen indirekten Vergleich nach Bucher über den Brückenkomparator Placebo durch (nachfolgend „indirekter Vergleich“). Den Effektschätzer des indirekten Vergleichs fasst er im zweiten Schritt mit dem Ergebnis der direkt vergleichenden TENERE Studie zu einer Gesamtaussage zusammen (nachfolgend „indirekter Vergleich + TENERE“).

Die 3 Studien unterscheiden sich insbesondere in den folgenden Aspekten:

- Die PRISMS Studie wurde mehr als 10 Jahre vor den Studien mit Teriflunomid durchgeführt. Die revidierten MS-Diagnose-Kriterien (PRISMS: Poser Kriterien; TEMPO: McDonald Kriterien; TOWER: revidierte McDonald Kriterien) sind sensitiver und machen Unterschiede in der Krankheitsschwere der Patienten wahrscheinlich. So hatten Patienten in der PRISMS Studie im Mittel etwa 3 Schübe in den 2 Jahren vor Studienbeginn, Patienten in den Studien TEMPO und TOWER nur ca. 2 Schübe.
- Die Dauer der Erkrankung (Zeit seit MS-Diagnose [Median Jahre]) unterschied sich zwischen den Studienpopulationen von PRISMS (Rebif: 6,4 / Placebo: 4,3) und den

Teriflunomid-Studien (TEMSO [Teriflunomid: 3,7 / Placebo: 3,5] / TOWER [Teriflunomid: 3,3 / Placebo: 3,2]).

- Der Anteil mit einer immunmodulatorischen Therapie vorbehandelten Patienten unterschied sich zwischen den Studienpopulationen von PRISMS (ca. 3 %) und TEMSO (ca. 30 %) / TOWER (ca. 30 %).
- Die Behandlungsdauer unterschied sich zwischen den Studien PRISMS / TEMSO (2 Jahre) und TOWER (48 bis 154 Wochen).

Bei der Zusammenfassung „indirekter Vergleich + TENERE“ unterscheidet sich die TENERE Studie wiederum wie folgt von den 3 Studien des indirekten Vergleichs:

- Dauer der Erkrankung: Patienten der TENERE hatten eine deutlich kürzere Erkrankungsdauer (im Median ca. 1 Jahr) verglichen mit allen 3 Studien aus dem indirekten Vergleich nach Bucher.
- Die Formulierungen der Rebif-Präparate unterschieden sich zwischen PRISMS und TENERE. In der TENERE Studie kam die neue Formulierung zur Anwendung (ohne humanes Serumalbumin oder fötales Kälberserum), wodurch laut pU u. a. die Verträglichkeit der Injektion verbessert wurde.
- Auch in der TENERE Studie wurden revidierte McDonald Kriterien zur MS-Diagnosestellung angewandt, was sonst nur bei der TOWER Studie der Fall war.
- Der Anteil der vorbehandelten Patienten lag in der TENERE Studie bei 18 %, also deutlich über dem in der PRISMS Studie und deutlich unter dem in den Studien TOWER und TEMSO.
- Die Behandlungsdauer der TENERE Studie lag im Median bei ca. 1,3 Jahren und damit deutlich unter der der beiden Studien PRISMS und TEMSO (jeweils 2 Jahre).

Statistische Methodik

Die im Dossier angewendete statistische Methodik (Meta-Analyse des Schätzers aus direkt vergleichender Studie und des Schätzers aus indirektem Vergleich nach Bucher [der placebo-kontrollierten Studien]) ist grundsätzlich geeignet [15]. Die vorliegende Heterogenität zwischen Studien zum Vergleich von Teriflunomid und Placebo (TEMSO, TOWER) im Rahmen der indirekten Vergleiche wurde jedoch nicht adäquat berücksichtigt.

Zusammenfassung

Zusammenfassend ist der indirekte Vergleich nicht geeignet, Aussagen zum Zusatznutzen von Teriflunomid aus dem direkten Vergleich zu stützen. Zum einen ist der Studienpool des indirekten Vergleichs aufgrund der vom pU gewählten Einschränkung auf das Wirkstoffpräparat IFN β -1a 44 μ g s. c. (Rebif) unvollständig. Zum anderen bestehen bei mehreren Punkten Zweifel an der inhaltlichen Ähnlichkeit der Studien. Darüber hinaus wurde die vorliegende Heterogenität zwischen den Studien TEMSO und TOWER im Rahmen der

indirekten Vergleiche nicht adäquat berücksichtigt. Schließlich wurden vom pU auch im indirekten Vergleich beim Endpunkt Therapieabbrüche wegen UE Ereignisse herangezogen, die nicht als UE respektive als patientenrelevant zu werten sind (siehe Abschnitt 2.7.2.4.3).

Da der indirekte Vergleich nicht geeignet ist, Aussagen zum Zusatznutzen von Teriflunomid aus dem direkten Vergleich zu stützen, werden in den folgenden Abschnitten ausschließlich Angaben kommentiert, die für die Bewertung der Ergebnisse der TENERE-Studie (direkter Vergleich) relevant sind.

2.7.2.4 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

2.7.2.4.1 Studiendesign und Population

Die Angaben des pU zum Design und zu den Patientenpopulationen der eingeschlossenen Studien befinden sich in Modul 4 (Abschnitt 4.3.1.2.1) des Dossiers.

Die Darstellung der Charakteristika der Studienpopulation und des Studiendesigns ist im Wesentlichen nachvollziehbar. Es fehlen jedoch einige Konkretisierungen bezüglich der Charakterisierung der Interventionen in der Studie TENERE. Des Weiteren fehlen bei der Charakterisierung der Studienpopulation Angaben zur Vorbehandlung mit (krankheitsmodifizierenden) MS-Wirkstoffen (siehe Abschnitt 2.7.2.2 [Patientencharakteristika]). Die fehlenden Informationen werden ergänzt. Der pU stellt auch die Anzahl der Schübe im vergangenen Jahr dar. Da sich die vom pU vorgelegten Subgruppenauswertungen auf das Merkmal Anzahl der Schübe in den letzten 2 Jahren beziehen (siehe Abschnitt 2.7.2.2 [Subgruppen / Effektmodifikatoren]), wird dieses Merkmal vorzugsweise in der Nutzenbewertung dargestellt.

2.7.2.4.2 Verzerrungspotenzial

Die Angaben des pU zum Verzerrungspotenzial auf Studienebene befinden sich in Modul 4 (Abschnitt 4.3.1.2.2) und für die einzelnen Endpunkte in Modul 4 (Abschnitt 4.3.1.3.1, Anhang 4-G) des Dossiers. Detailinformationen zu der eingeschlossenen Studie, die in die Bewertung des Verzerrungspotenzials eingehen, befinden sich in Modul 4 (Anhang 4-F).

Der pU bewertet das Verzerrungspotenzial auf Studienebene für die Studie TENERE als niedrig. Dieser Gesamtbewertung wird gefolgt, bei einzelnen Aspekten ergeben sich jedoch Unklarheiten, denn zur Erzeugung der Randomisierungssequenz finden sich keine expliziten Angaben. Aufgrund der zentralen Zuteilung der Patienten zu den Behandlungsarmen mittels eines Interactive Voice Response Systems (IVRS) wird die Randomisierung insgesamt jedoch nicht infrage gestellt.

Der Bewertung des pU hinsichtlich des Verzerrungspotenzials auf Endpunktebene wird nicht in allen Punkten gefolgt. Das Verzerrungspotenzial der betrachteten Endpunkte wird – mit Ausnahme der Gesamtmortalität – aufgrund der fehlenden Verblindung von Patient und

behandelnder Person sowie der grundsätzlich vorhandenen subjektiven Komponente bei der Erfassung der Endpunkte als potenziell hoch verzerrt eingestuft. Dies weicht von der Einschätzung des pU ab, da er das Verzerrungspotenzial der schubbezogenen Endpunkte und eines Teils der Endpunkte zu unerwünschten Ereignissen als niedrig einschätzt. Der pU begründet seine Einschätzung damit, dass die getroffenen Maßnahmen zur Beibehaltung der Verblindung des zur Bestätigung des Schubes (mittels EDSS) einbezogenen Neurologen ausreichend gewesen wären. So seien Patienten angehalten worden, die Injektionsstelle zu verdecken. Dies ist laut EMA eine der empfohlenen Maßnahmen zur Reduktion der Verzerrung durch ein unverblindetes Studiendesign [6]. Aufgrund der fehlenden Verblindung der untersuchten Patienten bei der Durchführung der Tests zur EDSS-Erhebung kann trotz verblindetem Neurologen eine unverzerrte Endpunkterhebung nicht sichergestellt werden. Im Gegensatz zu schubbezogenen Endpunkten schätzt der pU die post hoc ausgewerteten Ergebnisse zur Behinderungsprogression selbst als potenziell hoch verzerrt ein. Er begründet seine Einschätzung damit, dass bei den zugrunde liegenden EDSS-Messungen dem nicht verblindeten behandelnden Neurologen kein bestätigender Neurologe nachgeschaltet worden sei und nicht davon ausgegangen werden kann, dass die Erhebung verblindet erfolgt ist.

Seine Einschätzung der potenziell niedrig verzerrten Endpunkte zu unerwünschten Ereignissen begründet der pU damit, dass diese Endpunkte entweder klar definiert und mit einer objektivierbaren Komponente eingestuft wurden oder es sich um die Erfassung von Laborparametern handele, die in einem zentralen Labor erfasst und nicht durch eine subjektive Einschätzung der Patienten beeinflusst werden können. Diese Begründung ist so nicht nachvollziehbar. Selbst bei schwerwiegenden unerwünschten Ereignissen könnte beispielsweise die Entscheidung zur Hospitalisierung aufgrund eines UEs unter Kenntnis der Behandlung beeinflusst worden sein kann. Somit ist eine Verzerrung nicht ausreichend sicher auszuschließen.

2.7.2.4.3 Ergebnisse

Die Angaben des pU zu Studienergebnissen befinden sich in Modul 4 (Abschnitt 4.3.1.3) des Dossiers.

Die nachfolgenden Einschätzungen beziehen sich auf die Validität und Relevanz der dargestellten Endpunkte, die sich anhand des Methodikteils zunächst nicht abschließend bewerten ließen. Im Folgenden werden alle im Dossier (Ergebnisteil) dargestellten Endpunkte gelistet.

Berücksichtigte Endpunkte

EDSS-basierte Endpunkte: Schübe und Behinderungsprogression

Diese eingeschlossenen Zielgrößen zu Schüben und zur Behinderungsprogression werden in jeweils einem Komplex betrachtet und in der Gesamtschau als Morbiditätsendpunkte interpretiert. Dabei ist zu beachten, dass sich die einzelnen Zielgrößen konzeptionell überschneiden und nicht unabhängig interpretiert werden sollten. Bei der Interpretation dieser

Ergebnisse ist zu berücksichtigen, dass unklar ist, inwiefern durch die vorliegenden Definitionen der EDSS-Responsekriterien in den Ergebnissen der TENERE Studie klinische und patientenrelevante Änderungen zur Erfassung der Schübe und der Behinderungsprogression beigetragen haben. Bei Änderungen von einem EDSS-Ausgangswert von 0 auf 1 handelt es sich in fast allen funktionellen Systemen (FS) um eine dieser Änderung zugrunde liegende Verschlechterung der klinischen Zeichen, nicht jedoch einer spürbaren Symptomatik [11,12] (siehe Abschnitt 2.7.2.2 [Endpunkte]).

▪ Schübe

Schübe wurden in der Studie wie folgt definiert: Auftreten eines neuen klinischen Zeichens / Symptoms oder die klinische Verschlechterung eines vorhandenen Symptoms (eines, das mindestens 30 Tage lang stabil war), das mindestens 24 h anhielt und nicht von Fieber begleitet war. Im Vergleich zu der vorherigen stabilen Symptomeinstufung wurde als Schub erfasst, wenn

- 1) entweder eine Erhöhung um 1 Punkt in mindestens 2 FS-Funktionen oder eine Erhöhung um 2 Punkte in mindestens einer FS-Funktion (ausgenommen Darm / Blase und Gehirn) gemessen wurde
- 2) oder eine Erhöhung um mindestens 0,5 Punkte im EDSS-Wert (es sei denn, der EDSS war vorher 0, dann war eine Erhöhung um 1,0 Punkte gefordert) gemessen wurde.

Die Angaben in den Studienunterlagen ist bei der Bezeichnung der Messung der FS-Funktionen nicht eindeutig, da das Studienprotokoll die Arbeit von Cutter 1999 [16] zitiert, welche aber ein anderes Instrument behandelt (Multiple Sclerosis Functional Composite [MSFC]). Für die vorliegende Bewertung wird aufgrund der in diesem Zusammenhang in Studien gängigen Definitionen jedoch von der ausschließlichen Verwendung des EDSS ausgegangen.

- Patienten mit bestätigtem Schub: eingeschlossen
- Zeit bis zum bestätigten Schub: eingeschlossen
- jährliche Schubrate: eingeschlossen
- schubfreie Patienten: nicht eingeschlossen. Der Endpunkt ist durch den Endpunkt Patienten mit bestätigtem Schub ausreichend abgebildet.

▪ Behinderungsprogression

Für sein Dossier wertet der pU die in der TENERE Studie erhobenen EDSS-Daten unter Verwendung der in den Studien TEMSO und TOWER definierten Kriterien zur Behinderungsprogression post hoc aus. Diese wurde in den Studien wie folgt definiert: Eine Behinderungsprogression trat ein, wenn mindestens eine 1-Punkt-Erhöhung des EDSS-Wertes im Vergleich zu Studienbeginn (in den Fällen mit Ausgangswert $\leq 5,5$) oder eine Erhöhung um mindestens 0,5 Punkte (in den Fällen mit Ausgangswert $> 5,5$) stattfand.

Diese Erhöhung des EDSS-Werts musste für mindestens 12 Wochen bestehen bleiben (12 Wochen anhaltende Behinderungsprogression).

- Patienten mit (mindestens 12 Wochen anhaltender) Behinderungsprogression: eingeschlossen
- Zeit bis zur (mindestens 12 Wochen anhaltenden) Behinderungsprogression: eingeschlossen
- mittlere Änderung des EDSS-Werts: nicht eingeschlossen.

Die EDSS wird allgemein als nicht linear und die einzelnen Schritte auf der EDSS als nicht äquidistant angesehen. Das spiegelt sich auch schon in der Tatsache wider, dass laut der Vorgabe der EMA und der Definition im Studienprotokoll der TENERE Studie bei geringeren Behinderungen 1 Punkt als Behinderungsprogression angesehen wird, bei stärkerer Behinderung bereits eine Änderung um 0,5 Punkte ausreicht [6,17]. Die Behinderungsprogression ist darüber hinaus auch schon über den Anteil der Patienten mit Behinderungsprogression und die Zeit bis zur Behinderungsprogression ausreichend abgebildet. Deshalb wird der Endpunkt Zunahme der Behinderungsprogression über die mittlere Änderung des EDSS-Werts nicht in der Nutzenbewertung berücksichtigt.

Andere Endpunkte

- Therapieversagen
 - Zeit bis zum Therapieversagen (Kombinierter Endpunkt aus Zeit bis zum bestätigten Schub oder Behandlungsabbruch aufgrund jedweden Grunds): nicht eingeschlossen (siehe Abschnitt 2.7.2.2 [Endpunkte]).
- gesundheitsbezogene Lebensqualität
 - Fatigue Impact Scale (FIS) zur Erfassung der symptom-spezifischen (Erschöpfung) gesundheitsbezogenen Lebensqualität: eingeschlossen
- Patientenzufriedenheit
 - Treatment Satisfaction Questionnaire for Medication (TSQM): nicht eingeschlossen. Kein patientenrelevanter Endpunkt gemäß AM-NutzenV (siehe auch Abschnitt 2.7.2.2 [Endpunkte]).
- Mortalität:
 - Anzahl der Todesfälle: eingeschlossen.
- unerwünschte Ereignisse

Folgende Endpunkte zu unerwünschten Ereignissen wurden im Dossier angeführt und in der vorliegenden Dossierbewertung eingeschlossen:

- Gesamtrate SUE, Abbrüche wegen UE (siehe Kommentar zur Operationalisierung unten)
- Besondere unerwünschte Ereignisse:
 - Reaktionen an der Injektionsstelle (MedDRA HLT Injection site reactions)
 - grippeähnliche Symptome (MedDRA PT Influenza like illness)
 - Alopezie (MedDRA HLT Alopecia)
 - Diarrhö (MedDRA HLT Diarrhoea [excl. infective])
 - Übelkeit und Erbrechen (MedDRA HLT Nausea and vomiting symptoms)

Folgende Endpunkte wurden im Dossier angeführt, werden für die Dossierbewertung jedoch nicht als relevant eingestuft und nicht eingeschlossen:

- Gesamtrate UE (wird nur ergänzend dargestellt)
- Leberfunktionsstörung (inkl. Erhöhung des Alanin-Aminotransferase-[ALT]-Werts)
- Überempfindlichkeit / Hautreaktionen
- Neutropenie
- Infektionen und parasitäre Erkrankungen
- Bluthochdruck
- periphere Neuropathie
- Kopfschmerzen

Es wurden keine zusätzlich relevanten Endpunkte identifiziert.

Operationalisierung Abbruch wegen UE

In die vom pU vorgelegte Auswertung von Behandlungsabbrüchen wegen unerwünschter Ereignisse (Abbruch wegen UE) gingen sowohl neu aufgetretene Schwangerschaften als auch Untersuchungen bzw. Laborwertveränderungen ein. Eine Schwangerschaft ist im Rahmen einer Nutzenbewertung kein unerwünschtes Ereignis. Bei Ereignissen aufgrund von Untersuchungen bzw. Laborwertveränderungen ergibt sich die Patientenrelevanz nicht unmittelbar bzw. sind dies Surrogatendpunkte, zu denen der pU keine Validierung vorgelegt hat³.

Insgesamt werden diese Ereignisse in der Auswertung zum Endpunkt Abbruch wegen UE in der Nutzenbewertung nicht weiter betrachtet. Bei den Schwangerschaften (SOC Pregnancy, puerperium and perinatal conditions) handelte es sich um 2 Patienten mit Ereignis in der Behandlungsgruppe IFN β -1a 44 μ g (Rebif). Bei den Untersuchungen (SOC Investigations) handelte es sich um 5 Patienten mit Ereignis im Behandlungsarm Teriflunomid und 11

³ Der häufigste Abbruchgrund aufgrund von Untersuchungen war eine Erhöhung der ALT-Werte. Diese Ereignisse waren mehrheitlich von milder / moderater Intensität.

Patienten mit Ereignis bei IFN β -1a 44 μ g (Rebif). Patienten, die die Behandlung entweder aufgrund von Schwangerschaft oder aufgrund von Untersuchungen abbrachen, wiesen jeweils keine anderen Abbruchgründe auf. Darüber hinaus wurde eine Sensitivitätsanalyse durchgeführt, in der auch Ereignisse berücksichtigt wurden, die aufgrund von Untersuchungen (Laborwertveränderungen) zum Behandlungsabbruch führten, in der jedoch Schwangerschaft weiterhin nicht als relevantes Ereignis gewertet wurde.

Beobachtungsperioden

In der für den direkten Vergleich relevanten Studie TENERE wurde bei Patienten, die die Behandlung abbrachen oder die Studie wie geplant beendeten, aber nicht an der optionalen Extensionsstudie teilnahmen, ein Verfahren zur beschleunigten Elimination von Teriflunomid durchgeführt (siehe Abschnitt 2.3.2). Diese Wash-out Phase dauerte 11 Tage und war Teil des 16-wöchigen Follow-ups nach Beendigung der Studienmedikation.

Die vom pU in Modul 4 des Dossiers berichteten Auswertungen zu unerwünschten Ereignissen (inklusive Mortalität) in der TENERE Studie bezogen sich auf die „gesamte Behandlungsperiode“ („on-treatment period“), die sich aus der Periode während der Behandlung („during treatment period“) und der „Wash-out“ (beschleunigte Elimination) Periode zusammensetzten. Diese waren wie folgt definiert:

- Periode während der Behandlung („during treatment period“): Zeit von der ersten bis zur letzten Dosis der Studienmedikation (investigational medicinal products [IMP]).
- Beschleunigte Elimination „Wash-out“ Periode: Tag 1 nach Ende der Periode während der Behandlung („during treatment period“) bis zu 112 Tage nach der letzten Dosis der Studienmedikation.

Um eine Verzerrung der UE-Auswertung bezüglich des eingeleiteten Eliminationsverfahrens zu vermeiden, wurde für die Nutzenbewertung die Beobachtungsperiode während der Behandlung („during treatment period“) gewählt und entsprechende Ergebnisse dargestellt. Die Beobachtungsperiode der UE-Auswertung entspricht damit der Beobachtungsperiode zur Erhebung von Endpunkten zur Morbidität bzw. gesundheitsbezogenen Lebensqualität.

Subgruppenmerkmale und andere Effektmodifikatoren

Der pU legt zur TENERE Studie für einen relevanten Endpunkt (jährliche Schubrate) Analysen zu den Merkmalen Geschlecht, Alter, Region, EDSS zu Studienbeginn, Anzahl der Schübe in den vorangegangenen 2 Jahren und Vorbehandlung mit krankheitsmodifizierendem (MS) Wirkstoff vor. Die Ergebnisse der Interaktionstests zeigten keine relevanten Subgruppeneffekte bzw. Effektmodifikationen für diesen Endpunkt. Für alle anderen Endpunkte wurden keine Subgruppenauswertungen vorgelegt. Somit können Subgruppeneffekte insgesamt nicht beurteilt werden.

2.7.2.5 Kommentar zu Ergebnissen aus weiteren Unterlagen – indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

Im Dossier des pU wurde ein indirekter Vergleich randomisierter kontrollierter Studien vorgelegt, welcher den über den direkten Vergleich abgeleiteten Zusatznutzen unterstützen soll. Der indirekte Vergleich wird jedoch für die vorliegende Dossierbewertung nicht herangezogen (siehe Abschnitt 2.7.2.3.2).

2.7.2.6 Kommentar zu Ergebnissen aus weiteren Unterlagen – nicht randomisierte vergleichende Studien

Im Dossier des pU wurden keine nicht randomisierten vergleichenden Studien zur Beschreibung des Zusatznutzens von Teriflunomid herangezogen.

2.7.2.7 Kommentar zu Ergebnissen aus weiteren Unterlagen – weitere Untersuchungen

Im Dossier des pU wurden keine weiteren Untersuchungen zur Beschreibung des Zusatznutzens von Teriflunomid herangezogen.

2.7.2.8 Kommentar zur abschließenden Bewertung der Unterlagen zum Nachweis des Zusatznutzens

2.7.2.8.1 Beurteilung der Aussagekraft der Nachweise

Die Angaben des pU zur Beurteilung der Aussagekraft der Nachweise befinden sich in Modul 4 (Abschnitt 4.4.1) des Dossiers.

Der pU beurteilt die Aussagekraft der Nachweise anhand von Studienqualität, Validität der Endpunkte und Evidenzstufe.

Der pU schätzt die Ergebnissicherheit der Studie TENERE als hoch ein, da es sich um eine randomisierte kontrollierte Studie (Evidenzstufe 1b) handele, deren Verzerrungspotenzial auf Studienebene niedrig sei. Die Endpunkterhebung in der Studie habe mittels valider Messinstrumente stattgefunden und bei der Endpunkterhebung seien adäquate Maßnahmen getroffen worden, um das Verzerrungspotenzial niedrig zu halten. Des Weiteren wurden in der Studie mehrheitlich Patienten behandelt, die gemäß Zulassung für die Behandlung mit Teriflunomid infrage kommen.

Wie in den Abschnitten 2.4.2 und 2.7.2.4.2 beschrieben, wurde das Verzerrungspotenzial auf Studienebene zwar als niedrig bewertet, jedoch lag für alle Endpunkte (außer Mortalität) ein hohes Verzerrungspotenzial vor. Mehrheitlich haben demnach die vorliegenden Studienergebnisse eine mäßige Ergebnissicherheit. Lediglich für den Endpunkt Mortalität ist die Ergebnissicherheit hoch.

2.7.2.8.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß / Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht

Die Angaben des pU zum Zusatznutzen, einschließlich dessen Wahrscheinlichkeit und Ausmaß, und zu Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, befinden sich in Modul 4 (Abschnitte 4.4.2, 4.4.3) des Dossiers.

Der pU leitet auf Grundlage der in Modul 4 des Dossiers vorgelegten Ergebnisse insgesamt einen Beleg für einen beträchtlichen Zusatznutzen ab.

Das Ausmaß „beträchtlich“ ergibt sich aus der Sicht des pU durch die Vermeidung unerwünschter Ereignisse besonderen Interesses (Reaktionen an der Injektionsstelle und grippeähnliche Symptome werden hervorgehoben) sowie einer größeren Patientenzufriedenheit bei einer Behandlung mit Teriflunomid. Des Weiteren beansprucht der pU einen geringen Zusatznutzen aufgrund weniger unerwünschter Ereignisse, die zum Therapieabbruch führten.

Auf Grundlage des indirekten Vergleichs ermittelt der pU für 2 dieser Endpunkte (Abbruch wegen unerwünschter Ereignisse und für den Komplex besondere unerwünschte Ereignisse) Ergebnisse, die die Ergebnisse der TENERE Studie stützen sollen. So stuft er für diese Endpunkte die Wahrscheinlichkeit von einem „Hinweis“ zu einem „Beleg“ für einen Zusatznutzen herauf.

Aus den vorangegangenen Abschnitten 2.4, 2.7.2.2 und 2.7.2.4.3 und Kommentaren zum Dossier ergeben sich Abweichungen in der Einschätzung des Zusatznutzens für die vorliegende Bewertung. So wird von der Endpunktauswahl (u. a. bei der Patientenzufriedenheit sowie einzelner vom pU vorgelegten Endpunkte zu besonderen Nebenwirkungen) abgewichen. Des Weiteren war aufgrund der im Dossier verwendeten Operationalisierung eine Sensitivitätsanalyse zum Endpunkt Abbruch wegen UE erforderlich (siehe Abschnitt 2.7.2.4.3). Das Ergebnis dieser Sensitivitätsanalyse wird in Abschnitt 2.4.3 beschrieben. Der vorgelegte indirekte Vergleich war nicht geeignet, um Aussagen zum Zusatznutzen von Teriflunomid im Vergleich gegenüber der zweckmäßigen Vergleichstherapie treffen zu können (2.7.2.3.2).

Insgesamt ergibt sich kein Beleg für einen Zusatznutzen von Teriflunomid gegenüber der zweckmäßigen Vergleichstherapie. Die Bewertung zum Ausmaß und zur Wahrscheinlichkeit des Zusatznutzens von Teriflunomid ist Abschnitt 2.5 zu entnehmen.

2.7.2.9 Kommentar zur Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

2.7.2.9.1 Begründung für die Vorlage indirekter Vergleiche

Die Angaben des pU zur Begründung der Vorlage indirekter Vergleiche befinden sich in Modul 4 (Abschnitt 4.5.1) des Dossiers.

Die Angaben hierzu werden an anderer Stelle kommentiert (siehe Abschnitt 2.7.2.3.2).

2.7.2.9.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen

Im Dossier wurden keine nicht randomisierten vergleichenden Studien und weitere Untersuchungen zur Beschreibung des Zusatznutzens von Teriflunomid herangezogen.

2.7.2.9.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen

Im Dossier wurde nicht beschrieben, dass valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen können. Der pU legt im Dossier vielmehr selbst Ergebnisse zu patientenrelevanten Endpunkten vor.

2.7.2.9.4 Verwendung von Surrogatendpunkten

Der pU macht keine Angaben zu Surrogatendpunkten, die für die Bewertung des Zusatznutzens von Teriflunomid anhand des direkten Vergleichs relevant sind.

3 Kosten der Therapie

3.1 Kommentar zur Anzahl der Patienten mit therapeutisch bedeutsamem Zusatznutzen (Modul 3, Abschnitt 3.2)

Die Angaben des pU zur Anzahl der Patienten mit therapeutisch bedeutsamem Zusatznutzen befinden sich in Modul 3 (Abschnitt 3.2) des Dossiers. Der G-BA hat als zweckmäßige Vergleichstherapie für Patienten mit RRMS Beta-Interferon (1a oder 1b) oder Glatirameracetat festgelegt. Der pU wählt Interferon beta-1a (IFN β -1a) als zweckmäßige Vergleichstherapie, beschränkt seine Wahl jedoch auf IFN β -1a 44 μ g s. c. 3-mal pro Woche (Rebif), eines der Präparate aus der Wirkstoffgruppe.

3.1.1 Beschreibung der Erkrankung und Charakterisierung der Zielpopulation

Der pU beschreibt ausführlich und nachvollziehbar die schubförmig-remittierende Multiple Sklerose. Die Zielpopulation charakterisiert der pU korrekt gemäß der Fachinformation von Teriflunomid [3].

3.1.2 Therapeutischer Bedarf

Der pU gibt an, dass Teriflunomid gegenüber den derzeit als Basistherapeutika empfohlenen Wirkstoffen IFN β und Glatirameracetat besser verträglich und anwenderfreundlicher ist. Während IFN β und Glatirameracetat parenteral verabreicht werden, wird Teriflunomid oral eingenommen. Außerdem würden laut pU mit Teriflunomid grippeähnliche Symptome seltener auftreten als mit den beiden zweckmäßigen Vergleichstherapien.

3.1.3 Prävalenz und Inzidenz

GKV-Patienten in der Zielpopulation

Der pU stützt sich bei der Anzahl der GKV-Zielpopulation auf Daten aus dem morbiditätsorientierten Risikostrukturausgleich (Morbi-RSA). Der Morbi-RSA erfasst Multiple Sklerose (MS) und andere demyelinisierende Erkrankungen des zentralen Nervensystems in der hierarchisierten Morbiditätsgruppe 072 (HMG072). Für die Zuordnung zur HMG072 müssen die Patienten im Jahr 2011 2-mal eine gesicherte Diagnose gemäß den ICD-10-GM-Codes G35.0 bis G37.9 haben (jedoch in 2 unterschiedlichen Quartalen). Dem im Oktober 2012 vom Bundesversicherungsamt veröffentlichten Jahresausgleichsbescheid für das Jahr 2011 lässt sich entnehmen, dass 66 479 512 Versichertentage der HMG072 zugeordnet werden [18]. Der pU nimmt an, dass alle Patienten im Jahr 2011 durchgängig 365 Tage versichert waren. Somit ergeben sich 182 136 GKV-Patienten in der HMG072.

Dieses Vorgehen vergleicht der pU mit den Prävalenzdaten aus einer veröffentlichten Querschnittserhebung von 1997, durch die ein Wert von ca. 122 000 Patienten in Deutschland mit MS ermittelt wurde und auf die sich bisher die Angaben zur Prävalenz von MS in Deutschland bezogen [19]. Der pU begründet die bevorzugte Verwendung der Morbi-RSA-Daten mit einer höheren Repräsentativität im Vergleich zu der Stichprobenbefragung

(aufgrund der Vollerhebung) sowie damit, dass die Daten des Morbi-RSA aktueller sind als die Querschnittserhebung.

Tabelle 3-11 in Modul 3A des Dossiers (Seite 33) zeigt verständlich die methodische Herleitung der Anzahl von Patienten mit schubförmig-remittierender Multipler Sklerose. Hierbei wird die Zielpopulation über 3 weitere Schritte auf Grundlage der oben genannten 182 136 Patienten der GKV in der entsprechenden Gruppe des Morbi-RSA definiert:

- 1) Zunächst wird der Anteil der Patienten mit einer MS-Diagnose (ICD-10 G35.0-35.9) von denjenigen Patienten ermittelt, die der HMG072 zugewiesen wurden. Der pU geht davon aus, dass 96 % der Patienten der HMG072 eine solche MS-Diagnose aufweisen [20].
- 2) Von dieser Anzahl wird der Anteil erwachsener Patienten berechnet. In der Literatur wird der Anteil der Patienten, die bei Erkrankungsbeginn jünger als 16 Jahre sind, auf 3 bis 5 % geschätzt [21]. Der pU schätzt konservativ, dass 97 % der Patienten 18 Jahre und älter sind.
- 3) Zuletzt ermittelt der pU den Anteil der Patienten, die an schubförmig-remittierender MS erkrankt sind. Gemäß der Daten des Deutschen MS-Registers sind dies 57 % der Patienten mit MS [22].

Somit ergibt sich eine GKV-Zielpopulation von 96 675 Patienten. Aufgrund der Ergebnisunsicherheit gibt der pU eine Spanne für die GKV-Zielpopulation an. Diese berechnet er, indem er um die Angabe einen Unsicherheitskorridor von $\pm 10\%$ annimmt und anschließend auf die nächsten 5000 rundet. Somit ergibt sich eine geschätzte Spanne für die GKV-Zielpopulation von 85 000 bis 105 000 Patienten.

Bewertung des Instituts

Der pU geht bei der Ermittlung der Zielpopulation nachvollziehbar vor. Die Argumentation des pU für die Verwendung der Morbi-RSA-Daten im Gegensatz zu den bisher verwendeten Prävalenzdaten aus der Querschnittsbefragung von 1997 ist nachvollziehbar. Die für die Ermittlung der Zielpopulation verwendeten Parameter sind in allen 3 Schritten plausibel. Die angegebene Spanne des pU ist in ihrer Größenordnung plausibel.

Entwicklung der Prävalenz und Inzidenz

Der pU geht von einer konstanten Prävalenz der MS aus. Diese Einschätzung ist nachvollziehbar.

3.1.4 Anzahl der Patienten mit therapeutisch bedeutsamen Zusatznutzen

Diese Dossierbewertung stuft das Ausmaß und die Wahrscheinlichkeit des Zusatznutzens von Teriflunomid auf Basis der Daten als „nicht belegt“ ein (siehe Abschnitt 2.5).

3.2 Kommentar zu den Kosten der Therapie für die gesetzliche Krankenversicherung (Modul 3, Abschnitt 3.3)

Die Angaben des pU zu Kosten der Therapie für die gesetzliche Krankenversicherung befinden sich in Modul 3 (Abschnitt 3.3) des Dossiers.

3.2.1 Behandlungsdauer

Die Angaben des pU zur Behandlungsdauer sind nachvollziehbar, plausibel und entsprechen den Fach- und Gebrauchsinformationen [3].

3.2.2 Verbrauch

Die Angaben des pU zum Verbrauch sind plausibel und zudem nachvollziehbar dargestellt.

3.2.3 Kosten

Die Kosten von Teriflunomid stellt der pU nachvollziehbar dar. Bei der Kostendarstellung für die zweckmäßige Vergleichstherapie werden ausschließlich die Kosten für das Präparat Rebif aus der Wirkstoffklasse IFN β -1a dargestellt. Die Kosten für das Präparat Avonex aus derselben Wirkstoffklasse fehlen.

3.2.4 Kosten für zusätzlich notwendige Leistungen

Die ermittelten Kosten für die Leberfunktionstests und Tests der Schilddrüsenfunktion sind plausibel.

Bei der Ermittlung der Kosten für fiebersenkende Analgetika stützt sich der pU auf die deutsche Leitlinie und berücksichtigt die Einnahme von Ibuprofen. Jedoch empfiehlt die deutsche Leitlinie auch Paracetamol. Wird Paracetamol berücksichtigt, fallen hier keine Kosten für die GKV an, da es sich bei Paracetamol um ein OTC-Präparat handelt. Wird alternativ Ibuprofen berücksichtigt, sind die Angaben vom pU (50,06 € pro Jahr) in ihrer Größenordnung plausibel. Der pU vernachlässigt für Ibuprofen den Rabatt nach §130 a SGB V.

3.2.5 Jahrestherapiekosten

Die vom pU angegebenen Jahrestherapiekosten für das erste Therapiejahr (21 024,25 € für Teriflunomid und 20 124,88 € für Rebif) und die folgenden Therapiejahre (21 022,62 € für Teriflunomid und 20 556,94 € für Rebif) pro Patient sind plausibel. Die Kosten unterscheiden sich aufgrund der bei Rebif erforderlichen Aufitrationsphase und den zusätzlich anfallenden GKV-Leistungen im ersten Therapiejahr.

3.2.6 Versorgungsanteile

Der pU geht für Teriflunomid aufgrund der aktuellen MS-Therapeutika auf dem Markt von einem geschätzten Versorgungsanteil von 1500 bis 15 000 Patienten aus.

3.3 Konsequenzen für die Bewertung

Die angegebene Spanne für die GKV-Zielpopulation ist in ihrer Größenordnung plausibel. Die Jahrestherapiekosten für Teriflunomid und Rebif sind ebenfalls plausibel. Bei der Kostendarstellung für die zweckmäßige Vergleichstherapie werden ausschließlich die Kosten für das Präparat Rebif aus der Wirkstoffklasse IFN β -1a dargestellt. Eine Darstellung der Kosten für das Präparat Avonex aus derselben Wirkstoffklasse fehlt. Zu den beiden anderen vom G-BA als mögliche zweckmäßige Vergleichstherapie vorgeschlagenen Wirkstoffen macht der pU keine Angaben zu den Jahrestherapiekosten.

4 Kommentare zu sonstigen Angaben im Dossier des pharmazeutischen Unternehmers

4.1 Kommentar zur Zusammenfassung der Aussagen im Dossier (Modul 1)

Modul 1 stellt primär eine Zusammenfassung der Angaben aus anderen Modulen dar. Angaben zur Bewertung der Inhalte befinden sich in den Abschnitten 2.7, 3.1, 3.2, 4.2 und 4.3 sowie in Kapitel 5 dieses Dokuments.

4.2 Kommentar zu allgemeinen Angaben zum Arzneimittel und zu den zugelassenen Anwendungsgebieten (Modul 2)

Die allgemeinen Angaben des pU zum Arzneimittel und zugelassenen Anwendungsgebieten befinden sich in Modul 2 (Abschnitte 2.1 bis 2.4) des Dossiers.

Allgemeine Angaben zum Arzneimittel

In Abschnitt 2.1.1 und 2.1.2 des Dossiers macht der pU administrative Angaben zum Arzneimittel und beschreibt den Wirkmechanismus von Teriflunomid sowie von anderen im Anwendungsgebiet zugelassenen Arzneimitteln. Die Angaben sind an dieser Stelle ausreichend.

Zugelassene Anwendungsgebiete

Die Angaben zum zugelassenen Anwendungsgebiet in Deutschland entsprechen der Fachinformation. Darüber hinaus stellt der pU den internationalen Zulassungsstatus dar.

4.3 Kommentar zur qualitätsgesicherten Anwendung (Modul 3, Abschnitt 3.4)

Die Angaben des pU zur qualitätsgesicherten Anwendung befinden sich in Modul 3 (Abschnitt 3.4) des Dossiers.

Der pU zitiert in Abschnitt 3.4.1 die relevanten Abschnitte aus der Fachinformation (4.3 bis 4.7) von Teriflunomid [3]. Angaben zur Dosierung und Art der Anwendung (4.2) fehlen.

Gemäß den Angaben des pU in Abschnitt 3.4.2 liegen auf Basis des European Public Assessment Reports (EPARs) keine Einschränkungen in der Zulassung von Teriflunomid vor [23].

In Abschnitt 3.4.3 des Dossiers befindet sich eine deutsche Übersetzung der vom pU vorgeschlagenen Maßnahmen zur Risikominimierung („*proposed risk minimization activities*“) [24]. Die Angaben sind im Abgleich mit dem EU-Risk Management Plan, der im EPAR veröffentlicht ist, umfassend [23].

In Abschnitt 3.4.4 des Dossiers zu weiteren Anforderungen an eine qualitätsgesicherte Anwendung, benennt der pU keine weiteren über die unter Abschnitt 3.4.1, 3.4.2 und 3.4.3 benannten Anforderungen. Diese Angaben sind nachvollziehbar.

5 Zusammenfassung der Dossierbewertung

5.1 Zugelassene Anwendungsgebiete

Teriflunomid ist zur Behandlung erwachsener Patienten mit schubförmig-remittierender Multipler Sklerose (MS) angezeigt.

5.2 Medizinischer Nutzen und medizinischer Zusatznutzen im Verhältnis zur zweckmäßigen Vergleichstherapie

Tabelle 16 fasst die Ergebnisse der Bewertung zu Ausmaß und Wahrscheinlichkeit des Zusatznutzens von Teriflunomid gegenüber der zweckmäßigen Vergleichstherapie im zugelassenen Anwendungsgebiet zusammen.

Tabelle 16: Teriflunomid – Ausmaß und Wahrscheinlichkeit des Zusatznutzens

Anwendungsgebiet	Zweckmäßige Vergleichstherapie ^a	Ausmaß und Wahrscheinlichkeit des Zusatznutzens
Erwachsene Patienten mit schubförmig-remittierender Multipler Sklerose	Beta-Interferon (1a oder 1b) oder Glatirameracetat	Zusatznutzen nicht belegt
<p>a: Dargestellt ist jeweils die vom G-BA festgelegte zweckmäßige Vergleichstherapie. In den Fällen, in denen der pU aufgrund der Festlegung der zweckmäßigen Vergleichstherapie durch den G-BA aus mehreren Alternativen eine Vergleichstherapie auswählen kann, ist die entsprechende Auswahl des pU fett markiert. G-BA: Gemeinsamer Bundesausschuss</p>		

Das Vorgehen zur Ableitung einer Gesamtaussage zum Zusatznutzen stellt einen Vorschlag des IQWiG dar. Über den Zusatznutzen beschließt der G-BA.

5.3 Anzahl der Patienten in den für die Behandlung infrage kommenden Patientengruppen

Tabelle 17: Anzahl der GKV-Patienten in der Zielpopulation

Bezeichnung der Therapie (zu bewertendes Arzneimittel)	Anzahl der GKV-Patienten in der Zielpopulation	Kommentar des Instituts
Teriflunomid	85 000–105 000 ^a	Das Vorgehen des pU bei der Ermittlung der Zielpopulation ist nachvollziehbar. Die Angaben des pU sind in ihrer Größenordnung plausibel.
<p>a: Angaben des pU GKV: gesetzliche Krankenversicherung; pU: pharmazeutischer Unternehmer</p>		

5.4 Kosten der Therapie für die gesetzliche Krankenversicherung

Tabelle 18: Jahrestherapiekosten für die GKV für das zu bewertende Arzneimittel und die zweckmäßige Vergleichstherapie pro Patient

Bezeichnung der Therapie (zu bewertendes Arzneimittel, zweckmäßige Vergleichstherapie ^a)	Bezeichnung der Patientenpopulation	Jahrestherapiekosten pro Patient in €	Kommentar des Instituts
Teriflunomid	Erwachsene Patienten mit schubförmig-remittierender Multipler Sklerose	1. Jahr: 21 024,25 ^b Folgejahre: 21 022,62 ^b	Die vom pU angegebenen Jahrestherapiekosten pro Patient sind in ihrer Größenordnung plausibel. Die Kosten unterscheiden sich in den Therapiejahren aufgrund der zusätzlich anfallenden GKV-Leistungen im ersten Therapiejahr.
Beta-Interferon (1a oder 1b) oder Glatirameracetat	Erwachsene Patienten mit schubförmig-remittierender Multipler Sklerose	1. Jahr: 20 124,88 ^b Folgejahre: 20 556,94 ^b	Die vom pU angegebenen Jahrestherapiekosten für Beta-Interferon 1a pro Patient sind in ihrer Größenordnung plausibel. Jedoch fehlt die Berücksichtigung von Avonex aus der Wirkstoffklasse IFN β -1a. Die Kosten unterscheiden sich in den Therapiejahren aufgrund der bei Rebif erforderlichen Auftitrationsphase und den zusätzlich anfallenden GKV-Leistungen im ersten Therapiejahr. Zu den beiden anderen vom G-BA als mögliche zweckmäßige Vergleichstherapie vorgeschlagenen Wirkstoffen macht der pU keine Angaben hinsichtlich der Jahrestherapiekosten.

a: Dargestellt ist jeweils die vom G-BA festgelegte zweckmäßige Vergleichstherapie. In den Fällen, in denen der pU aufgrund der Festlegung der zweckmäßigen Vergleichstherapie durch den G-BA aus mehreren Alternativen eine Vergleichstherapie auswählen kann, ist die entsprechende Auswahl des pU **fett** markiert.

b: Angaben des pU

G-BA: Gemeinsamer Bundesausschuss; GKV: gesetzliche Krankenversicherung; IFN β -1a: Interferon beta-1a; pU: pharmazeutischer Unternehmer

5.5 Anforderungen an eine qualitätsgesicherte Anwendung

Nachfolgend werden die Angaben des pU aus Modul 1, Abschnitt 1.8 „Anforderungen an eine qualitätsgesicherte Anwendung“ ohne weitere Anpassung präsentiert. Details zur Bewertung dieser Inhalte sind Abschnitt 4.3 zu entnehmen.

„Die Behandlung ist von einem Arzt mit Erfahrung in der Behandlung der MS einzuleiten und zu überwachen.

Eine beschleunigte Elimination des Wirkstoffs aus dem Körper ist bei Bedarf möglich.

Bei Patienten mit schwer beeinträchtigtem Immunstatus (z. B. Aids) sollte Teriflunomid nicht angewendet werden.

1.8.1 Monitoring

Untersuchungen vor der Behandlung mit Teriflunomid:

-Blutdruck

-Alaninaminotransferase (SLT/SGPT)

-Großes Blutbild, einschließlich Differenzialblutbild

Untersuchungen während der Behandlung:

-Blutdruck

-Alaninaminotransferase (SLT/SGPT)

-Großes Blutbild, bei Anzeichen und Symptomen (z. b. Infektionen)

1.8.2 Leberfunktion

Eine starke Einschränkung der Leberfunktion ist in der Fachinformation als Kontraindikation aufgeführt.

Absetzen der Behandlung bei Verdacht auf Leberschädigung ist angezeigt.

Teriflunomid soll bei Patienten mit schwerer Hypoproteinämie nicht angewendet werden und bei Patienten, die viel Alkohol konsumieren, nur mit Vorsicht.

1.8.3 Infektionen

Ein Therapiebeginn mit Teriflunomid soll bei Patienten mit einer schweren aktiven Infektion bis zu deren Abklingen aufgeschoben werden.

Eine Unterbrechung der Teriflunomid-Behandlung sollte in Betracht gezogen werden, wenn ein Patient an einer schwerwiegenden Infektion erkrankt ist.

1.8.4 Schwangerschaft und Stillzeit

Bei Frauen im gebärfähigen Alter:

-Ausschluss einer Schwangerschaft vor Therapiebeginn.

-Empfehlung, eine zuverlässige Verhütungsmethode anzuwenden und Aufklärung über Risiken für den Fötus.

Mütter sollten während der Behandlung mit Teriflunomid nicht stillen.

1.8.5 Risiko von Wechselwirkungen mit anderen Arzneimitteln

Bei der gleichzeitigen Anwendung von Teriflunomid und anderen Arzneimitteln sollte die Möglichkeit von Wechselwirkungen in Betracht gezogen werden.“

6 Literatur

1. Bundesministerium für Gesundheit. Verordnung über die Nutzenbewertung von Arzneimitteln nach § 35a Absatz 1 SGB V für Erstattungsvereinbarungen nach § 130b SGB V (Arzneimittel-Nutzenbewertungsverordnung – AM-NutzenV). Bundesgesetzblatt Teil 1 2010; (68): 2324-2328.
2. Gemeinsamer Bundesausschuss. Verfahrensordnung des Gemeinsamen Bundesausschusses [online]. 26.11.2013 [Zugriff: 02.12.2013]. URL: http://www.g-ba.de/downloads/62-492-803/VerfO_2013-04-18_2013-11-26.pdf.
3. Genzyme. Aubagio 14 mg Filmtabletten: Fachinformation [online]. 08.2013 [Zugriff: 07.10.2013]. URL: <http://www.fachinfo.de>.
4. Merck. Rebif 44 Mikrogramm/0,5 ml Injektionslösung in einer Patrone: Fachinformation [online]. 02.2013 [Zugriff: 07.10.2013]. URL: <http://www.fachinfo.de>.
5. Polman CH, Reingold SC, Edan G, Filippi M, Hartung HP, Kappos L et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". Ann Neurol 2005; 58(6): 840-846.
6. European Medicines Agency. Guideline on clinical investigation of medicinal products for the treatment of multiple sclerosis [online]. 16.11.2006 [Zugriff: 07.10.2013]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003485.pdf.
7. European Medicines Agency. Guideline on clinical investigation of medicinal products for the treatment of multiple sclerosis: draft [online]. 20.09.2012 [Zugriff: 07.10.2013]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/10/WC500133438.pdf.
8. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden: Version 4.1 [online]. 28.11.2013 [Zugriff: 02.12.2013]. URL: https://www.iqwig.de/download/IQWiG_Methoden_Version_4-1.pdf.
9. Biogen Idec. Avonex 30 Mikrogramm (Interferon beta-1a): Fachinformation [online]. 07.2012 [Zugriff: 22.10.2013]. URL: <http://www.fachinfo.de>.
10. Cohen JA, Reingold SC, Polman CH, Wolinsky JS. Disability outcome measures in multiple sclerosis clinical trials: current status and future prospects. Lancet Neurol 2012; 11(5): 467-476.
11. Neurostatus. Neurostatus Scoring: definitions for a standardised, quantified neurological examination and assessment of Kurtzke's Functional Systems and Expanded Disability Status Scale in multiple sclerosis [online]. [Zugriff: 04.12.2013]. URL: http://www.neurostatus.net/media/specimen/Definitions_0410-2_s.pdf.
12. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an Expanded Disability Status Scale (EDSS). Neurology 1983; 33(11): 1444-1452.

13. Filippini G, Del Giovane C, Vacchi L, D'Amico R, Di Pietrantonj C, Beecher D et al. Immunomodulators and immunosuppressants for multiple sclerosis: a network meta-analysis. *Cochrane Database Syst Rev* 2013; (6): CD008933.
14. Freedman MS, Hughes B, Mikol DD, Bennett R, Cuffel B, Divan V et al. Efficacy of disease-modifying therapies in relapsing remitting multiple sclerosis: a systematic comparison. *Eur Neurol* 2008; 60(1): 1-11.
15. Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009; 338: b1147.
16. Cutter GR, Baier ML, Rudick RA, Cookfair DL, Fischer JS, Petkau J et al. Development of a multiple sclerosis functional composite as a clinical trial outcome measure. *Brain* 1999; 122(Pt 5): 871-882.
17. Río J, Nos C, Tintoré M, Téllez N, Galán I, Pelayo R et al. Defining the response to interferon-beta in relapsing-remitting multiple sclerosis patients. *Ann Neurol* 2006; 59(2): 344-352.
18. Bundesversicherungsamt. Jahresausgleichsbescheid-Risikogruppenanteile von 2011 [online]. 24.10.2012 [Zugriff: 10.12.2013]. URL: http://www.bundesversicherungsamt.de/fileadmin/redaktion/Risikostrukturausgleich/Info-Dateien%20und%20Auswertungen/JAB_2011.xls.
19. Hein T, Hopfenmüller W. Hochrechnung der Zahl an Multiple Sklerose erkrankten Patienten in Deutschland. *Nervenarzt* 2000; 71(4): 288-294.
20. Hapfelmeiner A, Larisch K, Nagel R, Seiffert A, Schmidt J, Trottmann M. Epidemiologische, medizinische und ökonomische Kenngrößen der Multiplen Sklerose in Deutschland: Studienbericht [unveröffentlicht]. 2012.
21. Stark W, Gärtner J. Multiple Sklerose (Enzephalomyelitis disseminata) des Kindes- und Jugendalters. *Monatsschrift Kinderheilkunde* 2009; 157(1): 67-80.
22. Deutsche Multiple Sklerose Gesellschaft. Viele MS-Patienten erhalten eine Immuntherapie. *Aktiv* 2009; (224): 8-9.
23. European Medicines Agency. AUBAGIO: European public assessment report [online]. 27.06.2013 [Zugriff: 10.12.2013]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/002514/WC500148684.pdf.
24. Sanofi. Teriflunomide (HMR1726): Aubagio; EU risk management plan [unveröffentlicht]. 2013.

Anhang A– Darlegung potenzieller Interessenkonflikte (externe Sachverständige sowie Patientinnen und Patienten bzw. Patientenorganisationen)

Externe Sachverständige

Diese Dossierbewertung wurde unter Einbindung externer Sachverständiger (einer medizinisch-fachlichen Beraterin / eines medizinisch-fachlichen Beraters) erstellt. Medizinisch-fachliche Berater/-innen, die wissenschaftliche Forschungsaufträge für das Institut bearbeiten, haben gemäß § 139b Abs. 3 Nr. 2 SGB V Gesetzliche Krankenversicherung „alle Beziehungen zu Interessenverbänden, Auftragsinstituten, insbesondere der pharmazeutischen Industrie und der Medizinprodukteindustrie, einschließlich Art und Höhe von Zuwendungen“ offenzulegen. Das Institut hat von der Beraterin / dem Berater ein ausgefülltes Formular „Offenlegung potenzieller Interessenkonflikte“ erhalten. Die Angaben wurden durch das speziell für die Beurteilung der Interessenkonflikte eingerichtete Gremium des Instituts bewertet. Es wurden keine Interessenkonflikte festgestellt, die die fachliche Unabhängigkeit im Hinblick auf eine Bearbeitung des vorliegenden Auftrags gefährden. Im Folgenden sind die Angaben zu Beziehungen zusammengefasst. Alle Informationen beruhen auf Selbstangaben der einzelnen Personen anhand des „Formblatts zur Offenlegung potenzieller Interessenkonflikte – Version ‚frühe Nutzenbewertung‘“. Das Formblatt ist unter www.iqwig.de abrufbar. Die in diesem Formblatt verwendeten Fragen befinden sich im Anschluss an diese Zusammenfassung.

Name	Frage 1	Frage 2 / Ergänzende Frage	Frage 3 / Ergänzende Frage	Frage 4 / Ergänzende Frage	Frage 5	Frage 6
Kieseier, Bernd	nein	ja / nein	ja / nein	ja / nein	nein	nein

Eingebundene Patientinnen und Patienten bzw. Patientenorganisationen

Im Folgenden sind die potenziellen Interessenkonflikte der eingebundenen Patientinnen und Patienten bzw. Patientenorganisationen dargestellt. Alle Informationen beruhen auf Selbstangaben der einzelnen Personen anhand des „Formblatts zur Offenlegung potenzieller Interessenkonflikte – Version ‚frühe Nutzenbewertung‘“. Das Formblatt ist unter www.iqwig.de abrufbar. Die in diesem Formblatt verwendeten Fragen befinden sich im Anschluss an diese Zusammenfassung.

Name; Institution	Frage 1	Frage 2 / Ergänzende Frage	Frage 3 / Ergänzende Frage	Frage 4 / Ergänzende Frage	Frage 5	Frage 6
Seestaedt, Gabriele; Deutsche Multiple Sklerose Gesellschaft, Bundesverband e. V.	nein	nein / nein	nein / nein	nein / nein	ja	nein

Im „Formblatt zur Offenlegung potenzieller Interessenkonflikte – Version ‚frühe Nutzenbewertung‘“ wurden folgende Fragen gestellt:

Frage 1: Sind oder waren Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor angestellt bei einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere bei einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband?

Frage 2: Beraten Sie oder haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor ein Unternehmen, eine Institution oder einen Interessenverband im Gesundheitswesen, insbesondere ein pharmazeutisches Unternehmen, einen Hersteller von Medizinprodukten oder einen industriellen Interessenverband, direkt oder indirekt beraten?

Ergänzende Frage zu Frage 2: Haben Sie darüber hinaus das von der Nutzenbewertung betroffene Unternehmen jemals im Zusammenhang mit der präklinischen oder klinischen Entwicklung des zu bewertenden Arzneimittels, direkt oder indirekt beraten?

Frage 3: Haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor direkt oder indirekt von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband, Honorare erhalten für Vorträge, Stellungnahmen oder Artikel?

Ergänzende Frage zu Frage 3: Haben Sie darüber hinaus von dem von der Nutzenbewertung betroffenen Unternehmen jemals im Zusammenhang mit der präklinischen oder klinischen Entwicklung des zu bewertenden Produkts Honorare erhalten für Vorträge, Stellungnahmen oder Artikel?

Frage 4: Haben Sie und / oder hat die Einrichtung⁴, die Sie vertreten, abseits einer Anstellung oder Beratungstätigkeit innerhalb des laufenden Jahres und der 3 Kalenderjahre davor von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband, finanzielle Unterstützung für Forschungsaktivitäten, andere wissenschaftliche Leistungen oder Patentanmeldungen erhalten?

Ergänzende Frage zu Frage 4: Haben Sie darüber hinaus persönlich abseits einer Anstellung oder Beratungstätigkeit jemals von dem von der Nutzenbewertung betroffenen Unternehmen im Zusammenhang mit der präklinischen oder klinischen Entwicklung des zu bewertenden Produkts finanzielle Unterstützung für Forschungsaktivitäten, andere wissenschaftliche Leistungen oder Patentanmeldungen erhalten?

⁴ Sofern Sie in einer ausgedehnten Institution tätig sind, genügen Angaben zu Ihrer Arbeitseinheit, zum Beispiel Klinikabteilung, Forschungsgruppe etc.

Frage 5: Haben Sie und / oder hat die Einrichtung⁴, bei der Sie angestellt sind bzw. die Sie vertreten, innerhalb des laufenden Jahres und der 3 Kalenderjahre davor sonstige finanzielle oder geldwerte Zuwendungen (z. B. Ausrüstung, Personal, Unterstützung bei der Ausrichtung einer Veranstaltung, Übernahme von Reisekosten oder Teilnahmegebühren ohne wissenschaftliche Gegenleistung) erhalten von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere von einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband?

Frage 6: Besitzen Sie Aktien, Optionsscheine oder sonstige Geschäftsanteile eines Unternehmens oder einer anderweitigen Institution, insbesondere von einem pharmazeutischen Unternehmen oder einem Hersteller von Medizinprodukten? Besitzen Sie Anteile eines „Branchenfonds“, der auf pharmazeutische Unternehmen oder Hersteller von Medizinprodukten ausgerichtet ist?