

IQWiG-Berichte – Nr. 163

Analytic Hierarchy Process (AHP) – Pilotprojekt zur Erhebung von Patienten- präferenzen in der Indikation Depression

Arbeitspapier

Version: 1.0
Stand: 08.05.2013

Impressum

Herausgeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema:

Analytic Hierarchy Process (AHP) – Pilotprojekt zur Erhebung von Patientenpräferenzen in der Indikation Depression

Auftraggeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

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

Dieser Bericht wurde unter Beteiligung externer Sachverständiger erstellt. Externe Sachverständige, die wissenschaftliche Forschungsaufträge für das Institut bearbeiten, haben gemäß § 139b Abs. 3 Nr. 2 Sozialgesetzbuch – Fünftes Buch – 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 jedem der Sachverständigen 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.

Externe Sachverständige:

- Marjan Hummel, University of Twente, School of Management and Governance, Dept. Health Technology & Services Research
- Maarten IJzerman, University of Twente, School of Management and Governance, Dept. Health Technology & Services Research
- Jeannette van Manen, University of Twente, School of Management and Governance, Dept. Health Technology & Services Research

An der Erstellung des Arbeitspapiers beteiligte Mitarbeiter des IQWiG¹:

- Marion Danner
- Andreas Gerber-Grote
- Fabian Volz
- Beate Wiegard

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

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	v
Abbildungsverzeichnis	vi
Abkürzungsverzeichnis.....	vii
Kurzfassung	ix
1 Hintergrund	1
1.1 Einführung in das Thema und Zielsetzung des Pilotprojekts	1
1.2 Die Indikation Depression	1
1.3 AHP zur Gewichtung von Endpunkten einer antidepressiven Behandlung	2
1.4 Grundlagen des AHP-Verfahrens	3
2 Methoden	7
2.1 Auswahl der Kriterien für die AHP-Entscheidungshierarchie	7
2.2 Strukturierung der Endpunkte.....	9
2.2.1 Operationalisierung der Behandlungsziele / Endpunkte	12
2.3 Ablauf des vorliegenden AHP-Verfahrens	13
2.3.1 Rekrutierung von Patienten	14
2.3.2 Rekrutierung von Experten.....	14
2.4 Vorabinformationen für die Teilnehmer	15
2.5 Ablauf der AHP-Befragungen	15
2.5.1 Einführung.....	15
2.5.2 Leitung der Befragungen.....	16
2.5.3 Bewertung und Diskussion der Paarvergleiche in den AHP-Befragungen.....	16
3 Ergebnisse.....	18
3.1 Teilnehmende Patienten	18
3.2 Teilnehmende Experten.....	18
3.3 Ergebnisse der Befragung der Patienten	19
3.4 Ergebnisse der Befragung der Experten.....	19
3.5 Unterschiede zwischen den Ergebnissen der Patienten- und der Expertenbefragung	20
3.6 Übersicht zum Vergleich der Befragungen	23
3.7 Ergebnisse zur alternativen Entscheidungshierarchie mit drei Endpunkten	24
3.8 Ergebnisse der Diskussionen in den Gruppen.....	25
4 Diskussion.....	27

4.1	Methodische Aspekte	27
4.2	AHP-Befragungen.....	30
4.3	Rekrutierung von Patienten und Experten	32
5	Fazit.....	33
6	Literatur	34
Anhang A – Vorabinformation der Patienten		38
Anhang B – Vorabinformation der Experten.....		41
Anhang C – Individuelle Werte der Befragung der Patienten und der Experten		44
Anhang D – Individuelle Werte der Befragung der Patienten und der Experten in der verkürzten Hierarchie.....		46

Tabellenverzeichnis

	Seite
Tabelle 1: Beispielmatrix	4
Tabelle 2: Gruppengewichte (geometrisches Mittel) und Werte der Konsistenzratio aus der Patientenbefragung.....	19
Tabelle 3: Gruppengewichte (geometrisches Mittel) und Werte der Konsistenzratio aus der Expertenbefragung	20
Tabelle 4: Priorisierung der Endpunkte in den Gruppen gemäß Tabellen 2 und 3.....	23
Tabelle 5: Gruppengewichte (geometrisches Mittel) aus der zweiten Priorisierungsrunde.....	24

Abbildungsverzeichnis

Abbildung 1: Darstellung eines paarweisen Vergleichs zweier Kriterien	4
Abbildung 2: AHP-Entscheidungsstruktur.....	11
Abbildung 3: Alternative AHP-Entscheidungsstruktur.....	12
Abbildung 4: Gegenüberstellung der aus der Patienten- und der Expertenbefragung resultierenden Prioritäten / Gewichte	21
Abbildung 5: Variation der Einschätzungen / Prioritäten innerhalb der Gruppen.....	22
Abbildung 6: Gewichte aus der zweiten Priorisierungsrunde	24

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AHP	Analytic Hierarchy Process
ANP	Analytic Network Process
BApK	Bundesverband der Angehörigen psychisch Kranker
BPtK	Bundespsychotherapeutenkammer
CA	Conjoint-Analyse
C. I.	Konsistenzindex
C. R.	Konsistenzratio
DEGAM	Deutsche Gesellschaft für Allgemeinmedizin und Familienmedizin
DGP	Deutsche Gesellschaft für Psychologie
DGPM	Deutsche Gesellschaft für Psychosomatische Medizin und Ärztliche Psychotherapie
DGPPN	Deutsche Gesellschaft für Psychiatrie, Psychotherapie und Nervenheilkunde
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders (Diagnostisches und Statistisches Handbuch Psychiatrischer Störungen), 4. Ausgabe
EbM	Evidenzbasierte Medizin
G-BA	Gemeinsamer Bundesausschuss
HAM-D	Hamilton Rating Scale for Depression
HTA	Health Technology Assessment
ICD 10	International Statistical Classification of Diseases and Related Health Problems (Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme), 10. Revision
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
KNB	Kosten-Nutzen-Bewertung
LVR	Landschaftsverband Rheinland
MADRS	Montgomery-Åsberg Depression Rating Scale
MAO	Monoamino-Oxidase-Hemmer
MTCA	Multi-Criteria Decision Analysis (multikriterielle Entscheidungsanalyse)
NDRI	Noradrenalin and Dopamine Reuptake Inhibitor (Noradrenalin- und Dopamin-Wiederaufnahme-Hemmer)
NVL	Nationale Versorgungsleitlinie
OR	Odds Ratio
R. I.	Random Index
SGB	Sozialgesetzbuch

Abkürzung	Bedeutung
SNRI	Serotonin-Noradrenalin Reuptake Inhibitor (Serotonin-Noradrenalin-Wiederaufnahme-Hemmer)
SSRI	Selective Serotonin Reuptake Inhibitor (selektiver Serotonin-Wiederaufnahme-Hemmer)

Kurzfassung

Hintergrund: Kosten-Nutzen-Bewertung, Effizienzgrenze und die Berücksichtigung von Patientenpräferenzen

Die Methoden zur Kosten-Nutzen-Bewertung (KNB) nach § 35b SGB V (in der Fassung, die vor dem 01.01.2011 galt), die das IQWiG in Zusammenarbeit mit einem internationalen Expertenpanel im Herbst 2009 publiziert hat, beruhen auf dem Konzept der Effizienzgrenze. Um eine Effizienzgrenze zu erstellen, werden Nutzen und Kosten möglichst aller alternativen Gesundheitstechnologien in einem Indikationsgebiet erhoben. Die nach Kosten und Nutzen effizientesten Technologien bilden dann die sogenannte Effizienzgrenze [1-3]. Gemäß den Methoden des IQWiG werden Effizienzgrenzen zunächst endpunktspezifisch erstellt. Um eine Aggregation der Ergebnisse endpunktspezifischer Effizienzgrenzen (z. B. zur Ermittlung zu empfehlender Erstattungspreise) zu ermöglichen, können die Ergebnisse auf Basis von Patientenpräferenzen gewichtet und aggregiert werden.

Voraussetzung für die Erstellung einer Effizienzgrenze sind auf der Nutzenseite die Ergebnisse von Studien, die nach den Kriterien der evidenzbasierten Medizin (EbM) bewertet werden. Dabei finden Ergebnisse zu patientenrelevanten Endpunkten Berücksichtigung. Patientenrelevante Endpunkte sind gemäß SGB V primär Endpunkte, die eine Wirkung auf die Mortalität, die Morbidität und die gesundheitsbezogene Lebensqualität der behandelten Personen abbilden [4,5].

Da Patienten sozusagen die „Endverbraucher“ von Gesundheitstechnologien und -leistungen sind, ist eine Berücksichtigung von Patientenpräferenzen innerhalb von Health Technology Assessments (HTA) selbst sowie auch innerhalb von HTA-basierten Entscheidungsprozessen (beispielsweise Erstattungsentscheidungen) von großer Bedeutung. HTA-Institutionen in vielen Ländern beziehen daher inzwischen regelhaft Patienten in ihre HTA-Prozesse ein, jedoch häufig noch in unzureichender Form. Quantitative Ansätze zur Messung von Patientenpräferenzen wie das Analytic-Hierarchy-Process(AHP)-Verfahren werden bislang nicht regelhaft eingesetzt. Im Rahmen des beschriebenen Effizienzgrenzenkonzepts im IQWiG besteht die Möglichkeit, endpunktspezifische Ergebnisse mithilfe von auf Patientenpräferenzen basierenden Gewichten zu aggregieren. Diese wiederum können mit verschiedenen Methoden der multikriteriellen Entscheidungsanalyse (MTCA) wie dem AHP-Verfahren erhoben werden.

Fragestellung

In diesem Pilotprojekt wurde geprüft, inwiefern das AHP-Verfahren im Rahmen der Identifikation, Gewichtung und Priorisierung multipler patientenrelevanter Endpunkte in der KNB in Deutschland angewendet werden kann. Die Anwendungsmöglichkeiten werden am Beispiel der Depression und ihrer pharmakologischen Therapieformen (Antidepressiva) untersucht.

Methode

In einem AHP-Verfahren wurden Patienten und die in deren Behandlung tätigen Experten direkt befragt. Dabei wurden die Teilnehmenden im Rahmen strukturierter Interviews gebeten zu entscheiden, welches/welcher von zwei Behandlungszielen / Endpunkten ihnen wichtiger und um wie viel wichtiger ihnen dieses Behandlungsziel bzw. dieser Endpunkt erschien. Die Ergebnisse aller so durchgeführten paarweisen Vergleiche stellten die Grundlage für die Berechnung des sogenannten „rechten Eigenvektors“ dar, der mithilfe der Matrixalgebra hergeleitet wird. Der rechte Eigenvektor enthält relative Gewichte für jeden der eingeschlossenen Endpunkte bzw. Therapieziele, hergeleitet aus den in den Bewertungen der paarweisen Vergleiche geäußerten Präferenzen.

Das AHP-Pilotprojekt wurde in zwei separaten Befragungen zur Präferenzenerhebung durchgeführt: eine mit Patienten und eine mit Experten (in der Behandlung depressiver Patienten). Zwischen den Paarvergleichen der einzelnen Ziele / Endpunkte wurden die individuellen Bewertungen in der jeweiligen befragten Gruppe diskutiert.

Ergebnisse

Das AHP-Pilotprojekt im IQWiG hat gezeigt, dass es sich hierbei um ein sehr strukturiertes und gut umsetzbares Verfahren handelt. Die Befragungen waren für die Patienten kognitiv gut handhabbar und konnten in einer ruhigen, entspannten Atmosphäre im Rahmen der Patientengruppe umgesetzt werden. Mithilfe der Gruppendiskussionen konnten Erkenntnisse über die Beweg- und Hintergründe der durch die Patienten bzw. Experten vorgenommenen Einschätzungen gewonnen werden.

Die separat durchgeführten Befragungen bei 12 Patienten und 7 Experten führten zu unterschiedlichen Gewichtungen. Während beide Gruppen dieselben 6 von 11 Endpunkten als die bedeutsamsten identifizierten, wichen die Gewichte innerhalb dieser 6 Endpunkte in den Gruppen voneinander ab. Die Patienten gewichteten das Ansprechen am höchsten ($w = 0,32$), wohingegen die Experten der Remission das höchste Gewicht gaben ($w = 0,48$). Auf das Ansprechen folgten in der Gruppe der Patienten die kognitive Funktionsfähigkeit ($w = 0,13$), die Verringerung von Angst ($w = 0,12$), die soziale Funktionsfähigkeit ($w = 0,11$), die Vermeidung eines Rückfalls ($w = 0,09$) und die Remission ($w = 0,09$). Bei den Experten folgten auf die Remission die Vermeidung eines Rückfalls ($w = 0,14$), die soziale Funktionsfähigkeit ($w = 0,09$), die kognitive Funktionsfähigkeit ($w = 0,06$), das Ansprechen ($w = 0,06$) und die Verringerung von Angst ($w = 0,05$). Die unerwünschten Ereignisse spielten eine eher geringere Rolle; nur in aggregierter Form, d. h. auf einer den einzelnen Endpunkten übergeordneten Ebene als kombinierter Endpunkt der Vermeidung unerwünschter Ereignisse (einschließlich der schwerwiegenden Ereignisse), kam er auf ein Gewicht von $w = 0,095$ bei den Patienten und $w = 0,08$ bei den Experten.

Fazit

Dieses Pilotprojekt hat gezeigt, dass das AHP-Verfahren als Methode sowohl mit Patienten als auch mit Experten durchgeführt werden kann. Die Befragung war durchgehend kognitiv gut zu bewältigen. Das strukturierte Verfahren der AHP-Methode ermöglichte es, Schritt für Schritt die Präferenzen von Individuen für bestimmte Behandlungsziele bzw. -endpunkte zu erheben und Gewichte für jeden dieser Endpunkte mittels der Matrixalgebra zu berechnen.

Die paarweisen Vergleiche der Endpunkte in Kombination mit den Gruppendiskussionen erlaubten einen intensiven Austausch von Wahrnehmungen, Informationen und Erfahrungen zwischen den befragten Personen und ließen Einsichten in die Beweg- und Hintergründe der abgegebenen Einschätzungen zu.

Bei der Durchführung des Pilotprojekts wurden einige methodische Herausforderungen des AHP-Verfahrens besonders offensichtlich, die in zukünftigen AHP-Befragungen auf jeden Fall berücksichtigt werden sollten. Zum einen korrelieren bzw. überlappen Endpunkte bzw. Therapieziele. Eine Erhebung von zu hohen Gewichten, die ggf. zum rank reversal führen könnte, lässt sich nie völlig vermeiden, müsste aber durch Sensitivitätsanalysen geprüft werden. Zum anderen bleibt weiterhin die Frage offen, der sich auch Verfahren zur Erhebung der QALYs bzw. andere Methoden der multikriteriellen Entscheidungsfindung stellen müssen: Welche Personen sollen befragt werden und wie übertragbar sind diese Ergebnisse auf eine gesamte Patientenpopulation?

Abgesehen davon, dass diese methodischen Fragen geklärt werden müssen, kann das AHP-Verfahren für verschiedene Fragestellungen im Prozess der HTA-Erstellung genutzt werden.

- Zur Identifizierung der für die Patienten relevant(est)en Endpunkte, um möglicherweise nur für diese Endpunkte Effizienzgrenzen zu erstellen.
- Gewichtung endpunktspezifischer Effektmaße (z. B. ORs) und Ableitung eines aggregierten gewichteten Effektmaßes für jede Behandlungsalternative.
- Aggregation der aus endpunktspezifischen Effizienzgrenzen resultierenden Ergebnisse und Ableitung eines aggregierten Maßes für die Kosten-Effektivität, z. B. kann ein aggregierter, gewichteter Erstattungspreis für Medikament A durch die Gewichtung der aus verschiedenen endpunktspezifischen Effizienzgrenzen abgeleiteten Erstattungspreise berechnet werden.

Darüber hinaus kann das AHP-Verfahren auch wichtige Hinweise liefern, welche Endpunkte in zukünftigen klinischen Studien primär berücksichtigt werden sollten, welche eher nachrangig und welche gegebenenfalls gar nicht.

Schlagwörter: Analytischer Hierarchieprozess, Analytic Hierarchy Process (AHP), Entscheidungsunterstützende Techniken, multikriterielle Entscheidungsanalyse, Depression, Patientenpräferenzen, Pilotprojekte

Keywords: Analytic hierarchy process (AHP), decision support techniques, multi criteria decision making, depression, patients' preferences, pilot projects

1 Hintergrund

1.1 Einführung in das Thema und Zielsetzung des Pilotprojekts

Health Technology Assessment (HTA), d. h. die systematische Bewertung medizinischer Verfahren und Technologien, wird zur Unterstützung von Anwendungs- und Erstattungsentscheidungen genutzt. Entscheidungsprozesse sind jedoch nicht nur vom klinischen und gesundheitsökonomischen Profil einer Technologie abhängig, sondern hängen von vielfältigen dynamischen Einflussfaktoren ab und beziehen die Expertise und Meinung zahlreicher Stakeholder und betroffener Gruppen ein. Da die Patienten die „Endverbraucher“ von Gesundheitstechnologien sind, sind deren Expertise und Einschätzungen von großer Bedeutung und sollten bei Entscheidungsprozessen berücksichtigt werden.

Die multikriterielle Entscheidungsanalyse (Multi-Criteria Decision Analysis, MCDA) kann eingesetzt werden, um komplexe Entscheidungssituationen im Gesundheitsbereich zu unterstützen und Patientenpräferenzen offenzulegen. Die verschiedenen Methoden bzw. Techniken der multikriteriellen Entscheidungsanalyse wurden im Rahmen des Operations Research, des Marketings und der Entscheidungsanalyse entwickelt. Zu diesen Techniken zählt auch das von dem Mathematiker Thomas L. Saaty in den 1970er-Jahren entwickelte „Analytic Hierarchy Process (AHP)“-Verfahren [2,3,6], das ursprünglich zur Unterstützung strategischer Unternehmensentscheidungen dienen sollte. Studien haben gezeigt, dass das AHP-Verfahren auch in unterschiedlichen Entscheidungsfindungssituationen im Gesundheitsbereich hilfreich sein kann [6-13].

In diesem Pilotprojekt wurde das AHP-Verfahren angewendet, um Patienten, die an einer Depression erkrankt sind, zu ermöglichen, Endpunkte einer Behandlung mit Antidepressiva zu priorisieren und zu gewichten. Zusätzlich wurden auch die endpunktbezogenen Präferenzen von in der Behandlung der Depression tätigen Experten erhoben.

Für die Antidepressiva Venlafaxin, Duloxetin, Bupropion, Mirtazapin und Reboxetin publizierte das IQWiG in den Jahren 2010 und 2011 die Ergebnisse zweier Nutzenbewertungen [14,15]. Zusätzlich wird derzeit durch das IQWiG eine gesundheitsökonomische Bewertung der genannten Antidepressiva durchgeführt. Die Priorisierung der in die Nutzenbewertungen eingeschlossenen Endpunkte einer Antidepressiva-Behandlung mithilfe des AHP-Verfahrens eröffnet zum einen die Möglichkeit, eine anschließende gesundheitsökonomische Bewertung auf die für die Patienten wichtigsten Endpunkte zu konzentrieren. Zum anderen können die für die einzelnen Endpunkte mittels AHP berechneten Gewichte zur Aggregation endpunktspezifischer Auswertungen und Preisschätzungen, die das Ergebnis einer Kosten-Nutzen-Modellierung sind, genutzt werden.

1.2 Die Indikation Depression

Die Depression ist eine häufige psychische Erkrankung mit einer Lebenszeitprävalenz zwischen 13 und 17 % in westlichen Ländern [16-19] und ca. 15 % in Deutschland [20]. Die Diagnose einer Episode der Depression gemäß DSM-IV- und ICD-10-Kriterien wird durch

eine Einschätzung von Symptomen getroffen, die in einer bestimmten Anzahl und Ausprägung für mindestens 2 Wochen vorhanden sein müssen. Eines der Symptome muss die depressive Verstimmung bzw. Freude- / Interesseverlust sein. Weitere einzuschätzende Symptome sind:

- Schlaflosigkeit / übertriebene Schläfrigkeit,
- Energieverlust,
- Appetitverlust / Gewichtsverlust,
- vermindertes Selbstwertgefühl / Schuldgefühle,
- psychomotorische Agitiertheit / Hemmung,
- Konzentrationsstörungen und
- wiederkehrende Gedanken an den Tod bzw. Suizidalität.

Auch körperliche Beschwerden wie Müdigkeit oder Schmerzen sind häufig [21]. Die Erkrankung impliziert oft starke Beeinträchtigungen im Alltag des Patienten, insbesondere der sozialen und beruflichen Funktions- und Leistungsfähigkeit sowie der gesundheitsbezogenen Lebensqualität [22,23].

Bei einer milden Depression kann versucht werden, die Erkrankung auch ohne pharmakologische Intervention zu behandeln (siehe [24,25]). Bei mittelgradiger bis schwerer Depression wird eine pharmakologische Behandlung empfohlen, gegebenenfalls in Kombination mit einer Verhaltenstherapie. Es gibt 5 Hauptgruppen von Antidepressiva: die selektiven Serotonin-Wiederaufnahme-Hemmer (SSRI), die tri- und tetrazyklischen Antidepressiva, die Monoamino-Oxidase-Hemmer (MAO-Hemmer), die Serotonin-Noradrenalin-Wiederaufnahme-Hemmer (SNRI) und weitere neuere Antidepressiva-Substanzen (wie Bupropion und Mirtazapin). In den Nutzenbewertungen des IQWiG wurden zum einen Venlafaxin und Duloxetine (beide SNRIs) sowie Bupropion (ein selektiver Noradrenalin- und Dopamin-Wiederaufnahme-Hemmer (NDRI)), Mirtazapin (ein Alpha-2-Rezeptor-Antagonist) und Reboxetin (ein selektiver Noradrenalin-Wiederaufnahmehemmer (NARI)) bewertet [14,15]. Die verschiedenen Antidepressiva unterscheiden sich zum einen nach dem Nutzen einer Behandlung (d. h. das Ausmaß der positiven Effekte der Behandlung auf verschiedene Endpunkte, wie z. B. das Ansprechen) sowie auch nach dem Ausmaß des durch die Behandlung verursachten Schadens (d. h. die negativen Effekte bzw. die Nebenwirkungen der Behandlung).

1.3 AHP zur Gewichtung von Endpunkten einer antidepressiven Behandlung

Das AHP-Verfahren nach Saaty unterstützt die Analyse von komplexen Entscheidungsproblemen, die bestimmte Trade-offs beinhalten. Ein solcher Trade-off ist beispielsweise auch die Abwägung zwischen den verschiedenen Nutzen- und Schadensaspekten einer Behandlung. Das AHP-Verfahren ermöglicht eine Messung, d. h. Quantifizierung der Bedeutung

verschiedener Entscheidungskriterien für eine bestimmte Entscheidung aus Sicht eines oder mehrerer Entscheider, und kann darüber hinaus angeben, wie gut verschiedene Entscheidungsalternativen diese Kriterien erfüllen. In diesem Pilotprojekt wurde das AHP-Verfahren ausschließlich eingesetzt, um die in den Nutzenbewertungen ausgewerteten Endpunkte auf Basis in Befragungen erhobener Patientenpräferenzen zu priorisieren und zu gewichten. Informationen zur Wirkung der Antidepressiva lagen aus den genannten Nutzenbewertungen vor, die auf einer systematischen Recherche nach und Auswertung von randomisierten kontrollierten klinischen Studien basierten.

1.4 Grundlagen des AHP-Verfahrens

Das AHP-Verfahren ermöglicht in direkten Befragungen die strukturelle und schrittweise Erhebung von individuellen Präferenzen für bestimmte Entscheidungskriterien und deren anschließende Auswertung mithilfe eines mathematischen Verfahrens. In strukturierten Interviews werden Individuen gebeten zu entscheiden, welches von zwei Kriterien ihnen wichtiger erscheint und um wie viel wichtiger ihnen dieses Kriterium im Vergleich zu dem anderen Kriterium erscheint. Die Ergebnisse aller so durchgeführten paarweisen Vergleiche stellen die Grundlage für die Berechnung des sogenannten „rechten Eigenvektors“ dar, der mithilfe der Matrixalgebra hergeleitet wird. Der rechte Eigenvektor besteht also aus den Gewichten für jedes der in einer Hierarchieebene eingeschlossenen Entscheidungskriterien, hergeleitet aus den in den Bewertungen der paarweisen Vergleichen geäußerten Präferenzen.

Gemäß der Methode des AHP-Verfahrens muss vor der Durchführung der paarweisen Vergleiche zunächst eine Entscheidungshierarchie entwickelt werden. Im Grundsatz sollte eine solche Hierarchie das Entscheidungsproblem, die Entscheidungskriterien (auf verschiedenen Ebenen) und die Handlungsalternativen auf verschiedenen Ebenen beinhalten. Die unterste Ebene, d. h. die Handlungsalternativen, sprich die konkreten Antidepressiva, war in der Entscheidungsstruktur zu diesem Pilotprojekt jedoch nicht relevant und wurde daher nicht einbezogen. Entscheidungskriterien, die sich auf derselben Hierarchieebene befinden, sollten eindeutig definiert sein, einander ausschließen, umfassend und miteinander vergleichbar sein.

- *Eindeutig definierte Kriterien:* Um zu vermeiden, dass es zu konzeptionellen Verzerrungen im Verfahren kommt, sollten Entscheidungskriterien sehr klar und eindeutig definiert sein und von allen Individuen, die am Verfahren teilnehmen, (gleich) verstanden werden.
- *Einander ausschließende Kriterien auf einer Hierarchieebene:* Die Kriterien auf einer Hierarchieebene sollten möglichst nicht überlappen und voneinander unabhängig sein. Überlappungen könnten zu einer Über- oder Unterschätzung einzelner Kriterien bei der Berechnung von Gewichten führen, da das AHP auf einer additiven Wertefunktion beruht. Außerdem könnte durch eine Abhängigkeit eines Kriteriums von einem anderen der Vergleich dieser beiden Kriterien erschwert oder gar unmöglich gemacht werden.
- *Umfassendes Spektrum an Kriterien:* Die Entscheidungsstruktur sollte alle Entscheidungskriterien beinhalten, die einen Einfluss auf die Präferenzen der Individuen für die

Alternativen haben. Wenn ein Kriterium nicht in der Struktur enthalten ist, das jedoch von großer Bedeutung für die befragten Individuen ist, so spiegeln die auf Basis dieser Entscheidungsstruktur berechneten Ergebnisse nicht sicher die Präferenzen der Individuen.

- *Miteinander vergleichbare Kriterien:* Alle Kriterien auf einer Hierarchieebene werden mithilfe einer 17-Punkte-Skala miteinander verglichen. Die Skala reicht von extremer Präferenz für Kriterium A (9 Punkte auf der Skala – linke Seite des Strahls in Abbildung 1) über Indifferenz zwischen beiden Kriterien (Mittelpunkt der Skala = Indifferenzpunkt) bis hin zur extremen Präferenz für Kriterium B (9 Punkte auf der Skala – rechte Seite des Strahls in Abbildung 1).

Basierend auf der beschriebenen Entscheidungsstruktur erfolgt die Quantifizierung von Präferenzen mithilfe von Paarvergleichsmatrizen, in denen alle paarweisen Vergleiche abgetragen sind. Die paarweisen Vergleiche geben hierbei die relative Bedeutung eines Kriteriums im Vergleich zu den jeweils anderen an. Obwohl die einzelnen Vergleiche auf einer ordinalen Skala gemessen werden (siehe Abbildung 1), gehen die entsprechenden Werte numerisch in die Matrizen und späteren Matrixberechnungen ein.

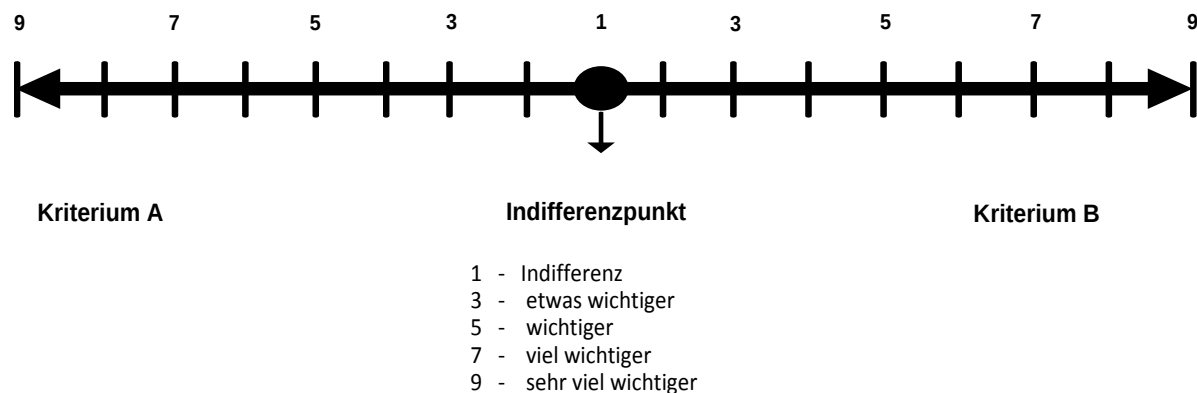


Abbildung 1: Darstellung eines paarweisen Vergleichs zweier Kriterien

Tabelle 1 zeigt beispielhaft eine resultierende Paarvergleichsmatrix $A = (a_{ij})$ mit n Entscheidungskriterien.

Tabelle 1: Beispielmatrix

Entscheidungs-kriterium	Kriterium 1	Kriterium 2	...	Kriterium n
Kriterium 1	a_{11}	a_{12}	...	a_{1n}
Kriterium 2	a_{21}	a_{22}	...	a_{2n}
...	...	...	...	...
Kriterium n	a_{n1}	a_{n2}	...	a_{nn}

Die Matrizen entsprechen den Annahmen der Reflexivität ($a_{ij} = 1$ für $i = j$ entlang der Diagonalen) und der Reziprozität ($a_{ij} = 1/a_{ji}$). Daher müssen nur Werte oberhalb der Diagonalen in die Matrix eingetragen werden. Dies impliziert, dass man für n Entscheidungskriterien $n(n-1)/2$ Einzelpaarvergleichswerte generiert.

Für jede Gruppe von Paarvergleichen kann im AHP-Verfahren ein Maß für die Konsistenz der abgegebenen Bewertungen berechnet werden, die sogenannte Konsistenzratio (C. R.). Dieses Maß überprüft, ob die einzelnen paarweisen Vergleiche konsistent mit den restlichen Vergleichen sind, d. h. die durch die Vergleiche offenbarten Präferenzen dem Kriterium der Transitivität genügen. Die C. R. basiert auf $\lambda_{\max}$, dem maximalen Eigenwert der Paarvergleichsmatrix. Dieser Eigenwert ist immer gleich oder größer als n für positive, reziproke Matrizen und ist gleich n , wenn die Matrix konsistent ist. Durch Normalisierung dieses Maßes auf die Größe der Matrix leitet Saaty zunächst den sogenannten Konsistenzindex (C. I.) her:

$$C.I. = \frac{(\lambda_{\max} - 1)}{(n - 1)}.$$

Für jede Matrix A werden 500 Zufallsmatrizen generiert und ihr durchschnittlicher C. I. berechnet – diesen bezeichnet man als einen „Random Index (R. I.)“. Die o. g. Konsistenzratio (C. R.) gibt nun an, wie weit die Matrix, die sich aufgrund der in der Befragung erhobenen Werte ergibt, von einer völlig zufälligen Paarvergleichsmatrix abweicht. Die C. R. ist folgendermaßen definiert:

$$C.R. = \frac{C.I.}{R.I.}.$$

10 % Inkonsistenz mit der Zufallsmatrix werden von Saaty als vertretbar betrachtet. Dies impliziert, dass man bei einer $C.R. \leq 0,1$ davon ausgehen kann, dass die Bewertungen so durchgeführt wurden, dass sie nur zu einem geringen Grad in sich widersprüchlich sind.

Mit dem maximalen Eigenwert ($\lambda_{\max}$) der Paarvergleichsmatrix wird der gesuchte Eigenvektor $w = \{w_1, \dots, w_n\}$ berechnet. Die Elemente w_i des Eigenvektors entsprechen den Gewichten der Entscheidungskriterien:

$$w_i = \frac{\sum_{j=1}^n a_{ij} \cdot w_j}{\lambda_{\max}} \text{ für alle } i = 1, 2, \dots, n.$$

Diese Berechnung mithilfe des Eigenvektors kann als einfache Durchschnittsberechnung betrachtet werden, in der die generierten Gewichte ein Durchschnitt aller möglichen Vergleiche der Kriterien miteinander sind. Es gibt verschiedene Berechnungsmethoden, die letztlich zu dem Eigenvektor führen, der die Gewichte für die Kriterien spiegelt. Erläuterungen zu den unterschiedlichen Berechnungsmethoden finden sich in [26] sowie [27].

Ein Gruppendurchschnitt kann entweder über jeden Paarvergleich oder über jedes gewichtete Entscheidungskriterium berechnet werden. Wenn ein Durchschnitt für jeden Paarvergleich gebildet wird, können die Gewichte für jedes Entscheidungskriterium basierend auf diesen Paarvergleichsdurchschnitten berechnet werden. Da die paarweisen Vergleiche relative Bewertungen sind, sollte hier der geometrische Mittelwert berechnet werden. Zur Berechnung der arithmetischen Mittelwerte der Gewichte für die Entscheidungskriterien sollten zunächst die Gewichte für jedes Individuum basierend auf dessen Paarvergleichsergebnissen berechnet werden. Der arithmetische Mittelwert aller berechneten personenbezogenen Gewichte für die Entscheidungskriterien entspricht dann dem für die Gruppe berechneten mittleren Gewicht für dieses Entscheidungskriterium.

Zur Berechnung eines Durchschnittswerts, der die Präferenzen der ganzen Gruppe für einen Paarvergleich spiegelt, wird die Berechnung des geometrischen Mittels aller Paarvergleiche empfohlen [28]. Dieser Empfehlung wird auch in diesem Pilotprojekt gefolgt, da es das Ziel der Studie ist, Gruppenpräferenzen für bestimmte Kriterien zu ermitteln. Wenn folglich das AHP bei einer Gruppe von x Individuen angewendet wird, werden die Zellen der Matrix mit aggregierten Paarvergleichsergebnissen, d. h. geometrischen Mittelwerten der individuellen Paarvergleiche (a_{ij}), gefüllt:

$$a_{ij} = (a_{ij(1)} \cdot a_{ij(2)} \cdot \dots \cdot a_{ij(x)})^{\left(\frac{1}{x}\right)} \text{ für alle } i, j = 1, 2, \dots, n.$$

Diese Ergebnisse geben keinen Einblick über die Streuung der Gewichte, die aus den individuellen Paarvergleichen resultieren. Daher können in einer zusätzlichen Auswertung der arithmetische Mittelwert der Gewichte und die Standardabweichung berechnet werden. Hierdurch können Einblicke gewonnen werden, welche individuellen Abweichungen es bei der Einschätzung und Bewertung der verschiedenen Entscheidungskriterien gab. Der arithmetische Mittelwert der individuellen Gewichte, z. B. für das Kriterium 1, über x befragte Individuen berechnet sich wie folgt:

$$W_1 = \frac{(w_{1(1)} + w_{1(2)} + \dots + w_{1(x)})}{x}.$$

Dieser arithmetische Mittelwert gibt Aufschluss darüber, für welche Entscheidungskriterien sich die Bewertungen der Gruppe am stärksten unterscheiden.

2 Methoden

Die (Entscheidungs-)Kriterien, die in diesem AHP-Pilotprojekt bewertet werden sollten, waren patientenrelevante Endpunkte einer Antidepressiva-Behandlung, welche die verschiedenen Effekte und Nebenwirkungen der Behandlung auf die Patienten messen. Die Kriterien wurden auf Hauptkriterien und diesen untergeordneten Subkriterien verteilt. Die Hauptkriterien waren konzeptionell voneinander abgegrenzte Kategorien von übergeordneten Behandlungszielen. Die Subkriterien waren Endpunkte der Behandlung, die als solche in den Nutzenbewertungen des IQWiG und den hierin eingeschlossenen randomisierten kontrollierten klinischen Studien ausgewertet wurden.

2.1 Auswahl der Kriterien für die AHP-Entscheidungshierarchie

Zunächst wurden alle in den IQWiG-Nutzenbewertungen ausgewerteten Endpunkte extrahiert und zusammengefasst bzw. geordnet [14,15]. Dabei wurden die Kategorien Mortalität, Morbidität und Lebensqualität berücksichtigt, die auch vom SGB V vorgegeben werden. Sowohl Nutzen- als auch Schadensendpunkte mussten in eine Hierarchie gebracht werden. Damit aber die befragten Patienten die AHP-Entscheidungsstruktur nachvollziehen und verstehen konnten, wurde für das konkrete AHP-Verfahren die folgende Nomenklatur gewählt:

Wirksamkeit bzw. Nutzen oder Zusatznutzen gemessen als Veränderung in den Endpunkten:

- Ansprechen, Remission, Rückfall, Rezidiv, Depressionsskalengesamtscore

Vermeidung von Nebenwirkungen bzw. Schaden gemessen als Veränderung in den Endpunkten:

- unerwünschte Ereignisse (als unerwünschte Ereignisse wurden beispielsweise die sexuelle Dysfunktion und die Hypertonie auch separat ausgewertet),
- schwerwiegende unerwünschte Ereignisse (als schwerwiegende unerwünschte Ereignisse wurde ebenfalls separat ausgewertet: die Suizidalität, einschließlich Suizid und versuchter Suizid),
- Therapieabbruch aufgrund unerwünschter Ereignisse.

Auswirkungen auf die Lebensqualität bzw. auf relevante die Lebensqualität abbildende einzelne Dimensionen von Lebensqualitätskonzepten oder einzelnen Symptomen:

- gesundheitsbezogene Lebensqualität (gemessen anhand verschiedener Fragebögen),
- separate begleitende und lebensqualitätsrelevante Symptome der Erkrankung, z. B. soziale Funktionsfähigkeit, kognitive Funktionsfähigkeit, Schmerz, Angst.²

In die Überlegungen, mit welchen Kriterien und Subkriterien die Entscheidungsstruktur gefüllt werden sollte und auf welchen Hierarchieebenen diese angesiedelt werden sollten, flossen unterschiedliche Aspekte ein. Zum einen wurde darauf geachtet, dass ein möglichst umfassendes Spektrum an Endpunkten abgebildet werden sollte. Endpunkte der Behandlung wurden in die Entscheidungsstruktur aufgenommen, wenn in den Nutzenbewertungen in ausreichendem Maß unterstützende Evidenz bzw. Belege / Hinweise für einen Nutzen / Schaden bzw. Zusatznutzen / -schaden für mindestens eine der Prüfsubstanzen vorlagen. Des Weiteren, wie oben angedeutet, mussten die Endpunkte, die als Messergebnisse über Skalen vorliegen, in für Patienten erfahrbare Dimensionen oder Aspekte von Lebensqualität überführt werden. Zusätzlich wurden Endpunkte aufgenommen, wenn aus der Hintergrundrecherche zu diesem Pilotprojekt hervorging, dass einzelne Endpunkte von besonderer Bedeutung für Patienten sind – auch wenn in den Nutzenbewertungen zu diesen Endpunkten nur unzureichende Evidenz bzw. keine Belege / Hinweise vorlagen. Gleichzeitig wurde auch der Aspekt, die Anzahl der Fragen nicht zu groß werden zu lassen, bei der Auswahl der Kriterien einbezogen, um der Gefahr vorzubeugen, dass die befragten Patienten und Patientinnen dann nur noch unüberlegt antworten würden.

Auswahl von Endpunkten, für die signifikante Ergebnisse vorlagen

Für folgende Endpunkte, die in die Entscheidungsstruktur aufgenommen wurden, wurde in den Nutzenbewertungen für mindestens eine der Prüfsubstanzen ein Beleg / Hinweis für einen (Zusatz-)Nutzen / -schaden nachgewiesen: Ansprechen, Remission, Rückfall, soziale Funktionsfähigkeit, unerwünschte Ereignisse, schwerwiegende unerwünschte Ereignisse und Therapieabbruch aufgrund unerwünschter Ereignisse, Angst.

Auswahl von weiteren Endpunkten, für die keine signifikanten Ergebnisse vorlagen

Weitere Endpunkte, die in die Entscheidungshierarchie aufgenommen wurden, da von ihnen vermutet wurde, dass sie von großer Relevanz für Patienten wären, obwohl aus den Nutzenbewertungen für diese Endpunkte nur unzureichende Evidenz bzw. keine Belege / Hinweise vorlagen waren: Schmerz, kognitive Funktionsfähigkeit, Suizidalität – einschließlich (versuchten) Suizids, sexueller Funktionsfähigkeit.

² In den Abbildungen und Tabellen in diesem Arbeitspapier werden Kurzbezeichnungen verwendet für die Kriterien, z. B. wird das Kriterium, das in der Vorabinformation für Patienten und Experten als „Verringerung von Angst“ erläutert wird, kurz mit Angst bezeichnet.

Schmerz und eine eingeschränkte kognitive Funktionsfähigkeit können die Lebensqualität depressiver Patienten stark einschränken und wurden daher als separate Endpunkte aufgenommen.

Die Suizidalität kann zum einen Symptom der depressiven Erkrankung sein. Sie kann jedoch auch durch die Einnahme von Antidepressiva ausgelöst bzw. verstärkt werden. Während für die Zielgruppe der Kinder und Jugendlichen Nachweise für eine erhöhte Suizidalität unter der Einnahme bestimmter Antidepressiva (insbesondere SSRI) vorliegen [29,30], ist die Datenlage für erwachsene Patienten weniger eindeutig [31]. In den Nutzenbewertungen des IQWiG wurde zur Bewertung des Therapieziels der Verhinderung bzw. Verringerung des Auftretens von Suizidalität die Operationalisierung über schwerwiegende unerwünschte Ereignisse bzw. Todesfälle gewählt [14,15]. Hierbei wurden diejenigen schwerwiegenden unerwünschten Ereignisse berücksichtigt, die anhand ihres Wortlauts Suizidalität wahrscheinlich erscheinen ließen. Die Anzahl der tatsächlichen Suizide wurde im Rahmen dieser Operationalisierung gemessen. Unter dieser Definition und den entsprechenden Ergebnissen der Nutzenbewertungen zeigte sich kein Hinweis oder Beleg für die Über- oder Unterlegenheit einer der Prüfsubstanzen bei jedoch unzureichender Datenlage. Dies galt sowohl für die Vergleiche gegenüber Placebo in der Kurz- und Langzeittherapie als auch für die Vergleiche zu aktiven Komparatorwirkstoffklassen und -substanzen. Aufgrund der unsicheren Datenlage empfehlen Behandlungsleitlinien eine engmaschige Beobachtung der Patienten auf eine möglicherweise auftretende / sich verstärkende Suizidalität in den ersten Behandlungswochen [32]. Da die Suizidalität trotz oder gerade wegen der unsicheren bzw. unzureichenden Datenlage jedoch ein höchst relevanter Behandlungsendpunkt ist und sein muss, wurde dieser ebenfalls in die Entscheidungshierarchie aufgenommen.

Obwohl in den Nutzenbewertungen des IQWiG keine entsprechenden Belege / Hinweise gefunden werden konnten, wird behauptet, dass Antidepressiva unterschiedliche Auswirkungen auf die sexuelle Funktionsfähigkeit depressiver Patienten haben. Dieser Endpunkt wurde daher in den vergangenen Jahren zunehmend in die entsprechenden klinischen Studien aufgenommen. Aufgrund einer möglicherweise starken Patientenrelevanz wurde daher auch dieser Endpunkt in die Entscheidungshierarchie aufgenommen.

2.2 Strukturierung der Endpunkte

Die Entscheidungsstruktur wurde in die zwei Ebenen der Haupt- und der Subkriterien unterteilt. Als Hauptkriterien wurden folgende konzeptionell unterschiedlichen übergeordneten Behandlungsziele in patientengerechter Sprache definiert:

- Wirksamkeit der Behandlung, die insbesondere den Nutzen oder Zusatznutzen gemessen als Veränderung in den Endpunkten Ansprechen, Remission, Rückfall, Rezidiv oder Depressionsgesamtscore umfasst,
- Vermeidung von Nebenwirkungen der Behandlung und

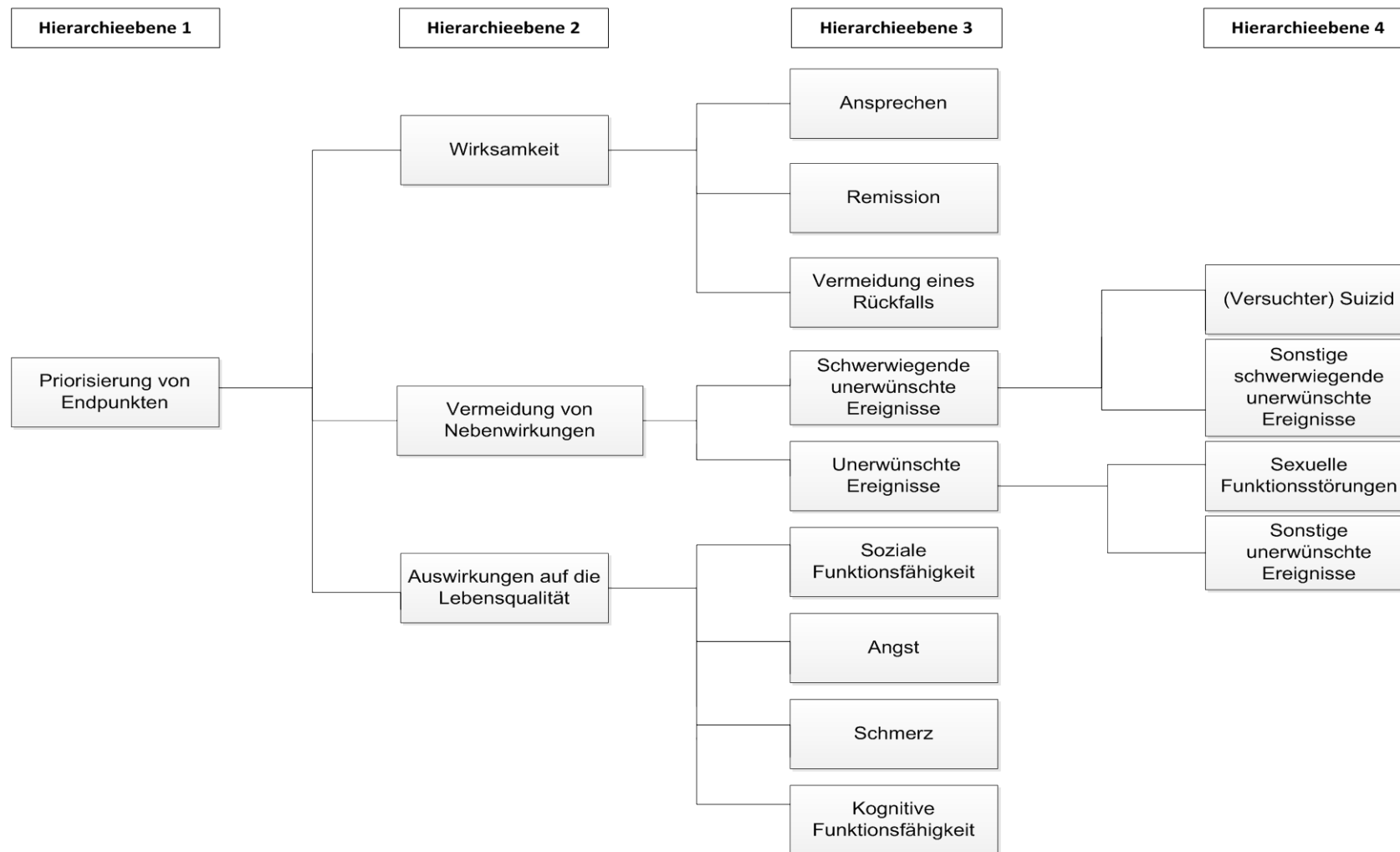
- Auswirkungen auf die Lebensqualität bzw. auf relevante die Lebensqualität abbildende einzelne Dimensionen von Lebensqualitätskonzepten oder einzelne Symptome.

Die positiven Effekte einer Behandlung als Nutzen bzw. Zusatznutzen wurden somit in den zwei Hauptkriterien „Wirksamkeit“ und „Auswirkungen auf die Lebensqualität“, die negativen Effekte einer Behandlung in dem Hauptkriterium „Vermeidung von Nebenwirkungen der Behandlung“ als Schaden bzw. Zusatzschaden zusammengefasst.

Bei der „Wirksamkeit“ steht die Messung der Auswirkungen einer Behandlung auf die depressive Symptomatik, gemessen anhand einer validierten Depressionsskala (HAM-D und MADRS), im Vordergrund. Die Auswertung der Depressionsskalengesamtscores ermöglicht wie folgt eine Operationalisierung der verschiedenen untergeordneten Endpunkte bzw. Subkriterien:

- Ansprechen $\Rightarrow$ Reduktion der depressiven Symptomatik um 50 % gemessen mit dem Gesamtscore einer Depressionsskala,
- Remission $\Rightarrow$ Unterschreitung eines Schwellenwertes auf der Depressionsskala (entsprechend keiner oder einer nur minimalen depressiven Symptomatik),
- Rückfall $\Rightarrow$ erneutes Überschreiten eines Schwellenwertes auf der Depressionsskala, Wiederauftreten der vollen depressiven Symptomatik nach vorheriger Remission bzw. vorherigem Ansprechen.

Bei den „Auswirkungen auf die Lebensqualität“ stehen die Auswirkungen einer Behandlung auf die Erkrankungssymptome, die nicht primär die Depressionserkrankung charakterisieren, diese jedoch begleiten, im Vordergrund. Hierzu zählen in der Entscheidungsstruktur die Endpunkte bzw. Subkriterien soziale Funktionsfähigkeit, Angst, Schmerz und kognitive Funktionsfähigkeit, die in den Nutzenbewertungen bzw. den diesen zugrunde liegenden klinischen Studien ebenfalls über entsprechende Fragebögen abgefragt und operationalisiert wurden. Abbildung 2 zeigt die resultierende primäre Entscheidungsstruktur.



Im Pilotprojekt wurde die Bezeichnung „Auswirkungen auf die Lebensqualität“ verwendet. Wie oben auf S. 10 dieses Dokuments erläutert, sind damit „Auswirkungen auf die Lebensqualität bzw. auf relevante die Lebensqualität abbildende einzelne Dimensionen von Lebensqualitätskonzepten oder einzelne Symptome“ gemeint.

Abbildung 2: AHP-Entscheidungsstruktur

In einer zweiten alternativen Entscheidungsstruktur wurden nur die drei Hauptkriterien Ansprechen, Remission und Therapieabbruch aufgrund unerwünschter Ereignisse einander gegenübergestellt. Diese Entscheidungsstruktur wurde als Alternativszenario gewählt, da gemäß den Nutzenbewertungen des IQWiG für diese Endpunkte jeweils für eine der Prüfsubstanzen ein Beleg für einen Zusatznutzen / -schaden vorlag und vermutet wurde, dass gerade diese drei Endpunkte eine sehr große Bedeutung für die Patienten haben könnten.

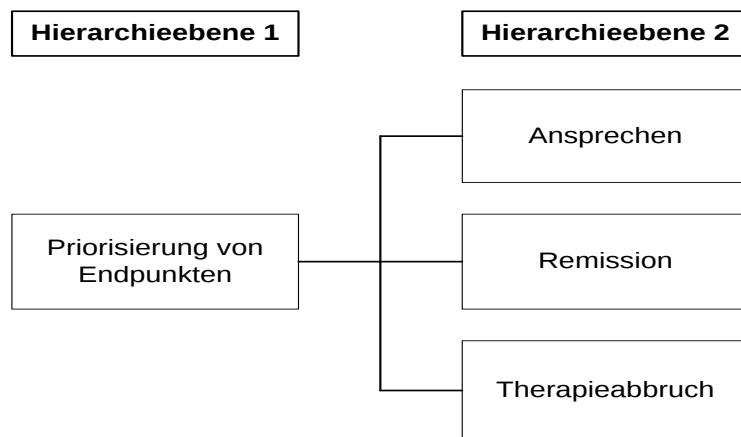


Abbildung 3: Alternative AHP-Entscheidungsstruktur

2.2.1 Operationalisierung der Behandlungsziele / Endpunkte

Die Endpunkte wurden folgendermaßen operationalisiert:

- 1) Wirksamkeit: erwünschte Wirkungen der Antidepressiva auf die Symptome der Depression einschließlich depressiver Verstimmung, Freude- / Interesseverlust, Schlaflosigkeit / übertriebene Schläfrigkeit, Energieverlust, Appetit- / Gewichtsverlust, vermindertes Selbstwertgefühl / Schuldgefühle, psychomotorische Agitiertheit / Hemmung, Konzentrationsstörungen und wiederkehrende Gedanken an den Tod.
 - a) Ansprechen: Halbierung von Anzahl / Ausmaß der depressiven Symptome, d. h. 50 %ige Verbesserung der auf einer Depressionsskala (e. g. HAM-D, MADRS) gemessenen Scores.
 - b) Remission: Reduktion der depressiven Symptome dergestalt, dass ein Patient nicht mehr die Kriterien einer depressiven Episode auf einer Depressionsskala erfüllt
 - c) Rückfall: Die depressiven Symptome müssen einen Schwellenwert auf einer Depressionsskala nach voller Remission überschreiten, sodass von einem Rückfall gesprochen werden muss.
- 2) Vermeidung von Nebenwirkungen mit der Unterteilung der Nebenwirkungen in schwerwiegende unerwünschte Nebenwirkungen und nicht schwerwiegende unerwünschte Nebenwirkungen. Schwerwiegende unerwünschte Nebenwirkungen sind lebensbedrohlich, führen zu schweren Behinderungen oder machen eine Hospitalisierung

erforderlich.³ Nicht schwerwiegende unerwünschte Nebenwirkungen sind nicht lebensbedrohlich und führen nicht zur Hospitalisierung.

- a) (Versuchter) Suizid: Hier ist die Suizidalität gemeint, die durch Antidepressiva ausgelöst wird, und nicht eine möglicherweise erhöhte krankheitsbedingte Suizidalität.
- b) Sonstige schwerwiegende Nebenwirkungen: Jedes unerwünschte Ereignis kann zu einem schwerwiegenden unerwünschten Ereignis werden, wenn es zum Tode führt, lebensbedrohlich wird, zu schweren Behinderungen oder zur Hospitalisierung führt.
- c) Sexuelle Funktionsstörungen: Verlust des Interesses am Geschlechtsverkehr und / oder Verlust / Einschränkungen der sexuellen Funktionsfähigkeit aufgrund der Behandlung mit Antidepressiva.
- d) Sonstige unerwünschte Nebenwirkungen: Beispiele hierfür sind Bluthochdruck, Unruhe, Schwindel, Übelkeit, trockener Mund, übermäßiges Schwitzen oder Gewichtsanstieg.

3) Auswirkungen auf die Lebensqualität:

- a) Soziale Funktionsfähigkeit: heißt, am sozialen Leben auf der Arbeit, in der Schule, zu Hause, in der Freizeit etc. teilnehmen zu können.
- b) Angst: Gefühle von Angst oder Anspannung.
- c) Schmerz: körperliche Schmerzen wie Kopf- oder Rückenschmerzen.
- d) Kognitive Funktionsfähigkeit: Fähigkeit, sich zu konzentrieren, logisch zu denken und alltägliche intellektuelle Aufgaben zu lösen.

4) Weiterer Endpunkt

- a) Therapieabbruch aufgrund unerwünschter Ereignisse: vorzeitiger Therapieabbruch aufgrund von Nebenwirkungen der Medikamente.

2.3 Ablauf des vorliegenden AHP-Verfahrens

Für die AHP-Befragungen in diesem Pilotprojekt wurden zum einen Patienten, zum anderen Experten rekrutiert, die Patienten behandeln, die an einer Depression leiden. Die befragten Patienten sollten Patienten sein, die bereits einmal eine depressive Episode erlebt haben und Erfahrungen mit der Einnahme von Antidepressiva gemacht haben. Die befragten Experten sollten eine ausgewogene Mischung von beruflichen Qualifikationen, Erfahrungen und Meinungen mitbringen. Darüber hinaus musste beachtet werden, dass die Gruppe nicht zu groß werden durfte, d. h. die Zahl von 10 bis 15 Teilnehmern sollte nicht überschritten werden. Bei größeren Gruppen (d. h. schon ab mehr als 8 Personen) nimmt die Bereitschaft

³ Die hier genannten Kriterien der schwerwiegenden unerwünschten Nebenwirkungen sind ein Auszug der in § 4 Abs. 13 AMG genannten Nebenwirkungen. Diese Auswahl wurde den Patienten vorgetragen.

der Teilnehmer ab, sich aktiv in die Gruppe einzubringen und eigene Erfahrungen oder Meinungen offenzulegen.

2.3.1 Rekrutierung von Patienten

Die Patienten wurden über die „Stabsstelle Patientenbeteiligung“ im Gemeinsamen Bundesausschuss (G-BA) und über diese in enger Zusammenarbeit mit dem „Bundesverband der Angehörigen psychisch Kranker“ (BApK) rekrutiert. Der BApK wiederum sprach 20 Selbsthilfegruppen und Organisationen an, die ihre Mitglieder bzw. interessierte Patienten zur Teilnahme an der IQWiG-AHP-Befragung aufriefen. Zusätzlich initiierte die BApK einen Aufruf zur Teilnahme an den Befragungen auf der Webseite des „Kompetenznetzes Depression, Suizidalität“.

2.3.2 Rekrutierung von Experten

Um sicherzustellen, dass die Experten eine ausgewogene Mischung von beruflichen Qualifikationen, Erfahrungen und Meinungen mitbringen, wurden sie über verschiedene Wege gesucht. Zunächst wurde die aktuelle deutsche Leitlinie zur Diagnose und Behandlung der Depression, die NVL, herangezogen, um deutsche Experten, die an der Entwicklung dieser Leitlinie beteiligt waren, zu identifizieren ([32,33]).

Zum Zweiten wurden die folgenden Webseiten großer deutscher wissenschaftlicher Institute und Organisationen, die sich mit der Depression befassen, auf der Suche nach den Namen weiterer Experten in der Behandlung depressiver Patienten durchsucht:

- Deutsche Gesellschaft für Psychiatrie, Psychotherapie und Nervenheilkunde (DGPPN)
- Kompetenznetz Depression, Suizidalität
- Deutsche Gesellschaft für Psychologie (DGP)
- Bundespsychotherapeutenkammer (BPtK)
- Deutsche Gesellschaft für Psychosomatische Medizin und Ärztliche Psychotherapie (DGPM)
- Deutsche Gesellschaft für Allgemeinmedizin und Familienmedizin (DEGAM)

Zuletzt wurden niedergelassene oder in Krankenhäusern angestellte spezialisierte Allgemeinmediziner, Psychiater, Psychologen und Psychotherapeuten, die im Großraum Köln / Aachen / Bonn / Düsseldorf praktizieren, über die folgenden Internetseiten identifiziert:

- Kassenärztliche Vereinigung Nordrhein („Suche nach Psychiatern und Psychotherapeuten“, <http://www.kvno.de/buerger/arztsuche/arztverz.html>)
- Psychotherapie-Informationsdienst (<http://www.psychotherapiesuche.de>), ebenfalls zugänglich über das Kompetenznetz Depression, Suizidalität (<http://www.kompetenznetz-depression.de/>)

- Gelbe Seiten
- Landschaftsverband Rheinland (LVR) zur Identifizierung von Krankenhausärzten, die in LVR-Kliniken angestellt sind

2.4 Vorabinformationen für die Teilnehmer

Im Vorfeld der AHP-Befragungen erhielten alle Teilnehmenden eine Beschreibung der Ziele der AHP-Befragung sowie eine Definition der Behandlungsziele und -endpunkte, die in den Befragungen bewertet werden sollten (Anhang A). Weder Patienten noch Experten erhielten darüber hinaus weitergehende Informationen über die Ergebnisse der Nutzenbewertungen des IQWiG zum Nutzen und Schaden der verschiedenen Antidepressiva. Das heißt, die Patienten und Experten gingen nur mit den ihnen bereits im Vorfeld bekannten Informationen und eigenen Erfahrungen mit Antidepressiva in die Befragungen und wurden ansonsten „informationsnaiv“ belassen. Der Hauptgrund für dieses Vorgehen war, dass man die Teilnehmer nicht durch die Ergebnisse der Nutzenbewertungen in der Bewertung der Bedeutung, die sie bestimmten Endpunkten beimessen würden, beeinflussen wollte. Außerdem sollten nicht zielführende Diskussionen über die Ergebnisse oder Methoden der Nutzenbewertungen des IQWiG vermieden werden, um nicht vom eigentlichen Ziel der AHP-Befragungen abzulenken, nämlich eine Priorisierung von Zielen und Endpunkten der Behandlung mit Antidepressiva vorzunehmen.

2.5 Ablauf der AHP-Befragungen

2.5.1 Einführung

In beiden AHP-Befragungen (Patienten und Experten) saßen die Teilnehmer an u-förmig zusammengestellten Tischen, der Moderator stand vor der Gruppe. Der Systemadministrator saß mit dem AHP-Auswertungssystem neben dem Moderator. Die Hardware des AHP-Systems bestand aus einem Laptop, auf dem Team Expert Choice installiert war, einem Projektor, einem Receiver und individuellen Bedien- bzw. Abstimmungssets für die Befragungsteilnehmer.

Vor Beginn der Befragungen wurden das Ziel und der Ablauf der AHP-Befragungen, die Definitionen der zu bewertenden Behandlungsziele und -endpunkte (die die Teilnehmer im Vorfeld bereits schriftlich erhalten hatten) und die verschiedenen Entscheidungsstrukturen, die ausgewertet werden sollten, vorgestellt und besprochen. Ein gutes und unter den Teilnehmern gleiches Verständnis der zu bewertenden Behandlungsziele und -endpunkte ist im AHP-Verfahren von großer Bedeutung (siehe Abschnitt 1.4). Vor Beginn der eigentlichen Bewertung von Paarvergleichen wurde in beiden Gruppen darüber hinaus eine informelle Vorstellungsrunde der Teilnehmer durchgeführt. Es wurden aber weder soziodemografische noch andere Informationen über die Teilnehmenden systematisch erfasst.

2.5.2 Leitung der Befragungen

Die Befragungen wurden separat in jeweils einer Gruppe von Patienten und einer Gruppe von Experten durchgeführt. Das Gruppensetting ermöglichte es, durch Erfahrungsaustausch und Diskussionen zwischen den einzelnen Paarvergleichsbewertungen Informationen über die Hinter- und Beweggründe der befragten Personen für deren Priorisierungsentscheidungen zu erlangen. Patienten und Experten wurden separat befragt, um sicherzustellen, dass zum einen Patienten offen und unbefangen ihre Erfahrungen und Ansichten mitteilen und ein aktiver Austausch zwischen den Betroffenen stattfinden kann. Zum anderen sollte auch den Experten ausreichend Gelegenheit zum fachlichen Informations- und Erfahrungsaustausch gegeben werden.

Ein Gruppenmoderator leitete die zwischen den einzelnen Paarvergleichsbewertungen stattfindenden Diskussionen in der Gruppe. Ein Systemverwalter bediente die zur Befragung genutzte AHP-Software. Der Moderator sollte die Teilnehmer zur Teilnahme an den Diskussionen ermutigen und diese moderieren, sich jedoch nicht mit persönlichen Ansichten in die Diskussionen einbringen. Der Systemverwalter sollte in Kenntnis der Befragungssoftware die Teilnehmer in Bezug auf die technischen Aspekte durch die Befragung führen.

2.5.3 Bewertung und Diskussion der Paarvergleiche in den AHP-Befragungen

Schrittweise wurden in den AHP-Befragungen immer zunächst auf der untergeordneten Hierarchieebene die Endpunkte in paarweisen Vergleichen miteinander verglichen, gefolgt von den Hauptkriterien auf der übergeordneten Hierarchieebene. Die Berechnung des rechten Eigenvektors, mit dessen Hilfe Gruppengewichte für jedes der in einer Hierarchieebene eingeschlossenen Entscheidungskriterien generiert werden, erfolgte mithilfe des Programms „Expert Choice 2011“.

Die Paarvergleiche wurden wie folgt durchgeführt: Dem wichtigeren Ziel / Endpunkt zweier Ziele / Endpunkte wurde ein Wert von > 1 bis ≤ 9 zugeteilt. Wenn beide Endpunkte einer Person gleich wichtig erschienen, wurde die 1 vergeben, die für Indifferenz steht. Die 9 bedeutet eine extrem größere Bedeutung des einen im Vergleich zu dem anderen Endpunkt. Die Befragten gaben für jeden paarweisen Vergleich zweier Ziele / Endpunkte, die jeweils separat an die Wand projiziert wurden, mithilfe ihres persönlichen Abstimmungssets ihre Bewertung ab. Die Gesamtheit aller Einzelbewertungen wurde dann ebenfalls projiziert, sodass jeder Teilnehmer die Bewertungen aller anderen Teilnehmer sehen konnte. In der Patientenbefragung blieben diese Einzelbewertungen anonym, um sicherzustellen, dass Patienten offen und unbefangen ihre Bewertung abgeben und die Ergebnisse anschließend entsprechend in der Gruppe diskutieren konnten. Bei den Experten blieben die Bewertungen nicht anonym, um eine zielgerichtete und expertenbezogene Diskussion einzelner Bewertungen zu ermöglichen. Wenn beispielsweise ein Experte in seiner Bewertung stark von der Bewertung der anderen abwich, konnte dieser gezielt von den anderen Teilnehmern befragt werden.

Der Moderator ermutigte die Befragungsteilnehmer zum einen, ihre Erfahrungen und Motivationen mit den anderen Teilnehmern zu teilen, betonte jedoch auch, dass voneinander abweichende Bewertungen akzeptiert werden sollten. Nach den Gruppendiskussionen, die sich an jede Bewertungsrunde eines Paarvergleichs anschlossen, konnte jeder Befragungsteilnehmer seine Bewertung noch einmal verändern. Für jeden Paarvergleich wurden die individuellen Bewertungen der Gruppe sofort im Anschluss zu einem geometrischen Gruppenmittelwert zusammengefasst. Nach Beendigung jeder Bewertungsrunde auf einer Hierarchieebene wurde die Konsistenzratio für die abgegebenen Bewertungen berechnet. Im Falle einer C. R. $> 0,1$ sollten die Teilnehmer ihre Bewertung noch einmal überprüfen. Die finalen Gewichte für die jeweiligen Behandlungsziele / -endpunkte wurden aus diesen Gruppenmittelwerten berechnet. Die Ergebnisse für jede Hierarchieebene der Entscheidungsstruktur wurden den Teilnehmern gezeigt und dann ebenfalls in der Gruppe diskutiert.

3 Ergebnisse

3.1 Teilnehmende Patienten

Mithilfe der BApK wurden schätzungsweise 300 bis 500 Patienten erreicht – entweder über die Ankündigung der AHP-Befragung auf den verschiedenen Webseiten oder über die direkte Kontaktierung von Patienten per E-Mail. 14 Patienten erklärten sich per E-Mail zur Teilnahme an der IQWiG-Befragung bereit, von denen letztlich 12 Patienten tatsächlich teilnahmen – 9 Frauen und 3 Männer. Da die Hürden für die befragten Patient(inn)en möglichst niedrig sein sollten, wurden, wie in den Methoden beschrieben, keine weiteren soziodemografischen Daten erhoben. Die teilnehmenden externen und internen an diesem Projekt Beteiligten schätzten die befragten Patienten nach Alter zwischen 30 und 70 Jahren, mehrheitlich zwischen 40 und 50. Aus der informellen Vorstellungsrunde wurde der Eindruck gewonnen, dass ein Großteil der Patienten einen überdurchschnittlichen Ausbildungsstand hatte. Außerdem schienen die Patienten durchgehend mindestens eine mittelschwere bis schwere depressive Episode durchlitten zu haben und befanden sich zum Zeitpunkt der Befragung in einer Phase der Remission oder Heilung.

3.2 Teilnehmende Experten

108 in der Behandlung der Depression tätige Experten wurden per E-Mail eingeladen, an der IQWiG-Befragung teilzunehmen. Hiervon waren ca. 60 % spezialisierte Allgemeinärzte, Psychiater, Psychologen und Psychotherapeuten, die im Raum Köln / Aachen / Bonn / Düsseldorf entweder als niedergelassene Ärzte oder in Krankenhäusern praktizierten. Die anderen 40 % waren Experten, die über die Entwicklergruppe der NVL-Leitlinie oder über andere wissenschaftliche Fachgesellschaften identifiziert worden waren (wie in Abschnitt 2.3.2 beschrieben). Die Responserate auf die Befragungseinladung (d. h. Personen, die überhaupt auf die E-Mail reagierten) lag bei ca. 28 % in der ersten Gruppe (niedergelassene oder im Krankenhaus praktizierende Ärzte / Psychiater, Psychologen und Psychotherapeuten) und bei ca. 48 % in der zweiten Gruppe (über die NVL oder über Fachgesellschaften identifizierte Experten).

Trotz der recht hohen Responserate waren letztlich nur 7 Experten bereit, an der Befragung teilzunehmen.

In den Antworten auf die Befragungseinladung wurde die Nichtteilnahme im Wesentlichen mit 2 Aspekten begründet:

- keine Bezahlung, nur Erstattung der Reisekosten durch das IQWiG, und
- terminliche Probleme (da ein fixer Befragungstermin vorgegeben war, mit fester Datums- und Zeitvorgabe).

3.3 Ergebnisse der Befragung der Patienten

Die von den 12 Patienten vergebenen Prioritäten bzw. Gewichte für die Behandlungsziele und -endpunkte sind in Tabelle 2 dargestellt, ausgehend von den Ergebnissen der Befragung zu den auf der untersten Ebene der Entscheidungshierarchie dargestellten Endpunkten. Die Konsistenzratio für die AHP-Patientenbefragung lag bei 0,04.

Tabelle 2: Gruppengewichte (geometrisches Mittel) und Werte der Konsistenzratio aus der Patientenbefragung

Behandlungsziele / Endpunkte der Behandlung	Gruppengewicht nach Hierarchieebene				C. R.
	1	2	3	4	
Summe Gruppengewicht	1,000				0,04
Wirksamkeit	0,500				0,10
Ansprechen	0,324				
Remission	0,085				
Vermeidung eines Rückfalls	0,091				
Vermeidung von Nebenwirkungen	0,095				0,00
Vermeidung unerwünschter Ereignisse	0,030				0,00
sexuelle Funktionsstörungen	0,007				
sonstige unerwünschte Ereignisse	0,023				
Vermeidung schwerwiegender unerwünschter Ereignisse	0,065				0,00
(versuchter) Suizid	0,026				
sonstige schwerwiegende unerwünschte Ereignisse	0,039				
Auswirkungen auf die Lebensqualität	0,405				0,02
Verbesserung der sozialen Funktionsfähigkeit	0,107				
Verringerung von Angst	0,118				
Verringerung von Schmerz	0,054				
Verbesserung der kognitiven Funktionsfähigkeit	0,125				
C. R.: Konsistenzratio					

3.4 Ergebnisse der Befragung der Experten

An der Expertenbefragung nahmen 7 Experten teil. Die Gruppengewichte für die Behandlungsziele / -endpunkte sind in Tabelle 3 zusammengefasst. Die Konsistenz lag bei einer C. R. von 0,05.

Tabelle 3: Gruppengewichte (geometrisches Mittel) und Werte der Konsistenzratio aus der Expertenbefragung

Behandlungsziele / Endpunkte der Behandlung	Gruppengewicht nach Hierarchieebene				C. R.
	1	2	3	4	
Summe Gruppengewicht	1,000				0,05
Wirksamkeit	0,681				0,03
Ansprechen	0,061				
Remission	0,475				
Vermeidung eines Rückfalls	0,144				
Vermeidung von Nebenwirkungen	0,080				0,00
Vermeidung unerwünschter Ereignisse	0,037				0,00
sexuelle Funktionsstörungen	0,007				
sonstige unerwünschte Ereignisse	0,029				
Vermeidung schwerwiegender unerwünschter Ereignisse	0,043				0,00
(versuchter) Suizid	0,022				
sonstige schwerwiegende unerwünschte Ereignisse	0,020				
Auswirkungen auf die Lebensqualität	0,240				0,08
Verbesserung der sozialen Funktionsfähigkeit	0,090				
Verringerung von Angst	0,054				
Verringerung von Schmerz	0,033				
Verbesserung der kognitiven Funktionsfähigkeit	0,062				
C. R.: Konsistenzratio					

3.5 Unterschiede zwischen den Ergebnissen der Patienten- und der Expertenbefragung

Abbildung 4 stellt die Ergebnisse der Patienten- und der Expertenbefragungen in grafischer Form einander gegenüber. Insgesamt wird aus dieser Darstellung ersichtlich, dass Patienten und Experten sich primär in ihren Einschätzungen der Endpunkte zur Wirksamkeit unterscheiden, d. h. für das Ansprechen, die Remission und die Vermeidung von Rückfällen. Hier ist für die Patienten das Ansprechen der wichtigste Endpunkt, wohingegen die Experten mehr Gewicht insbesondere auf die Remission, aber auch auf den Rückfall verlagern. Den Nebenwirkungen messen sowohl Patienten als auch Experten insgesamt und auch für die einzelnen Endpunkte eine relativ geringe Bedeutung bei. Den Auswirkungen auf die Lebensqualität messen sowohl Patienten als auch Experten eine mittlere relative Bedeutung bei, wobei die Patienten insgesamt diese Endpunkte leicht höher gewichten als die Experten.

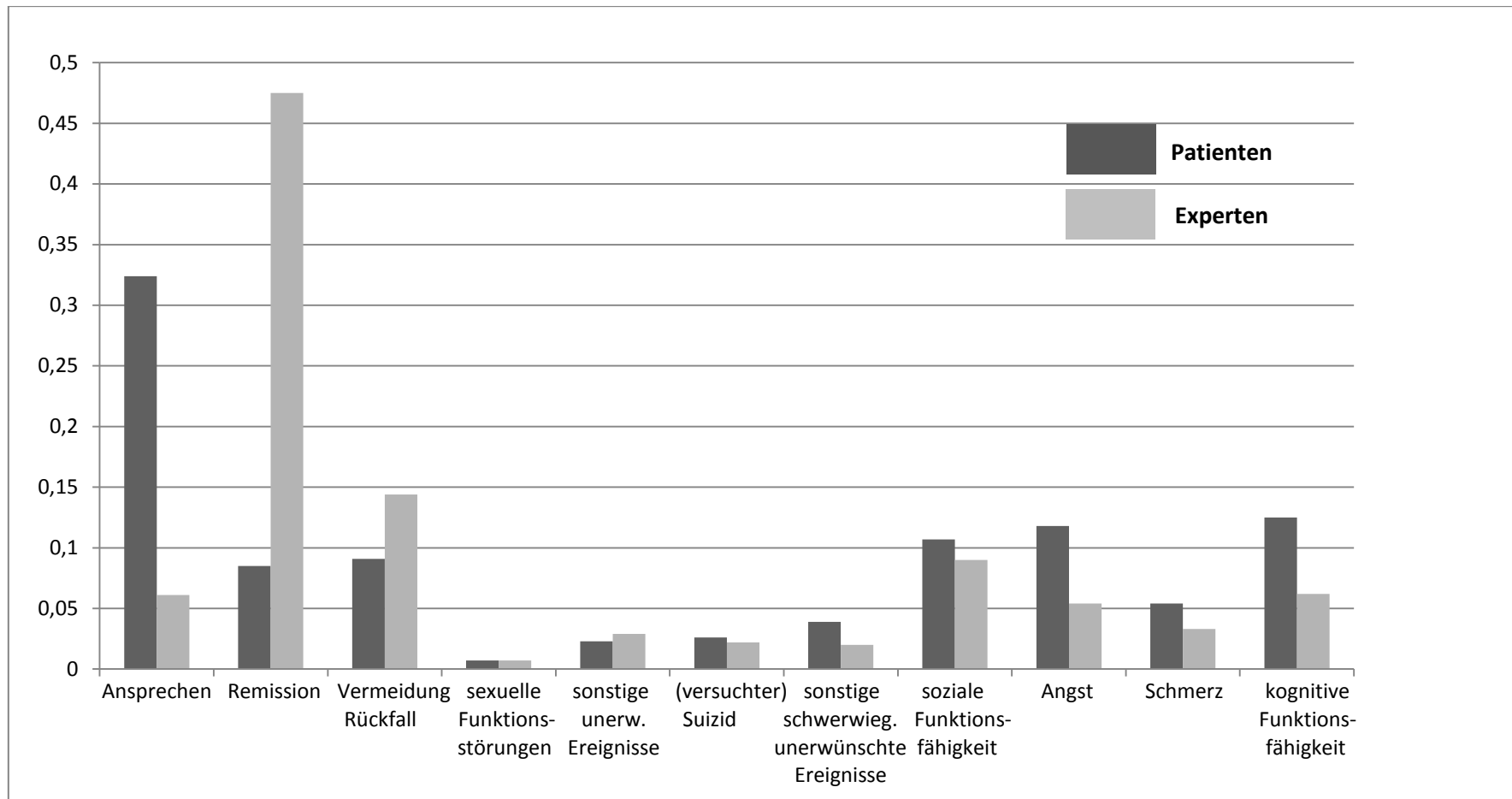


Abbildung 4: Gegenüberstellung der aus der Patienten- und der Expertenbefragung resultierenden Prioritäten / Gewichte

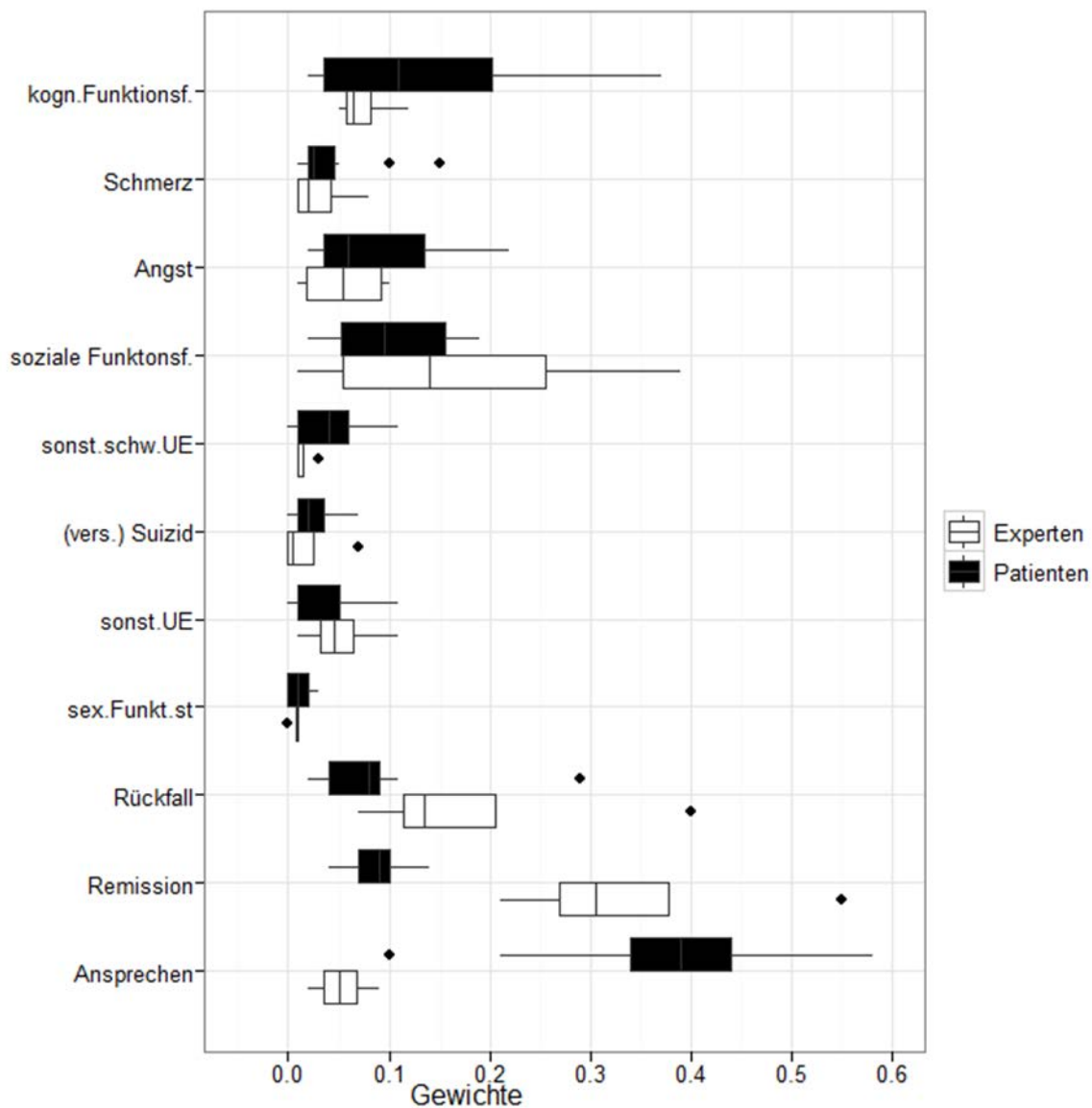


Abbildung 5: Variation der Einschätzungen / Prioritäten innerhalb der Gruppen

Die Kästen zeigen das untere und obere Quartil. Als Linie sichtbar in den weißen Kästen ist das arithmetische Mittel. Die sog. Whiskers geben die Standardabweichung an. Ausreißer sind mit Punkten gekennzeichnet.

In der Gruppe der Patienten und der Experten variierten die Einschätzungen jeweils nur relativ wenig bei der Bedeutung, die die Individuen den verschiedenen unerwünschten Wirkungen einer Behandlung beimessen. Für die Endpunkte, die der Wirksamkeit einer Behandlung auf die Depressionssymptomatik und der Verringerung sonstiger unerwünschter Krankheitssymptome untergeordnet waren, variierten die Einschätzungen stärker. Bei den Patienten variierten die Einschätzungen im Kriterium Auswirkungen auf die Lebensqualität für die Subkriterien Angst und kognitive Funktionsfähigkeit am stärksten, jedoch auch für die

Endpunkte Schmerz und soziale Funktionsfähigkeit. Aus den Diskussionen zwischen den Patienten bei den einzelnen Paarvergleichen ging hervor, dass es bei den Patienten offensichtlich eine bimodale Verteilung für Personen gab, die die Angst bzw. den Schmerz jeweils als dominantes Erkrankungssymptom wahrnahmen. Bei den Experten hingegen variierten bei diesen Endpunkten die Priorisierungen am stärksten bei dem Subkriterium soziale Funktionsfähigkeit. Bei den Patienten variierten die Einschätzungen zu den Endpunkten der Wirksamkeit am stärksten beim Endpunkt Ansprechen, bei den Experten respektive beim Endpunkt Remission.

3.6 Übersicht zum Vergleich der Befragungen

Tabelle 4 zeigt die Rangfolge für die verschiedenen Endpunkte, die sich aus den Gruppengewichten jeweils für die Gruppe der Patienten und der Experten ableiten lassen.

Tabelle 4: Priorisierung der Endpunkte in den Gruppen gemäß Tabellen 2 und 3

Behandlungsendpunkte	Rang Patientengewichtung	Rang Expertengewichtung
Remission	6	1
Ansprechen	1	5
Vermeidung eines Rückfalls	5	2
Verbesserung der sozialen Funktionsfähigkeit	4	3
Verbesserung der kognitiven Funktionsfähigkeit	2	4
Verringerung von Angst	3	6
Verringerung von Schmerz	7	7
Vermeidung sonstiger schwerwiegender unerwünschter Ereignisse	8	10
Vermeidung sonstiger unerwünschter Ereignisse	10	8
Vermeidung von Suiziden	9	9
Vermeidung sexueller Funktionsstörungen	11	11

Sowohl in der Gruppe der Patienten als auch in der Gruppe der Experten wurden dieselben 6 Endpunkte als die wichtigsten Behandlungsendpunkte angesehen. Dies sind zum einen die Endpunkte zum Behandlungsziel Wirksamkeit, d. h. Ansprechen, Remission und Vermeidung eines Rückfalls, zum anderen die folgenden Endpunkte zum Behandlungsziel Auswirkungen auf die Lebensqualität: soziale Funktionsfähigkeit, kognitive Funktionsfähigkeit und Angst. Diese sechs Endpunkte decken in der Gruppe der Patienten 85 % der Gesamtgewichte ab. Nebenwirkungen spielten eine eher untergeordnete Rolle und wurden auf der Ebene der Endpunkte relativ gering gewichtet. Nur in aggregierter Form, d. h. auf einer der einzelnen

Endpunkten übergeordneten Ebene als kombinierter Endpunkt der Vermeidung von Nebenwirkungen (einschließlich der schwerwiegenden unerwünschten Ereignisse), kamen sie auf ein Gewicht von 0,095 bei den Patienten und 0,08 bei den Experten.

3.7 Ergebnisse zur alternativen Entscheidungshierarchie mit drei Endpunkten

In einer zusätzlichen zweiten Priorisierungsrunde wurden nur drei Behandlungsendpunkte anhand einer alternativen Entscheidungshierarchie zur Wahl gestellt. Tabelle 5 und Abbildung 6 zeigen die Ergebnisse hierzu. In Appendix B2 sind die aus der zweiten Priorisierungsrunde resultierenden individuellen Gewichte, arithmetischen Mittelwerte und Standardabweichungen zusammengefasst.

Tabelle 5: Gruppengewichte (geometrisches Mittel) aus der zweiten Priorisierungsrunde

Endpunkte der Behandlung	Gruppengewicht Patienten	Gruppengewicht Experten
Ansprechen	0,716	0,081
Remission	0,194	0,654
Vermeidung eines Therapieabbruchs aufgrund unerwünschter Ereignisse	0,090	0,265

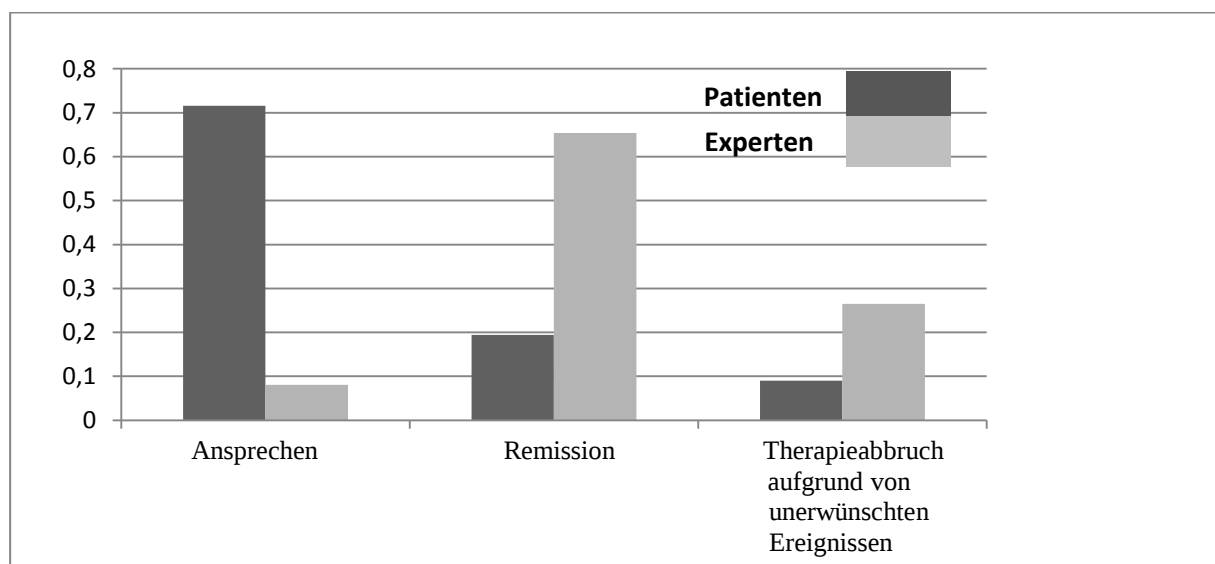


Abbildung 6: Gewichte aus der zweiten Priorisierungsrunde

Die Werte der Konsistenzratio für die Ergebnisse aus der Patienten- bzw. Expertenbefragung waren 0,02 und 0,00. Im vorliegenden Fall ist das Verhältnis der Gewichte für Remission und Ansprechen zueinander bei den Patienten und den Experten zwischen den beiden alternativen Hierarchiestrukturen relativ konstant geblieben. Haben die Patienten in der ersten Entscheidungsstruktur Ansprechen 3,8-mal wichtiger als Remission bewertet, war es in der verkürzten Entscheidungsstruktur 3,7-mal. Bei den Experten wurde in der komplexen

Entscheidungsstruktur Remission für 7,9-mal wichtiger als Ansprechen, hier 8,1-mal wichtiger eingestuft. Im Rahmen des Verfahrens spricht dies für eine gewisse Konsistenz und Reproduzierbarkeit der Ergebnisse.

3.8 Ergebnisse der Diskussionen in den Gruppen

Nach der Erstbewertung jedes Paarvergleichs konnten die Teilnehmer ihre Bewertungen in der Gruppe diskutieren. In den Diskussionen wurden zum einen die Beweggründe für die Bewertungsentscheidungen diskutiert, zum anderen wurden persönliche Erfahrungen zu einem Endpunkt ausgetauscht. Die Diskussionen ergaben wichtige Erkenntnisse, warum bestimmte Endpunkte wichtiger für Patienten bzw. Experten waren als andere.

Während Patienten in einer akuten depressiven Episode bei der Behandlung mit Antidepressiva ein möglichst schnelles Ansprechen und eine Besserung des eigenen Zustands als absolut wichtigsten Endpunkt beschrieben, stellten die Experten eher das Ziel einer Remission und der Vermeidung eines Rückfalls in den Vordergrund.

Die Ausführungen der Patienten waren hier sehr eindeutig dahin gehend, dass der Zustand der akuten Depression als derart unerträglich empfunden wurde, dass eine Besserung und ein erstes „Herausreißen“ aus diesem Zustand durch die Behandlung als überragendes Behandlungsziel angesehen wurde. Diese Ausführungen sollten jedoch nicht dahin gehend missverstanden werden, dass die Patienten kein Interesse an einer Remission haben. Jedoch sahen die Patienten die volle Remission eher als schwer, wenn nicht gar unerreichbares Fernziel an und berichteten, dass das Erreichen einer vollen Remission als sehr unwahrscheinlich betrachtet wird. Man könnte auch sagen, dass Patienten ein Medikament, das schnelle und effektive Besserung bringt, immer einem Medikament vorziehen würden, welches weniger schnell und effektiv Besserung erzielt, obwohl Letzteres ggf. schneller und häufiger zur Remission führt. Die folgende Aussage eines Patienten illustriert die Motivation der Patienten: „Ich würde lieber den Rest meines Lebens mit einer milden Depression leben, als in einer akuten mittelschweren bis schweren Episode keine Hoffnung zu haben, dass es ein Medikament gibt, welches mir in der akuten Phase Besserung zu bringen vermag.“

Die Experten hatten andere Beweggründe, warum sie das Ansprechen im Vergleich zur Remission und Vermeidung eines Rückfalls für weniger wichtig hielten. Eine Argumentationslinie war: „Das Ansprechen ist nicht der wichtigste Endpunkt. Wichtig ist es, eine volle Remission zu erzielen. Wenn keine volle Remission erreicht werden kann, ist die Langzeitprognose für den Patienten nicht gut.“

Die Diskussionen in den jeweiligen Gruppen zeigten auch, dass die Nebenwirkungen einer Behandlung im Vergleich zu den anderen Endpunkten als nicht so wichtig wahrgenommen wurden. Die Nebenwirkungen wurden entweder als nicht so einschneidend (insbesondere bei den einfachen Nebenwirkungen einer Behandlung wie Übelkeit, Schwindel etc.) angesehen oder als sehr selten, sehr unwahrscheinlich und damit keine unmittelbare Bedrohung darstellend (insbesondere die schwerwiegenden unerwünschten Ereignisse wie beispielsweise

der Suizid) eingestuft. Für den Endpunkt sexuelle Funktionsfähigkeit zeigten die Gruppendiskussionen, dass weder Patienten noch Experten diesen Endpunkt für wichtig hielten. Die Aussage eines Patienten hierzu war sehr aufschlussreich: „Wenn ich nicht in der Lage bin, überhaupt mit meinem Partner zu kommunizieren, stellt meine sexuelle Funktionsfähigkeit ein eher untergeordnetes Problem dar.“

4 Diskussion

Das in diesem Arbeitspapier dargestellte AHP-Pilotprojekt ermöglichte es, Gewichte für verschiedene Endpunkte einer Behandlung mit Antidepressiva, zum einen basierend auf den von depressiven Patienten und zum anderen auf den von den in der Behandlung depressiver Patienten tätigen Experten geäußerten Präferenzen zu generieren. Die getesteten AHP-Entscheidungsstrukturen umfassten eine Reihe wichtiger klinischer Behandlungsziele und -endpunkte, für die im Rahmen zweier IQWiG-Nutzenbewertungen zuvor Effektschätzer berechnet und Aussagen zu Nutzen und Schaden der eingeschlossenen Antidepressiva gemacht worden waren. Während das Pilotprojekt einerseits zeigen konnte, dass das AHP-Verfahren sowohl mit Patienten als auch mit Experten erfolgreich durchgeführt werden konnte und konsistente und zum Teil reproduzierbare Ergebnisse lieferte, offenbarten sich andererseits verschiedene methodische Herausforderungen des Verfahrens und Umsetzungsprobleme, die im Folgenden weiter diskutiert werden.

4.1 Methodische Aspekte

Die Konstruktion einer Entscheidungshierarchie

Ein erster wichtiger Schritt vor Durchführung des AHP-Verfahrens war die Erstellung einer AHP-Entscheidungshierarchie, die nicht oder nur minimal überlappende Behandlungsziele und -endpunkte beinhalten sollte. Eine Überlappung kann durch konzeptionelle, biologische und statistische Abhängigkeiten zwischen zwei Endpunkten verursacht werden. In diesem speziellen Fall der Endpunkte einer Behandlung von Patienten mit Depression war es eine Herausforderung, möglichst umfassend die Endpunkte abzubilden, die in den Nutzenbewertungen des IQWiG bewertet wurden, und gleichzeitig zu berücksichtigen, dass die Instrumente oder Fragebögen, die zur Messung dieser Endpunkte eingesetzt wurden, wie HAM-D, oftmals konzeptionell überlappen. Dies ist bei Instrumenten, d. h. Fragebögen oder Skalen, die in der Psychiatrie zur Operationalisierung verschiedener Endpunkte genutzt werden, häufig der Fall. Beispielsweise erfassen Fragebögen zur Operationalisierung der gesundheitsbezogenen Lebensqualität verschiedene der sonstigen Erkrankungssymptome wie Angst oder Schmerz, die zugleich in den Nutzenbewertungen bzw. in den diesen zugrunde liegenden randomisierten klinischen Studien als separate Endpunkte gemessen werden, und dies wiederum mithilfe separater Fragebögen. Darüber hinaus gibt es oftmals biologische Abhängigkeiten zwischen den Endpunkten, d. h. z. B., ein gesteigertes Maß an Angst beeinflusst auch die soziale Funktionsfähigkeit des Patienten. Wenn abhängige Endpunkte auf einer Ebene in eine Entscheidungshierarchie eingebunden werden, ist eine Überlappung unvermeidlich. In Bezug auf statistische Abhängigkeiten zwischen den Endpunkten ist eine Kausalität nicht zwingend erforderlich. Schon eine einfache Assoziation zwischen den Endpunkten kann die paarweisen Vergleiche und damit auch die Interpretation der Befragungsergebnisse erschweren.

Nichtsdestotrotz gilt ein gewisses Maß an Überlappung im AHP-Verfahren als akzeptabel und muss, wie in diesem Pilotprojekt gezeigt, bis zu einem gewissen Grad als unvermeidbar

hingenommen werden, wenn das Ziel ist, ein möglichst umfassendes Spektrum an wichtigen Behandlungszielen und -endpunkten zu gewichten. Die Wahl der in die Entscheidungshierarchie eingeschlossenen Ziele und Endpunkte sollte jedoch grundsätzlich von einem unabhängigen Experten überprüft werden, um Überlappungen auf einer Hierarchieebene auf ein mögliches Minimum zu beschränken. Darüber hinaus können Sensitivitätsanalysen helfen zu analysieren, inwiefern bestehende Überlappungen möglicherweise die Ergebnisse der AHP-Befragungen verzerren.

Wenn die Aufnahme relevanter, jedoch voneinander abhängiger Endpunkte unerlässlich ist, gibt es auch die Möglichkeit der Nutzung des sogenannten „Analytic Network Process (ANP)“. Das ANP-Verfahren ist eine Erweiterung des AHP-Verfahrens, welches explizit die bestehenden Abhängigkeiten zwischen einzelnen Endpunkten berücksichtigt. Weitere Informationen zum ANP finden sich beispielsweise in [34].

Hinzufügen / Herausnahme von Endpunkten aus der Entscheidungsstruktur

Aus verschiedenen Gründen könnten auch nach den Befragungen noch einzelne Endpunkte aus der Entscheidungshierarchie herausgenommen bzw. hinzugefügt werden müssen. In letzterem Fall müsste die Befragung wiederholt werden. In der Literatur wird hierzu darauf hingewiesen, dass unwichtige, für die befragten Personen nicht relevante Endpunkte unter der Annahme akzeptabler Konsistenz herausgenommen werden können [33]. Jedoch ist es theoretisch immer möglich, dass die nachträgliche Herausnahme (oder auch das Hinzufügen) von Endpunkten zum sogenannten „rank reversal“ führt, d. h. einer Umkehrung von abgegebenen Bewertungen. Dies wird wahrscheinlicher, wenn ein Endpunkt ergänzt oder herausgenommen wird, der eine große Bedeutung für die befragten Personen hat. Eine der eigentlichen AHP-Bewertung vorangehende (qualitative) Befragung der Zielgruppe der Befragung könnte genutzt werden, um sicherzustellen, dass alle bzw. nur relevante Endpunkte eingeschlossen sind. Eine Vorabprüfung auf Überlappungen könnte darüber hinaus vermeiden, dass im Nachgang ein Endpunkt aufgrund zu großer Überlappungen herausgenommen werden muss und damit möglicherweise nachträglich ein rank reversal der Bewertungen induziert wird.

Neben den in randomisierten klinischen Studien bewerteten Behandlungsendpunkten besteht darüber hinaus die Möglichkeit, dass Patienten und Experten weitere Endpunkte für sehr wichtig halten, die (noch) nicht in die klinischen Studien einbezogen wurden. Daher könnte im Vorfeld eines AHP-Verfahrens auch beispielsweise eine Pilot- oder Fokusgruppenbefragung (einschließlich Brainstorming) stattfinden, um Endpunkte zu identifizieren, die von den befragten Personen, insbesondere Patienten, als besonders wichtig erachtet werden. Zwischenzeitlich sollte man sich darüber im Klaren sein, dass eine AHP-Struktur, die sich auf die in klinischen Studien analysierten Endpunkte beschränkt, möglicherweise nicht umfassend ist und daher eingeschränkt übertragbare Ergebnisse generieren kann.

Definition und Bedeutung der Behandlungsziele / -endpunkte

In den Gesprächen / Diskussionen mit den Patienten und den Experten in diesem Pilotprojekt wurde deutlich, dass eine sehr präzise und eindeutige Formulierung und Definition der Ziele und Endpunkte extrem wichtig ist. Patienten und Experten müssen ein einheitliches Verständnis eines Endpunkts haben, insbesondere müssen die zeitliche und erkrankungsbezogene Dimension der abzugebenden Bewertung verdeutlicht werden. Beispielsweise zeigte es sich in den Befragungen, dass es für den Vergleich der Endpunkte Ansprechen und Remission sehr wichtig war zu verstehen, dass die Bewertung für eine akute mittelschwere bis schwere Episode vorgenommen werden sollte, d. h. unter der Annahme, dass der Patient sich in einer solchen akuten Phase der Erkrankung befindet.

Die Diskussionen mit den Patienten und Experten halfen darüber hinaus dabei, die Bedeutung der einzelnen Endpunkte für die Befragten besser zu verstehen. Beispielsweise wurde aus den Diskussionen sehr schnell ersichtlich, dass der Endpunkt der sexuellen Funktionsfähigkeit bei der Behandlung mit Antidepressiva weder für Patienten noch für Experten eine große Bedeutung hat. Die Wirkung eines Antidepressivums auf die sexuelle Funktionsfähigkeit beeinflusst offensichtlich weder das Verschreibungsverhalten von Ärzten noch die Präferenz eines Patienten für ein bestimmtes Antidepressivum. Würde sich eine solche Einschätzung in einer repräsentativen Befragung mittels AHP bestätigen, könnte man z. B. darauf verzichten, diesen Endpunkt in eine Nutzen- bzw. Kosten-Nutzen-Bewertung aufzunehmen, wenn er für die Entscheidung für oder gegen ein konkretes Arzneimittel kaum eine Rolle spielt. Darüber hinaus wurde für einige Endpunkte auch deutlich, dass ihre Bewertung in klinischen Studien entweder auf einem höheren oder geringeren Aggregationslevel erfolgen sollte. Beispielsweise legt diese Studie nahe, den kombinierten Endpunkt „Therapieabbruch aufgrund unerwünschter Ereignisse“ zu betrachten, der eine Teilmenge sowohl der schwerwiegenden als auch der nicht schwerwiegenden unerwünschten Ereignisse abdeckt, und nicht die jeweils einzelnen Endpunkte in diesen Kategorien, denen einzeln jeweils nur ein sehr geringes Gewicht zukommt.

Weitere methodische Aspekte des AHP-Verfahrens

Zusätzlich zu den bereits diskutierten Herausforderungen des AHP-Verfahrens wurden in verschiedenen Studien bereits weitere methodische Schwächen des Verfahrens offengelegt und besprochen (z. B. [35-37]). Verschiedene Ansätze zum Umgang mit diesen liegen vor oder werden diskutiert [38-42].

In diesem Pilotprojekt wurde die Bedeutung der Behandlungsziele bzw. -endpunkte anhand einer 17-Punkte-Skala bewertet. Diese Skala wurde mithilfe empirischer Untersuchungen entwickelt und validiert. Es handelt sich hierbei einerseits um eine Ordinalskala, d. h., den Werten 1 bis 9 (in jede Richtung des paarweisen Vergleichs) sind Einschätzungskategorien (indifferent, etwas wichtiger, wichtiger, viel wichtiger, sehr viel wichtiger) zugeteilt, zwischen denen eine Rangordnung besteht. Andererseits werden die numerischen Werte zur Berechnung von Gewichten weiter genutzt. Fraglich bleibt in diesem Kontext, ob diese Werte

tatsächlich die Bewertung der befragten Personen angemessen wiedergeben. Die Entwicklung neuer und ggf. geeigneterer Skalen steht im Vordergrund verschiedener Untersuchungen (z. B. [43]). Derzeit erscheint jedoch keine andere Skala so akzeptiert und breit genutzt wie die in diesem Pilotprojekt verwendete [42,44], was dennoch nicht die allgemeine Diskussion hinfällig macht, inwiefern sich auf Basis von ordinalen Einschätzungen kardinale Werte ergeben können.

Darüber hinaus wird die Validität des AHP-Verfahrens auch im Kontext der Nutzentheorie diskutiert und beispielsweise mit der nutzentheoretisch basierten Conjoint-Analyse (CA) verglichen [44,45]. AHP legt keine explizite Nutzenfunktion zugrunde wie z. B. das CA-Verfahren. Jedoch wird das AHP-Verfahren auch gerade aus diesem Grunde von einigen Autoren bevorzugt, da die Methode der paarweisen Vergleiche eine prozedurale („eingeschränkte“) Rationalität verlangt und damit den menschlichen Entscheidungsprozessen sehr entgegenkommt [45]. Wichtig im Vergleich des CA- und AHP-Verfahrens ist zudem, dass es wenig Unterschiede in der Vorhersagevalidität zwischen CA und AHP zu geben scheint [46].

4.2 AHP-Befragungen

Bedeutung der Gruppendiskussionen

Die Gruppendiskussionen ermöglichten es, die Beweg- und Hintergründe der einzelnen individuellen Bewertungsentscheidungen zu den paarweisen Vergleichen besser zu verstehen. Sowohl die Patienten als auch die Experten diskutierten separat in ihren Gruppen ihre jeweiligen Bewertungsentscheidungen. Insbesondere die großen Abweichungen in den Einschätzungen der Patienten im Gegensatz zu den Experten wurden so verständlich – dies war in diesem Pilotprojekt bei der unterschiedlichen Bewertung der Endpunkte Ansprechen und Remission der Fall. Die Erkenntnisse sind hilfreich, um zukünftige AHP-Befragungen zielgerichteter und, unter Kenntnis der genannten Herausforderungen der Methoden, auch vorausschauender planen zu können. Die Gegenüberstellung der Ergebnisse aus den Patienten- bzw. Expertenbefragungen, nicht nur in Bezug auf die generierten Gewichte, sondern ganz besonders auch in Bezug auf die Ergebnisse der Gruppendiskussionen, war insoweit hilfreich, als sich daraus Erkenntnisse für zukünftige AHP-Befragungen ableiten lassen. Während die Erhebung von Patientenpräferenzen in Bezug auf die Ergebnisnutzung das eigentliche Ziel der AHP-Befragung war, trugen die Expertenbefragungen dazu bei, die Endpunktgewichtungen besser einschätzen und interpretieren zu können. Die Gegenüberstellung hob auch wichtige Aspekte, die es bei der Definition der Endpunkte zu beachten gilt, klar hervor – beispielsweise die Wichtigkeit einer eindeutigen Definition der zeit- und krankheitsbezogenen Dimension der abzugebenden Bewertungen.

Gemäß den Vorgaben der AHP-Methode spielen die Gruppendiskussionen eine wichtige Rolle im AHP-Verfahren. Die regelmäßige Überprüfung der Konsistenz der abgegebenen Bewertungen auf jeder Ebene der Entscheidungshierarchie erlaubt es, bei unzulässig hoher Inkonsistenz einen Austausch und eine Klärung in der Gruppendiskussion herbeizuführen, um

Bewertungen, die auf Missverständnissen beruhen, zu korrigieren[47]. In diesem Pilotprojekt war eine Neubewertung bzw. Wiederholung eines paarweisen Vergleichs aufgrund der insgesamt guten Konsistenz der Bewertungen nicht notwendig. Die Werte für die Konsistenzratio auf jeder Ebene der paarweisen Vergleiche waren akzeptabel und lagen unterhalb von 0,1. Dies spricht dafür, dass die Bewertungen in beiden Gruppen einem verlässlichen und rationalen Entscheidungsverhalten entsprechen (d. h. die Präferenzen der Befragten den Anforderungen der Transitivität und Reziprozität genügen).

Struktur und Moderation der Befragungen / Gruppendiskussionen

Die Gruppendiskussionen in der Patientengruppe fanden in einer ruhigen und vertrauensvollen Atmosphäre statt. Die Patienten hoben mehrfach hervor, dass sie das Interesse des IQWiG an ihren Meinungen und Einschätzungen zu schätzen wüssten und den mit dem Pilotprojekt eingeleiteten Versuch, diese systematisch zu erfassen, befürworteten. Da die Bewertungen der Patienten immer nur in anonymisierter Form für alle sichtbar an die Wand projiziert wurden, fühlten die Patienten sich nicht zu einer Rechtfertigung ihrer Einschätzung verpflichtet, sondern sagten nur dann etwas zu den Bewertungen, wenn sie das Bedürfnis hatten. Insgesamt erläuterten und diskutierten die Patienten ihre Bewertungen mit sehr großer Offenheit.

Die Atmosphäre in der Expertenbefragung war weniger entspannt als in der Patientenbefragung. An verschiedenen Stellen wurde die Aufmerksamkeit der Experten vom eigentlichen Ziel der Befragung abgelenkt, da von einzelnen Teilnehmenden die Diskussion der Methoden der Nutzenbewertungen des IQWiG in den Vordergrund gerückt wurde. In den Diskussionen war die Dominanz eines Experten sehr stark wahrnehmbar. Jedoch schien dieser Experte die Bewertungen der anderen Experten nicht offensichtlich zu beeinflussen. Zwei kritische Aspekte wurden deutlich. Erstens folgte aus der nicht anonymisierten Form der Befragung, dass die Experten jeweils die Namen der anderen Experten und deren Bewertungen sehen konnten, was grundsätzlich dazu verleiten konnte, sich der Einschätzung eines Meinungsführers anzuschließen. Zweitens war es so, dass die Bewertungen der ersten Experten bereits an die Wand projiziert wurden, bevor die anderen ihre Bewertung überhaupt abgegeben hatten. Dies erhöhte die Wahrscheinlichkeit einer gewissen Beeinflussung einer Bewertung durch bereits projizierte Ergebnisse anderer Bewerter. Daraus folgt, dass die AHP-Befragungen daher auch für Experten analog zur Patientenbefragung anonymisiert durchgeführt werden sollten und die Ergebnisse einer Bewertungsrunde erst für alle sichtbar gemacht werden sollten, wenn alle ihre Bewertung abgegeben haben.

Insbesondere der Moderator der Gruppendiskussionen in der Expertengruppe sollte eine sehr gute fachliche und menschliche Autorität in der Gruppe haben und in der Moderation der Gruppendiskussion ausgebildet sein. Darüber hinaus hat sich gezeigt, dass die Befragungen auch eine starke kognitive Beanspruchung der Moderatoren erfordern und daher idealerweise nicht an einem Tag stattfinden sollten, wie dies in dieser Pilotstudie der Fall war (jede Befragung dauerte ca. 3 Stunden).

4.3 Rekrutierung von Patienten und Experten

In diesem Pilotprojekt wurden Patienten rekrutiert, ohne dass vorab ein wirkliches „Profil“ der Patienten, die in die Studie aufgenommen werden sollten, definiert worden war. Der Aufruf zur Teilnahme an der Befragung, der primär auf verschiedenen Webseiten von Selbsthilfegruppen veröffentlicht wurde, war an Patienten mit der Vorgeschichte einer Depression adressiert, ohne weitere Einschränkung von Schweregrad der Erkrankung, Alter, Geschlecht oder sozio-ökonomischem Status. In einem zukünftigen AHP-Verfahren sollten die aus epidemiologischer Sicht wichtigen soziodemografischen Charakteristika wie Alter, Geschlecht und Ausbildungsstand sowie die Schweregrade der Erkrankung – mittelgradig und schwer – angemessen in der Gruppe der Patienten vertreten sein und daher auch erhoben werden. Ist dies nicht im Rahmen einer Befragung möglich, sollten mehrere und gegebenenfalls stratifizierte Befragungen durchgeführt werden, um die Vielfältigkeit der betroffenen Patienten angemessen abzubilden. Eine weitere Möglichkeit wäre die Durchführung von Online-AHP-Befragungen mit einer Abfrage der genannten Charakteristika, um nach Subgruppen differenziert auswerten zu können.

Die Einschätzung der kognitiven Funktionsfähigkeit war in der Patientenbefragung in diesem Pilotprojekt beispielsweise ein Endpunkt, der sehr unterschiedlich, jedoch insgesamt vergleichsweise hoch bewertet wurde. Eine Begründung könnte sein, dass in dieser Gruppe die Patienten mit hohem Ausbildungsstand überrepräsentiert waren, was auch aus der informellen Vorstellungsrunde, die der Befragung voranging, hervorging bzw. aus den freiwillig gegebenen Informationen abgeleitet werden konnte. Die Vermutung, dass die Bedeutung, die die Patienten ihrer kognitiven Leistungsfähigkeit beimessen, von ihrem Ausbildungs- und Berufsstand abhängt, führt zu der Schlussfolgerung, dass die Bedeutung dieses Endpunkts in dieser Befragung überschätzt worden sein könnte.

Auch wenn weder die soziodemografischen Daten noch Daten zu Krankheitsdauer und -verlauf der befragten Patienten erhoben wurden, so ergab sich aus den Diskussionen während der Befragung, dass die Gruppe heterogen nach Schweregrad der Erkrankung und der Symptome war. Während bspw. einige Patienten das Erkrankungssymptom Angst als besonders wichtig empfanden, stand bei anderen der Schmerz im Vordergrund. Auch hier muss darauf geachtet werden, dass beide Patientengruppen in einer Befragung angemessen repräsentiert werden, um nicht den einen oder anderen Endpunkt zu über- oder zu unterschätzen. Wenn beispielsweise epidemiologische Studien nahelegen, dass ca. 60 % der depressiven Patienten zusätzlich primär unter dem Symptom Angst leiden und ca. 30 % unter Schmerzen, dann sollten die Gruppe der befragten Patienten sowie auch die Befragung selbst hierauf abgestellt sein.

Derartige Aspekte stellen große Herausforderungen an alle vergleichbaren Entscheidungsverfahren, also z. B. AHP und CA, und die entsprechenden Rekrutierungs- und Befragungskonzepte. Wiederholte und im Hinblick auf die genannten soziodemografischen, sozioökonomischen und krankheitsbezogenen Aspekte stratifizierte Befragungen (bzw. Befragungen in Subgruppen von Patienten) stellen hier mögliche Lösungsansätze dar, sollten jedoch im Vorfeld auch selbst immer einer Kosten-Nutzen-Abwägung unterzogen werden.

5 Fazit

Dieses Pilotprojekt hat gezeigt, dass das AHP-Verfahren als Methode sowohl mit Patienten als auch mit Experten durchgeführt werden kann. Die Befragung war durchgehend kognitiv gut zu bewältigen. Das strukturierte Verfahren der AHP-Methode ermöglichte es, Schritt für Schritt die Präferenzen von Individuen für bestimmte Behandlungsziele bzw. -endpunkte zu erheben und Gewichte für jeden dieser Endpunkte mittels der Matrixalgebra zu berechnen.

Die paarweisen Vergleiche der Endpunkte in Kombination mit den Gruppendiskussionen erlaubten einen intensiven Austausch von Wahrnehmungen, Informationen und Erfahrungen zwischen den befragten Personen und ließen Einsichten in die Beweg- und Hintergründe der abgegebenen Einschätzungen zu.

Bei der Durchführung des Pilotprojekts wurden einige methodische Herausforderungen des AHP-Verfahrens besonders offensichtlich, die in zukünftigen AHP-Befragungen auf jeden Fall berücksichtigt werden sollten. Zum einen korrelieren bzw. überlappen Endpunkte bzw. Therapieziele. Eine Erhebung von zu hohen Gewichten, die ggf. zum rank reversal führen könnte, lässt sich nie völlig vermeiden, müsste aber durch Sensitivitätsanalysen geprüft werden. Zum anderen bleibt weiterhin die Frage offen, der sich auch Verfahren zur Erhebung der QALYs bzw. andere Methoden der multikriteriellen Entscheidungsfindung stellen müssen: Welche Personen sollen befragt werden und wie übertragbar sind diese Ergebnisse auf eine gesamte Patientenpopulation?

Abgesehen davon, dass diese methodischen Fragen geklärt werden müssen, kann das AHP-Verfahren für verschiedene Fragestellungen im Prozess der HTA-Erstellung genutzt werden.

- Zur Identifizierung der für die Patienten relevant(est)en Endpunkte, um möglicherweise nur für diese Endpunkte Effizienzgrenzen zu erstellen.
- Gewichtung endpunktspezifischer Effektmaße (z. B. ORs) und Ableitung eines aggregierten gewichteten Effektmaßes für jede Behandlungsalternative.
- Aggregation der aus endpunktspezifischen Effizienzgrenzen resultierenden Ergebnisse und Ableitung eines aggregierten Maßes für die Kosten-Effektivität, z. B. kann ein aggregierter, gewichteter Erstattungspreis für Medikament A durch die Gewichtung der aus verschiedenen endpunktspezifischen Effizienzgrenzen abgeleiteten Erstattungspreise berechnet werden.

Darüber hinaus kann das AHP-Verfahren auch wichtige Hinweise liefern, welche Endpunkte in zukünftigen klinischen Studien primär berücksichtigt werden sollten, welche eher nachrangig und welche gegebenenfalls gar nicht.

6 Literatur

1. Saaty TL. A scaling method for priorities in hierarchical structures. J Math Psychol 1977; 15(3): 234-281.
2. Saaty TL. The Analytic Hierarchy Process: planning, priority setting, resource allocation. New York: McGraw-Hill; 1980.
3. Saaty TL. Group decision making and the AHP. In: Golden BL, Wasil EA, Harker PT (Ed). The analytic hierarchy process: applications and studies. Berlin: Springer; 1989. S. 59-67.
4. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden: Version 4.0 [online]. 23.09.2011 [Zugriff: 20.01.2012]. URL: https://www.iqwig.de/download/IQWiG_Methoden_Version_4_0.pdf.
5. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden zur Bewertung von Verhältnissen zwischen Nutzen und Kosten [online]. 12.10.2009 [Zugriff: 20.01.2012]. URL: https://www.iqwig.de/download/Methodik_fuer_die_Bewertung_von_Verhaeltnissen_zwischen_Kosten_und_Nutzen.pdf.
6. Dolan JG. Medical decision making using the Analytic Hierarchy Process: choice of initial antimicrobial therapy for acute pyelonephritis. Med Decis Making 1989; 9(1): 51-56.
7. Dolan JG, Bordley DR. Individualized decision-making using the Analytic Hierarchy Process (AHP): reliability, validity and clinical usefulness. Med Decis Making 1991; 11(4): 322-322.
8. Dolan JG. Are patients capable of using the Analytic Hierarchy Process and willing to use it to help make clinical decisions? Med Decis Making 1995; 15(1): 76-80.
9. Dolan JG, Bordley DR. Using the Analytic Hierarchy Process (AHP) to develop and disseminate guidelines. QRB Qual Rev Bull 1992; 18(12): 440-447.
10. Dolan JG. Shared decision-making: transferring research into practice; the Analytic Hierarchy Process (AHP). Patient Educ Couns 2008; 73(3): 418-425.
11. Benaim C, Perennou DA, Pelissier JY, Daures JP. Using an Analytical Hierarchy Process (AHP) for weighting items of a measurement scale: a pilot study. Rev Epidemiol Sante Publique 2010; 58(1): 59-63.
12. Sloane EB, Liberatore MJ, Nydick RL. Medical decision support using the Analytic Hierarchy Process. J Healthc Inf Manag 2002; 16(4): 38-43.
13. Hummel JM, IJzerman MJ. The use of the Analytic Hierarchy Process in health care decision making. Enschede: University of Twente; 2009.

14. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen: Abschlussbericht; Auftrag: A05-20A; Version 1.1 [online]. 18.08.2010 [Zugriff: 07.05.2012]. (IQWiG-Berichte; Band 55). URL: https://www.iqwig.de/download/A05-20A_Abschlussbericht_SNRI_bei_Patienten_mit_Depressionen_V1-1.pdf.
15. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Bupropion, Mirtazapin und Reboxetin bei der Behandlung der Depression: Abschlussbericht; Auftrag A05-20C; Version 1.1 [online]. 30.05.2011 [Zugriff: 07.05.2012]. (IQWiG-Berichte; Band 68). URL: https://www.iqwig.de/download/A05-20C_Abschlussbericht_1-1_Bupropion_Mirtazapin_und_Reboxetin_bei_Depressionen.pdf.
16. Bloom BS. Prevalence and economic effects of depression. *Manag Care* 2004; 13(6 Suppl Depression): 9-16.
17. O'Connor EA, Whitlock EP, Beil TL, Gaynes BN. Screening for depression in adult patients in primary care settings: a systematic evidence review. *Ann Intern Med* 2009; 151(11): 793-803.
18. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA* 2003; 289(23): 3095-3105.
19. Alonso J, Angermeyer MC, Bernert S, Bruffaerts R, Brugha TS, Bryson H et al. Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl* 2004; 109(420): 21-27.
20. Wittchen HU, Pfister H, Schmidtkunz B, Winter S, Müller N. Zusatzsurvey "Psychische Störungen" (Bundesgesundheitsurvey 98): Häufigkeit, psychosoziale Beeinträchtigungen und Zusammenhänge mit körperlichen Erkrankungen; Anlagen (Anlagen zum Schlussbericht BMBFBW 01 EH 9701/8). München: Max-Planck-Institut für Psychiatrie, klinische Psychologie und Epidemiologie; 2000.
21. American Psychiatric Association (Ed). Diagnostic criteria from DSM-IV-TR. Washington: APA; 2000.
22. Masand PS, Gupta S. Long-term side effects of newer-generation antidepressants: SSRIS, venlafaxine, nefazodone, bupropion, and mirtazapine. *Ann Clin Psychiatry* 2002; 14(3): 175-182.
23. Nuevo R, Leighton C, Dunn G, Dowrick C, Lehtinen V, Dalgard OS et al. Impact of severity and type of depression on quality of life in cases identified in the community. *Psychol Med* 2010; 40(12): 2069-2077.

24. Deutsche Gesellschaft für Psychiatrie, Psychotherapie und Nervenheilkunde, Bundesärztekammer, Kassenärztliche Bundesvereinigung, Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften, Arzneimittelkommission der deutschen Ärzteschaft, Bundespsychotherapeutenkammer et al. S3-Leitlinie/Nationale VersorgungsLeitlinie: unipolare Depression; Langfassung; Version 1.3 [online]. 01.2012 [Zugriff: 26.07.2012]. URL: http://www.versorgungsleitlinien.de/themen/depression/pdf/s3_nvl_depression_lang.pdf.
25. Williams JW Jr, Mulrow CD, Chiquette E, Noel PH, Aguilar C, Cornell J. A systematic review of newer pharmacotherapies for depression in adults: evidence report summary. Ann Intern Med 2000; 132(9): 743-756.
26. Saaty TL. Highlights and critical points in the theory and application of the Analytic Hierarchy Process. Eur J Oper Res 1994; 74(3): 426-447.
27. Dolan JG, Isselhardt BJ Jr, Cappuccio JD. The Analytic Hierarchy Process in medical decision making: a tutorial. Med Decis Making 1989; 9(1): 40-50.
28. Forman E, Peniwati K. Aggregating individual judgments and priorities with the Analytic Hierarchy Process. Eur J Oper Res 1998; 108(1): 165-169.
29. Bridge JA, Iyengar S, Salary CB, Barbe RP, Birmaher B, Pincus HA et al. Clinical response and risk for reported suicidal ideation and suicide attempts in pediatric antidepressant treatment: a meta-analysis of randomized controlled trials. JAMA 2007; 297(15): 1683-1696.
30. Laughren TP. Memorandum: overview for December 13 meeting of Psychopharmacologic Drugs Advisory Committee (PDAC) [online]. 16.11.2006 [Zugriff: 02.04.2013]. URL: <http://www.fda.gov/ohrms/dockets/ac/06/briefing/2006-4272b1-01-fda.pdf>.
31. Stone M, Laughren T, Jones ML, Levenson M, Holland PC, Hughes A et al. Risk of suicidality in clinical trials of antidepressants in adults: analysis of proprietary data submitted to US Food and Drug Administration. BMJ 2009; 339: b2880.
32. Arzneimittelkommission der deutschen Ärzteschaft. Stellenwert von Antidepressiva in der Depressionsbehandlung [online]. 17.04.2008 [Zugriff: 26.01.2012]. URL: <http://www.akdae.de/Stellungnahmen/Weitere/20080417.pdf>.
33. Wijnmalen DJD, Wedley WC. Non-discriminating criteria in the AHP: removal and rank reversal. Journal of Multi-Criteria Decision Analysis 2009; 15(5-6): 143-149.
34. Barzilai J, Golany B. AHP rank several, normalization and aggregation rules. INFOR 1994; 32(2): 57-64.
35. Bana e Costa CA, Vansnick JC. A critical analysis of the eigenvalue method used to derive priorities in AHP. Eur J Oper Res 2008; 187(3): 1422-1428.

36. Finan JS, Hurley WJ. The Analytic Hierarchy Process: can wash criteria be ignored? *Computers and Operations Research* 2002; 29(8): 1025-1030.
37. Liang L, Wang G, Hua Z, Zhang B. Mapping verbal responses to numerical scales in the Analytic Hierarchy Process. *Socioecon Plann Sci* 2008; 42(1): 46-55.
38. Lin JSJ, Chou SY, Chouhuang WT, Hsu CP. Note on "Wash criterion in Analytic Hierarchy Process". *Eur J Oper Res* 2008; 185(1): 444-447.
39. Saaty TL. Decision making with the Analytic Hierarchy Process. *International Journal of Services Sciences* 2008; 1(1): 83-98.
40. Saaty TL, Vargas LG. The Analytic Hierarchy Process: wash criteria should not be ignored. *International Journal of Management and Decision Making* 2006; 7(2/3): 180-188.
41. Finan JS, Hurley WJ. Transitive calibration of the AHP verbal scale. *Eur J Oper Res* 1999; 112(2): 367-372.
42. Miller GA. The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychol Rev* 1956; 63(2): 81-97.
43. Kallas Z, Lambarraa F, Gil JM. A stated preference analysis comparing the analytical hierarchy process versus choice experiments. *Food Qual Prefer* 2011; 22(2): 181-192.
44. IJzerman MJ, Van Till JA, Snoek GJ. Comparison of two multi-criteria decision techniques for eliciting treatment preferences in people with neurological disorders. *Patient* 2008; 1(4): 265-272.
45. Kinoshita E. Why we need AHP/ANP instead of utility theory in today's complex world: AHP from the perspective of bounded rationality [online]. 07.2005 [Zugriff: 31.08.2012]. URL: http://www.isahp.org/2005Proceedings/Papers/KinoshitaE_AHP&ANPvsUtilityTheory.pdf.
46. Mulye R. An empirical comparison of three variants of the AHP and two variants of conjoint analysis. *J Behav Decis Mak* 1998; 11(4): 263-280.
47. Henig MI, Buchanan JT. Solving MCDM problems: process concepts. *Journal of Multi-Criteria Decision Analysis* 1996; 5(1): 3-21.

Anhang A – Vorabinformation der Patienten

Einführende Informationen zur AHP-Befragung

Das IQWiG hat im Auftrag des Gemeinsamen Bundesausschusses (G-BA) im vergangenen Jahr systematisch Informationen zur Wirksamkeit und zu den Nebenwirkungen verschiedener Antidepressiva zusammengetragen. Das heißt, es wurden klinische Studien ausgewertet, die die erwünschten und unerwünschten Wirkungen von Antidepressiva auf den Körper und die Psyche (Morbidität), die Sterblichkeit (Mortalität) und die Lebensqualität untersucht haben.

Klinische Studien untersuchen die Wirkungen eines Medikamentes immer im Vergleich zu keinem oder anderen Medikamenten für bestimmte Endpunkte wie z. B. die Verbesserung von Symptomen der Depression oder die von den Studienteilnehmenden berichteten Nebenwirkungen der Behandlung. Diese Endpunkte sind jedoch für Patientinnen und Patienten nicht unbedingt alle gleich wichtig. Das IQWiG testet derzeit verschiedene wissenschaftliche Methoden, um herauszufinden, welche der Endpunkte aus den klinischen Studien für Patientinnen und Patienten die wichtigsten sind. Mithilfe dieser Informationen könnten zukünftige Kosten-Nutzen-Bewertungen durch das IQWiG speziell Ergebnisse zu den Endpunkten in den Mittelpunkt stellen, die für Patientinnen und Patienten am wichtigsten sind.

In der Befragung am 26.03.2010 möchten wir daher von Ihnen erfahren, wie bedeutsam die verschiedenen Endpunkte einer Behandlung mit Antidepressiva für Sie sind. Zur Vorbereitung auf diese Befragung geben wir Ihnen im Vorfeld einige Informationen zu den Endpunkten, zu deren Gewichtung wir an diesem Termin Ihre Einschätzung benötigen.

1 Wirksamkeit der Depressionsbehandlung

Mit „Wirksamkeit“ sind hier die erwünschten Wirkungen der Antidepressiva auf die Symptome der Depression gemeint. Die depressive Symptomatik wird unter anderem durch Befragung der Patientin oder des Patienten mithilfe von Depressionsfragebögen ermittelt. Die Anzahl bzw. das Ausmaß der berichteten Symptome hilft dem Arzt zu entscheiden, ob und wie stark die Patientin oder der Patient an einer Depression erkrankt ist. In der Wissenschaft hat man sich darauf geeinigt, die Effekte einer Therapie eindeutig zu definieren, damit man die unterschiedlich starke Wirkung von verschiedenen Behandlungen vergleichen kann.

Folgende Endpunkte werden hierzu genutzt:

1.1 Ansprechen auf die Behandlung (Response)

Von einem Ansprechen wird gesprochen, wenn die Anzahl bzw. das Ausmaß der durch die Patientin oder den Patienten berichteten depressiven Symptome halbiert werden kann. D. h. beispielsweise, dass die Beantwortung eines Depressionsfragebogens durch die Patientin oder den Patienten nicht mehr einen Gesamtwert von 20 Punkten ergibt, sondern sich nur noch auf 10 Punkte summiert.

1.2 Erreichen von Symptomfreiheit (volle Remission)

Von Symptomfreiheit spricht man, wenn die Symptome so deutlich reduziert werden können, dass die Patientin oder der Patient nicht mehr die Kriterien einer Depression erfüllt. D. h., die Patientin oder der Patient berichtet dann nur noch über wenige, keine oder nur schwach ausgeprägte depressive Symptome.

1.3 Vermeidung eines Rückfalls

Man spricht davon, dass ein Rückfall durch eine Behandlung vermieden wird, wenn eine Patientin / ein Patient innerhalb von bis zu 12 Monaten nach Symptomfreiheit (Remission) nicht wieder die volle depressive Symptomatik zeigt.

2 Unerwünschte Wirkungen (Nebenwirkungen) der Depressionsbehandlung

Nebenwirkungen sind die unerwünschten Wirkungen einer Behandlung. In klinischen Studien wird immer versucht, möglichst viele von Patientinnen und Patienten berichtete Nebenwirkungen zu erfassen, die dann meist auch in ihrer Gesamtheit als „unerwünschte Wirkungen“ untersucht werden.

2.1 Unerwünschte Wirkungen

Unerwünschte Wirkungen einer Arzneimitteltherapie sind nicht lebensbedrohlich und erfordern keine Einweisung ins Krankenhaus.

2.1.1 Unerwünschte Wirkungen / Ereignisse

Es handelt sich hierbei um Nebenwirkungen eines Medikamentes wie beispielsweise den Anstieg des Blutdrucks, Übelkeit / Erbrechen, Unruhe / Schlafstörungen, starkes Schwitzen, Gewichtszunahme, Hautausschläge, Schwindel etc.

2.1.2 Sexuelle Funktionsstörungen

Die Behandlung mit bestimmten Antidepressiva kann zu sexuellen Funktionsstörungen führen. Bei Frauen und Männern können Erregungs- und Orgasmusstörungen sowie ein vermindertes Lustempfinden auftreten. Bei Männern kann es zudem zu Ejakulations- und Potenzstörungen kommen.

2.2 Schwerwiegende unerwünschte Wirkungen

Schwerwiegende unerwünschte Wirkungen sind unerwünschte Wirkungen, die lebensbedrohlich sind, zu schweren Behinderungen führen oder eine Krankenhauseinweisung erforderlich machen.

2.2.1 (Versuchte) Selbsttötung

Gedanken an Selbsttötung und Selbsttötungsversuche gehören zu den möglichen Symptomen einer bestehenden Depression. Jedoch können Antidepressiva das Risiko für eine Selbsttötung bzw. einen Selbsttötungsversuch erhöhen, wenn sie den Antrieb verbessern, sodass Selbsttötungsgedanken häufiger in die Tat umgesetzt werden.

2.2.2 Sonstige schwerwiegende unerwünschte Wirkungen

Jede unerwünschte Wirkung eines Medikamentes kann zu einem schwerwiegenden unerwünschten Ereignis werden, wenn sich die Wirkung so verschlimmert, dass sie lebensbedrohlich wird, zur Krankenhauseinweisung bzw. zu schwerer Behinderung oder gar zum Tod führt. Beispielsweise können durch Medikamente verursachte Herzrhythmusstörungen, die zu einer Krankenhauseinweisung führen, ein solches schwerwiegendes Ereignis sein.

2.3 Therapieabbruch aufgrund unerwünschter Wirkungen

Von einem Therapieabbruch spricht man, wenn eine Behandlung mit Antidepressiva aufgrund unerwünschter Nebenwirkungen des Medikamentes vorzeitig beendet wird.

3 Begleitende Auswirkungen auf die Lebensqualität

Antidepressiva können die Lebensqualität depressiver Patientinnen und Patienten beeinflussen.

3.1 Positive Auswirkungen auf soziale Funktionsbereiche

Verbesserung der Teilnahme am sozialen Leben und am Familienleben; Beruf und Freizeitaktivitäten können wieder besser ausgeübt werden.

3.2 Verringerung von Angst

Gegebenenfalls können Angstzustände / Ängste verringert werden.

3.3 Verringerung von Schmerzen

Mögliche Verringerung von Schmerzen (unterschiedlicher Art und Ursprungs).

3.4 Verbesserung / Erhaltung der intellektuellen Leistungsfähigkeit

Möglicherweise Verbesserung / Erhaltung der Denk- bzw. Konzentrationsfähigkeit.

Anhang B – Vorabinformation der Experten

Einführende Informationen zur AHP-Befragung

Das IQWiG hat im Auftrag des Gemeinsamen Bundesausschusses (G-BA) im vergangenen Jahr zwei Nutzenbewertungen zu Antidepressiva durchgeführt, in denen systematisch Informationen zur Wirksamkeit und zu den Nebenwirkungen verschiedener Antidepressiva zusammengetragen wurden. In diesen Nutzenbewertungen wurden die Effekte der eingeschlossenen Antidepressiva auf verschiedene Endpunkte innerhalb der Zielgrößen Morbidität, Mortalität und Lebensqualität untersucht. Bei den eingeschlossenen Antidepressiva handelte es sich zum einen um die Gruppe der SNRI (Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer) und zum anderen um die drei Substanzen Bupropion, Mirtazapin und Reboxetin. Basierend auf den Ergebnissen dieser Nutzenbewertungen soll in diesem Jahr nun eine Kosten-Nutzen-Bewertung (KNB) durchgeführt werden, die die genannten Substanzen im Vergleich zu weiteren Komparatoren hinsichtlich ihrer Kosten-Effektivität überprüft.

Da die in die Nutzenbewertungen eingeschlossenen Endpunkte jedoch für Patientinnen und Patienten von unterschiedlicher Bedeutung sind, testet das IQWiG derzeit verschiedene wissenschaftliche Methoden, wie die für die Patientinnen und Patienten wichtigsten Endpunkte identifiziert werden können. Mithilfe dieser Informationen könnten zukünftige Kosten-Nutzen-Bewertungen durch das IQWiG Ergebnisse zu den Endpunkten in den Mittelpunkt stellen, die für Patientinnen und Patienten und in deren Behandlung tätige Expertinnen und Experten als am wichtigsten eingeschätzt werden.

In der Befragung am 26. März 2010 möchten wir daher erfahren, wie Sie die Bedeutung der verschiedenen Endpunkte einer Depressionsbehandlung für Ihre Patientinnen und Patienten einschätzen. Zur Vorbereitung auf diese Befragung geben wir Ihnen im Vorfeld einige Informationen zu den Endpunkten, zu deren Gewichtung wir an diesem Termin Ihre Einschätzung benötigen.

1 Wirksamkeit der Depressionsbehandlung

Mit Wirksamkeit sind die erwünschten Wirkungen der Antidepressiva auf die Symptome der Depression gemeint. Die depressive Symptomatik wird, wie Sie wissen, in klinischen Studien mithilfe von Depressionsfragebögen wie zum Beispiel der Hamilton-Depressions-Skala ermittelt.

Folgende Endpunkte werden im Rahmen der Wirksamkeit unterschieden und flossen in die IQWiG-Berichte ein. Auch wenn diese Informationen für Sie nichts Neues sind, stellen wir sie zusammen, damit auch alle Expertinnen und Experten unter den jeweiligen Endpunkten tatsächlich dasselbe verstehen.

1.1 Ansprechen auf die Behandlung (Response)

Von einem Ansprechen wird gesprochen, wenn die Anzahl bzw. das Ausmaß der durch die Patientin oder den Patienten berichteten depressiven Symptome halbiert werden kann, d. h. eine 50%ige Verbesserung auf der Depressionsskala erzielt wird.

1.2 Erreichen von Symptombfreiheit (volle Remission)

Symptombfreiheit bedeutet, dass die Symptome so deutlich reduziert werden, dass die Patientin / der Patient nicht mehr die Kriterien einer depressiven Episode erfüllt, d. h. einen bestimmten Schwellenwert auf einer Depressionsskala unterschreitet.

1.3 Vermeidung eines Rückfalls

Man spricht davon, dass ein Rückfall durch eine Behandlung vermieden wird, wenn ein(e) Patient(in) innerhalb von bis zu 12 Monaten nach Symptombfreiheit (Remission) nicht wieder die volle depressive Symptomatik zeigt.

2 Unerwünschte Wirkungen (Nebenwirkungen) der Depressionsbehandlung

Nebenwirkungen sind die unerwünschten Wirkungen einer Behandlung. In klinischen Studien werden Nebenwirkungen meist als Aggregat untersucht, in das verschiedene unerwünschte Ereignisse einfließen. Aufgrund der Datenlage in den Studien zu den Antidepressiva mussten die folgenden Kategorien gebildet werden, die nicht direkt die unterschiedlichen Effekte aller Wirkstoffe umfassen.

2.1 Unerwünschte Wirkungen

Unerwünschte Wirkungen einer Arzneimitteltherapie sind nicht lebensbedrohlich und erfordern keine Einweisung ins Krankenhaus.

2.1.1 Unerwünschte Wirkungen / Ereignisse

Es handelt sich hierbei um Nebenwirkungen wie beispielsweise den Anstieg des Blutdrucks, Übelkeit / Erbrechen, Unruhe / Schlafstörungen etc.

2.1.2 Sexuelle Funktionsstörungen

Die Behandlung mit bestimmten Antidepressiva kann bei Männer und Frauen zu sexuellen Funktionsstörungen führen.

2.2 Schwerwiegende unerwünschte Wirkungen

Schwerwiegende unerwünschte Wirkungen sind unerwünschte Wirkungen, die lebensbedrohlich sind, zu schweren Behinderungen führen oder eine Krankenhauseinweisung erforderlich machen.

2.2.1 (Versuchter) Suizid

Gedanken an Selbsttötung und Selbsttötungsversuche gehören zu den möglichen Symptomen einer bestehenden Depression. Jedoch können antriebssteigernde Antidepressiva auch das Risiko einer Selbsttötung bzw. eines Selbsttötungsversuch erhöhen.

2.2.2 Sonstige schwerwiegende unerwünschte Wirkungen

Jede unerwünschte Wirkung eines Medikamentes kann zu einem schwerwiegenden unerwünschten Ereignis werden, wenn sich die Wirkung so verschlimmert, dass sie lebensbedrohlich wird, zur Krankenhauseinweisung bzw. zu schwerer Behinderung oder gar zum Tod führt.

2.3 Therapieabbruch aufgrund unerwünschter Wirkungen

Von einem Therapieabbruch spricht man, wenn eine Behandlung mit Antidepressiva aufgrund unerwünschter Nebenwirkungen des Medikamentes vorzeitig beendet wird.

3 Begleitende Auswirkungen auf die Lebensqualität

Antidepressiva können die Lebensqualität depressiver Patientinnen und Patienten beeinflussen.

3.1 Positive Auswirkungen auf soziale Funktionsbereiche

Verbesserung der Teilnahme am sozialen Leben und am Familienleben; Beruf und Freizeitaktivitäten können wieder besser ausgeübt werden.

3.2 Verringerung von Angst

Gegebenenfalls können Angstzustände / Ängste verringert werden.

3.3 Verringerung von Schmerzen

Mögliche Verringerung von Schmerzen (unterschiedlicher Art und Ursprungs).

3.4 Verbesserung / Erhaltung der intellektuellen Leistungsfähigkeit

Möglicherweise Verbesserung / Erhaltung der Denk- bzw. Konzentrationsfähigkeit.

Anhang C – Individuelle Werte der Befragung der Patienten und der Experten

Patienten	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	MW	SD
Wirksamkeit	0,540	0,327	0,272	0,761	0,463	0,500	0,620	0,559	0,715	0,464	0,474	0,189	0,490	0,169
Ansprechen	0,379	0,209	0,086	0,582	0,101	0,386	0,439	0,428	0,522	0,338	0,138	0,138	0,312	0,172
Remission	0,129	0,100	0,077	0,087	0,070	0,035	0,070	0,088	0,135	0,050	0,072	0,015	0,077	0,035
Vermeidung eines Rückfall	0,032	0,018	0,109	0,092	0,292	0,078	0,111	0,042	0,058	0,075	0,264	0,036	0,101	0,088
Vermeidung von Nebenwirkungen	0,163	0,260	0,067	0,073	0,056	0,073	0,056	0,089	0,067	0,255	0,053	0,054	0,106	0,077
Vermeidung unerwünschter Ereignisse	0,020	0,130	0,010	0,008	0,007	0,015	0,045	0,074	0,056	0,128	0,005	0,005	0,042	0,046
sexuelle Funktionsstörungen	0,015	0,032	0,002	0,002	0,001	0,002	0,011	0,009	0,028	0,018	0,001	0,001	0,010	0,011
sonstige unerwünschte Ereignisse	0,005	0,097	0,007	0,006	0,006	0,013	0,034	0,065	0,028	0,109	0,004	0,005	0,032	0,038
Vermeidung schwerwiegender unerwünschter Ereignisse	0,143	0,130	0,057	0,065	0,049	0,058	0,011	0,015	0,011	0,128	0,047	0,049	0,064	0,046
(versuchter) Suizid	0,071	0,065	0,010	0,022	0,007	0,052	0,002	0,004	0,008	0,021	0,024	0,024	0,026	0,024
sonstige schwerwiegende unerwünschte Ereignisse	0,071	0,065	0,048	0,043	0,042	0,006	0,009	0,011	0,003	0,106	0,024	0,024	0,038	0,031
Auswirkungen auf die Lebensqualität	0,297	0,413	0,661	0,166	0,481	0,427	0,324	0,352	0,218	0,281	0,474	0,757	0,404	0,173
soziale Funktionsfähigkeit	0,026	0,060	0,175	0,019	0,194	0,079	0,058	0,159	0,128	0,049	0,141	0,054	0,095	0,061
Verringerung von Angst	0,210	0,060	0,093	0,024	0,049	0,205	0,215	0,017	0,028	0,145	0,055	0,414	0,126	0,118
Verringerung von Schmerz	0,045	0,146	0,024	0,099	0,017	0,096	0,018	0,026	0,011	0,013	0,029	0,140	0,055	0,051
Verbesserung der kognitiven Funktionsfähigkeit	0,015	0,146	0,369	0,024	0,223	0,048	0,033	0,150	0,051	0,074	0,249	0,148	0,128	0,109
MW: arithmetisches Mittel (Mittelwert); SD: Standardabweichung														

Experten	E1	E2	E3	E4	E5	E6	E7	MW	SD
Wirksamkeit	0,301	0,742	0,740	0,472	0,779	0,777	0,759	0,653	0,189
Ansprechen	0,020	0,039	0,161	0,040	0,063	0,039	0,087	0,064	0,048
Remission	0,209	0,545	0,536	0,292	0,317	0,606	0,546	0,436	0,158
Vermeidung eines Rückfalls	0,072	0,158	0,042	0,140	0,399	0,132	0,126	0,153	0,116
Vermeidung von Nebenwirkungen	0,072	0,055	0,056	0,084	0,147	0,057	0,068	0,077	0,033
Vermeidung unerwünschter Ereignisse	0,058	0,046	0,006	0,008	0,118	0,006	0,051	0,042	0,041
sexuelle Funktionsstörungen	0,012	0,007	0,001	0,002	0,012	0,003	0,010	0,007	0,005
sonstige unerwünschte Ereignisse	0,046	0,040	0,005	0,006	0,106	0,003	0,041	0,035	0,036
Vermeidung schwerwiegender unerwünschter Ereignisse	0,014	0,009	0,050	0,075	0,029	0,051	0,017	0,035	0,024
(versuchter) Suizid	0,002	0,001	0,045	0,068	0,004	0,046	0,009	0,025	0,027
sonstige schwerwiegende unerwünschte Ereignisse	0,012	0,008	0,006	0,008	0,026	0,005	0,009	0,011	0,007
Auswirkungen auf die Lebensqualität	0,626	0,203	0,204	0,444	0,074	0,166	0,172	0,270	0,194
Verbesserung der sozialen Funktionsfähigkeit	0,389	0,121	0,051	0,214	0,010	0,021	0,068	0,125	0,136
Verringerung von Angst	0,087	0,014	0,104	0,098	0,013	0,069	0,017	0,057	0,041
Verringerung von Schmerz	0,034	0,015	0,038	0,075	0,006	0,041	0,014	0,032	0,023
Verbesserung der kognitiven Funktionsfähigkeit	0,116	0,051	0,011	0,058	0,046	0,036	0,073	0,056	0,033
MW: arithmetisches Mittel (Mittelwert); SD: Standardabweichung									

Anhang D – Individuelle Werte der Befragung der Patienten und der Experten in der verkürzten Hierarchie

Patienten	P1	P2	P3	P4	P5	P6	P7	P9	P10	P11	P12	P13	MW	SD
Ansprechen auf die Behandlung	0,714	0,540	0,757	0,796	0,324	0,804	0,796	0,722	0,742	0,799	0,467	0,809	0,689	0,159
Remission	0,143	0,297	0,189	0,125	0,602	0,122	0,079	0,227	0,203	0,105	0,467	0,094	0,221	0,162
Therapieabbruch aufgrund unerwünschter Wirkungen	0,143	0,163	0,054	0,079	0,075	0,074	0,125	0,051	0,055	0,096	0,067	0,097	0,090	0,036
MW: arithmetisches Mittel (Mittelwert); SD: Standardabweichung														

Experten	E1	E2	E3	E4	E5	E6	E7	MW	SD
Ansprechen auf die Behandlung	0,078	0,045	0,047	0,057	0,097	0,058	0,138	0,074	0,034
Remission	0,750	0,770	0,775	0,218	0,264	0,791	0,790	0,623	0,261
Therapieabbruch aufgrund unerwünschter Wirkungen	0,171	0,185	0,178	0,725	0,639	0,151	0,072	0,303	0,263
MW: arithmetisches Mittel (Mittelwert); SD: Standardabweichung									