Aus dem Institut für Sozialmedizin und Epidemiologie

der Medizinischen Fakultät der Universität Lübeck

Direktor: Prof. Dr. med. Alexander Katalinic

und

dem Institut für Allgemeinmedizin

der Medizinischen Fakultät der Universität Hamburg

Direktor: Prof. Dr. med. Martin Scherer

Entspricht die zeitliche Entwicklung des Wissensstandes dem Erkenntnisbedarf zur empirisch fundierten Entscheidung über die Anwendung von Gesundheitstechnologien?

INAUGURALDISSERTATION

zur Erlangung der Doktorwürde

der Universität zu Lübeck

- Aus der Sektion Medizin -

vorgelegt von

Philipp Georg Heinrich Storz-Pfennig

(geb. Storz)

aus Dresden

Lübeck 2014

I. Berichterstatter: Prof. Dr. med. Martin Scherer

II. Berichterstatter/in: Priv.-Doz. Dr. med. Matthias Nagel

Tag der mündlichen Prüfung: 14.01.2015

Zum Druck genehmigt. Lübeck, den 14.01.2015

– Promotionskommission der Sektion Medizin –

Inhaltsverzeichnis

Tabellenverzeichnis **5**

Abbildungsverzeichnis **6**

Abkürzungsverzeichnis **7**

I. Einleitung und Fragestellung **8**

I.1 Hintergrund 8
I.2 Spezifischer Problemkontext 11
I.3 Konkretisierung der Fragestellung 13

II. Material und Methoden **15**

II.1 Datengrundlage 16

II.1.1 Auswahl 16
II.1.2 Patienten, Interventionen/Kontrollinterventionen und Zielgrößen in der zugrunde liegenden systematischen Übersichtsarbeit 17
II.1.3 Studien in der zugrunde liegenden systematischen Übersichtsarbeit 18
II.1.4 Besonderheiten bei der Verwendung des Datensatzes der zugrunde liegenden systematischen Übersichtsarbeit für die vorliegende Untersuchung 18

II.2 Bewertung der Erkenntnisgewinnung im Zeitverlauf 20

II.2.1 Zielgrößen 20
II.2.2 Ermittlung des Erkenntnisstandes zu einem Zeitpunkt 21
II.2.2.1. Mixed Treatment Comparison (MTC: Lu & Ades 2004, Wells 2009) 22
II.2.2.2. Ergebnisse aus der MTC zu Behandlungseffekten 24
II.2.2.3. Gegenüberstellung direkter und gemischter Vergleiche 25
II.2.3 Verbleibende Erkenntnisunsicherheit und Analyse des Zeitverlaufs 26

II.3 Bewertung einer alternativen Erkenntnisgewinnung im Zeitverlauf 29

II.3.1 Bewertung der Studien im Hinblick auf weiteren Erkenntnisgewinn 29
II.3.2 Durchführung und Auswertung der Analyse 30

II.4 Verwendete Analyse- und Datenverarbeitungsprogramme 32

III. Ergebnisse **33**

III.1 Erkenntnisgewinnung im Zeitverlauf 33

III.1.1 In die Analyse eingeschlossene Studien-, Patienten- und Wirkstoffzahlen 33
III.1.2 Verlauf der verbleibenden Erkenntnisunsicherheit 35
III.1.3 Wahrscheinlichkeit, die besten Wirkstoffe identifiziert zu haben im Verlauf 37

III.1.4 Gegenüberstellung der Ergebnisse direkter und gemischter Vergleiche 44

III.2 Faktischer und alternativer Erkenntnisgewinn im Zeitverlauf 45

III.2.1 In die Analyse eingeschlossene Studien-, Patienten- und Wirkstoffzahlen 45
III.2.2 Vergleichender Verlauf der verbleibenden Erkenntnisunsicherheit 46
III.2.3 Ergebnisse für einzelne Wirkstoffe 49
III.2.4 Ermittlung des alternativen Verlaufs der Erkenntnisgewinnung 49

III.3 Ergebniszusammenfassung 53

IV. Diskussion 54

IV.1 Limitationen und Ergebnisse anderer Analysen 55

IV.1.1 Datengrundlage, Indikations- und Interventionskontext 55
IV.1.2 Erkenntnisstände und Zeitverlauf der Erkenntnisgewinnung 59

IV.2 Fazit und Ausblick 64

V. Zusammenfassung 67

VI. Literaturverzeichnis 70

VII. Anhänge 74

VII.1 Programm zur MTC-Analyse 74
VII.2 Anzahl der Studien, Patienten und Wirkstoffe im Zeitverlauf 75
VII.3 Weitere Ergebnisse zur Erkenntnisgewinnung im Zeitverlauf 78
VII.4 Gegenüberstellung direkter und indirekter/gemischter Vergleiche 80
VII.5 Ergebnisse direkt vergleichender Meta-Analysen im Zeitverlauf 81
VII.6 Ergebnisse indirekter/gemischter Vergleiche im Zeitverlauf 90
VII.7 Ergebnisse indirekter/gemischter Vergleiche im Zeitverlauf (alternativer Verlauf der Erkenntnisgewinnung) 93
VII.8 Weitere Ergebnisse zur Erkenntnisgewinnung im Zeitverlauf (alternativer Verlauf der Erkenntnisgewinnung) 96

VIII. Danksagungen 97

IX. Lebenslauf 98

Tabellenverzeichnis

Tabelle 1 Anzahl einbezogener Studien, Patienten und Wirkstoffe (betreffend der Zielgröße korrigiertes Therapieansprechen) 34

Tabelle 2 Ränge der Wirkstoffe in Bezug auf die Wahrscheinlichkeit, dass es sich bei dem jeweiligen Wirkstoff um den jeweils besten handelt im Zeitverlauf (erste 4 Ränge, alle Zielgrößen) 40

Tabelle 3 Anzahl einbezogener Studien, Patienten und Wirkstoffe (Vergleich der faktischen und alternativen Studienreihenfolge, Zielgröße korrigiertes Therapieansprechen) 45

Tabelle 4 Informationswerte (W-Werte) für Wirkstoffkombinationen (bezogen auf 10.000 fiktive Behandlungsepisoden), Publikation und Verwendung von Studien im Zeitverlauf (absteigend sortiert nach dem W-Wert 2002) 52

Tabelle 5 Anzahl einbezogener Studien, Patienten und Wirkstoffe (betreffend der Zielgröße Therapieansprechen) 75

Tabelle 6 Anzahl einbezogener Studien, Patienten und Wirkstoffe (betreffend der Zielgröße Therapieabbruch) 76

Tabelle 7 Anzahl einbezogener Studien, Patienten und Wirkstoffe (betreffend der Zielgröße korrigiertes Therapieansprechen) 77

Tabelle 8 Ergebnisse der kumulativen MTC-Meta-Analyse und der Informationswertermittlung (Zielgröße korrigiertes Therapieansprechen) 90

Tabelle 9 Ergebnisse der kumulativen MTC-Meta-Analyse und der Informationswertermittlung (Zielgröße Therapieansprechen) 91

Tabelle 10 Ergebnisse der kumulativen MTC-Meta-Analyse und der Informationswertermittlung (Zielgröße Therapieabbruch) 92

Tabelle 11 Ergebnisse der kumulativen MTC-Meta-Analyse und der Informationswertermittlung (Zielgröße korrigiertes Therapieansprechen) – alternativer Verlauf der Erkenntnisgewinnung 93

Tabelle 12 Ergebnisse der kumulativen MTC-Meta-Analyse und der Informationswertermittlung (Zielgröße Therapieansprechen) – alternativer Verlauf der Erkenntnisgewinnung 94

Tabelle 13 Ergebnisse der kumulativen MTC-Meta-Analyse und der Informationswertermittlung (Zielgröße Therapieabbruch)) – alternativer Verlauf der Erkenntnisgewinnung 95

Tabelle 14 In die Ermittlung des alternativen Verlaufs der Erkenntnisgewinnung einbezogene Studien mit Publikationsjahr (2003 und später) und Jahr der Verwendung bei der Ermittlung des alternativen Verlaufs der Erkenntnisgewinnung 96

Abbildungsverzeichnis

Abbildung 1 Wert des maximal möglichen, weiteren Erkenntnisgewinns (V-Werte) 35

Abbildung 2 Erwartete Erfolgsraten (korrigiertes Therapieansprechen, E_{max}, $E_{aktuell}$) pro 10.000 Episoden 37

Abbildung 3 Wahrscheinlichkeit, dass der beste Wirkstoff einer Gruppe von Wirkstoffen bestimmter Größe (1-6 Wirkstoffe) angehört im Zeitverlauf (Zielgröße korrigiertes Therapieansprechen) 38

Abbildung 4 Wahrscheinlichkeit, dass es sich bei einem jeweiligen Wirkstoff um den besten handelt, im Zeitverlauf (Zielgröße korrigiertes Therapieansprechen) 41

Abbildung 5 Ereignisrate (T-Werte) der Wirkstoffe im Zeitverlauf (Zielgröße korrigiertes Therapieansprechen) 42

Abbildung 6 Ereignisrate (T-Werte) 2009 und Veränderung der Ereignisrate (T-Werte) der Wirkstoffe im Zeitverlauf (2002-2009) (Zielgröße korrigiertes Therapieansprechen, die 2009 besten Wirkstoffe sind farblich hervorgehoben.) 43

Abbildung 7 Ereignisrate (T-Werte) ausgewählter Wirkstoffe im Zeitverlauf (Zielgröße korrigiertes Therapieansprechen) 43

Abbildung 8 Gegenüberstellung gemischter (Punktschätzer) und direkter Vergleiche (Konfidenzintervall) für die Zielgröße Therapieabbruch (rote Punkte: gemischter Wert außerhalb des 95%-Konfidenzintervalles für den direkten Schätzer) 44

Abbildung 9 Wert des maximal möglichen, weiteren Erkenntnisgewinns (Vergleich der faktischen und alternativen Studienreihenfolge, Zielgröße korrigiertes Therapieansprechen 47

Abbildung 10 Erwartete Erfolgsraten (korrigiertes Therapieansprechen) pro 100.000 Episoden (perfekte Information T: E(max) und aktuelle Information T: E(aktuell)) in der Gegenüberstellung von faktischem und alternativem Verlauf 48

Abbildung 11 Ereignisrate (T-Werte) ausgewählter Wirkstoffe im Zeitverlauf, vergleichend faktischem gegenüber alternativem Erkenntnisgewinn im Zeitverlauf (Zielgröße korrigiertes Therapieansprechen) 49

Abbildung 12 Bewertung von Wirkstoffvergleichen in Studien und Anzahl verfügbarer Studien (alternativer Verlauf der Erkenntnisgewinnung) 50

Abbildung 13 Erwartete Erfolgsraten (Therapieansprechen, E_{max}, $E_{aktuell}$) pro 10.000 Episoden 78

Abbildung 14 Erwartete Erfolgsraten (Therapieabbruch, E_{max}, $E_{aktuell}$) pro 10.000 Episoden 78

Abbildung 15 Wahrscheinlichkeit, dass der beste Wirkstoff einer Gruppe von Wirkstoffen bestimmter Größe (1-6 Wirkstoffe) angehört im Zeitverlauf (Zielgröße Therapieansprechen) 79

Abbildung 16 Wahrscheinlichkeit, dass der beste Wirkstoff einer Gruppe von Wirkstoffen bestimmter Größe (1-6 Wirkstoffe) angehört im Zeitverlauf (Zielgröße Therapieabbruch) 79

Abbildung 17 Gegenüberstellung gemischter (Punktschätzer) und direkter Vergleiche (Konfidenzintervall) für die Zielgröße Therapieansprechen 80

Abkürzungsverzeichnis

AWMF	Arbeitsgemeinschaft wissenschaftlicher medizinischer Fachgesellschaften
BUPRO	Bupropion (Arzneimittelwirkstoff)
CGI	Clinical Global Impression
CITAL	Citalopram (Arzneimittelwirkstoff)
DIMDI	Deutsches Institut für medizinische Dokumentation und Information
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders, 4. Auflage
DULOX	Duloxetin (Arzneimittelwirkstoff)
ESCIT	Escitalopram (Arzneimittelwirkstoff)
F32	(ICD-10-)Klassifikation für eine depressive Episode
F33	(ICD-10-)Klassifikation für eine rezidivierende depressive Störung
FLUOX	Fluoxetin (Arzneimittelwirkstoff)
FLUVO	Fluvoxamin (Arzneimittelwirkstoff)
G-BA	Gemeinsamer Bundesausschuss
HDRS	Hamilton Depression Rating Scale
HTA	Health Technology Assessment
ICD-10	Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme, 10. Revision
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
lnOR	Natürlicher Logarhitmus (ln) der OR (Odds-Ratio)
MADRS	Montgomery-Åsberg Depression Rating Scale
MILNA	Milnacipran (Arzneimittelwirkstoff)
MIRTA	Mirtazapin (Arzneimittelwirkstoff)
MTC	Mixed-Treatment-Comparison
OR	Odds-Ratio
PAROX	Paroxetin (Arzneimittelwirkstoff)
RCT	Randomized controlled trial (randomisierte, kontrollierte Interventionsstudie)
REBOX	Reboxetin (Arzneimittelwirkstoff)
SE	Standard Error
SERTR	Sertralin (Arzneimittelwirkstoff)
VBA	Visual Basic of Applications
VENLA	Venlafaxin (Arzneimittelwirkstoff)
VOI	Value of Information (Informationswert)

I. Einleitung und Fragestellung

I.1 Hintergrund

Die Bewertung therapeutischer oder diagnostischer Interventionen zur Verhütung oder Bekämpfung von Erkrankungen nimmt in den Gesundheitswissenschaften einen bedeutenden Platz ein. Die Gewinnung verlässlicher Erkenntnisse über Nutzen und Schaden medizinischer Interventionsmöglichkeiten kann sowohl zu besseren, alltäglich zu treffenden Entscheidungen in konkreten klinisch-medizinischen Behandlungssituationen, als auch zu besseren Entscheidungen über die Anwendbarkeit dieser Interventionen auf der Ebene von Gesundheits- oder Krankenversicherungssystemen beitragen.

Bis weit in das 20. Jahrhundert hinein waren, zumindest nach heutigen Maßstäben eines naturwissenschaftlich geprägten Kausalitätsverständnisses, nur wenige wirksame, therapeutische Möglichkeiten bekannt. Eine generell illustrative, wenn auch nicht auf historische Genauigkeit abzielende, Charakterisierung des Verständnisses der medizinischen Möglichkeiten kann so formuliert werden: "Fifty years ago, physicians were little more than diagnosticians, their activities being essentially 'limited to identification of … illness, the prediction of the likely outcome, and then the guidance of the patient and his family while the illness ran its full, natural course" (Weisbrod 1991: 526).

Spätestens ab etwa der Mitte des 20. Jahrhunderts wurden jedoch, auf der Grundlage eines biologisch-naturwissenschaftlich geprägten Krankheitsverständnisses eine zunehmende Zahl therapeutischer Technologien, nicht zuletzt eine inzwischen große Zahl von Arzneimitteln, entwickelt, die einen Anspruch auf wirksame Bekämpfung vieler Erkrankungen erhoben und erheben. In Verbindung mit demographischen und allgemeinen Veränderungen der Lebensumstände (möglicherweise im Zeitverlauf auch aufgrund bereits erzielter medizinisch-therapeutischer Erfolge) hat sich im selben Zeitraum das Erkrankungsspektrum von typischen Akut- zu chronischen Erkrankungen (insbesondere z.B. Herz-Kreislauf-Erkrankungen, Diabetes, Muskel-Skelett- und psychischen Erkrankungen) verschoben. Zumindest gilt dies für, im globalen Vergleich, reiche Gesellschaften.

Diese Entwicklungen haben säkular die gesellschaftliche bzw. sozioökonomische Bedeutung der Medizin, gemessen z. B. an den globalen Maßstäben von Arztzahlen oder Anteilen der Gesundheitsausgaben am Sozialprodukt, erheblich wachsen lassen. Ein historisch unbekannt hoher Bedarf trifft auf eine historisch ebenfalls unbekannte Fülle an Behandlungsmöglichkeiten. Für viele Erkrankungen, bzw. Stadien oder Erkrankungszustände im Rahmen eines generellen Krankheitsbildes, stehen gegenwärtig u. U. eine Vielzahl von

Behandlungsmöglichkeiten bereit, deren jeweils einzelne – vor allem aber vergleichende – Bewertung eine erhebliche Herausforderung darstellt.

Beruhend auf einer Reihe von Gründen, insbesondere eines zu vermutenden Selbstvergewisserungsbedarfes der medizinischen Profession sowie der offenkundigen generellen ökonomischen Bedeutung der Medizin in der Gegenwart, sind im Laufe des (wesentlich des späteren) 20. Jahrhunderts eine Reihe von Initiativen auf klinischen, wissenschaftlichen und politischen Felder ergriffen worden, um das generelle Problem der Erkenntnisgewinnung zur Anwendung angemessener Therapien anzugehen. Als emblematischer Bezugspunkt wird hier auf Archibald Cochranes "Effectiveness and Efficiency: random reflections on health services" (Cochrane 1972) verwiesen, in der dieser eine systematische Erforschung der Wirksamkeit von Therapieverfahren – insbesondere basiert auf randomisierten Studien – forderte. Auch in der Gegenwart verkörpert insbesondere die "Cochrane Collaboration" den Anspruch therapeutische Entscheidung so weit wie möglich auf gesicherter, wissenschaftlicher Grundlage treffen zu wollen. Dieses Bestreben kann auch als Ausgangspunkt der "evidenzbasierten Medizin" (*evidence based medicine*) angesehen werden, die einen besonderen, zusätzlichen Schwerpunkt auf die Anwendbarkeit der wissenschaftlichen Erkenntnisse in der klinischen Praxis legt (Sackett 1996). Demgegenüber erfolgt die systematische Bewertung von "Gesundheitstechnologien" (hiermit werden vornehmlich, wenn auch nicht ausschließlich, spezifische Therapie- und Diagnostikverfahren bezeichnet) im Rahmen von Gesundheitstechnolgiebewertungen (*Health Technology Assessments*, HTA) mit der Zielrichtung die Eignung bestimmter Verfahren auf der Ebene von Gesundheits- oder Krankenversicherungssystemen zu bewerten. Solche Verfahren werden mittlerweile in vielen Gesundheitssystemen, auch in Deutschland (z. B. durch das „Deutsche Institut für medizinische Dokumentation und Information“ – DIMIDI, sowie durch das „Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen“ – IQWiG), auf sehr unterschiedlicher Grundlage, Reichweite und mit unterschiedlicher Geltung, Verbindlichkeit und Fokussierung angewandt. Auch bei Gesundheitstechnologiebewertungen bildet die systematische Bewertung der therapeutischen Wirksamkeit (bzw. des Nutzens und Schadens) ein Kernelement, ohne das die Bewertung ökonomischer, sozialer, ethischer u. a. Sachverhalte von möglicher Relevanz bei der Bewertung bedeutungslos bliebe. Dies gilt ebenso im Rahmen der Leitlinienentwicklung, die auch als ein Instrument der evidenzbasierten Medizin angesehen werden kann (für Deutschland ist hier insbesondere das Leitlinienprogramm der Arbeitsgemeinschaft wissenschaftlicher medizinischen Fachgesellschaften – AWMF – einschlägig) sowie generell bei der Erstellung und Publikation systematischer Übersichtsarbeiten/Reviews – die wiederum häufig

als Bestandteile bzw. Grundlagen evidenzbasierter Medizin, von HTAs und Leitlinien anzutreffen sind.

Ein all diesen Bestrebungen bzw. resultierden Produkten gemeinsames Kernelement stellt die systematische Sichtung, Bewertung und zusammenfassende Analyse geeigneter, wissenschaftlicher Studien dar. Das systematische Vorgehen besteht, neben der Formulierung einer klaren Fragestellung, klarer Kriterien für die zu untersuchenden Patientengruppen, die Interventionen und Zielgrößen der Analyse insbesondere darin, möglichst keine in diesem Rahmen relevante Studie bzw. Publikation zu übersehen, da andernfalls eine Verzerrung des Ergebnisses (*bias*) befürchtet werden muss. Zudem sollten möglichst nur Studien der höchsten Ergebnissicherheit herangezogen werden. Gegenwärtig ist die randomisierte, kontrollierte Interventionsstudie (*randomized controlled trial*, RCT) hier als der methodische "Goldstandard" für die adäquate Bewertung therapeutischer Interventionen etabliert und daher als Maßstab der Eignung von Studien für die Bewertung der therapeutischen Wirksamkeit im Rahmen systematischer Bewertungsverfahren anzusehen. In nahezu allen Fällen bedarf es einer speziellen Begründung, weshalb solche Studien nicht benötigt werden, nicht vorliegen oder zukünftig durchgeführt werden könnten bzw. weshalb auch andere Studien als zur Bewertung hinreichend erachtet, oder überhaupt zur Bewertung herangezogen werden sollten. Die Bevorzugung von RCT ist wesentlich darin begründet, dass nur mit Hilfe einer zufälligen Zuteilung auf die Vergleichsgruppen (Randomisierung) in einer vergleichenden Studie gewährleistet werden kann, dass auch unbekannte Determinanten des Therapieerfolges zwischen den Vergleichsgruppen (zumindest im Grenzwert, d. h. bei einer genügend großen Zahl an Studienteilnehmern) gleichverteilt sind und somit nicht z. B. fälschlich günstigere Ergebnisse einer Therapie zugeschrieben werden, obwohl diese in Wirklichkeit lediglich auf günstigeren prognostischen Faktoren der Studiengruppe beruht, die diese Therapie erhielt.

Zur systematischen Bewertung gehört über die Berücksichtigung ergebnissicherer Studien hinaus jedoch auch eine weitere (methodische) Qualitätsbewertung der ggf. zu berücksichtigenden bzw. einzuschließenden Studien, die deren Potenzial für eine Reihe von Ergebnisverzerrungen, trotz Randomisierung, zu erfassen sucht und damit wichtige Anhaltspunkte für die Verläßlichkeit der Ergebnisse liefert (Higgins & Green 2011). Ergebnisunterschiede zwischen Vorgehensweisen können z. B. dadurch ein verzerrtes Bild von den tatsächlichen, auf die Eigenschaften der verglichenen Verfahren zurückzuführende Unterschiede darstellen, dass die Gruppen zu Beginn unterschiedlich waren (*selection bias*), unterschiedlich behandelt wurden (*performance bias*), die Erhebung der Zielgrößen und deren Publikation unterschiedlich erfolgte (*detection/reporting bias*).

Schließlich kann, als ein weiterer Analyseschritt, eine quantifizierte Zusammenfassung der Ergebnisse in Form eine Meta-Analyse versucht werden, mit deren Hilfe die ggf. voneinander abweichenden Ergebnisse aus einer Anzahl vergleichender Studien zur Berechnung eines Gesamteffektschätzers zusammengefasst werden können. Hierbei sind jedoch spezifische (Homogenitäts-)Kriterien zu beachten bzw. im Nachgang zu untersuchen (Higgins & Green 2011; Deeks 2008). Meta-Analytische Methoden, die wesentlich durch die konsistente Anwendung von Meta-Analysen durch die Arbeit der Cochrane Collaboration weite Verbreitung gefunden haben, sind insbesondere dann notwendig, oder zumindest sehr nahe gelegt, wenn eine größere Zahl hinreichend in Bezug auf inhaltliche (Interventionen, Zielgrößen und Patientengruppen bzw. Behandlungsindikationen) vergleichbare Studien vorliegen. Insbesondere wenn einzelne Studien im Hinblick auf die Absicherung gegen Zufallsfehler (besonders mangelnde statistische "Power" zum Nachweis von Effektunterschieden) zu klein sind, um ein deutliches Bild der relativen Effekte der Interventionen zu geben, können Meta-Analysen u. U. hier einen in keiner der einzelnen Studien allein sichtbaren Effekt zeigen. Dies kann, auch im Hinblick auf eine Betrachtung der Erkenntnisgewinnung im Zeitverlauf hilfreich sein („kumulative Meta-Analyse“). So wurde z. B. (retrospektiv) gezeigt, dass bereits 1981 der Nachweis des Nutzens der Therapie mit Beta-Blockern in der Sekundärprävention nach Myokardinfarkt im Grundsatz vorlag (Egger 2001, Lau 1992). Durch eine prospektive, zeitnah wiederholte meta-analytische Synthese vorliegender Erkenntnisse im Zeitverlauf kann frühzeitig ein erreichter Wissensstand ermittelt werden. In idealisierter Betrachtung kann dadurch verhindert werden, dass Therapien aufgrund von veralteten Erkenntnissen in der Praxis angewendet werden. Dies kann bedeuten, dass bisher praktizierte Verfahren nicht mehr angewandt werden, da sie anderen, neueren Therapienformen unterlegen sind. Es kann aber ebenso dazu führen, dass eine Aufnahme von neuen Therapieverfahren unterbleibt, insofern diese bestehenden Verfahren nicht überlegen bzw. unterlegen sind.

I.2 Spezifischer Problemkontext

Wie oben dargestellt kann die Bewertung therapeutischer Möglichkeiten gegenwärtig auf in methodischer, institutioneller und organisatorischer Hinsicht entwickelte Strukturen und Prozesse (Breite Anerkennung der Notwendigkeit ergebnissicherer Studien, evidenzbasierter Leitlinien, HTA) aufbauen. Damit ist jedoch keinesfalls gewährleistet, dass heute etablierte und neu angebotene ("innovative") therapeutische Möglichkeiten tatsächlich umfassend und valide einer Prüfung ihrer Eignung in den genannten Kontexten, sei es des HTA und der Nutzenbewertung, der Erarbeitung klinischer Leitlinien bzw. zur direkten, alltäglichen Entscheidung in der Behandlungssituation umstandslos zugeführt werden. Eher wird häufig das Gegenteil der Fall sein und wesentliche Defizite weiter bestehen

bleiben, die schon bei der Begründung bzw. Entwicklung der genannten Institutionen bzw. Organisationen und Methodiken vorlagen. Der wesentliche, zunächst sichtbare, Grund liegt nach wie vor im Fehlen der adäquaten Studien, die überhaupt der Bewertung zugrunde gelegt werden können. Dies ist nahezu universell, oder zumindest überaus häufig, in den Schlussfolgerungen und Diskussionsabschnitten entsprechender Publikationen und bewertender Berichte nachzulesen. Typische Probleme stellen die nicht selten kurze Beobachtungszeit, die Analyse selektiver Patientenpopulationen (welche nicht der Gruppe derjenigen entsprechen, die die Therapie in der Versorgung erhalten), die fehlende Erhebung relevanter Zielgrößen u. a. dar. Ein besonderes Problem ist häufig dadurch gegeben, dass keine (direkt) vergleichenden Studien zwischen möglichen Therapiealternativen vorliegen. Insbesondere auch im Arzneimittelbereich liegen häufig vor allem Placebo-kontrollierte Studien vor, die im Rahmen der Arzneimittelzulassung durchgeführt wurden und die eine unmittelbar vergleichende Aussage zu bereits etablierten Arzneimitteltherapien nicht erlauben (z. B. Evans 2011:115 ff., Glaeske 2010:60 ff.).

Neben der übergreifenden Frage welche Ursachen für die Fortdauer dieser Mängel verantwortlich sind (z. B. fehlende Ressourcen der Forschungsförderung oder mangelndes bzw. sogar entgegenstehendes Interesse der Entwickler und Anbieter von Therapieverfahren, z. B. von Arzneimitteln) stellt sich bei praktischen Bewertungsprozessen die Frage, ob und in welchem Maße die Vielzahl ja doch durchgeführter Studien tatsächlich geeignet ist, eine methodisch adäquat ansetzende Nutzenbewertung zu informieren. In diesem Zusammenhang stellt sich auch die Frage, ob die Menge jeweils vorliegender Studien, als Ergebnis von Forschungsprozessen in bestehenden Strukturen, in ihrer Gesamtheit und Abfolge für eine solche Bewertung geeignet ist. Zweifel daran, dass dies einfach bejaht werden kann, erscheinen angebracht. So erscheint es u. a. zweifelhaft, ob Studienautoren von publizierten klinischen Studien vor Durchführung der jeweiligen Studien eine umfassende, systematische Prüfung des Erkenntnisstandes tatsächlich vorgenommen haben – um den Nutzen der geplanten Studie vor deren Beginn sicherzustellen (Clarke & Chalmers 1998, Clarke 2002). Dieser kritische Befund wurde recht kürzlich erneut bestätigt (Clarke 2007) und es lässt sich auch zeigen, dass Studienautoren vorangehende, relevante Arbeiten häufig zumindest nicht zitieren (Robinson & Goodman 2011). Aufgrund dieser Befunde erscheint es als höchst unsicher, ob jeweils durchgeführte Studien zielgenau den Erkenntnisbedarf in Bezug auf spezifische Fragestellungen bzw. Indikationen beheben können.

Gegenwärtig richtet sich eine zunehmende Aufmerksamkeit auf analytische Verfahren, die im Rahmen von Bewertungsprozessen eingesetzt werden können, um Ergebnisunsicherheiten für die Bewertung insgesamt zu quantifizieren und darauf aufbauend den weiteren

Forschungsbedarf zu beschreiben (z. B. Goeree & Levin 2006; Scott 2008). Zudem werden Konzepte der bedingten Erstattung durch Finanzierungsträger, unter der Auflage weiterer Erforschung (*coverage with evidence development*) verschiedentlich zumindest fallweise angewandt (Stafiniski 2010) um die zur endgültigen Entscheidung zur Erstattung von Leistungen in Gesundheitssystemsicht fehlenden Erkenntnisse nach Möglichkeit zu gewinnen.

Zur Bewertung des erwarteten Wertes weiterer Studien bei einem gegebenen Entscheidungsproblem wurden in der Vergangenheit ein statistisch-entscheidungstheroretisches Konzept (Ermittlung des Informationswertes, *value of information*) entworfen (Claxton 2004, Briggs 2006, Meltzer 2011), das für die vorliegende Arbeit herangezogen wurde (vgl. Abschnitt II.2). Die Erkenntnisgewinnung, besonders im Rahmen direkter Vergleichsstudien, findet, ausgehend von einigen grundsätzlichen Erkenntnissen zur Anwendbarkeit der Technologie, die bereits bei deren Zulassung bzw. Einführung vorliegen, gegenwärtig in der Regel in der Zeit statt, während derer die Technologie bereits eine mehr oder minder große Verbreitung erfährt. Hier wird häufig ein Dilemma in Bezug auf den geeigneten Zeitpunkt der Durchführung einer umfassenden Bewertung konstatiert: Eine zu frühe Bewertung basiert auf zu wenigen verlässlichen Erkenntnissen, eine zu späte Bewertung kann u. U. die Verbreitung der Technologie bzw. Vorgehensweise nicht mehr entscheidend beeinflussen. Im Übrigen stellt die Schnelligkeit, mit der Klarheit über den relativen Wert unterschiedlicher Therapiemöglichkeiten bei einer gegebenen Indikation erzielt werden kann, einen hohen Wert für Patientinnen und Patienten dar. Wenigstens wenn tatsächlich ein Unterschied in der Wirksamkeit (Nutzen bzw. Schaden) zwischen Therapiealternativen existiert, erscheint es beinahe unvermeidlich, dass, während die Unklarheit andauert, zumindest ein Teil der Patienten nicht die später als überlegen erwiesene Therapie erhält.

I.3 Konkretisierung der Fragestellung

Die vorliegende Arbeit verfolgt das Ziel, einen Beitrag zur Entwicklung einer Methodik zu liefern, die es zukünftig erlauben könnte, aus einer Anzahl möglicher Studien diejenigen zu identifizieren, die dazu beitragen könnten, schneller Klarheit über den Wert unterschiedlicher Therapieoptionen zu gewinnen. Hierzu werden zwei spezifische Fragen gestellt:

- Wie ist der Verlauf der Erkenntnisgewinnung im Hinblick auf eine Reduktion von Erkenntnisunsicherheit durch im Zeitverlauf erzielte Studienergebnisse zu bewerten?

- Lässt sich ein alternativer Verlauf der Erkenntnisgewinnung konzipieren, dessen Anwendung zu einer schnelleren Reduktion von Erkenntnisunsicherheit im Zeitverlauf geführt haben würde?

Ausdrücklich nicht das Ziel der vorliegenden Arbeit ist es hingegen, den aktuellen Erkenntnisstand in dem exemplarisch ausgewählten Therapiegebiet zu ermitteln oder wiederzugeben. Dessen Kenntnis wird vielmehr, auf Grund der notwendig retrospektiven Anlage der Analyse, bereits prinzipiell vorausgesetzt. Ziel der Arbeit ist darzustellen, wie dieser – zumindest in wichtigen Teilaspekten – erreicht worden ist und ob nicht Verfahren gefunden werden können, mit dem dieser Stand schneller hätte erreicht werden können.

II. Material und Methoden

Die Fragestellung wird exemplarisch an einer therapeutischen Indikation geprüft, für die eine Reihe von Behandlungsmöglichkeiten grundsätzlich zur Verfügung steht. An einem, notwendig retrospektiv vorliegenden Studienkollektiv, das grundsätzlich die zur vergleichenden Bewertung der Therapieoptionen geeigneten Studien umfasst wird die Methodik erprobt. Damit wird gleichzeitig geprüft, ob im exemplarischen Fall Defizite tatsächlich, wie vermutet, vorliegen, ein schnellerer Abbau von Unsicherheit mit demselben Aufwand also möglich gewesen wäre. Das Studienkollektiv wird dabei aus einer systematischen Übersichtsarbeit übernommen, die entsprechend der oben gegebenen Schilderung (vgl. Abschnitt I.1) RCT zu der spezifischen Fragestellung systematisch ermittelt und auswertet und somit als adäquate Grundlage dienen kann (Abschnitt II.1).

Die Beschreibung und Bewertung des Erkenntnisstandes erfordert zunächst eine Festlegung, wie dieser zu bestimmten, unterschiedlichen Zeitpunkten, im Hinblick auf den erwartbaren Nutzen der Therapie für Patienten anhand klinisch relevanter Parameter zu erfassen ist (Abschnitt II.2 insbesondere Abschnitt II.2.2). Daran anschließend wird das angewandte Verfahren zur Ermittlung der jeweils zu einem bestimmten Zeitpunkt verbleibenden Ergebnisunsicherheit bezogen auf diesen Erkenntnisstand beschrieben (Abschnitt II.2.3) und wie dieses im Zeitverlauf bewertet und analysiert werden soll. Abschnitt II.3 beschreibt, wie ein alternativer Verlauf der Erkenntnisgewinnung konzipiert wird und wie die Ergebnisse, vergleichend zum faktischen Verlauf, bewertet werden sollen. In Abschnitt II.4 werden die jeweils verwendeten Analyse- und Datenverarbeitungsprogramme angegeben.

II.1 Datengrundlage

In Abschnitt II.1.1 wird die Auswahl der exemplarisch untersuchten Indikation, der Therapiealternativen und der systematischen Übersichtsarbeit, die die entsprechenden Ergebnisse enthält, beschrieben. Eine genauere Charakterisierung der Patientengruppe, der Therapiealternativen und Zielgrößen erfolgt in Abschnitt II.1.2, während in Abschnitt II.1.3 die Studien näher charakterisiert werden. Abschnitt II.1.4 beschreibt weitere Charakteristika bezogen auf das Anwendungsfeld, die im Rahmen der vorliegenden Arbeit von Bedeutung sind.

II.1.1 Auswahl

Der vorliegenden Arbeit wurde die Fragestellung zur Wirksamkeit der Arzneimitteltherapie bei (unipolar) depressiven Erkrankungsepisoden zu Grunde legt. Studien und Studienergebnisse hierzu werden der systematischen Übersichtsarbeit Cipriani 2009a entnommen, in die direkt vergleichende RCT für 12 neuere antidepressive Arzneimittelwirkstoffe eingeschlossen und deren Wirksamkeit im Bezug die Häufigkeit des Therapieansprechen (Response) und die Häufigkeit des Therapieabbruchs (Dropout), unabhängig von der jeweiligen ggf. aufgezeichneten Begründung, untersucht wurden.

Für die vorliegende, exemplarische Analyse wurde ein Therapiegebiet gesucht für das das Vorliegen vieler unterschiedlicher Therapieoptionen und vieler, auch direkt vergleichender, Studien erwartet werden konnte. Dies ist u. a. bei der medikamentösen Behandlung (unipolarer) Depressionen der Fall. Letztlich entscheidend für die Auswahl war, dass mit der Arbeit Cipriani 2009a bereits eine qualitativ hochwertige, aktuelle, systematische Übersichtsarbeit vorliegt, die (nicht zuletzt mit Hilfe der auch in der vorliegenden Arbeit eingesetzten meta-analytischen Methodik zum simultanen Vergleich einer Reihe von Therapieoptionen. vgl. Abschnitt II.2.2.1) das primäre Ziel verfolgt den oder die bestgeeigneten Wirkstoffe bei der gegebenen Indikation zu identifizieren.

Obwohl die Auswahl jedoch nicht in erster Linie und nicht direkt von epidemiologischen Gesichtspunkten im Hinblick auf die Krankheitslast geprägt war, so wurde mit der exemplarischen Auswahl doch ein bedeutendes Erkrankungsgebiet einbezogen. (Unipolare) Depressionen sind sehr häufige Erkrankungen die, insbesondere bei vorliegendem höherem Schweregrad, eine hohe Belastung für die Betroffenen bedeuten.

II.1.2 Patienten, Interventionen/Kontrollinterventionen und Zielgrößen in der zugrunde liegenden systematischen Übersichtsarbeit

Alle Analysen die im Rahmen der vorliegenden Untersuchung durchgeführt werden, basieren auf den Ergebnissen der in die systematische Übersichtsarbeit von Cipriani 2009a eingeschlossenen Studien. Bei der Arbeit von Cipriani et al.[1] handelt es sich um eine systematische Übersichtsarbeit, in die ausschließlich randomisierte, kontrollierte Studien (RCT) eingeschlossen wurden in denen mindestens zwei Wirkstoffe aus einer Gruppe von zwölf antidepressiven Wirkstoffen miteinander verglichen wurden (Monotherapie). Dabei handelt es sich um die folgenden Wirkstoffe (alphabetische Reihenfolge): Buproprion, Citalopram, Duloxetin, Escitalopram, Fluoxetin, Fluvoxamin, Milnacipran, Mirtazapin, Paroxetin, Reboxetin, Sertralin und Venlafaxin. In dieser Auswahl wurden somit wesentliche, "neuere" antidepressive Arzneimittel(-wirkstoffe) berücksichtigt.

In die Arbeit von Cipriani et al. wurden Studien zur Behandlung einer depressiven Episode entsprechend der Kriterien der ICD-10 (Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme, 10. Revision) - Klasse F32* (bzw. im Rahmen einer Episode im Rahmen einer rezidivierenden Depression, F33*) bzw. entsprechend der Kriterien des DSM-IV (*Diagnostic and Statistical Manual of Mental Disorders, 4. Auflage*) oder entsprechender Kriterienkataloge, die im Rahmen von wissenschaftlichen Untersuchungen Anwendung finden, bei erwachsenen Patienten (18 Jahre oder älter) beiderlei Geschlechts berücksichtigt.

Grundlage der Bewertung waren die Ergebnisse acht Wochen nach Beginn der Therapie. Beim Fehlen von Ergebnissen nach acht Wochen in den Studien bzw. den Studienpublikationen wurden ggf. ersatzweise Werte von minimal sechs bis maximal zwölf Wochen nach Behandlungsbeginn herangezogen. Als Zielgrößen für die Analysen von Cipriani et al. (und auch der vorliegenden Arbeit) wurden der Anteil der jeweils randomisierten Patienten bei denen ein Therapieansprechen (*Response*) beobachtet werden konnte, sowie der Anteil der Patienten die die Therapie abbrachen (Therapieabbruch bzw. das Ausscheiden aus der Studie (*Dropout*)) gewählt. Das Therapieansprechen wurde von Cipriani et al. definiert als Erreichen einer Verbesserung des vor Therapiebeginn erzielten Punktwertes von mindestens 50% auf der Hamilton Depression Rating Scale (HDRS) oder der Montgomery-Åsberg Depression Rating Scale (MADRS). Waren diese Angaben in den

---

[1] Hier liegen zwei Publikationen bzw. publizierte Dokumente vor: Cipriani 2009a (die Vollpublikation der Übersichtsarbeit – einschließlich des Datensatzes, der die relevanten Daten aller Studien enthält, die der Übersichtsarbeit zugrunde liegen) sowie Cipriani 2009b (das Protokoll der Übersichtsarbeit, mit weiteren methodischen Ausführungen).

Studienpublikationen nicht verfügbar, so wurde alternativ die CGI-Bewertungen (*Clinical Global Impression*) herangezogen und die Bewertungen "much improved" sowie "very much improved" im Rahmen der CGI-Einstufung ebenfalls als Therapieansprechen bewertet. Als Therapieabbruch/Dropout wurde jedes Ausscheiden aus der Studie bis zum Messzeitpunkt (acht Wochen), unabhängig von der ggf. beschriebenen Begründung, gezählt (Cipriani 2009a, Cipriani 2009b).

II.1.3 Studien in der zugrunde liegenden systematischen Übersichtsarbeit

Im Ergebnis eines differenzierten, mehrstufigen Such- und Auswahlverfahren durch Cipriani et al., das auch die Suche nach bisher nicht publizierten Studien einschloss, wurden schließlich insgesamt 117 Studien, mit insgesamt 25.928 Teilnehmern, eingeschlossen. In 115 Studien wurden jeweils zwei der in die Analyse eingeschlossenen Wirkstoffe verglichen, in zwei Studien jeweils drei (Fluoxetin, Paroxetin und Sertralin). Ein Überblick über einige Charakteristika des Studienpools und welche Wirkstoffe direkt in Studien verglichen wurden und wie viele Studien zu den jeweils untersuchten Vergleichen vorlagen, sind in Cipriani 2009a enthalten. Es handelte sich ausschließlich um randomisierte, kontrollierte Studien an Patienten mit depressiver Episode (Schwangerschaftsdepressionen wurden ausgeschlossen) in denen die Behandlung in zumindest zwei Studienarmen mit zwei verschiedenen der in die Analyse einbezogenen Wirkstoffen erfolgte und bei denen das Therapieansprechen oder der vorzeitige Therapieabbruch (möglichst beides) als Zielgrößen verwendet werden konnten.

II.1.4 Besonderheiten bei der Verwendung des Datensatzes der zugrunde liegenden systematischen Übersichtsarbeit für die vorliegende Untersuchung

Da der Publikationszeitpunkt für die zentrale Fragestellung der vorliegenden Untersuchung von wesentlicher Bedeutung ist, wurde das Publikationsjahr der Ergebnisse der nach der Arbeit von Cipriani et al. berücksichtigten Studien gesondert erfasst. Dies geschah auf der Grundlage des Datensatzes der Arbeit durch Extraktion der Kürzel der eindeutigen Studienbezeichnung (z. B. das Jahr 2002 aus der Bezeichnung "Fava2002"). Bei Studien ohne Jahreszahl in der Studienbezeichnung wurde von unpublizierten Studien(-ergebnissen) ausgegangen. Diesen wurde für die Analysen der vorliegenden Untersuchung das Jahr 2009 (Publikationsjahr der Arbeit von Cipriani et al.) zugeordnet. Da in der vorliegenden Untersuchung keine unterjährigen Analysen betreffend des Zeitverlaufes angestellt werden, wurde auf eine genauere (unterjährige) Erfassung des Publikationszeitpunktes verzichtet. Die in einzelnen Analysen ggf. technisch notwendige, sukzessive

Berücksichtigung mehrerer, in demselben Jahr publizierter Studien, ist daher nicht aussagekräftig, da die hierbei implizierte Reihenfolge der Publikation von Studien innerhalb eines Jahres arbiträr ist.

Wie oben (vgl. Abschnitt II.1.3) beschrieben, handelte es sich bei der überwiegenden Zahl (115 von 117) der zu berücksichtigenden Studien um zwei-armige Studien, bzw. um Studien die im Rahmen der Analyse von Cipriani et al. sowie der vorliegenden Analyse als zweiarmige Studien zu betrachten waren, da nur die Ergebnisse von zwei Behandlungsarmen einbezogen wurden, da nur in diesen eine Behandlung mit einem der eingeschlossenen Wirkstoffe erfolgte. In zwei Studien wurden jedoch jeweils drei Wirkstoffe (Fluoxetin, Paroxetin und Sertralin) aus der Menge der in der Analyse insgesamt berücksichtigten Wirkstoffe untersucht. Da die Berücksichtigung von mehrarmigen Studien in den Analysen der vorliegenden Untersuchung erhebliche, insbesondere rechentechnische Vorkehrungen erfordert hätte (vgl. Ades 2007) und lediglich die genannten beiden Studien betrifft, wurden die beiden dreiarmigen Studien zum Zweck der vorliegenden Analyse in jeweils drei "fiktive" zweiarmige Studien umgewandelt bzw. aufgespalten und deren Ergebnisse als solche in die Analyse aufgenommen. Daraus resultiert zwar die mehrfache Verwendnung von Ergebnissen einiger Studienteilnehmer der zwei Studien in der Analyse, was aber vor dem Hintergrund der vergleichsweise geringen Zahl und daraus resultierender geringer Verzerrung im Kontext der vorliegenden Gesamtanalyse hinnehmbar erschien.

Der ggf. notwendig Verweis auf einzelne in der Arbeit von Cipriani et al. eingeschlossene Studien erfolgt, wie z. B. oben betreffend die Studie von Fava et al., durch Angabe eines eindeutigen Studienkürzels (z. B. "Fava 2002") entsprechend der Publikation bzw. des Datensatzes der Übersichtsarbeit. Der genaue Literaturnachweis kann dem Literaturverzeichnis von Cipriani et al. entnommen werden und ist im Literaturverzeichnis der vorliegenden Arbeit nicht enthalten. Es wird bei der Erwähnung einer solchen Studienpublikation in der vorliegenden Arbeit jeweils eindeutig kenntlich gemacht, dass es sich um eine solche, von Cipriani et al. eingeschlossene, Studie handelt.

Die Ergebnisse aller von Cipriani et al. eingeschlossenen und in der vorliegenden Arbeit verwendeten Studien wurden entsprechend des öffentlich verfügbaren elektronischen Anhanges der Arbeit von Cipriani et al. in Bezug auf die beiden verwendeten Zielgrößen (Response, Dropout) extrahiert.

II.2 Bewertung der Erkenntnisgewinnung im Zeitverlauf

Die Analyse beruht auf allen vorliegenden, direkten Vergleichsstudien (vgl. Abschnitt II.1 oben). Mit Hilfe der im Folgenden beschriebenen Methodik soll ein umfassendes Bild des jeweiligen Erkenntnisstandes sowohl zu den jeweils einzelnen Zeitpunkten als auch im Zeitverlauf bezogen auf relevante Zielgrößen (Abschnitt II.2.1) gezeichnet werden. Um Konsistenz und einheitliche Quantifizierung möglichst umfassend zu erreichen, ist die Integration der Bewertung der Wirkstoffe auf der Grundlage einer einheitlichen Zielgröße sinnvoll, besonders in Bezug auf die Entwicklung und Bewertung eines alternativ möglichen Verlaufs der Erkenntnisgewinnung (Abschnitt II.3). Daher wird zusätzlich ein Verfahren angewandt, mit dessen Hilfe die beiden separaten Zielgrößen der Analyse, Therapieansprechen und Therapieabbruch, jeweils pro Studie und Studienarm (Wirkstoff) in eine neue Zielgröße einfließen, die das Therapieansprechen korrigiert um das Ausmaß des Therapieabbruchs wiedergibt. Das hierbei verwendete Verfahren wird im Anschluss (Abschnitt II.2.1) beschrieben.

Für jeden Zeitpunkt werden die Wirkstoffe in eine eindeutige Rangfolge in Bezug auf ihre therapeutischen Aussichten bezogen auf die jeweilige Zielgröße gebracht und das Ausmaß der Unterschiedlichkeit darauf basierend quantifiziert (Abschnitt II.2.2). Zudem kann auf dieser Grundlage für jeden Zeitpunkt auch quantitativ das Ausmaß der noch bestehenden Unsicherheiten in der Bewertung abgeschätzt werden. Dies wiederum ermöglicht es, die ggf. erreichte Reduktion der Unsicherheit durch die im Zeitverlauf publizierten Studien zu ermessen (Abschnitt II.2.3).

II.2.1 Zielgrößen

Zur Erfassung des Behandlungseffektes werden in der vorliegenden Analyse, wie in der zugrundeliegenden, systematischen Übersichtsarbeit (Cipriani 2009a) zwei Zielgrößen verwendet. Zum einen handelt es sich dabei um das Therapieansprechen (*Response*), sowie den Therapieabbruch (*Dropout*).

Zudem wird die eine zusammenfassende Zielgröße, im folgenden als "(um den Therapieabbruch korrigierte) Response" oder als "(um den Therapieabbruch korrigiertes) Therapieansprechen" (kurz: "korrigiertes Therapieansprechen") ermittelt. Deren Berechnung erfolgt hier entsprechend eines Vorschlages von Boada 2008 für vergleichbare Situationen, in denen die Bewertung einer Intervention sich wesentlich auf eine Zielgröße für das Erreichen des Behandlungszieles und eine weitere Zielgröße für den Abbruch der Therapie stützt. Impliziert ist hierbei die Unabhängigkeit der beiden Größen, insbesondere wird

nicht angenommen, dass z. B. der Anteil der Patienten mit Therapieabbruch in der Gruppe der Patienten mit Therapieansprechen geringer ist.

Die zusammenfassende Zielgröße wird entsprechend folgender Berechnungsvorschrift jeweils pro Studienarm berechnet:

$$n_{R(korr)} := n_R \times [(N-n_D) / N]$$

$n_{R(korr)}$ Anzahl Responder (korrigierte Response)
n_R Anzahl Responder
n_D Anzahl Dropout
N Anzahl Patienten in Studienarm

Das korrigierte Therapieansprechen ist somit das Produkt der Anzahl aller Responder (n_R) mit dem Anteil der Patienten, die die Therapie nicht abgebrochen haben ([(N-n_D) / N). Im Idealfall (keine Therapieabbrüche) entspricht die korrigierte der unkorrigierten Response. Im ungünstigsten Fall ist die Anzahl der Responder, korrigiert, gleich Null falls alle Patienten die Therapie abgebrochen haben.

II.2.2 Ermittlung des Erkenntnisstandes zu einem Zeitpunkt

Bei der Anwendung üblicher meta-analytischer Methoden, die direkte Vergleiche zwischen zwei Therapien oder Vorgehensweisen zum Gegenstand haben, können erhebliche Unsicherheiten und Lücken verbleiben. Erstens fehlt hier häufig die Möglichkeit zu allen Therapievergleichen Aussagen zu treffen, sofern nicht für alle Vergleiche Studien vorliegen. Dieses Problem ist hier gegeben, da nicht für alle Wirkstoffpaare vergleichende Studien vorliegen (so liegt z. B. keine vergleichende Studie für die Wirkstoffe Escitalopram und Mirtazapin vor). Zweitens sind die Aussagemöglichkeiten durch das statistische Signifikanzkriterium, das Aussagen nur bei einer – wie immer im einzelnen festgelegten und häufig konventionell bestimmten Situation, – als hinreichend erachteten Absicherung gegen zufällige Einflüsse erlaubt eingeschränkt. Schließlich erlauben solche Analysen es nicht, zumindest nicht ohne weiteres, eine Rangfolge der Wirkstoffe in Bezug auf die Zielgrößen festzulegen bzw. die relative Wirksamkeit aller Wirkstoffe untereinander festzulegen. Insbesondere kann nicht der jeweils erwartet beste Wirkstoff bezogen auf die einzelnen Zeitpunkte ermittelt werden, da immer nur "1:1"- oder "1:n"-Vergleiche möglich sind. Stellt sich im Ergebnis solcher Analysen beispielsweise heraus, dass eine Intervention (die Anwendung eines Wirkstoffes) A einer Intervention B überlegen ist, eine Intervention C einer Intervention B ebenfalls überlegen ist, so ist damit für das Verhältnis von A zu C noch keine Aussage möglich. Auch wenn eine Intervention A einer Intervention B überle-

gen ist und diese wiederum C überlegen ist, so scheint auch ohne eine weitere Analyse eine Überlegenheit von A gegenüber C im Sinne einer relationalen Transitivität zwar nahegelegt. Sie ist aber noch nicht empirisch gezeigt und auch zunächst nicht quantifizierbar.

Um diesen Problemen, insbesondere dem Problem des Fehlens direkt vergleichender Studien, zu begegnen, wurden eine Reihe von Methoden vorgeschlagen, die indirekte Vergleiche von Therapien ermöglichen, wenn für diese z. B. Erkenntnisse aus Therapievergleichen mit einem gemeinsamen *tertium* (häufig ein Vergleich zur Placebo-Behandlung) vorliegen. In der vorliegenden Arbeit wurde zu diesem Zwecke die "Mixed-Treatment-Comparison" (MTC)-Methodik ausgewählt (Lu & Ades 2004, Ades 2007, Wells 2009, Schöttker 2009). Diese weist gegenüber anderen methodischen Zugängen für den Zweck der vorliegenden Untersuchung Vorteile auf. Es ist mit Hilfe der MTC-Methodik einfacher als bei Verwendung anderer Methoden möglich, viele Therapievergleiche in einer Analyse simultan zu betrachten und es ist insbesondere möglich direkte und indirekte Therapievergleiche zu integrieren (hierauf verweist die Bezeichnung als „gemischt“ bzw. "mixed"), was allerdings auch Auswirkungen auf den geschätzten, relativen Effekt zwischen Interventionen haben kann für die direkte Vergleiche vorliegen (Abschnitt II.2.2.3). Ein weiterer Vorteil der MTC-Methode im Rahmen der vorliegenden Untersuchung besteht darin, dass die Ergebnisse direkt zur Ermittlung der zu einem Zeitpunkt jeweils verbleibenden Ergebnisunsicherheit (Abschnitt II.2.3) und damit zur Bewertung des Erkenntnisgewinns im Zeitverlauf genutzt werden können. Die MTC-Methodik wurde überdies auch für die Analyse in der Publikation von Cipriani 2009a genutzt, deren Daten der vorliegenden Untersuchung zugrunde liegen.

II.2.2.1. Mixed Treatment Comparison (MTC: Lu & Ades 2004, Wells 2009)

Für jede der verwendeten Zielgrößen (Therapieansprechen, Therapieabbruch, korrigiertes Therapieansprechen) wird als Datengrundlage die Zahl der beobachteten Fälle (r_{ix} – Ereignisse in der Studie i bezogen auf die Therapie / den Wirkstoff / den Studienarm X) aus der Zahl aller in die Analyse jeweils einbezogenen Patienten (n_{ix}) verwendet. Aus diesen lassen sich alle weiteren interessierenden Größen, insbesondere die relativen Therapieeffekte (δ_{xy} – Differenz der Therapieeffekte der Therapien bzw. Wirkstoffe X und Y) ermitteln. Die Therapieeffekte werden mit Hilfe eines logistischen (Logit-)Modells geschätzt. Jeweils für die i-te Studie in der (zumindest) x und y untersucht wurden wird δ_{xy} geschätzt als (Lu & Ades 2004, Wells 2009):

$$\delta_{ixy} = \text{logit}(p_{ix}) - \text{logit}(p_{iy})$$

$p_{i[x/y]}$ Wahrscheinlichkeit des Ereignisses (Binominalverteilung von $r_{i[x/y]}/n_{i[x/y]}$)

Die δ_{ixy} sind die logarithmierten Odds-Ratios der Therapieeffektdifferenz. Für die Meta-Analyse aller einbezogenen Studien wird eine Normalverteilung der δ_{ixy} zu Grunde gelegt, um die δ_{xy} zu schätzen:

$$\delta_{ixy} \sim N(\delta_{xy};\sigma^2)$$

σ^2 Varianzterm, abhängig vom meta-analytischen Modell.

Für die vorliegenden Analyse wurde ein Modell mit zufälligen Effekten verwendet, ohne Korrektur für die Korrelationsstruktur drei- oder mehrarmiger Studien (Lu & Ades 2004, Ades 2007 – *simple random effects model*)[2].

Therapien bzw. Wirkstoffe x und y, für die keine direkten Vergleiche aus Studien verfügbar sind, können auf Grund indirekter Vergleiche geschätzt werden. Seien A, B und C unterschiedliche Therapien bzw. Wirkstoffe und liegen vergleichende Studien A-C und B-C, jedoch keine vergleichenden Studien A-B vor, so ergibt sich z. B.:

$$\delta_{AB} = \delta_{CB} - \delta_{CA}$$

Die Therapieeffektschätzer, für die direkt vergleichende Studien vorliegen (im Beispiel δ_{CB} und δ_{CA}), werden im Rahmen der Analyse auch als Basisparameter bezeichnet. Die Therapieeffektschätzer, für die solche nicht vorliegen (im Beispiel δ_{AB}), werden als (abgeleitete) funktionelle Parameter bezeichnet. Auf diese Weise werden indirekte Vergleiche umgesetzt. Hinzu kommt, dass indirekte und direkte Vergleiche simultan berücksichtigt werden. Läge, in Ergänzung des obigen Beispiels, zusätzlich eine direkt vergleichende Studien A-B vor, so ginge diese, bzw. der auf dieser basierende Effektschätzer, in die Analyse des Gesamtergebnisses (δ_{AB}) ein. Dies ergibt sich dadurch, dass alle δ_{xy} immer auf der Grundlage aller paarweisen Vergleiche (δ_{ixy}) ermittelt werden, die auf der Grundlage der eingeschlossenen Studien möglich sind. Auch wenn zu bestimmten paarweisen Vergleichen direkte Studien vorliegen, so kann doch der Schätzer für die vergleichende Wirkung im Rahmen einer MTC-Analyse zusätzlich von indirekten Vergleichen beeinflusst sein. Daher wird zusätzlich geprüft, ob signifikante Abweichungen zwischen der nur auf direkten Vergleichen beruhenden paarweisen Ergebnisse von den paarweisen Ergebnissen

[2] Unter den eingeschlossenen Studien finden sich lediglich zwei dreiarmige Studien. Diese wurden für die Zwecke der vorliegenden Analyse jeweils in drei "fiktive" zweiarmige Studien (die jeweils den möglichen paarweisen Vergleichen entsprechen) zerlegt.

der MTC-Analyse in einem erheblichen Umfang beobachtet werden (Abschnitt II.2.2.3). Bei signifikanten Abweichungen im für die Gesamtergebnisse erheblichen Umfang, wären deren Ergebnisse entsprechend vorsichtiger zu interpretieren.

Die relevanten Parameter des Modells, insbesondere die δ_{xy}, werden mit Hilfe einer viele Male wiederholten Simulationsrechnung geschätzt. In jeder einzelnen Simulationsrechnung wird aus den zugrundeliegenden Verteilungen für die grundlegenden Parameter jeweils mit Hilfe eines Zufallszahlengenerators ein zufälliger Wert gezogen, und auf dieser Grundlage die Ergebnisse für den jeweiligen Simulationslauf ermittelt. Begonnen wird dabei mit „uninformativen“ Prior-Verteilungen, d. h. dass kein spezifisch bereits vorliegender Erkenntnisstand vorausgesetzt wird. Sind alle einzelnen Simulationsrechnungen durchgeführt, werden die Ergebnisse gemittelt und die Streuung bzw. die darauf basierenden Vertrauensbereiche berechnet. Grundsätzlich werden hier symmetrische 95%-Vertrauensbereiche so ermittelt, dass jeweils die 2,5% höchsten bzw. 2,5% niedrigsten Werte gerade außerhalb dieses Bereiches liegen.

Die Simulationsrechnungen in der vorliegenden Arbeit wurden mit Hilfe des Programmes WinBUGS, Version 1.4 (vgl. Abschnitt II.4) durchgeführt. Für jede Simulation wurden zunächst 10.000 Iterationen (einzelne Simulationsrechnungen) durchgeführt, deren Ergebnisse jedoch nicht verwendet wurden ("*Burn-In*"). Anschließend wurden 100.000 weitere Iterationen durchlaufen, aus denen die Ergebnisse ermittelt wurden. Auf eine Konvergenzprüfung der Simulationsprogramme wurde, angesichts bzw. zugunsten der hohen Zahl an Iterationen, verzichtet. Eines der inhaltlich gleichförmig aufgebauten WinBUGS-Programme (die separat für alle Zielgrößen und alle Analysezeitpunkte erstellt wurden) ist im Anhang (Abschnitt VII.1) dargestellt.

II.2.2.2. Ergebnisse aus der MTC zu Behandlungseffekten

Zusätzlich zu den δ_{xy} (X,Y ε Menge aller maximal 12 Wirkstoffe, für die zu einem analysierten Zeitpunkt überhaupt zumindest eine vergleichende Studie zur einem der anderen Wirkstoffe vorliegt, mit $x \neq y$), mithin den Odds-Ratios für alle Wirkstoffvergleiche, sind besonders zwei weitere Ergebnisse relevant. So kann auf der Grundlage der Analyseergebnisse jeweils ein "absoluter" Behandlungseffekt pro Wirkstoff (T_x), als bei der Anwendung des jeweiligen Wirkstoffs zu erwartender Ereignisrate, geschätzt werden. Dies erfolgt mit dem Ziel, die Relevanz der relativen Größen bzw. Unterschiede zwischen den Wirkstoffen zu ermessen. Die Abschätzung basiert hier auf einer unterstellten, vorgegebenen "Baseli-

ne"-Rate für einen der Wirkstoffe[3]. Für die anderen Wirkstoffe wird dann mit Hilfe der jeweiligen Odds-Ratio im Vergleich zu diesem Wirkstoff ein größerer oder kleinerer absoluter Behandlungseffekt geschätzt. In der vorliegenden Analyse wurde ein "fiktiver", logarithmierter Wert der Odds von 0, entsprechend einer Rate jeweils pro Zielgröße von 0,5 (d. h. dass dieser Wert unterstellt, dass für die Hälfte aller Studienteilnehmer jeweils ein Ereignis beobachtet wurde) als Grundlage angenommen und die absoluten Ereignisraten der anderen Wirkstoffe relativ zu diesem Wert ermittelt.

Jede MTC-Analyse pro Zeitpunkt und Zielgröße gibt für jeden Wirkstoff eine Wahrscheinlichkeit an, mit dem es sich bei diesem Wirkstoff um den besten, zweitbesten etc. bis hin zum schlechtesten handelt. Insbesondere die Wahrscheinlichkeit, dass es sich pro Zielgröße jeweils um den besten Wirkstoff handelt ($p_{x,bester}$ - höchste Rate des Therapieansprechens bzw. des korrigierten Therapieansprechens, geringste Rate des Therapieabbruchs) ist hier von Bedeutung. Gleichwohl wird durch die Analyse für jeden Zeitpunkt, jede Zielgröße und jeden Rangplatz (maximal 12, entsprechend der Anzahl der eingeschlossenen Wirkstoffe pro Zeitpunkt) eine Wahrscheinlichkeit ermittelt, mit der dieser Rang von dem jeweiligen Wirkstoff eingenommen wird. Dabei summieren sich die Wahrscheinlichkeiten pro Rang über alle Wirkstoffe jeweils zu 1. Die Wahrscheinlichkeiten, dass ein bestimmter Wirkstoff jeweils einen bestimmten Rang einnimmt, werden als Auswertungen aus den durchgeführten Simulationsläufen (s. o.) ermittelt. So wird z. B. um den p_{bester}-Wert eines Wirkstoffes X zu ermitteln die Zahl der Simulationsläufe, in denen der Wirkstoff X das beste Ergebnis erzielte, durch die Zahl der Simulationsläufe geteilt und die Wahrscheinlichkeit diesem Anteil gleichgesetzt. Entsprechend wird mit den übrigen Wirkstoffen und betreffend der übrigen Rangplätze verfahren.

II.2.2.3. Gegenüberstellung direkter und gemischter Vergleiche

Für alle Wirkstoffkombinationen und Zeitpunkte für die direkte Vergleiche möglich sind, werden diese außerhalb der MTC ermittelt und den entsprechenden Ergebnissen aus den MTC gegenübergestellt. Diese Gegenüberstellung erfolgt bezogen auf die Zielgrößen Therapieansprechen und Therapieabbruch. Dies erfolgt, da generell bei der Anwendung indirekt vergleichender und gemischter Verfahren empfohlen wird, deren Ergebnisse den Ergebnissen aus direkt vergleichenden Studien gegenüberzustellen. Werden Diskrepan-

---

[3] Auf Grund der kontigenten Zuordnung des Wirkstoffes Fluoxetin (mit der Wirkstoff-Nr. 1) als präferierter Referenzkategorie unterstellt dies fiktiv (aber vermutlich in der Größenordnung nicht unrealistisch) jeweils, pro Zielgröße, diesem Wirkstoff die Ereignisrate von 0,5. Ereignisraten für die übrigen Wirkstoffe sind als Größen relativ zu dieser Referenzkategorie zu betrachten. Dies hat einschränkende Bedeutung für die Interpretation dieser Größen.

zen festgestellt, so verlangt dies eine genauere Prüfung. Nicht in allen Fällen (wenn z. B. nur wenig direkt vergleichende Ergebnisse vorliegen) sind die Ergebnisse aus indirekten/gemischten Vergleichen als weniger zuverlässig anzusehen. Die direkten Vergleiche wurden in der vorliegenden Arbeit mit Hilfe von Standardmethoden der Meta-Analyse berechnet (Deeks 2008), wobei die OR als Therapieeffektschätzer im Bezug auf beide Zielgrößen verwendet wurde. Diese OR bzw. deren logarithmierte Werte ($\ln OR_{direkt}$) können unmittelbar den Ergebnissen der MTC ($\delta = \ln OR_{indirekt}$) gegenübergestellt werden.

Zur Prüfung der Übereinstimmung ($\ln OR_{direkt} = \ln OR_{indirekt}$) oder Abweichung direkt und gemischter ermittelter Werte wurde jeweils geprüft, ob der gemischt ermittelte Wert im 95%-Konfidenzintervall des direkt ermittelten Wertes lag. Ein gemischt ermittelter Wert außerhalb des Konfidenzintervalls des direkt ermittelten Wertes, würde auf eine erhebliche Diskrepanz der beiden Werte hinweisen.

II.2.3 Verbleibende Erkenntnisunsicherheit und Analyse des Zeitverlaufs

Jede Erkenntnis zu den Therapieeffekten aus empirischen Studien ist mit der Unsicherheit verbunden, dass die Ergebnisse auch tatsächlich die wirklichen Therapieeffekte abbilden. Neben systematischen Fehlern (Verzerrungen, *biases*) sind hierfür Zufallsfehler verantwortlich, die sich aus den notwendig begrenzten Stichprobengrößen ergeben. Die Durchführung weiterer Studien kann die Erkenntnissicherheit erhöhen, zumindest dann, wenn keine weiteren, systematischen Fehler auftreten. Entsprechend der Fragestellung der vorliegenden Arbeit wird der Versuch unternommen, diese Entwicklung im Zeitverlauf nachzuzeichnen und zu ermitteln, ob und in welchem Ausmaß die Erkenntnisunsicherheit in Bezug auf die Therapieeffekte der in die Analyse eingeschlossenen Wirkstoffe sich im Zeitverlauf tatsächlich reduziert hat.

Hierzu bedarf es eines Verfahrens zur Quantifizierung der jeweils zu einem Zeitpunkt noch verbleidenden Erkenntnisunsicherheit (Briggs 2006 Kapitel 6 – mit weiteren Literaturangaben auch zu den Grundlagen der statistischen Entscheidungstheorie, auf der das Verfahren basiert sowie zu bisherigen Anwendung im Rahmen der Bewertung von Therapieverfahren). Das Verfahren basiert hier auf den Ergebnissen der oben beschriebenen MTC-Analyse, genauer auf den jeweiligen wirkstoff-bezogenen Wahrscheinlichkeiten p_{bester} bzw. den Ereignisraten T jeweils bezogen auf die Zielgrößen, für die die verbleibenden Erkenntnisunsicherheiten geschätzt werden (vgl. Abschnitt II.2.2.2). Die Erkenntnisunsicherheit wird dabei als derjenige maximale "Informationswert" (V, *value of (perfect) information*) operationalisiert, der durch weitere Studien äußerstenfalls ausgeschöpft bzw.

weiter reduziert werden kann. Er wird zu jedem Zeitpunkt, bezogen auf jede Zielgröße, bestimmt als:

$V = E_{max} - E_{aktuell}$ ermittelt als

$E_{max} = 1 \times \max(T_x)$ und
$E_{aktuell} = \sum(i=1 \ldots i_{max})\ p_{bester,i} \times T_i$

T_x Maximale (bzw. optimale Ereignisrate) aller Wirkstoffe (jeweilige Zielgröße)
$i \ldots i_{max}$ Alle zu einem jeweiligen Zeitpunkt berücksichtigten Wirkstoffe (max. 12)
$p_{bester,i}$ Wahrscheinlichkeit, dass es sich bei Wirkstoff i um den besten handelt
T_i Ereignisrate des Wirkstoffes i

Der Wert V (entsprechend dem *Expected value of perfect information*, EVPI – Briggs 2006, Kapitel 6) ist als Differenz zweier Erwartungswerte über den Therapieerfolg bestimmt. Der erste Wert, $\max(T_x)$ bezeichnet dabei den erwarteten Therapieerfolg bei "perfekter" Information, d. h. der – unterstellten – Sicherheit ("p=1"), dass die Therapie mit der höchsten, beobachteten Erfolgsquote (eben $\max(Tx) = E_{max}$) tatsächlich die beste der verglichenen Therapieoptionen ist und sich diese Erfolgsquote also realisiert. Der zweite Wert stellt den Erwartungswert des Therapieerfolges bei tatsächlich jeweils aktuell zu einem Zeitpunkt gegebener Unsicherheit dar. Er errechnet sich als mit den Wahrscheinlichkeiten (das die jeweilige Therapieoption tatsächlich die beste ist) gewichtete Summe der Erfolgsquoten. Ein einfaches Beispiel mag dies erläutern. Angenommen es wären zwei Therapieoptionen A und B mit folgenden Ergebnissen verfügbar: $p_{bester,A} = 0{,}4$; $T_a = 0{,}5$; $p_{bester,B} = 0{,}6$; $T_b = 0{,}7$. Für angenommene 10.000 Behandlungsepisoden resultierte dann mit angenommener "perfekter" Information ein Wert von $E_{max} = \max(T_B) = 1 \times 0{,}7 \times 10.000 = 7.000$ erfolgreichen Behandlungen. Mit aktuell gegebener Unsicherheit dagegen lediglich ein Wert für $E_{aktuell} = \sum(i=1 \ldots i_{max})\ p_{bester,i} \times T_i = [\ (0{,}4 \times 0{,}5) + (0{,}6 \times 0{,}7)] \times 10.000 = 0{,}55 \times 10.000 = 5.500$. Der verbleibende Informationswert V beliefe sich somit auf 1.500. Es ist erkennbar, dass sich mit sowohl mit veränderten Differenzen der jeweiligen Therapieeffekte als auch mit veränderten Wahrscheinlichkeiten von deren Realisierung andere verbleibende Infomationswerte verbinden. Änderten sich z. B. die Wahrscheinlichkeiten p_A und p_B durch hinzukommende Informationen auf die Werte 0,3 bzw. 0,7, so resultierte nunmehr für $E_{aktuell} = \sum(i=1 \ldots i_{max})\ p_{bester,i} \times T_i \times 10.000 = 5.700$ und entsprechend ergäbe sich für V nunmehr der Wert 1.300. Mithin wäre der Veränderung der Wahrscheinlichkeiten, die eine nunmehr höhere Erkenntnissicherheit wiedergäbe, ein Wert von (1.500-1.300)=200/10.000 Behandlungsepisoden zuzuschreiben.

Wie beschrieben, wird für jeden Zeitpunkt (Jahr) und für jede Zielgröße der Erkenntnisstand in Gestalt eines verbleibenden Informationswertes V ermittelt. Auf dieser Grundlage wird die zeitliche Entwicklung des Erkenntnisstandes durch die zeitliche Folgen von V-Werten (sowie für deren wesentliche Komponenten E_{max} und $E_{aktuell}$) für die eingeschlossenen Zielgrößen beschrieben. Zudem kann ermittelt werden, welche Studien oder Gruppe von Studien die jeweils zwischen zwei Zeitpunkten zusätzlich in die Analyse aufgenommen wurden welche Veränderungen der V-Werte zuzuordnen sind, mithin deren Beitrag zur Veränderung des Erkenntnisstandes.

II.3 Bewertung einer alternativen Erkenntnisgewinnung im Zeitverlauf

Über die Beschreibung des tatsächlichen Verlaufs der Erkenntnisgewinnung hinaus, liegt der vorliegenden Arbeit die Frage zugrunde, wie diese Entwicklung auch im Hinblick auf alternativ mögliche Verläufe der Erkenntnisgewinnung bewertet werden kann. Da faktisch keine alternativen Verläufe zu beobachten sind, muss ein hypothetischer Weg beschrieben werden einen alternativen Verlauf zu konstruieren und die von diesem zu erwartenden Erkenntnisgewinne im Zeitverlauf den tatsächlich beobachteten gegenüberzustellen – unter der Leitfrage, ob auf der Grundlage dieser Gegenüberstellung ggf. Defizite des tatsächlichen Verlaufs benannt und Hinweise auf Verbesserungsmöglichkeiten abgeleitet werden können. Dabei muss Sorge getragen werden, dass ein beschriebener, alternativer Verlauf ausgehend von einem bestimmten Zeitpunkt ausschließlich vom Erkenntnisstand zu diesem Zeitpunkt und nicht von erst später erlangtem Wissen abhängt.

Grundsätzlich beruht der hier entwickelte, mögliche alternative Verlauf der Erkenntnisgewinnung auf einer angenommenen, zeitlich veränderten Reihenfolge der Studiendurchführung bzw. Publikation. Für die Beschreibung (insbesondere der zeitlichen Veränderung) des verbleibenden Informationswertes V anhand des Erkenntnisstandes zu einem bestimmten Zeitpunkt (s. Abschnitt II.2.3) bis zu dem jeweiligen Folgezeitpunkt werden jeweils alle Studien, deren Ergebnisse zwischen den beiden Zeitpunkt publiziert wurden, zusätzlich herangezogen. Dieses Vorgehen wird zur Beschreibung des alternativ möglichen Verlaufes der Erkenntnisgewinnung gezielt verändert. Ausgehend von einem Basiszeitpunkt werden bis zum Folgezeitpunkt nunmehr zunächst gezielt diejenigen Studien herangezogen, von denen ein Erkenntnisgewinn im besonderen Maße erwartet werden kann. Dieses Prozedere wird über die weiteren Folgezeiträume hinweg wiederholt angewandt. Von dieser Vorgehensweise bzw. der resultierenden Studienfolge und damit verbundenen Veränderung des Erkenntnisstandes wird eine im Zeitverlauf schneller mögliche Reduktion der Erkenntnisunsicherheit erwartet. Dieser schnellere Rückgang bemisst sich anhand des Vergleichs zweier Reihen von V-Werten: der Folge der jeweils verbleibenden Informationswerte V im alternativen Verlauf im Vergleich zur entsprechenden Folge im faktischen Verlauf.

II.3.1 Bewertung der Studien im Hinblick auf weiteren Erkenntnisgewinn

Von allen bekannten Charakteristika der zu einem bestimmten Zeitpunkt noch nicht in die Analyse einbezogenen Studien wird lediglich berücksichtigt, welche Wirkstoffe in diesen Studien verglichen wurden. Lediglich auf diesem Charakteristikum basiert die Bewertung,

in welchem Ausmaß von den jeweils verbleibenden Studien ein Erkenntnisgewinn in besonderem Maße erwartet wird. Das Kriterium, bzw. die Zurechnung des zu erwartenden Informationswertes gilt daher streng genommen für die jeweiligen beiden zu vergleichenden Wirkstoffe[4] und wird auf die Studien, die jeweils genau diese beiden Wirkstoffe untersuchen, übertragen. Der erwartete Informationswert wird wie folgt ermittelt:

$$W_{x,y} = [(p_x + p_y) * \max(T_x,T_y)] - [(p_x \times T_x) + (p_y \times T_y)]$$

p_x, p_y Wahrscheinlichkeit, dass es sich bei x bzw. y um den besten Wirkstoff handelt

T_x, T_y Ereignisrate von x bzw. y

Die Bestimmung des W erfolgt in Analogie zur Bestimmung des Informationswertes V für alle Wirkstoffe zu einem jeweiligen Zeitpunkt, beschränkt auf die jeweiligen beiden zu vergleichenden Wirkstoffe, für die der Wert weiterer Erkenntnisse abgeschätzt werden soll (wären alle Wahrscheinlichkeiten weiterer Wirkstoffe außer p_x und p_y gleich 0 so gilt W=V). Die Ermittlung des W-Wertes wird mit Hilfe der Ergebnisse zur Zielgröße des korrigierten Therapieansprechens durchgeführt. Der Wert von W ist umso höher, je größer die Summe der Wahrscheinlichkeiten p_x,p_y ausfällt. Er ist umso geringer, je geringer die Differenz dieser Wahrscheinlichkeiten sowie die Differenz der jeweiligen Ereignisraten T_x, T_y ausfällt. Die Bewertung des Vergleichs von jeweils zwei Wirkstoffen unterstellt, dass von den Ergebnissen vergleichender Studien zwischen den jeweiligen Wirkstoffen keine Effekte auf die Bewertung anderer Wirkstoffe ausgehen.

II.3.2 Durchführung und Auswertung der Analyse

Als Basiszeitpunkt von dem ausgehend ein alternativer Verlauf der Erkenntnisgewinnung konzipiert wird, wird das Jahr 2002 (das erste Jahr, ab dem alle 12 Wirkstoffe in die Analyse eingeschlossen werden können) gewählt. Ab diesem Zeitpunkt sind noch 51 Studien mit Publikationsdatum ab dem Jahr 2003 oder unpublizierte Studien zu berücksichtigen, die in den Erkenntnisstand des Jahres 2002, der den Ausgangspunkt bildet, noch nicht einbezogen wurden. Diese werden sukzessive in die Bewertung einbezogen. Abweichend von der faktischen Publikationsreihenfolge der Studien in den Jahren 2003-2007 bzw. dem Jahr "2009" (unpublizierten Studien), wird die alternative Reihenfolge basierend auf dem Informationswertkriterium W (s. Abschnitt II.3.1) bestimmt. Beginnend mit dem Jahr

[4] Es werden nur zweiarmige Studien in Betracht gezogen, da beginnend mit dem Jahr 2003 nur zweiarmige Studien vorliegen, die Gegenstand der Konstruktion einer alternativen Abfolge sind.

2003 werden demnach nicht alle tatsächlich im Jahr 2003 publizierten Studien im nächsten Schritt (2002-2003) eingeschlossen, sondern die gleiche Zahl derjenigen Studien mit Publikationsjahr 2003-2007 bzw. 2009 die die Summe des Informationswertes W jeweils maximieren. Daraufhin wird eine neuer, alternativer, Informationsstand für das Jahr 2003 mit Hilfe der MTC-Analyse ermittelt und die angegebenen Schritte für die nächsten Folgezeiträume (2003-2004, 2004-2005, 2005-2006, 2006-2007, 2007-2009) wiederholt. Zum Ende des Betrachtungszeitraum (2009) entspricht der alternativ ermittelte Erkenntnisstand wieder dem faktischen, da die Menge der einbezogenen Studien (zu diesem Zeitpunkt: aller Studien) wiederum dieselbe ist. Die Analyse wird ausschließlich für die zusammenfassende Zielgröße des korrigierten Therapieansprechens durchgeführt. Damit werden inkonsistente Bewertungen von Wirkstoffvergleichen bzw. Studien vermieden.

Für alle Jahre, für die ein alternativer Erkenntnisstand ermittelt wird (2003-2007) werden die Abweichungen der jeweils verbleibenden Ergebnisunsicherheit im faktischen Verlauf V mit den Werten des alternativen Verlaufs $V_{alternativ}$ verglichen. Der alternative Verlauf ist auf dieser Grundlage als mit einem günstigeren Erkenntnisverlauf verbunden zu kennzeichnen, wenn gilt:

$$\sum(j=2003 \ldots 2007)\ V_{j,alternativ} < \sum(j=2003 \ldots 2007)\ V_{j,faktisch}$$

d. h. die Summe der V-Werte aller Jahre im Analysezeitraum im alternativen Verlauf muss geringer sein als die Summe der entsprechenden V-Werte im faktischen Verlauf. Ist dies der Fall, so ist der alternative Verlauf als besser als der faktische Verlauf zu bewerten.

II.4 Verwendete Analyse- und Datenverarbeitungsprogramme

Die Daten der zugrundeliegenden Studien aus der systematischen Übersichtsarbeit Cipriani 2009a wurden entsprechend der ebenfalls publizierten Datenextraktion in einer Microsoft Access®-Datenbank (Datenbank-Anwendungsversion Microsoft® Office Access 2003/2010) erfasst und dort für weitere Analysen aufbereitet.

Sämtliche MTC-Analysen wurden in dem speziellen statistischen Modellierungs-, Simulations- und Analysesoftware WinBUGS (Version 1.4[5]) durchgeführt. Die einzelnen WinBUGS-Programme (es wurde jeweils pro Zeitpunkt und Zielgröße bzw. pro Analyse – faktischer und alternativer Verlauf der Erkenntnisgewinnung separate Programme erstellt, vgl. Abschnitt VII.1 im Anhang für ein Beispielprogramm) wurden durch entsprechend programmierte Routinen in der verwendeten Datenbank automatisch erstellt und dann mit WinBUGS verarbeitet. Ergebnisse anderer Meta-Analysen zu direkten Wirkstoffvergleichen (vgl. Abschnitt II.2.2.3) wurden entsprechend der in Deeks 2008 angegebenen Rechenvorschriften, in Form von VBA (Visual-Basic for Applications)-Programmen in der Microsoft Access®-Datenbank implementiert und die Analysen dort erstellt.

Die Ergebnisse aller Analysen wurden in das Tabellenkalkulationsprogramm Microsoft® Excel 2003/2010 übernommen und dort wurden weitere Analysen und Aufbereitungen vorgenommen.

[5] http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml

III. Ergebnisse

Im folgenden Abschnitt (III.1) werden die Ergebnisse der Analysen zum Erkenntnisgewinn im Zeitverlauf wiedergebeben. Daran anschließend (Abschnitt III.2) werden die Ergebnisse, die bei einem hypothetisch alternativen Verlauf der Erkenntnisgewinnung zu verzeichnen gewesen wären, wiedergegeben und dem faktischen Verlauf der Erkenntnisgewinnung gegenübergestellt. In Abschnitt III.3 werden die Ergebnisse zusammenfassend dargestellt.

III.1 Erkenntnisgewinnung im Zeitverlauf

Im folgenden Abschnitt wird zunächst die Anzahl der in die Analyse eingeschlossenen Studien, Patienten und Wirkstoffe angegeben (Abschnitt III.1.1). Darauf folgend (Abschnitt III.1.2) werden die Ergebnisse zum Verlauf der jeweils verbleibenden Erkenntnisunsicherheit bzw. des Wertes der maximal noch zu gewinnenden weiteren Erkenntnisse berichtet. Daran schließt sich eine Darstellung der Ergebnisse im Zeitverlauf für die einzelnen Wirkstoffe an (Abschnitt III.1.3). In Abschnitt III.1.4 werden die Ergebnisse der Gegenüberstellung direkter und gemischter Vergleich dargestellt.

III.1.1 In die Analyse eingeschlossene Studien-, Patienten- und Wirkstoffzahlen

Insgesamt wurden in die Analyse zum Erkenntnisgewinn im Zeitverlauf die Ergebnisse von 110 Studien (an insgesamt 23.756 Patienten) zu allen 12 berücksichtigten Wirkstoffen, bezogen auf die Zielgröße korrigiertes Therapieansprechen, einbezogen (Tabelle 1). Die etwas abweichenden Zahlen für die weiteren Zielgrößen (Therapieabbruch und Therapieansprechen) sind im Anhang (Abschnitt VII.2) dargestellt. Abweichungen ergeben sich, da nicht allen Studien Ergebnisse zu allen Zielgrößen zu entnehmen sind. Inbesondere die Zielgröße korrigiertes Therapieansprechen erfordert, dass nur solche Studien eingeschlossen werden, denen sowohl Ergebnisse zum Therapieansprechen als auch zum Therapieabbruch zu entnehmen sind. Schon zum ersten betrachteten Zeitpunkt wurden die 6 Wirkstoffe FLUOX, BUPRO, SERTR, PAROX, CITAL, FLUVO einbezogen. Die übrigen Wirkstoffe kamen im Zeitverlauf hinzu (VENLA und MILNA 1994, MIRTA und REBOX 1997, ESCIT 2000 und DULOX 2002). Die Gruppe der jeweils jahresbezogen in der Analyse berücksichtigten Wirkstoffe unterscheidet sich nicht zwischen den auf die einzelnen Zielgrößen bezogenen Analysen.

Tabelle 1 Anzahl einbezogener Studien, Patienten und Wirkstoffe (betreffend der Zielgröße korrigiertes Therapieansprechen)

Jahr	Studien (Jahr) [n]***	Patienten (Jahr) [n]***	Studien kumuliert [n]***	Patienten kumuliert [n]***	Wirkstoffe [n]	Wirkstoffe zusätzlich
1993*	7	912	7	912	6	
1994	3	380	10	1.292	8	VENLA, MILNA
1995	4	648	14	1.940	8	
1996	2	360	16	2.300	8	
1997	8	2.028	24	4.328	10	MIRTA,REBOX
1998	7	1.592	31	5.920	10	
1999	8	1.758	39	7.678	10	
2000	14	2.328	53	10.006	11	ESCIT
2001	4	823	57	10.829	11	
2002	10	2.033	67	12.862	12	DULOX
2003	6	608	73	13.470	12	
2004	5	1.134	78	14.604	12	
2005	8	1.820	86	16.424	12	
2006	9	2.320	95	18.744	12	
2007	4	1.246	99	19.990	12	
2009**	11	3.766	110	23.756	12	

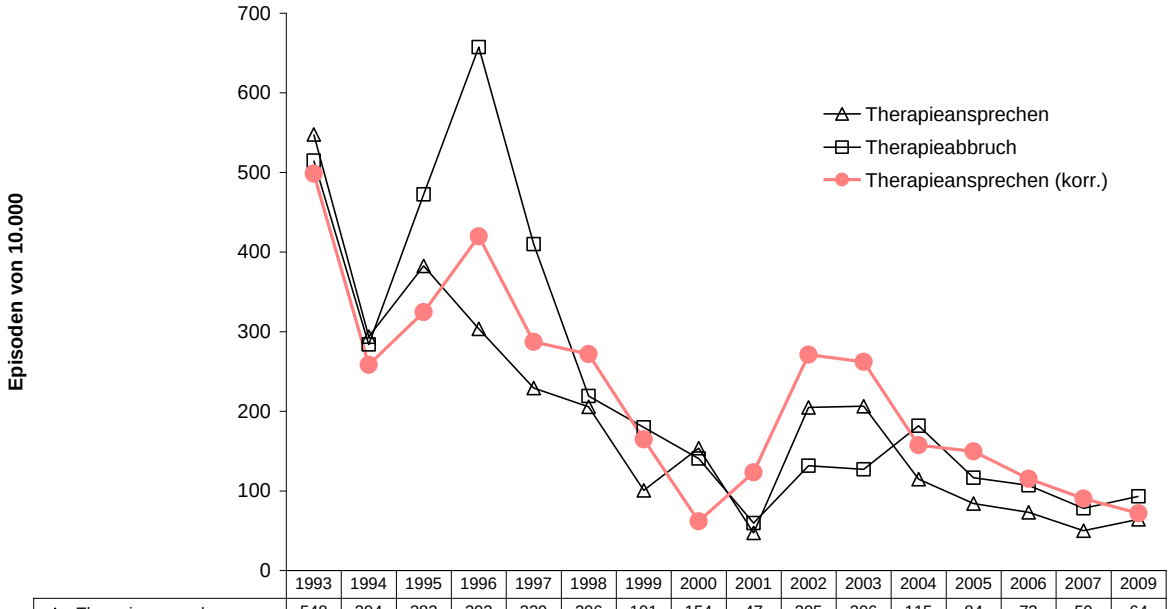
*Die Jahre 1991 und 1993 (1992 wurden keine Studien publiziert) wurden zusammengefaßt analysiert, da im Jahr 1991 nur eine einzige Studie publiziert worden war.

**Alle unpublizierten Studien wurden unter dem Jahr 2009 (Publikationsjahr von Cipriani2009a) erfasst.

*** Die Anzahl der Studien (und entsprechend die Zahl der Patienten), für die Ergebnisse zur Zielgröße "Therapieansprechen" verfügbar waren, zählt die jeweils zusätzlichen "Studien", die aus der Zerlegung der beiden dreiarmigen Studien resultieren (vgl. Abschnitt II.1.4) mit.

III.1.2 Verlauf der verbleibenden Erkenntnisunsicherheit

Der Verlauf, als zeitliche Reihung von Werten des jeweils nach einem Zeitpunkt (Jahr) verbleibenden, maximal möglichen weiteren Erkenntnisgewinns (V-Werte, entsprechend Abschnitt II.2.3) ist in Abbildung 1 für die eingeschlossenen Zielgrößen dargestellt.



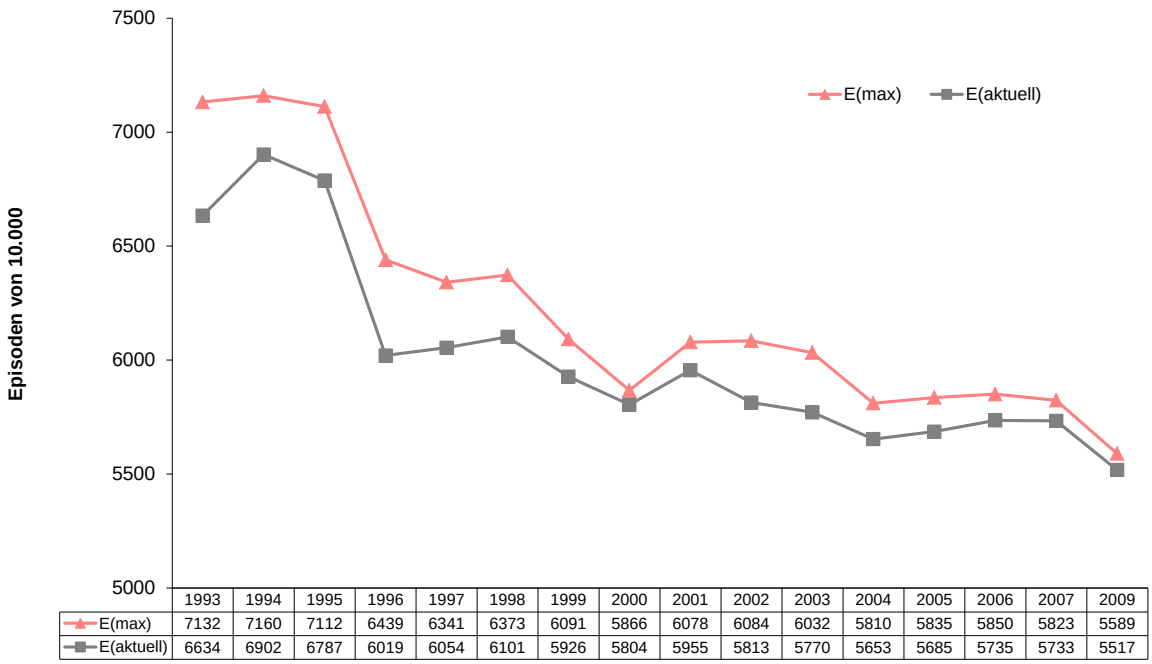
	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2009
Therapieansprechen	548	294	382	303	229	206	101	154	47	205	206	115	84	73	50	64
Therapieabbruch	515	284	472	657	410	219	180	141	59	131	127	182	117	107	78	93
Therapieansprechen (korr.)	498	258	325	420	287	272	165	62	123	271	262	157	150	115	90	72

Abbildung 1 Wert des maximal möglichen, weiteren Erkenntnisgewinns (V-Werte)

Der Wert maximal verbleibenden Erkenntnisgewinns (in Episoden/10.000[6]) sank im Zeitverlauf von 498/10.000 Episoden nach dem Erkenntnisstand des Jahres 1993 auf 72/10.000 Episoden nach dem Erkenntnisstand des Jahres 2009 (entsprechend dem Jahr 2007 zuzüglich der unpublizierten Studien, die Teil der Datengrundlage bilden) bezogen auf die Zielgröße des korrigierten Therapieansprechens. Erkennbar ist eine tendenziell deutliche Reduktion im Zeitverlauf, mit jedoch erheblichen Schwankungen. Der Verlauf für die anderen Zielgrößen (Therapieabbruch und Therapieansprechen) stellt sich grundsätzlich ähnlich dar, wobei jedoch im Zeitverlauf ebenfalls deutliche – und andere – Schwankungen zu beobachten sind. Im späteren Zeitverlauf (etwa ab dem Jahr 2002) ist allerdings eine gewisse Stabilisierung, mit nur noch vergleichsweise geringeren Veränderungen und einem insgesamt homogeneren Verlauf zu beobachten.

[6] Die Darstellung wird hier sowie im Folgenden generell auf die angenommene Zahl von 10.000 Behandlungsepisoden bezogen. Dies erfolgt lediglich aufgrund der besseren Lesbarkeit.

Die im Zeitverlauf zu beobachtenden Schwankungen und inbesondere auch das erneute Ansteigen des Wertes des noch möglichen Erkenntnisgewinns beruhen darauf, dass der Wert des jeweils verbleibenden, weiteren möglichen Erkenntnisgewinns jeweils, pro Zeitpunkt, als Differenz angenommener Sicherheit (die Wahrscheinlichkeit, dass es sich bei dem jeweils besten Wirkstoff mit Sicherheit, $p_{bester} = 1$) mit der tatsächlich zu diesem Zeitpunkt bestehender Unsicherheit (alle Wirkstoffe sind mit einer bestimmten Wahrscheinlichkeit als beste anzusehen) handelt (vgl. Abschnitt II.2.3). Die Analyse zeigt, dass beide Werte sich im Zeitverlauf erheblich ändern (Abbildung 2). Die in Abbildung 1 dargestellten Werte zum korrigierten Therapieansprechen sind, jeweils pro Zeitpunkt, den Differenzen der in Abbildung 2 wiedergegebenen Werte gleich. Nicht nur die Differenz (der maximal verbleibende Informationswert) sondern auch dessen Komponenten verändern sich im Zeitverlauf. Insbesondere verändert sich auch der Wert des jeweils maximal erreichbaren im Hinblick auf die jeweilige Zielgröße (E_{max} in Abbildung 2, in Bezug auf das korrigierte Therapieansprechen). Er sinkt im Zeitverlauf deutlich ab, von einem initialen Wert (1993) von geschätzten 7.132 erfolgreich zu behandelten Episoden auf einen Wert von 5.589 geschätzten erfolgreich zu behandelnden Episoden (2009, entsprechend dem Stand von 2007 einschließlich unpublizierter Studien). Dies führt z. B. dazu, dass der Wert für das Jahr 1998 von 6.101, der den aktuellen Erkenntnisstand bemisst, über allen Werten, die den bestmöglichen Erkenntnisstand bemessen (6.091 für 1999, 5.866 für 2000 usw.) liegt. Die Darstellung in Abbildung 2 bezieht sich auf die Zielgröße des korrigierten Therapieansprechens. Die Ergebnisse der weiteren Zielgrößen Therapieabbruch und Therapieansprechen sind im Anhang (Abschnitt VII.2, Abbildung 13, Abbildung 14) dargestellt und ergeben wesentlich denselben Befund. Insbesondere zeigen sich auch höhere erwartete Therapieabbruchraten (Abbildung 14) zu späteren gegenüber früheren Zeitpunkten. Detaillierte tabellarisch dargestellte Ergebnisse zu allen drei Zielgrößen in Zeitverlauf finden sich im Anhang (Abschnitt VII.6, Tabelle 8, Tabelle 9, Tabelle 10).

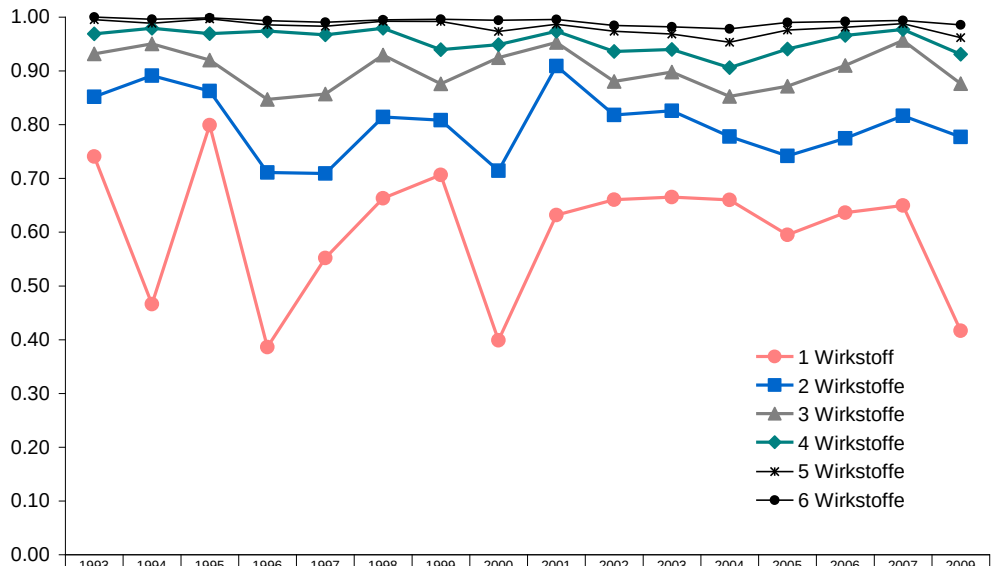


	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2009
E(max)	7132	7160	7112	6439	6341	6373	6091	5866	6078	6084	6032	5810	5835	5850	5823	5589
E(aktuell)	6634	6902	6787	6019	6054	6101	5926	5804	5955	5813	5770	5653	5685	5735	5733	5517

Abbildung 2 Erwartete Erfolgsraten (korrigiertes Therapieansprechen, E_{max}, $E_{aktuell}$) pro 10.000 Episoden

III.1.3 Wahrscheinlichkeit, die besten Wirkstoffe identifiziert zu haben im Verlauf

In Abbildung 3 finden sich unterschiedliche Wahrscheinlichkeiten, dass es sich bei unterschiedlichen Wirkstoffen bzw. Gruppen von Wirkstoffen um die jeweils besten handelt, bezogen auf die zusammengefaßte Zielgröße des korrigierten Therapieansprechens (Ergebnisse zu den anderen Zielgrößen finden sich im Anhang, Abschnitt VII.3, Abbildung 15, Abbildung 16). Für einzelne Wirkstoffe ("Gruppe" mit einem Wirkstoff) schwankt die Wahrscheinlichkeit im Zeitverlauf zwischen rund 0,39 und 0,80 (Jahre 1996/2000 bzw. Jahr 1995), dass es sich bei dem jeweiligen Wirkstoff um den besten aus der Gruppe der eingeschlossenen Wirkstoffe handelt. Zwar zeichnet sich, etwa ab dem Jahr 2001, hier eine Stabilisierung ab (die Wert zwischen 2001 und 2007 liegen alle bei einer Wahrscheinlichkeit von rund 0,6). Der Wert für das Jahr 2009 (unpublizierte Studien) liegt jedoch wiederum nurmehr bei einem Wert von rund 0,42.



	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2009
1 Wirkstoff	0.741	0.466	0.799	0.386	0.552	0.663	0.706	0.399	0.632	0.660	0.665	0.660	0.595	0.636	0.650	0.417
2 Wirkstoffe	0.852	0.891	0.863	0.711	0.709	0.814	0.808	0.714	0.909	0.818	0.826	0.778	0.742	0.774	0.816	0.777
3 Wirkstoffe	0.932	0.950	0.920	0.847	0.857	0.930	0.876	0.925	0.953	0.880	0.898	0.852	0.871	0.910	0.957	0.876
4 Wirkstoffe	0.969	0.979	0.969	0.974	0.967	0.979	0.939	0.949	0.974	0.936	0.940	0.906	0.940	0.966	0.977	0.931
5 Wirkstoffe	0.995	0.989	0.997	0.985	0.983	0.992	0.992	0.973	0.987	0.974	0.968	0.953	0.976	0.981	0.988	0.962
6 Wirkstoffe	1.000	0.996	0.998	0.993	0.990	0.995	0.996	0.994	0.996	0.984	0.982	0.978	0.990	0.992	0.994	0.985

Abbildung 3 Wahrscheinlichkeit, dass der beste Wirkstoff einer Gruppe von Wirkstoffen bestimmter Größe (1-6 Wirkstoffe) angehört im Zeitverlauf (Zielgröße korrigiertes Therapieansprechen)

Für Gruppen von Wirkstoffen ergeben sich notwendig jeweils pro Zeitpunkt (Jahr) höhere Wahrscheinlichkeiten, dass sich in der jeweiligen Gruppe der Wirkstoff mit den größten Erfolgsaussichten gemessen an der Zielgröße befindet (da die Wahrscheinlichkeit der Gruppenmitglieder hierfür addiert werden – für die 12er-Gruppe aller Wirkstoffe ist die Wahrscheinlichkeit, dass sich unter dieser der Beste befindet immer gleich 1). Die hier dargestellen Wahrscheinlichkeiten für Gruppen wurden gebildet, indem z. B. in den Gruppe von jeweils zwei Wirkstoffen pro Zeitpunkt (Jahr) immer diejenigen Wirkstoffe mit der höchsten und der zweihöchsten Wahrscheinlichkeit, dass es sich um den besten Wirkstoff handelt zusammengefaßt und deren Wahrscheinlichkeiten addiert wurden. Für Gruppen mit einer höheren Anzahl von Wirkstoffen wurde entsprechend verfahren. Zu berücksichtigen ist zum einem, dass die Anzahl der insgesamt in die Analye einbezogenen Wirkstoffe zu den einzelnen Analysezeitpunkten unterschiedlich ist (vgl. Tabelle 1), sodass sich z. B. für das Jahr 1996 in der Summe der Wahrscheinlichkeiten für eine Gruppe von 6 Wirkstoffen notwendig der Wert 1 ergibt, da zu diesem Zeitpunkt nur 6 Wirkstoffe einbezogen sind und es sich also um die Gesamtgruppe aller Wirkstoffe handelt, die somit notwendig auch den relativ besten Wirkstoff enthält. Zum anderen ist zu berücksichtigen, dass im Zeitverlauf die Gruppen jeweils anders zusammengesetzt sein können, da die

Zuordnung ja auf den jeweils beobachteten Erfolgswahrscheinlichkeiten basiert, die sich im Zeitverlauf ändern können.

Erkennbar ist generell, dass mit zunehmender Gruppengröße eine wesentliche Erhöhung der Wahrscheinlichkeit, dass der beste Wirkstoff sich in der jeweiligen Gruppe befindet, verbunden ist. Werden 6 Wirkstoffe in die Gruppenbildung einbezogen, so erreichen die Wahrscheinlichkeiten zu allen Zeitpunkten, auch in den Jahren ab 2002 in denen alle 12 Wirkstoffe einbezogen sind, Werte um 0,99. Die Wahrscheinlichkeit, dass sich der jeweils beste Wirkstoff in einer dieser Gruppen befindet ist also sehr hoch bzw. ist die Wahrscheinlichkeit, dass es sich bei einem der übrigen Wirkstoffe um den besten handelt sehr gering. Auch bei kleineren Gruppengrößen werden bereits höhere Wahrscheinlichkeiten von rund 0,80 (2 Wirkstoffe), rund 0, 90 (3 Wirkstoffe) und rund 0,95 (4 Wirkstoffe) erzielt, dass sich der beste Wirkstoff in der Gruppe der jeweiligen Größe befindet.

Die Zusammensetzung dieser Gruppen ändert sich im Zeitverlauf allerdings erheblich (Tabelle 2). Ab dem Jahr 2002 (ab dem erstmals alle Wirkstoffe in die Analyse einbezogen werden können) sind jedoch nur noch 6 Wirkstoffe im weiteren Verlauf auf den ersten 4 Rängen, bezogen auf die Zielgrößen zum Therapieansprechen, vertreten: ESCIT, MIRTA, DULOX, VENLA, BUPRO und SERTR (BUPRO und DULOX sind 2009 nicht unter den ersten 4 Wirkstoffen). Entsprechend gilt bezogen auf den Therapieabbruch, dass folgende 5 Wirkstoffe vertreten sind: ESCIT, BUPRO, DULOX, CITAL und SERTR (DULOX ist 2009 nicht unter den 4 ersten Wirkstoffen).

Tabelle 2 Ränge der Wirkstoffe in Bezug auf die Wahrscheinlichkeit, dass es sich bei dem jeweiligen Wirkstoff um den jeweils besten handelt im Zeitverlauf (erste 4 Ränge, alle Zielgrößen)

Zielgröße	Jahr	Rang (1)	Rang (2)	Rang (3)	Rang (4)
Th.-An.	1993	SERTR	BUPRO	CITAL	FLUVO
Th.-An.	1994	VENLA	SERTR	CITAL	BUPRO
Th.-An.	1995	VENLA	BUPRO	CITAL	SERTR
Th.-An.	1996	VENLA	SERTR	BUPRO	CITAL
Th.-An.	1997	MIRTA	VENLA	BUPRO	SERTR
Th.-An.	1998	MIRTA	BUPRO	VENLA	SERTR
Th.-An.	1999	BUPRO	SERTR	MIRTA	VENLA
Th.-An.	2000	MIRTA	ESCIT	BUPRO	VENLA
Th.-An.	2001	ESCIT	MIRTA	VENLA	SERTR
Th.-An.	2002	ESCIT	MIRTA	DULOX	VENLA
Th.-An.	2003	ESCIT	MIRTA	DULOX	VENLA
Th.-An.	2004	ESCIT	MIRTA	VENLA	DULOX
Th.-An.	2005	ESCIT	MIRTA	DULOX	VENLA
Th.-An.	2006	ESCIT	MIRTA	DULOX	BUPRO
Th.-An.	2007	ESCIT	MIRTA	BUPRO	VENLA
Th.-An.	2009	MIRTA	ESCIT	VENLA	SERTR
Th.-Ab.	1993	SERTR	CITAL	BUPRO	PAROX
Th.-Ab.	1994	VENLA	SERTR	CITAL	BUPRO
Th.-Ab.	1995	VENLA	SERTR	CITAL	BUPRO
Th.-Ab.	1996	SERTR	CITAL	BUPRO	VENLA
Th.-Ab.	1997	MIRTA	SERTR	BUPRO	VENLA
Th.-Ab.	1998	MIRTA	SERTR	BUPRO	CITAL
Th.-Ab.	1999	BUPRO	CITAL	SERTR	MILNA
Th.-Ab.	2000	BUPRO	ESCIT	MIRTA	CITAL
Th.-Ab.	2001	ESCIT	BUPRO	MIRTA	CITAL
Th.-Ab.	2002	ESCIT	DULOX	BUPRO	CITAL
Th.-Ab.	2003	ESCIT	BUPRO	DULOX	CITAL
Th.-Ab.	2004	BUPRO	DULOX	ESCIT	CITAL
Th.-Ab.	2005	BUPRO	ESCIT	DULOX	SERTR
Th.-Ab.	2006	ESCIT	BUPRO	DULOX	CITAL
Th.-Ab.	2007	ESCIT	BUPRO	CITAL	SERTR
Th.-Ab.	2009	ESCIT	BUPRO	SERTR	CITAL
Th-An. (k.)	1993	SERTR	CITAL	BUPRO	PAROX
Th-An. (k.)	1994	SERTR	VENLA	CITAL	BUPRO
Th-An. (k.)	1995	VENLA	SERTR	CITAL	BUPRO
Th-An. (k.)	1996	SERTR	VENLA	CITAL	BUPRO
Th-An. (k.)	1997	MIRTA	SERTR	VENLA	BUPRO
Th-An. (k.)	1998	MIRTA	SERTR	BUPRO	VENLA
Th-An. (k.)	1999	BUPRO	SERTR	MIRTA	VENLA
Th-An. (k.)	2000	ESCIT	MIRTA	BUPRO	SERTR
Th-An. (k.)	2001	ESCIT	MIRTA	BUPRO	VENLA
Th-An. (k.)	2002	ESCIT	DULOX	MIRTA	BUPRO
Th-An. (k.)	2003	ESCIT	DULOX	BUPRO	MIRTA
Th-An. (k.)	2004	ESCIT	BUPRO	MIRTA	DULOX
Th-An. (k.)	2005	ESCIT	MIRTA	BUPRO	DULOX
Th-An. (k.)	2006	ESCIT	BUPRO	MIRTA	DULOX
Th-An. (k.)	2007	ESCIT	BUPRO	MIRTA	SERTR
Th-An. (k.)	2009	ESCIT	MIRTA	SERTR	VENLA

Th.-An.: Therapieansprechen, Th.-Ab.: Therapieabbruch, Th.-An. (k.): korrigiertes Therapieansprechen

Die Ergebnisse für die Wahrscheinlichkeiten, dass es sich bei einem der Wirkstoffe um den jeweils besten handelt zeigt Abbildung 4 im Zeitverlauf (1993-2009) für das korrigierte Therapieansprechen.

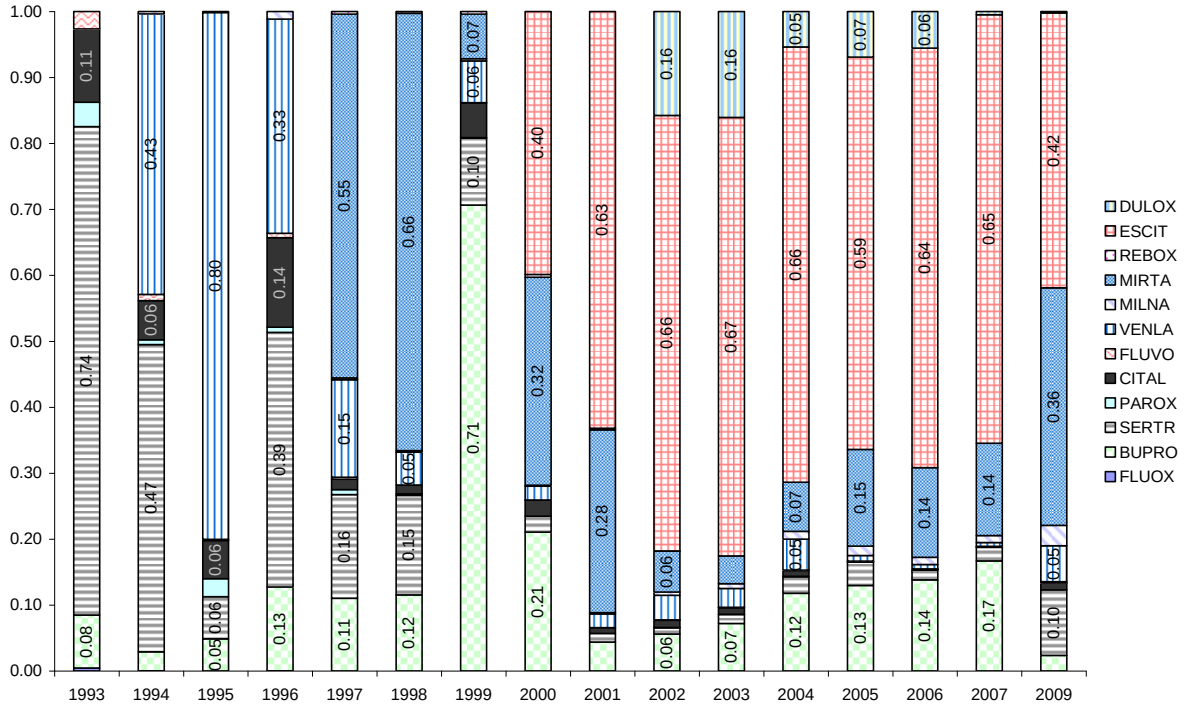


Abbildung 4 Wahrscheinlichkeit, dass es sich bei einem jeweiligen Wirkstoff um den besten handelt, im Zeitverlauf (Zielgröße korrigiertes Therapieansprechen)

Im Endergebnis (2009, alle bis 2007 publizierten und unpublizierten Studien) ist die Wahrscheinlichkeit, dass es sich jeweils um den besten Wirkstoff handelt, für ESCIT (0,42), MIRTA (0,36), SERTR (0,10) und VENLA (0,05) am höchsten. In der Summe ergibt sich für die Wahrscheinlichkeit, dass der beste Wirkstoffe in dieser Gruppe liegt ein Wert von 0,93. Im Zeitverlauf ist zu beobachten, dass für die Mehrzahl der betrachteten Zeitpunkte (Jahre), mit Ausnahme der Jahre 1994, 1996 und 2000 sowie 2009, ein Wirkstoff mit einer Wahrscheinlichkeit > 0,5 als vermutlich bester Wirkstoff, im Rahmen der vorliegenden Analyse, bezeichnet werden kann. Insbesondere für die Jahre ab 2001 gilt, dass für diese immer ein Wirkstoff eine entsprechende Wahrscheinlichkeit von > 0,5 aufwies, dies jedoch für das Jahr 2009, d. h. unter Berücksichtigung der unpublizierten Studien, nicht mehr gilt.

Die Betrachtung der in Relation der Wirkstoffe zueinander tatsächlich zu erwartenden Therapieeffekte (T-Werte, Ereignisraten, bezogen auf die Zielgrößen) zeigt, dass im Zeitverlauf (insbesondere in den ersten Jahren) generell eine Verringerung der Unterschiede zwischen den Wirkstoffen zu beobachten war (Abbildung 5), wie dies auch in der nicht

nach Wirkstoffen differenzierten Analyse zum Verlauf des Erkenntnisgewinns zu beobachten war (vgl. Abbildung 2).

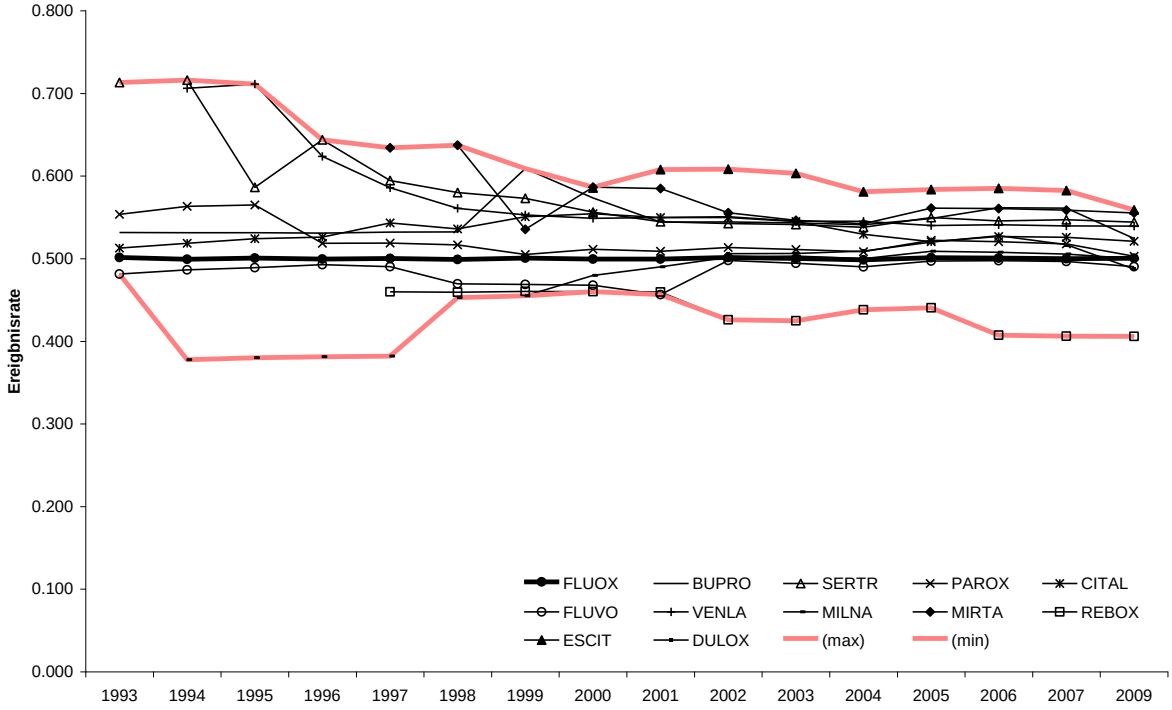


Abbildung 5 Ereignisrate (T-Werte) der Wirkstoffe im Zeitverlauf (Zielgröße korrigiertes Therapieansprechen)

Dies wird auch durch die Betrachtung der Veränderungen im Zeitverlauf seit 2002 deutlich (Abbildung 6). Für die Mehrzahl der Wirkstoffe resultiert zu Beginn des Zeitraums eine höhere erwartete Ereignisrate (korrigiertes Therapieansprechen) als am Ende dieses Zeitraums. Auch die Betrachtung der 2009 besten Wirkstoffe über den gesamten Untersuchungszeitraum zeigt eine charakteristische Veränderung im Zeitverlauf (Abbildung 7). Für die beiden Wirkstoffe, die bereits seit Beginn des Untersuchungszeitraumes (1993) in die Analyse eingeschlossen wurden (SETRA und VENLA), ist in den ersten Jahren (etwa bis zum Jahr 2001) ein deutlicher Rückgang der Ereignisrate zu beobachten, in den Folgejahren jedoch allenfalls noch vergleichsweise sehr geringe Veränderungen. Dies gilt in ähnlicher Weise auch für MIRTA, für das Ergebnisse ab 1997 einbezogen wurden. Lediglich für ESCIT ist ein anderer Verlauf erkennbar, mit jedoch vergleichsweise geringen Veränderungen im dem Zeitverlauf seit 2002, für den betreffend dieses Wirkstoffes erstmals Ergebnisse einbezogen wurden.

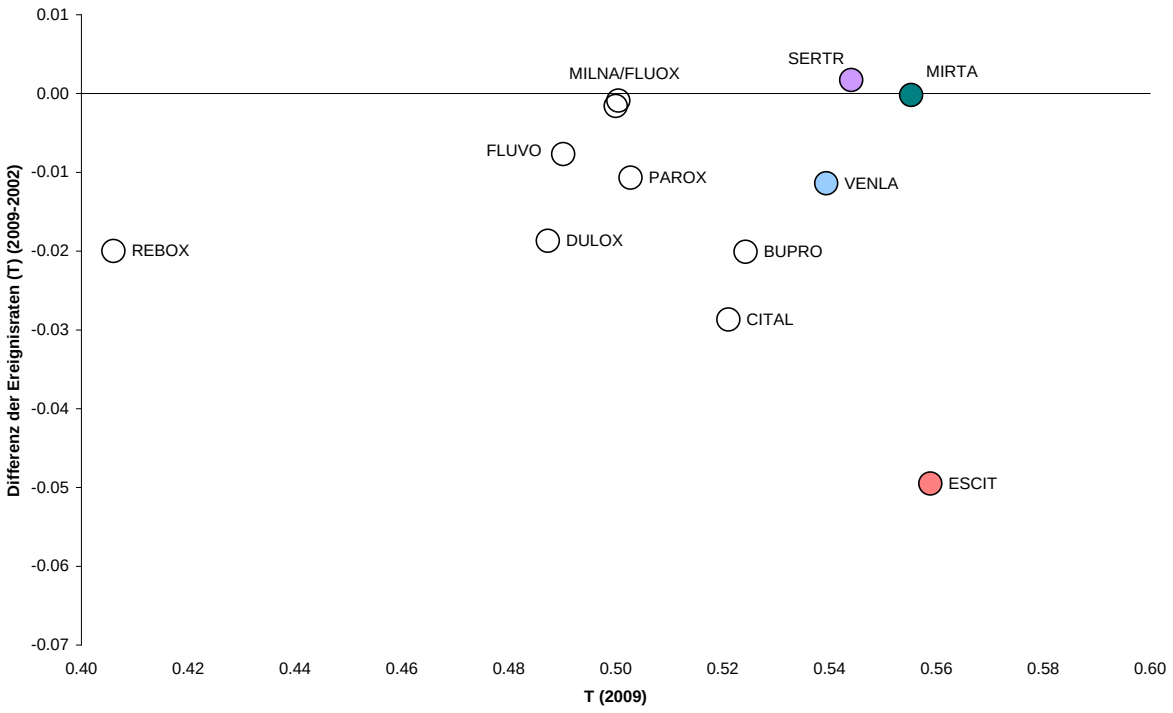


Abbildung 6 Ereignisrate (T-Werte) 2009 und Veränderung der Ereignisrate (T-Werte) der Wirkstoffe im Zeitverlauf (2002-2009) (Zielgröße korrigiertes Therapieansprechen, die 2009 besten Wirkstoffe sind farblich hervorgehoben.)

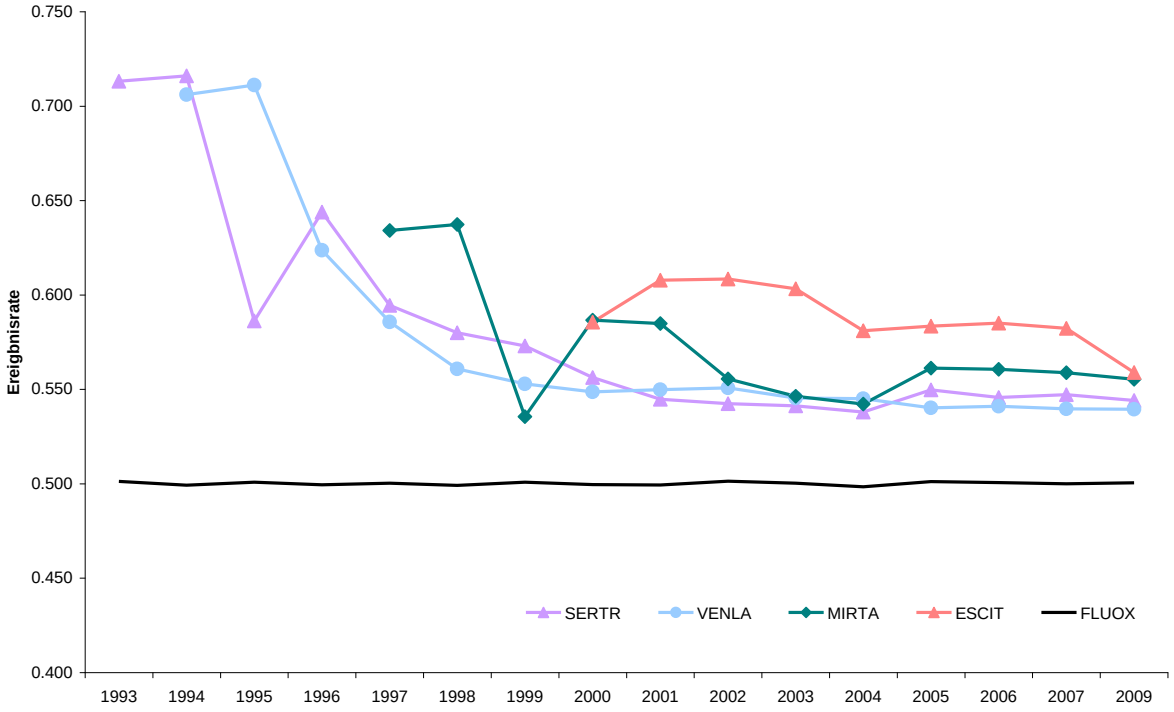


Abbildung 7 Ereignisrate (T-Werte) ausgewählter Wirkstoffe im Zeitverlauf (Zielgröße korrigiertes Therapieansprechen)

III.1.4 Gegenüberstellung der Ergebnisse direkter und gemischter Vergleiche

Die Gegenüberstellung der Ergebnisse von direkten und gemischten Vergleichen erfolgte auf der Basis aller paarweise möglichen Wirkstoffvergleiche zu allen Zeitpunkten in Bezug auf die beiden Zielgrößen Therapieansprechen und Therapieabbruch. Dabei ergaben sich insgesamt 186 Wertepaare. Die Prüfung erfolgte daraufhin, ob der gemischt ermittelte Wert ($\ln OR_{indirekt}$) im 95%-Konfidenzintervall des direkt ermittelten Wertes ($\ln OR_{direkt}$) lag. Diese Prüfung zeigte lediglich in 2 von 186 Vergleichen eine erhebliche Abweichung von direkten und gemischten Werten (Zielgröße Therapieabbruch, Vergleich von CITAL und SERTR 1997; Zielgröße Therapieabbruch, CITAL und FLUOX im Jahr 2003, s. Abbildung 8, vgl. Abschnitt II.2.2.3.). Die entsprechenden Ergebnisse für die Zielgröße Therapieansprechen (Abbildung 17) finden sich im Anhang (Abschnitt VII.4). Die Ergebnisse aller direkt vergleichenden Meta-Analysen finden sich ebenfalls im Anhang (Abschnitt VII.5).

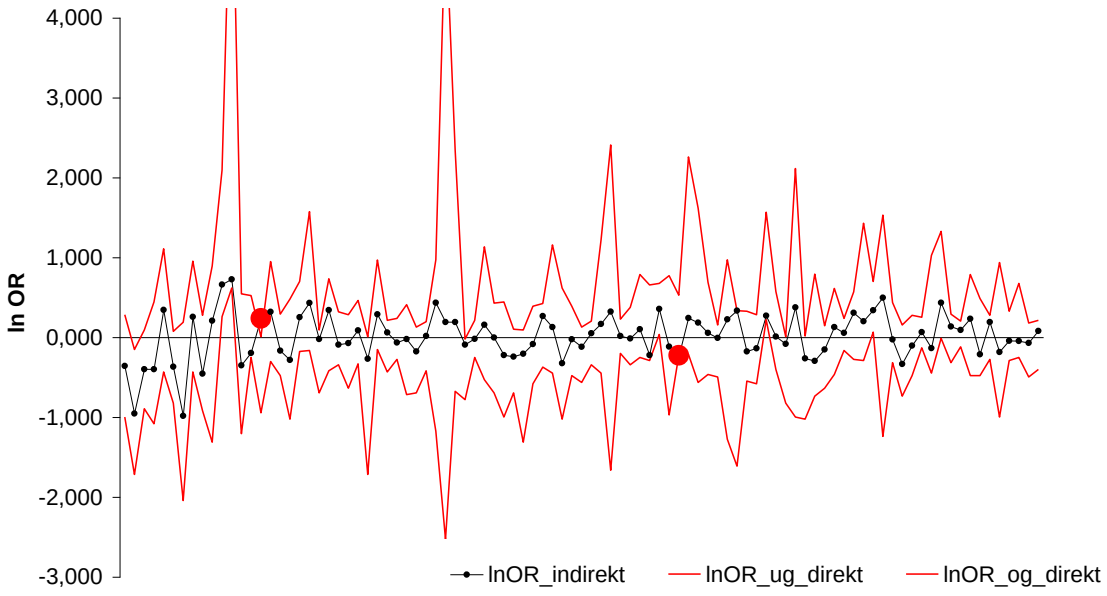


Abbildung 8 Gegenüberstellung gemischter (Punktschätzer) und direkter Vergleiche (Konfidenzintervall) für die Zielgröße Therapieabbruch (rote Punkte: gemischter Wert außerhalb des 95%-Konfidenzintervalles für den direkten Schätzer)

III.2 Faktischer und alternativer Erkenntnisgewinn im Zeitverlauf

Im folgenden Abschnitt wird zunächst die Anzahl der in die Analyse eingeschlossenen Studien, Patienten und Wirkstoffe angegeben (Abschnitt III.2.1). Darauf folgend (Abschnitt III.2.2) werden die Ergebnisse zum Verlauf der jeweils verbleibenden Erkenntnisunsicherheit berichtet. Daran schließt sich eine Darstellung der Ergebnisse für die einzelnen Wirkstoffe an (Abschnitt III.2.3). In Abschnitt III.2.4 werden einzelne Ergebnisse der Ermittlung des alternativen Verlaufs der Erkenntnisgewinnung beschrieben.

III.2.1 In die Analyse eingeschlossene Studien-, Patienten- und Wirkstoffzahlen

Insgesamt wurden in die vergleichende Analyse von faktischem und alternativem Erkenntnisverlauf die Ergebnisse von 43 weiteren Studien (Publikation ab 2003 und unpublizierte Studien) aller 12 berücksichtigten Wirkstoffe, bezogen auf die Zielgröße des korrigierten Therapieansprechens, einbezogen (3). Eine detaillierte Auflistung der Studien ist im Anhang (Abschnitt VII.8, Tabelle 14) wiedergegeben.

Tabelle 3 Anzahl einbezogener Studien, Patienten und Wirkstoffe (Vergleich der faktischen und alternativen Studienreihenfolge, Zielgröße korrigiertes Therapieansprechen)

Jahr	Studien (Jahr) [n]**	Studien kumuliert [n]**	Patienten (Jahr) [n]** - faktisch	Patienten (Jahr) [n]** - alternativ	Patienten kumuliert [n]** - faktisch	Patienten kumuliert [n]** - alternativ
bis 2002	67	67	12,862	12,862	12,862	12,862
2003	6	73	608	1,886	13,470	14,748
2004	5	78	1,134	1,193	14,604	15,941
2005	8	86	1,820	2,480	16,424	18,421
2006	9	95	2,320	2,068	18,744	20,489
2007	4	99	1,246	699	19,990	21,188
2009*	11	110	3,766	2,568	23,756	23,756

*Alle unpublizierten Studien wurden unter dem Jahr 2009 (Publikationsjahr von Cipriani2009a) erfasst.
**Die Anzahl der Studien (und entsprechend die Zahl der Patienten), für die Ergebnisse zur Zielgröße "Therapieansprechen" verfügbar waren, zählen die jeweils zusätzlichen "Studien", die aus der Zerlegung der beiden dreiarmigen Studien resultieren (vgl. Abschnitt II.1.4) mit.

Die Anzahl der jeweils in einem weiteren Zeitschritt (Jahr) ab 2003 einzubeziehenden Studien entsprach genau der Zahl, die auch im faktischen Verlauf in diesem Zeitschritt einbezogen wurden (z. B. kamen zwischen 2003 und 2004 genau 5 Studien hinzu, da sie in diesem Zeitraum publiziert worden waren. Somit wurden bei der Gestaltung des alter-

nativen Verlaufs auch genau 5 zusätzliche Studien in diesem Zeitschritt einbezogen). Es ergeben sich im Zeitverlauf allerdings unterschiedliche Patientenzahlen, da die jeweils gleiche Zahl von Studien im alternativen Verlauf, da es sich ja um andere Studien handelte bzw. handeln konnte, in der Summe andere Patientenzahlen aufwiesen. Die Zahl der Patienten wächst demnach im alternativen Verlauf der Erkenntnisgewinnung zunächst schneller als im faktisch zu beobachtenden Verlauf. Insbesondere ist dies für das Jahr 2003 der Fall. Dieses Verhältnis kehrt sich im weiteren Zeitverlauf (ab 2006) um. Studien, denen entsprechend der bei der Erstellung des alternativen Verlaufs ein höherer, erwarteter Erkenntnisgewinn zugeordnet wurde, hatten demnach im Mittel höhere Patientenzahl eingeschlossen als Studien, denen ein relativ geringerer, erwarteter Erkenntnisgewinn zugeordnet wurde.

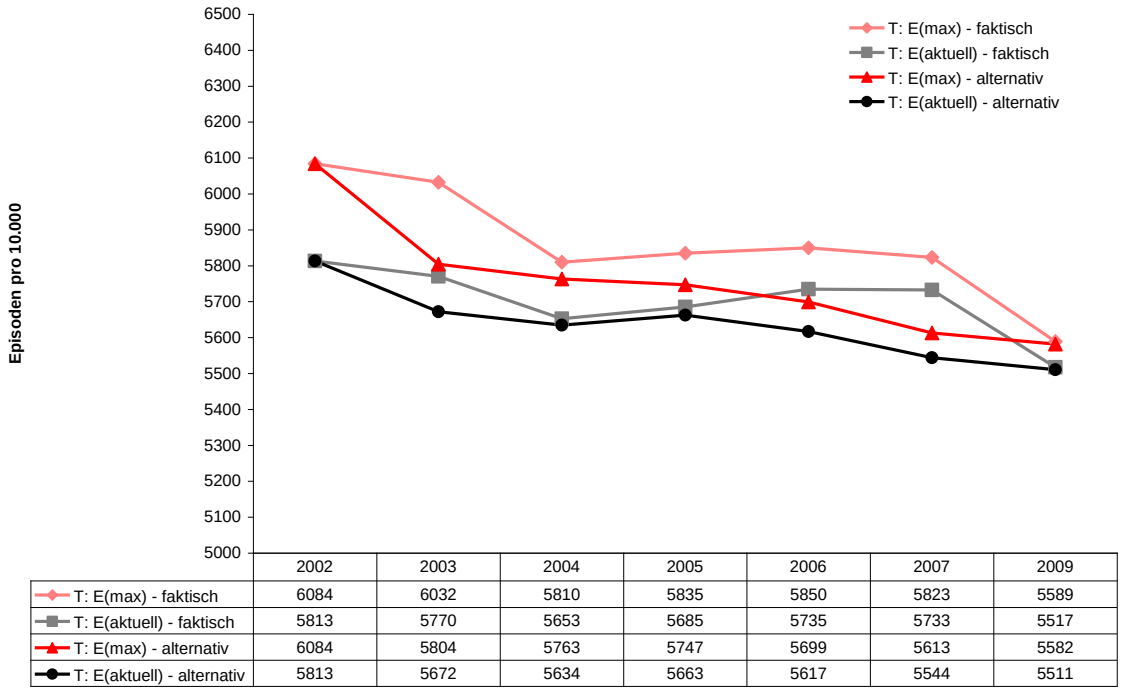
III.2.2 Vergleichender Verlauf der verbleibenden Erkenntnisunsicherheit

Der Verlauf, als zeitliche Reihung von Werten des jeweils nach einem Zeitpunkt (Jahr) verbleibenden, maximal möglichen weiteren Erkenntnisgewinns (V-Werte, entsprechend Abschnitt II.2.3) ist in Abbildung 9 vergleichend für die faktische (Publikations-)Reihenfolge der eingeschlossenen Studien gegenüber einer alternativ erarbeiteten Reihenfolge (vgl. Abschnitt II.3) für die Zielgröße des korrigierten Therapieansprechens dargestellt. Dabei ist zu berücksichtigen, dass gemäß der Methodik ein alternativer Verlauf lediglich für den Verlauf des Erkenntnisgewinns nach dem Jahr 2002 ermittelt wurde, sodass nur Daten für die Jahre 2003 und folgende dargestellt sind.

Abbildung 9 Wert des maximal möglichen, weiteren Erkenntnisgewinns (Vergleich der faktischen und alternativen Studienreihenfolge, Zielgröße korrigiertes Therapieansprechen

Für alle Zeitpunkte (Jahre) die in die vergleichende Analyse einbezogen waren und somit für den gesamten Zeitverlauf und in der Summe der Werte des jeweils noch verbleibenden Potenzials möglichen Erkenntnisgewinns (V-Werte), zeigt sich ein schnellerer und somit günstigerer Verlauf des Erkenntnisgewinns im alternativen Verlauf.

Die Veränderungen im Zeitverlauf können auch hier, analog zur Darstellungen des faktischen Verlaufs der Erkenntnisgewinnung (vgl. Abschnitt III.1.2, insbesondere Abbildung 2) vor dem Hintergrund des Verlaufs der jeweils erwarteten Ergebnisse bei unterstellter "perfekter" gegenüber aktuell verfügbarer Erkenntnis betrachtet werden (Abbildung 10). Das Absinken der erwarteten Werte (korrigiertes Therapieansprechen) findet sich sowohl bezogen auf den faktischen, als auch bezogen auf den alternativen Verlauf. Sowohl für die erwarteten Ergebnisse bei unterstellter perfekter Erkenntnis als auch für die erwarteten Ergebnisse bei jeweils aktueller Erkenntnis zeigt sich dabei ein schnellerer Rückgang im alternativen im Vergleich zum faktischen Verlauf. Für die Jahre 2006 und 2007 hat dies z. B. zur Folge, dass die möglichen Ergebnisse (unterstellter, perfekter Erkenntnisstand) entsprechend des alternativen Verlaufes bereits ungünstiger (5.699 bzw. 5.613) sind, als die nach aktuellem Erkenntnisstand (noch) erreichbaren Erfolge (5.735 bzw. 5.733).



	2002	2003	2004	2005	2006	2007	2009
T: E(max) - faktisch	6084	6032	5810	5835	5850	5823	5589
T: E(aktuell) - faktisch	5813	5770	5653	5685	5735	5733	5517
T: E(max) - alternativ	6084	5804	5763	5747	5699	5613	5582
T: E(aktuell) - alternativ	5813	5672	5634	5663	5617	5544	5511

Abbildung 10 Erwartete Erfolgsraten (korrigiertes Therapieansprechen) pro 100.000 Episoden (perfekte Information T: E(max) und aktuelle Information T: E(aktuell)) in der Gegenüberstellung von faktischem und alternativem Verlauf

Differenzierte Ergebnisse zum alternativen Verlauf der Erkenntnisgewinnung in Bezug auf alle Zielgrößen sind im Anhang (Abschnitt VII.8) dargestellt.

III.2.3 Ergebnisse für einzelne Wirkstoffe

Für die Wirkstoffe die 2009 (bezogen auf das korrigierte Therapieansprechen) die ersten 4 Ränge einnehmen zeigt Abbildung 11 den Verlauf der Ereignisrate vergleichend zwischen alternativem und faktischem Verlauf der Erkenntnisgewinnung.

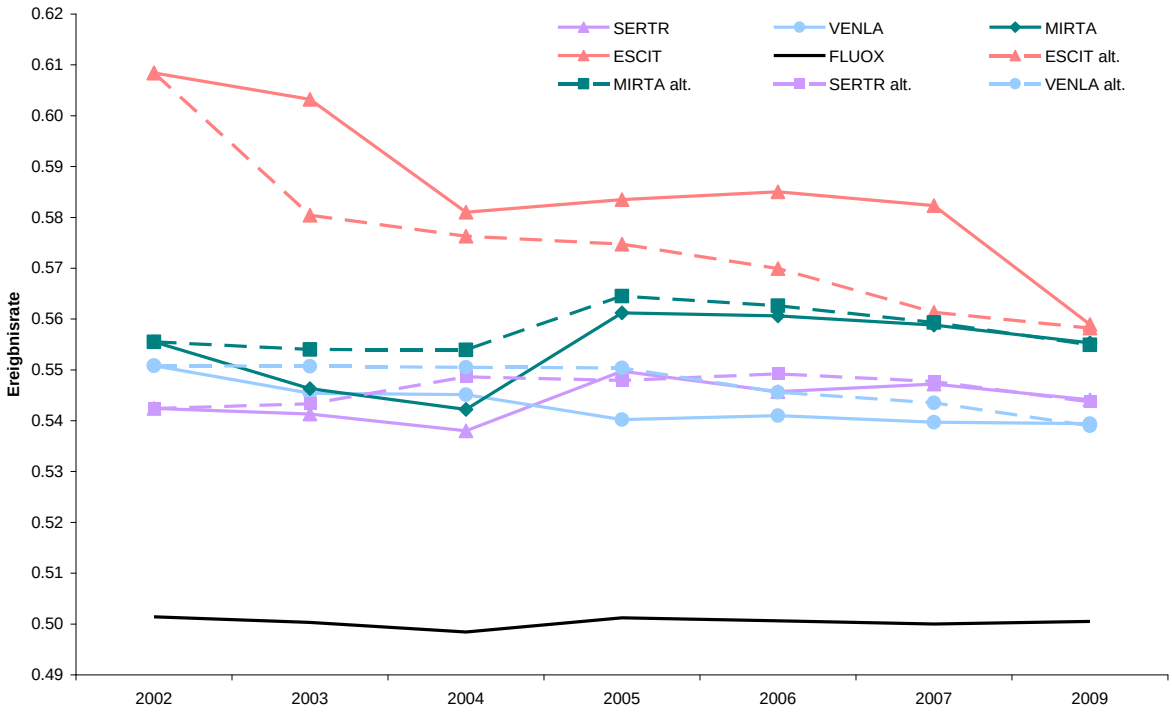


Abbildung 11 Ereignisrate (T-Werte) ausgewählter Wirkstoffe im Zeitverlauf, vergleichend faktischem gegenüber alternativem Erkenntnisgewinn im Zeitverlauf (Zielgröße korrigiertes Therapieansprechen)

Dabei ist erkennbar, dass für die Wirkstoffe ESCIT und MIRTA, inbesondere für Escitalopram, ein alternativer Verlauf resultiert, der im Mittel früher im Zeitverlauf zu Werten näher am Wert für 2009 führt als der faktische Verlauf.

III.2.4 Ermittlung des alternativen Verlaufs der Erkenntnisgewinnung

Die Ermittlung des alternativen Verlaufs der Erkenntnisgewinnung basiert auf der bevorzugten Berücksichtigung zunächst der Studien, die auf Grund der in ihnen jeweils untersuchten Wirkstoffvergleiche insbesondere einen weiteren Erkenntnisgewinn versprechen. Dazu wird der Wert aller möglichen paarweisen Wirkstoffvergleiche (Informationswert der paarweisen Vergleiche, W-Wert) ermittelt (s. Abschnitt II.3). In Abbildung 12 sind alle W-Werte für alle paarweisen Vergleiche aller Jahre (Erkenntnisstände jeweils nach 2002) gegen die Anzahl der Studien, in denen der jeweilige Vergleich untersucht wurde, aufgetragen.

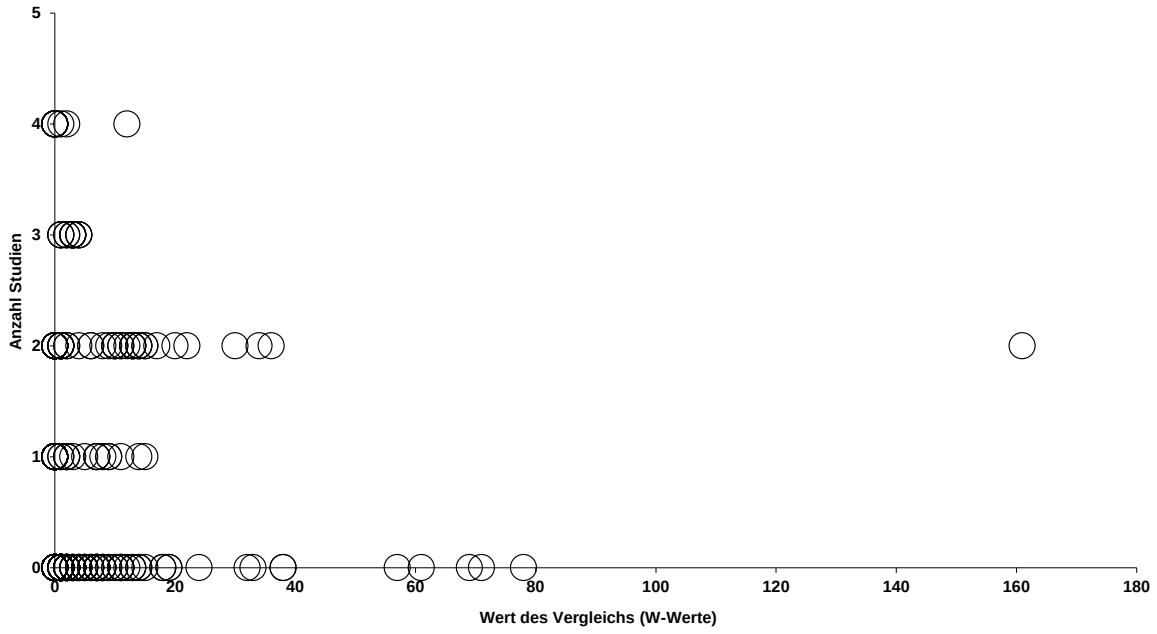


Abbildung 12 Bewertung von Wirkstoffvergleichen in Studien und Anzahl verfügbarer Studien (alternativer Verlauf der Erkenntnisgewinnung)

Es wird deutlich wie einerseits für viele Vergleiche, auch relativ hohen Informationswertes (größere Werte auf der horizontalen Achse) keine Studien zur Verfügung (Wert 0 auf der vertikalen Achse) standen. Andererseits untersuchen auch viele Studien Vergleiche, denen kein hoher Informationswert zuzuordnen war. Insgesamt ist das sich bietende Bild von einer „Idealsituation“, in der insbesondere viele Studien für Vergleiche mit hohem Informationwert vorlägen, erkennbar deutlich entfernt.

Eine detaillierte Darstellung im Hinblick auf den zeitlichen Verlauf der Entwicklung des Informationswertes einzelner paarweise Wirkstoffvergleiche und das Vorliegen für Studien zeigt Tabelle 4. Eine zusätzliche, detaillierte Auflistung der Studien, mit Publikationsjahr und Jahr der Verwendung bzw. Einbeziehung in die Ermittlung des alternativen Verlaufs der Erkenntnisgewinnung ist im Anhang (Abschnitt VII.8, Tabelle 14) wiedergegeben.

Zunächst zeigt sich, dass die Informationswerte der paarweisen Vergleiche sich jeweils erheblich voneinander unterscheiden. Für mehr als die Hälfte alle Vergleiche ergibt sich schon für das Jahr (nach) 2002 praktisch kein Informationswert mehr und es tritt auch keine erhebliche Änderung im Zeitverlauf mehr ein. Für die Wirkstoffvergleiche mit (höherem) Informationswert (nach) 2002 ist ein erheblicher Rückgang besonders unmittelbar

nach dem Jahr 2002 zu verzeichnen. Im Hinblick auf die Anzahl der Studien, die jeweils einen der paarweisen Vergleiche untersuchten zeigt auch in der differenzierten Betrachtung kein Zusammenhang zwischen dem Vorliegen von Studien (s. letzte Spalte von Tabelle 4) überhaupt oder deren Anzahl und dem Informationswert des untersuchten Vergleiches. Allein 21 der 43 nach 2002 publizierten (oder unpublizierten) Studien, somit knapp die Hälfte, gilt Vergleichen ohne relevanten Informationswert schon 2002 und ohne relevante Veränderung im Zeitverlauf.

Der insgesamt schnellere Rückgang verbleibender Unsicherheit bzw. die schnellere Reduktion des verbleibenden Informationswertes spiegelt sich auch in Tabelle 4. Bei den Vergleichen von relativ hohem Wert (oben in der Tabelle) konnten Studien im Rahmen der Konstruktion des alternativen Verlaufs „vorgezogen“ werden. Dies ist in der Tabelle durch das Verhältnis der grünen zu den grauen Hinterlegungen ablesbar. Im unteren Teil der Tabelle, in dem die Vergleiche mit geringem bzw. nahezu keinem Informationswert wiedergegeben sind, verhält es sich tendenziell umgekehrt.

Zugleich ist auch in Tabelle 4 deutlich erkennbar, dass für viele Vergleiche mit relativ hohem Informationswert gar keine Studien vorlagen. Zunehmend mussten im Zeitverlauf der Konstruktion des alternativen Verlaufes Studien einbezogen werden, die Wirkstoffvergleichen mit allenfalls sehr geringem Informationswert galten.

Die Ergebnisse auch dieser „diagnostischen“ Untersuchung in Bezug auf den Versuch einen alternativen Erkenntnisgewinnungsprozess zu konstruieren, nährt somit die Vermutung, dass der günstigere, mögliche Verlauf der Erkenntnisgewinnung durch eine hypothetisch andere zeitliche Abfolge der Einbeziehung der tatsächlich durchgeführten Studien, wie hier dargestellt, lediglich einen Teil möglicher Verbesserungen der Erkenntnisgewinnung im Zeitverlauf abzubilden vermag. Ein womöglich noch größerer Teil bleibt ungenutzt, da Studien zur Erkenntnisgewinnung zu insbesondere interessanten Wirkstoffvergleichen nicht, oder nicht in ausreichendem Maße, vorliegen und stattdessen Studien durchgeführt wurden, die zumindest zu dem hier unterstellten primären Zweck, den besten Wirkstoff zu finden, allenfalls wenig beitragen.

Tabelle 4 Informationswerte (W-Werte) für Wirkstoffkombinationen (bezogen auf 10.000 fiktive Behandlungsepisoden), Publikation und Verwendung von Studien im Zeitverlauf (absteigend sortiert nach dem W-Wert 2002)

Informationswerte für Wirkstoffkombinationen (W-Werte) und Studien

		Jahr							Studien
Wirkstoff (A)	**Wirkstoff (B)**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2009**	**gesamt**
ESCIT	DULOX	161	9	10	1	1	2	2	2
MIRTA	DULOX	78	6	7	1	1	2	2	0
VENLA	DULOX	71	5	7	1	1	1	1	0
CITAL	DULOX	69	4	3	0	0	0	1	0
BUPRO	DULOX	61	6	7	1	1	1	1	0
SERTR	DULOX	57	4	7	1	1	1	2	0
BUPRO	ESCIT	36	30	34	13	15	15	8	2
MIRTA	ESCIT	33	38	38	32	24	7	12	0
VENLA	ESCIT	22	17	20	12	10	10	11	2
PAROX	DULOX	12	1	2	0	0	0	0	4
BUPRO	MIRTA	6	2	2	9	11	14	8	0
SERTR	ESCIT	6	13	13	11	14	13	14	2
CITAL	ESCIT	6	14	1	1	1	1	4	2
MILNA	ESCIT	5	10	11	13	15	19	19	0
BUPRO	VENLA	4	3	3	3	3	4	4	3
BUPRO	CITAL	3	5	1	1	1	1	0	0
MILNA	MIRTA	3	7	8	11	14	18	18	0
BUPRO	MILNA	2	7	8	7	8	11	8	0
CITAL	MILNA	2	5	3	3	3	2	7	0
VENLA	MILNA	2	6	7	8	10	13	13	0
SERTR	MILNA	2	5	7	8	11	14	15	1
VENLA	MIRTA	2	2	3	7	7	9	9	1
SERTR	CITAL	1	0	1	1	1	1	2	0
SERTR	MIRTA	1	4	2	7	9	11	11	0
PAROX	MILNA	1	1	2	2	1	2	1	0
CITAL	MIRTA	1	4	1	1	1	1	3	0
FLUVO	ESCIT	1	1	1	1	1	1	2	0
SERTR	VENLA	1	3	1	1	2	2	3	3
FLUOX	SERTR	0	0	0	0	0	0	0	0
FLUOX	FLUVO	0	0	0	0	0	0	0	0
FLUOX	MILNA	0	0	0	0	0	0	0	0
FLUOX	DULOX	0	0	0	0	0	0	0	0
BUPRO	SERTR	0	4	3	2	5	7	5	0
BUPRO	PAROX	0	0	0	0	0	0	0	0
BUPRO	FLUVO	0	1	1	1	1	1	1	0
BUPRO	REBOX	0	0	0	0	0	1	0	0
SERTR	PAROX	0	0	0	0	0	0	0	0
SERTR	FLUVO	0	0	1	1	1	1	1	0
PAROX	VENLA	0	0	0	0	0	0	0	0
PAROX	REBOX	0	0	0	0	0	0	0	0
CITAL	FLUVO	0	0	0	0	0	0	1	0
CITAL	VENLA	0	3	1	1	1	1	2	0
FLUVO	VENLA	0	1	1	1	1	1	1	0
FLUVO	MILNA	0	0	0	0	0	0	0	0
FLUVO	MIRTA	0	1	1	1	1	1	1	0
FLUVO	REBOX	0	0	0	0	0	0	0	0
FLUVO	DULOX	0	0	0	0	0	0	0	0
MILNA	REBOX	0	0	0	0	0	0	0	0
MILNA	DULOX	0	0	1	0	0	0	0	0
MIRTA	REBOX	0	0	0	0	0	1	0	0
REBOX	ESCIT	0	0	0	0	1	1	0	0
REBOX	DULOX	0	0	0	0	0	0	0	0
FLUOX	BUPRO	0	0	0	0	0	0	0	1
FLUOX	CITAL	0	0	0	0	0	0	0	1
FLUOX	VENLA	0	0	0	0	0	0	0	1
FLUOX	REBOX	0	0	0	0	0	0	0	1
SERTR	REBOX	0	0	0	0	0	1	0	1
PAROX	CITAL	0	0	0	0	0	0	0	1
PAROX	FLUVO	0	0	0	0	0	0	0	1
PAROX	MIRTA	0	0	0	0	0	0	0	1
VENLA	REBOX	0	0	0	0	0	1	0	1
FLUOX	MIRTA	0	0	0	0	0	0	0	2
FLUOX	ESCIT	0	0	0	0	0	0	0	2
PAROX	ESCIT	0	1	0	0	0	0	0	2
CITAL	REBOX	0	0	0	0	0	0	0	2
FLUOX	PAROX	0	0	0	0	0	0	0	4

Studie(-n) verwendet

Studie(-n) publiziert, insofern abweichend vom Jahr der Verwendung

III.3 Ergebniszusammenfassung

Der Wert maximal verbleibenden Erkenntnisgewinns sank im Zeitverlauf von 498/10.000 Episoden nach dem Erkenntnisstand des Jahres 1993 auf 72/10.000 Episoden nach dem Erkenntnisstand des Jahres 2009 bezogen auf das korrigierte Therapieansprechen (Abbildung 1), mit ähnlichen Ergebnissen für die anderen Zielgrößen (Therapieansprechen, Therapieabbruch). Erkennbar ist eine deutliche Reduktion im Zeitverlauf, mit jedoch z. T. erheblichen Schwankungen und zeitweiligem Wiederansteigen des Wertes verbleibenden, möglichen Erkenntnisgewinns. Dies kann im Zusammenhang mit der ebenfalls im Zeitverlauf deutlich veränderten, von 7.132/10.000 auf 5.589/10.000 zurückgehenden Größe der als maximal erreichbar anzunehmenden (bzw. in Bezug auf die Therapieabbruchrate minimal) Rate des Therapieansprechens betrachtet werden (Abbildung 2). Zugleich nähern sich die Ereignisraten der Wirkstoffe im Zeitverlauf tendenziell einander an (Abbildung 5). Insgesamt ist im Zeitverlauf festzustellen, dass die maximale Wahrscheinlichkeit bezogen auf einen Wirkstoff, dass es sich bei diesem um den besten handelt, Werte um 0,6 erreichen (wobei in Bezug auf den letzten Beobachtungszeitpunkt, das "Jahr" 2009, nur 0,4 erreicht werden). Für Gruppen von z. B. 2, 4 und 6 Wirkstoffen ergeben sich Wahrscheinlichkeiten von rund 0,8 bzw. 0,95 und 0,99 (Abbildung 3).

Der alternativ konzipierte Erkenntnisverlauf, der die nach 2002 publizierten und unpublizierten Studien nach aufgrund des in diesen untersuchten Wirkstoffvergleichen und dessen geschätztem Erkenntniswert priorisiert in den Verlauf einbezieht, führt zu einem schnelleren Erkenntnisgewinn als der faktische Verlauf (Abbildung 9). Für die Wirkstoffe die 2009 (bezogen auf das korrigierte Therapieansprechen) die ersten 4 Ränge einnehmen ist dabei erkennbar, dass ein alternativer Verlauf resultiert, der im Mittel früher im Zeitverlauf zu Werten näher am Wert für 2009 führt als der faktische Verlauf. Die Ermittlung des alternativen Verlaufs der Erkenntnisgewinnung basiert auf der bevorzugten Berücksichtigung zunächst der Studien, die auf Grund der in ihnen jeweils untersuchten Wirkstoffvergleiche insbesondere einen weiteren Erkenntnisgewinn versprechen. Die Analyse zeigt hier zudem durchgängig, dass für viele als im Bezug auf weiteren Erkenntnisgewinn besonders relevant bewertete Wirkstoffvergleiche keine direkt vergleichenden Studien verfügbar sind. Die faktische Situation zur Verfügbarkeit von Studien erscheint von einer „Idealsituation“, in der Studien vor allem für als relevant zu wertende Vergleiche relativ zeitnah zur Verfügung stehen weit entfernt.

IV. Diskussion

Die Erkenntnisgewinnung zur relativen Wirksamkeit von 12 neueren Antidepressiva bei der Behandlung akuter, depressiver Episoden zwischen 1993 und 2009 hat nicht zu höherem erwarteten Nutzen für Patienten geführt. Während der Informationswert verbleibender Unsicherheit erheblich reduziert wurde, ist zugleich das Ausmaß der Unterschiedlichkeit zwischen den Wirkstoffen in späteren im Vergleich zu früheren Jahren ebenfalls geringer – die Reduktion des Informationswertes basiert wesentlich nicht auf einer höheren Sicherheit bei der Wirkstoffauswahl, sondern auf einer geringeren Bedeutung dieser Wahl im Maß der zu erwartenden Wirksamkeitsunterschiede. Dies kann als wesentliches Ergebnis der Erkenntnisgewinnung im Zeitverlauf betrachtet werden. Auch dieses eher "negative" Ergebnis hätte allerdings im Zeitverlauf schneller erzielt werden können, wenn eine am jeweiligen Erkenntnisstand orientierte Durchführung von Studien in einem hypothetisch-alternativen Verlauf der Erkenntnisgewinnung im Zeitverlauf angenommen wird. Dabei wurde auch deutlich, dass viele Studien, die nach diesem Kriterium insbesondere wichtig gewesen wären, nicht durchgeführt wurden – demgegenüber eine ganze Reihe von Studienergebnissen vorliegen, die kaum zur Erkenntnisgewinnung im Sinne der vorliegenden Analyse beitragen. Es sind demnach noch immer viele Fragen zur vergleichenden Wirksamkeit innerhalb der Gruppe der 12 einbezogenen Wirkstoffe unbeantwortet. Diese können, würden weitere Studien mit relevanten Komparatoren aus der Gruppe der betrachteten Wirkstoffe durchgeführt, sich vielleicht klären lassen. Es muss gegenwärtig als unklar betrachtet werden, ob das Ergebnis unter dem Eindruck einer im Zeitverlauf tendenziell geringeren Differenz zwischen den Wirkstoffen eher zu der Erkenntnis führen würde, dass die Wirkstoffe im Wesentlichen ergebnisgleich sind, oder ob noch (allerdings eher geringe) Differenzen nachgewiesen werden können.

Eine abschließende Entscheidung zur vergleichenden Wirksamkeit der untersuchten Antidepressiva ist daher noch (immer) nicht möglich (Ioannidis 2009), mit Ausnahme der Erkenntnis, dass Reboxetin anderen Wirkstoffen unterlegen ist (Eyding 2010), was auch die Ergebnisse der vorliegenden Analyse deutlich zeigen. Hervorzuheben ist auch, dass es nicht das primäre Ziel der Studie war, die Frage nach der vergleichenden Wirksamkeit abschließend zu beantworten, sondern vielmehr in genereller Perspektive den Weg der Erkenntnisgewinnung exemplarisch in Bezug auf die in der Behandlung von depressiven Episoden verwendeten Arzneimittelwirkstoffe nachzuvollziehen und zu bewerten.

IV.1 Limitationen und Ergebnisse anderer Analysen

IV.1.1 Datengrundlage, Indikations- und Interventionskontext

Aufgrund der Komplexität der Untersuchung, der Zusammenführung vieler Studien und methodischer Ansätze in einer integrierten Analyse ergeben sich eine Reihe von Limitationen der Aussagekraft und Ergebnissicherheit.

Übersichtsarbeit, Studien und Qualität

Die Übersichtsarbeit von Cipriani et al., als Datengrundlage der vorliegenden Analyse, erforderte von deren Autoren eine Reihe methodischer Entscheidungen (Cipriani 2009a). Diese Übersichtsarbeit entspricht hohen, methodischen Qualitätsanforderungen, da eine systematische Suche nach Studien durchgeführt wurde, klare Ein- und Ausschlusskriterien angewandt wurden, eine systematische Studienbewertung und Datenextraktions- und Analyseprozesse geplant und durchgeführt wurde. Diese wurde dokumentiert in dem publizierten Review-Protokoll (Cipriani 2009b). Insbesondere wurde das Risiko der Ergebnisverzerrung (*bias*) in einem dreifach gegliederten Klassifikationsschema (A-, B- oder C-Klassifikation) erfasst. In die Analyse wurden nur Studien mit einer A- oder B-Bewertung eingeschlossen, sodass Studien mit besonders hohem Risiko einer Ergebnisverzerrung für die Analysen nicht berücksichtigt wurden. Allerdings ist auch aus dem publizierten Datensatz keine detaillierte Zuordnung zu dieser Klassifikation auf der Basis von Einzelstudienbewertungen verfügbar. Eine Sensitivitätsanalyse, die z. B. hätte prüfen können, ob eine Beschränkung der vorliegenden Analyse auf Studien mit einer "A"-Bewertung zu anderen Ergebnissen geführt hätte, war somit nicht möglich. Auch wenn Cipriani et al. den Versuch unternommen haben, auch unpublizierte (bzw. nicht im Volltext in Zeitschriften mit *peer review*-Verfahren publizierte) Studien zu berücksichtigen, so gibt es doch ebenfalls Hinweise, dass dies nicht vollständig gelungen sein könnte (Turner 2009).

Zielgrößen

Entsprechend der Cipriani et al.-Arbeit wurden Therapieansprechen (*Response*) und Therapieabbruch (*Dropout*) als grundlegende Zielgrößen verwendet. Während eine Korrelation des Therapieansprechens (definiert als wesentliche Symptomreduktion) zwar mit der für den eigentlichen Therapieerfolg noch bedeutenden Zielgröße des Erreichens einer Remission (wesentliche Symptomfreiheit) als deutlich assoziiert vermutet werden kann, so könnten sich doch die Wirkstoffe hierbei unterscheiden. Die Wahl der Zielgröße zum Therapieansprechen war offenbar darin begründet, dass so die Ergebnisse möglichst vieler,

ansonsten geeigneter Studien einbezogen werden konnten. Insbesondere in Bezug auf die Zielgröße zum Therapieabbruch kann vermutet werden, dass hier in der klinischen Praxis deutlich abweichende (vermutlich höhere) Abbruchraten auftreten. Ob dies in der Praxis zu differentiellen, zwischen den Wirkstoffen unterschiedlichen, Raten führt, kann hier nicht beurteilt werden. Die Wahl einer Zielgröße zum Therapieabbruch, unabhängig von der ggf. erfragten bzw. erfassten Begründung basiert ebenfalls darauf, dass diese Rate in unterschiedlichen Studien am ehesten gleichsinnig erfass- und berichtbar ist. Werden Gründe für die Abbrüche einbezogen, so ist ggf. mit recht unterschiedlichen Klassifikationen und zusätzlichen Unsicherheiten betreffend deren Erfassung zu rechnen. Ein möglicher Nachteil der Wahl dieser Zielgröße kann jedoch sein, dass insbesondere ein mangelnder Wirkeintritt sowohl zu einem Therapieabbruch als auch zum Nichterreichen eines Therapieansprechens führen kann.

Wirkstoffe und Therapieverfahren

Im klinischen Kontext der Behandlung depressiver Episoden werden neben den 12 in die Analyse eingeschlossenen "neueren" Antidepressiva bzw. Wirkstoffen auch eine Reihe älterer Wirkstoffe (insbesondere die sogenannten "Trizyklika") angewandt, sowie weitere neue Wirkstoffe (z. B. Desvenlafaxin). Zudem stehen unterschiedliche Formen der Psychotherapie sowie Kombinationstherapien zur Verfügung. Eine Reihe anderer Aspekte der Depressionsbehandlung wurden nicht berücksichtigt, z. B. bezogen auf das Versorgungssetting, auf Formen integrierter Versorgung, die Rolle der Arzneimitteltherapie in der Aufrechterhaltung der Remission und der Prävention weiterer Episoden (Geddes 2003, Kaymaz 2008) oder Zweit-Linientherapien (Rush 2006, Trivedi 2006). Der Geltungsbereich der Analysen ist daher entsprechend eingeschränkt, im Hinblick auf eine initiale Therapie eine Episode mit einer Arzneimittelmonotherapie, ohne Berücksichtigung besonderer Behandlungssettings und weiterreichender Therapieziele.

Placebo-Studien

Kirsch 2008 hatte argumentiert, Antidepressiva zeigten überhaupt nur für Patienten höherer Symptomschwere eine über Placeboeffekte hinausgehende Wirksamkeit. Dies hat für die vorliegende Analyse zwar keine unmittelbare Bedeutung da Ergebnisse unter Placebo nicht einbezogen wurden. Indirekt kann die Frage allerdings durchaus von Bedeutung sein, da es den Ausschluss von Studien bzw. Patienten mit geringerem initialen Schweregrade der Symptomatik aus der Analyse nahelegte. Dies wäre aber wiederum aus praktischen Gründen schwierig, da davon auszugehen ist, dass in vielen Studien Patienten unterschiedlicher Schweregrad eingeschlossen werden und nach Schweregrade differen-

zierte Analysen nicht immer verfügbar sind. Cipriani et al. (Cipriani 2009a) hatten Placebo-Studien (Studien die lediglich einen aktiven Komparator untersuchten) ausgeschlossen, da solche Studien aufgrund ihrer Durchführung im Zulassungskontext häufig (insbesondere in jüngeren Jahren) besonderer Vorkehrungen bedürften und somit gegenüber aktiv-kontrollierten Studien (zumindest zwei aktive Komparatoren) vermutlich sehr weit von der Versorgungssituation entfernt seien. Zudem wurde auf Ergebnisse verwiesen, die eine säkulare Veränderung des Ansprechens auf Placebo nahelegen. Die Vor- und Nachteile der Berücksichtigung von placebo-kontrollierten Studien wurden direkt bezogen auf die Arbeit von Cipriani et al. nachfolgend diskutiert (Turner 2009, Ioannidis 2009) ohne dass verläßliche Schlussfolgerungen möglich scheinen. Einerseits erlaubte es die verwendete MTC-Methodik durchaus, auch placebo-kontrollierte Studien einzubeziehen ("Placebo" wäre hier formal einfach wie ein weiterer, dann 13. „Wirkstoff“ zu behandeln). Andererseits wird bei der Einbeziehung solcher Studien, die häufig und wesentlich dem Zulassungskontext entstammen, auch das Risiko einer höheren Verzerrung gesehen. Diese Überlegung stützt sich insbesondere auf Erkenntnisse (Ioannidis 2009), die zeigen, dass die Publikation von Zulassungsstudien solche Verzerrungen (Darstellung günstigerer Ergebnisse in den Publikation gegenüber den zur Zulassung eingereichen Studiendossiers) aufweisen. Trotz der Vorbehalte gegenüber der Einbeziehung von placebokontrollierten Studien erscheint eine erweiterte Analyse durchaus lohnenswert, wenn es gelingt die relevanten Studien eindeutig zu identifizieren und deren tatsächliche, möglichst unverzerrte, Ergebnisse zu verwenden. Dabei ist jedoch zu bedenken, dass dies einen sehr großen Aufwand nach sich zöge, da mit einer Vielzahl solcher, potenziell zu berücksichtigender, Studien zu rechnen sein dürfte. Im Übrigen ist es unter diesem Aspekte allerdings auch nicht unwahrscheinlich, dass die Bewertung der "Effizienz" der Erkenntnisgewinnung im Zeitverlauf noch erheblich kritischer ausfallen dürfte, als dies ohnehin auf Grund der Ergebnisse der vorliegenden Analyse bereits der Fall ist.

Relevanz für Therapieentscheidungen

Die Antwort auf die übergeordnete Frage, ob und dann ggf. welches antidepressive Arzneimittel bzw. welcher Wirkstoff zur Behandlung einer gegebenen Episode bei einem gegebenen Patienten eingesetzt werden kann und sollte, kann nicht nur von der generellen, vergleichenden Wirksamkeits- bzw. Nutzenerkenntnis (verstanden im integralen Sinne einer Zusammenschau relevanter positiver und negativer Effekte) abhängen. Es müssen vielmehr auch weitere spezifische, insbesondere patientenseitige, Faktoren wie z. B. bereits vorliegenden Erfahrungen mit dem Ansprechen oder Nicht-Ansprechen auf bestimmte Wirkstoffe oder Kontraindikationen berücksichtigt werden. In einem Brief bezogen auf die Publikation von Cipriani 2009a (Gartlehner 2009) wurde eben dies angemerkt. In ihrer

Reaktion haben Cipriani et al. demgegenüber zu Bedenken gegeben, dass sehr detaillierte diesbezügliche Abwägungen nach ihrer Erfahrung in der klinischen Praxis eher die Ausnahme darstellen und somit eine generelle Empfehlung durchaus sinnvoll sei (Cipriani 2009c). Grundsätzlich kann in solchen Fragen versucht werden, eine Klärung durch entsprechende Stratifizierung von Interventionen, Indikationen und Patienten auf eine beliebig detaillierte Ebene vorzunehmen. Dies wird jedoch häufig, zumindest in einer retrospektiven Analyse, daran scheitern (müssen), dass derart detaillierte Daten nicht verfügbar sind. Es stellt sich daher immer die Frage der Gültigkeit von Studienergebnissen über die gesamte untersuchte Patientengruppe als Stichprobe aus einer Population sowie der Übertragbarkeit der Ergebnisse auf nicht in Studien eingeschlossene Patientengruppen. Einerseits ist damit ein Risiko von Fehlschlüssen verbunden, andererseits jedoch auch das Risiko ggf. relevante Ergebnisse nicht zu berücksichtigen. Eine Frage in bezug auf die letztgenannte Konstellation kann z. B. dann eintreten, wenn eine spezifische Patientengruppen nur in einer sehr kleinen Zahl von Studien mit entsprechend (statistisch) hoher Unzuverlässigkeit untersucht wurde, Patienten dieser Gruppe aber auch Teil der Studienpopulation einer sehr viel größeren Zahl von zuverlässigen Studien war, ohne dass hier detaillierte Ergebnisse für eine solche spezifische Gruppe vorliegen – und sich die Ergebnisse widersprechen oder zumindest graduell wesentlich unterscheiden. Ohne die sich hier ggf. anschließenden weiteren Überlegungen im Detail zu diskutieren, spricht einiges dafür, die Ergebnisse für eine generelle Patientenpopulation, sofern sie hinreichend zuverlässig sind, zur Bewertungsgrundlage zu machen. Zumindest kann dies gelten solange belastbare Ergebnisse für sehr spezifische Gruppen fehlen und wenn für unterschiedliche Patientengruppen ein relativ einheitlicher Wirkmechanismus nicht wesentlich in Frage gestellt werden muss. Für spezifische Gruppen, z. B. solche Patienten, die deutliche Kontraindikationen aufweisen sind dann jeweils definierte und begründete Ausnahmen vorzusehen. Für solche Patienten sind die generellen Ergebnisse dann nicht gültig bzw. irrelevant.

Cipriani et al. hatten aus den Ergebnissen ihrer Analyse den Schluss gezogen, dass zumindest einige Hinweise zur Anwendung bestimmter Wirkstoffe gegeben werden können. Dies betrifft die Empfehlung, dass Escitalopram und Sertralin in der Regel als gute Wahl angesehen werden könnten ("might be the best choice", Cipriani 2009a: 753), während die Ergebnisse von Reboxetin, Fluvoxamin, Paroxetin und Duloxetin als ungünstig bewertet wurden ("making them less favourable options"). Einer solchen Schlussfolgerung wurde jedoch, in einer anderen aktuellen Übersichtsarbeit, widersprochen (Gartlehner 2011, Gartlehner 2012). Solche Schlussfolgerungen seien nicht zuverlässig. Tatsächlich sind die unterschiedlichen Positionen auch im Rahmen der Diskussion der Arbeit von Cipriani et

al. formuliert worden (Gartlehner 2009, Cipriani 2009c). Im Rahmen der vorliegenden Arbeit wurde eine genauere Analyse der für diese sehr unterschiedlichen Ergebnisse verantwortlichen Vorgehensweisen und methodischen Entscheidungen nicht vorgenommen. Dies wurde vor dem Hintergrund des Ziels der vorliegenden Untersuchung auch nicht für notwendig gehalten, da keine definitive Aufklärung über die Frage des oder der besten Wirkstoffe intendiert war. Die Ergebnisse der Untersuchung tragen jedoch zur Beantwortung der Frage mit bei, insofern sie zeigen, dass im Zeitverlauf zunächst (scheinbar) erkennbare größere Unterschiede aufgrund der Einbeziehung später publizierter Ergebnisse eher in Frage zu stellen waren.

IV.1.2 Erkenntnisstände und Zeitverlauf der Erkenntnisgewinnung

Über die im vorangegangen Abschnitt hinaus diskutierten Limitationen und Aspekte der Methodik, insbesondere bezogen auf den klinischen Kontext und methodische Fragen der zugrunde gelegten Übersichtsarbeit von Cipriani et al., hinaus, werden im weitere methodische Fragen diskutiert, die sich insbesondere auf die Vorgehensweise bei der Ermittlung der Erkenntnisstände, der Beschreibung und Bewertung der Entwicklung des Erkenntnisstandes und der Vorgehensweise bei der Charakterisierung eines hypothetisch-alternativen Verlaufs ergeben.

Integrierte Zielgröße: korrigiertes Therapieansprechen

Aufgrund der Fragestellung und Zielsetzung der Untersuchung ergab sich die Notwendigkeit, die Erkenntnisstände zu bestimmten Zeitpunkten (bzw. Jahresperioden) möglichst einheitlich zu charakterisieren. Die Verwendung von mehr als einer Zielgröße, hier von zunächst den beiden Zielgrößen "Therapieansprechen" und "Therapieabbruch", bildet dabei ein grundsätzliches Problem: Falls die Ergebnisse zu einem Zeitpunkt bezogen auf die Zielgrößen zu unterschiedlichen Bewertungen der Relevanz weiterer Forschungen nahelegen – an welcher Zielgröße orientiert soll dann die Bewertung erfolgen?

Zwar ist eine generelle Beschreibung der Reduktion des Wertes weiterer Erkenntnisse für einzelne Zielgrößen getrennt möglich und diese führt auch in der vorliegenden Untersuchung zu recht ähnlichen Ergebnissen (vgl. Abbildung 1). Dies ist aber keineswegs analytisch sichergestellt, sondern als kontigentes Ergebnis anzusehen. Es ist ja durchaus denkbar, dass die auf unterschiedlichen Zielgrößen basierende Beschreibung zu sehr unterschiedlichen Ergebnissen führen kann, die auch eine unterschiedliche Bewertung des zeitlichen Verlaufs der Erkenntnisgewinnung nahelegen könnten. Betrachtet man einzelne Wirkstoffe bzw. Gruppen von Wirkstoffen, so wird deutlich, dass diese sich in Bezug

auf die Zielgrößen sowohl zu einzelnen Zeitpunkten als auch im Zeitverlauf deutlich unterscheiden (vgl. Tabelle 2).

Cipriani et al. hatten zwar nicht den Versuch unternommen, die Ergebnisse der Zielgrößen formal bzw. quantitativ zu integrieren. Sie hatten jedoch im Sinne einer Abwägung, die zumindest um Zustimmung in Form von Plausibilität wirbt, auf der Grundlage einer solch informellen Zusammenschau Empfehlungen formuliert. – Tatsächlich ist, eine wie immer informelle oder implizite Zusammenfassung nicht vermeidbar (bzw. zumindest retrospektiv aufgrund faktisch getroffener Entscheidungen rekonstruierbar), da ja letztlich nicht bestimmte Effekte separat sondern nur in ihrer spezifisch verwirklichten Kombination bestimmter Interventionen gewählt werden können.

Die Verwendung einer ähnlich abwägenden, eher informellen, Vorgehensweise wie Cipriani et al. sie vorgenommen hatten, stößt allerdings in der vorliegenden Analyse an ihre Grenzen. Das Vorgehen musste in der vorliegenden Analysen über die ganze Reihe der Zeitpunkte bzw. Jahre hinweg konstant und reproduzierbar gestaltet werden.. Die Notwendigkeit ein quantitatives Verfahren zur Integration der relevanten Zielgrößen vorab zu bestimmen, war insbesondere im Hinblick auf die Definition von Kriterien gegeben, anhand derer ein hypothetisch-alternativer Forschungsverlauf entwickelt wurde. Hier war es notwendig, ein quantitatives Kriterium, basierend auf einer einheitlichen Zielgröße, anzugeben, welche Therapievergleiche im Verlauf insbesondere als relevant anzusehen waren und daher zeitlich prioritär behandelt wurden (zum methodischen Vorgehen bei der Entwicklung eines hypothetisch-alternativen Verlaufs siehe im Folgenden).

Zu diesem Zweck wurde auf das von Boada 2008 vorgeschlagene Vorgehen zurückgegriffen und eine aus Therapieansprechen und Therapieabbrüchen kombinierte Zielgrößen des "um die Therapieabbrüche korrigierten Therapieansprechens" jeweils pro Studienarm gebildet (naturgemäß konnten in Analysen beruhend auf dieser kombinierten Zielgröße nur solche Studien einbezogen werden, für die für beide Zielgrößen Ergebnisse vorlagen – was aber für die große Mehrzahl der Studien der Fall war). Die Methode der Bildung der kombinierten Zielgröße impliziert eine Unabhängigkeit beider Komponenten-Zielgrößen, was im vorliegenden Fall (s. die Diskussion zu den verwendeten Zielgrößen in Abschnitt IV.1.1) allerdings nicht sicher gewährleistet ist, da ein Teil der Therapieabbrüche durch mangelnde Wirksamkeit verursacht gewesen sein könnte. Die vergleichende Betrachtung der Ergebnisse der kombinierten Zielgröße mit den Einzelzielgrößen zeigen, dass die Ergebnisse der kombinierten Zielgröße (korrigiertes Therapieansprechen) recht ähnlich den Ergebnissen zum (nicht korrigierten) Therapieansprechen sind und von den Ergebnissen zum Therapieabbruch erkennbar abweichen. In Bezug auf die Bewertung des Erkenntnis-

verlauf legt die vergleichende Betrachtung der Ergebnisse zu den Zielgrößen nahe, dass eine Bewertung der Reduktion der Erkenntnisunsicherheit im Zeitverlauf basierend auf der Zielgröße eher zu (noch) ungünstigeren Ergebnissen geführt hätte, da hier insgesamt (anschaulich z. B. in Abbildung 1) eine höhere Streuung der Ergebnisse vorliegt. Einschränkend ist anzumerken, dass dies jedoch nicht formal überprüft wurde.

Methodik indirekter/gemischter Vergleiche

Ob die Methoden indirekter Vergleiche bzw. solcher Verfahren, die direkte und indirekte Vergleiche integrieren („gemischt“, MTC-Methodik wie hier bzw. in der Review von Cipriani et al. angewandt) nützlich sind und zu verlässlichen Ergebnissen führen, wird aktuell intensiv diskutiert (u. a. Wells 2009, Song 2009, Song 2011, Mills 2011). Die bisherigen Ergebnisse führen jedoch nicht zu einer Zurückweisung indirekter oder direkte und indirekte Ergebnisse integrierender (gemischter) Vergleichsmethoden. Eher läßt sich die Wertung generell als "vorsichtig optimistisch" bezeichnen, wobei die Erfahrungen bisher als noch begrenzt betrachtet werden. Der bisherige methodische Forschungsstand gibt auch keinen Hinweis darauf, dass indirekte Vergleichsmethoden, die nur punktuell nicht vorliegende direkte Vergleichsergebnisse durch bestimmte indirekte Vergleiche ersetzen, gegenüber der alle direkten und indirekten Vergleiche einbeziehenden MTC-Methodik zu bevorzugen wären. Ein solches Vorgehen wäre zwar im vorliegenden Fall bezogen auf die Datengrundlage möglich, aber aus praktischen Erwägungen heraus sehr aufwändig, da für alle analysierten Zeitpunkte (Jahre) zunächst alle direkten Vergleichsmöglichkeiten ermittelt und die entsprechend fehlenden indirekten Vergleiche berechnet werden müssten. Zudem wären hier Inkonsistenzen im Zeitverlauf zu erwarten. So würde z. B. sehr wahrscheinlich der Fall eintreten, dass betreffend eines bestimmten Vergleichs in einer früheren Periode noch keine direkt vergleichenden Studienergebnisse vorliegen und also durch indirekte Vergleiche ersetzt werden müssten, während zu einem späteren Zeitpunkt direkt vergleichende Ergebnisse vorlägen, die dann zur Grundlage der Analyse gemacht würden.

Zur Bewertung der Verlässlichkeit indirekter und gemischter Vergleichsmethodiken empfiehlt es sich, zu prüfen, ob sich deren Ergebnisse von den Ergebnissen direkter Vergleiche unterscheiden, sofern direkte Vergleiche jeweils ebenfalls möglich sind. Es wurden solche Vergleiche in Übersichtsarbeiten als Maßstab der Eigung indirekter und gemischter Vergleichsmethoden verwendet (Schöttker 2009, Wells 2009). Eine solche Analyse war auch hier möglich, da für eine Vielzahl von Vergleichen direkte Ergebnisse aus Studien vorliegen (vgl. Abbildung 2 in Cipriani 2009a), und wurde unternommen, indem für alle Zeitpunkte und alle Vergleiche zwischen jeweils zwei Wirkstoffen, für die direkte Ergeb-

nisse vorlagen, die Ergebnisse (OR) aus ledig den direkten Vergleichsstudien mit den Ergebnissen aus den MTC-Analysen (OR) verglichen wurden. Im Ergebnis (vgl. Abschnitt III.1.4) zeigten sich keine Abweichungen im relevanten Ausmaß, sodass die Ergebnisse der MTC-Analysen aus dieser Sicht als verläßlich angesehen werden können. Generell sei jedoch angemerkt, dass die Bewertungsmaßstäbe in diesem methodischen Feld nicht unbedingt als eindeutig angesehen werden können. So kann z. B. eine indirekt vergleichende Analyse auch dann sinnvoll sein, wenn eine Reihe verlässlicher, aussagesicherer Studien in eine indirekte Analyse einbezogen werden können und nur wenige, ggf. weniger aussagesichere Studien zum direkten Vergleich vorliegen. Weichen die Ergebnisse voneinander ab, so ist in solch einer Konstellation keineswegs sichergestellt, dass die indirekten bzw. gemischten Ergebnisse weniger verläßlich sind als die Ergebnisse des direkten Vergleichs. Welche Analyse eine stärkere Verzerrung und damit eine größere Abweichung von der wahren Differenz der verglichenen Vorgehensweisen aufweist wäre in einen solchen Fall vielmehr als unsicher zu betrachten.

Es ist davon auszugehen, dass die Methoden indirekter und gemischter Vergleiche in Zukunft verstärkt genutzt werden dürften, da viele relevante Bewertungsfragen sowohl etablierter als auch neu hinzukommender Behandlungsmethoden heute insbesondere im Kontext bereits bestehender, mehrerer Therapiemöglichkeiten thematisiert werden müssen. Letzteres ist zum Einen unter dem Aspekt anzunehmen, dass sich die Entwicklung immer neuer Therapien bisher relativ ungebrochen fortsetzt. Auch gelingt es eher selten, bestehende Therapien als eindeutig unterlegen zu erweisen sodass diese nicht mehr angewandt werden und vergleichend damit nicht mehr berücksichtigt werden müssten. Zum Anderen ist die zunehmende Forderung nach einer "evidenzbasierten" Gesundheitsversorgung, auch unter Kostengesichtspunkten, zu berücksichtigen, die z. B. auch in dem großen US-amerikanischen *Comparative Effectiveness Research*-Programm (IOM 2009) Ausdruck findet.

Bewertung der Erkenntnisgewinnung im Zeitverlauf

Die Methodik kumulativer Meta-Analysen wurde in der Literatur verschiedentlich beschrieben bzw. angewandt (Lau 1992, Jüni 2004, Trikalinos 2004 u. a.), jedoch nicht in Zusammenhang mit Meta-Analysen auf der Basis indirekter bzw. gemischter Vergleiche oder Informationswertanalysen. Es wurden zwar Vorschläge zur Bewertung der kumulierten Erkenntnisentwicklung (im Zeitverlauf) in Bezug auf Meta-Analysen gemacht (Mueller-leile 2006, Bollen 2006, Mullen 2001, Pogue & Yusuf 1997). Diese erschienen auf die vorliegende, komplexere Analyse nicht anwendbar, insbesondere, da in den zitierten Arbeiten in der Regel der Versuch gemacht wird einen Zeitpunkt bzw. Erkenntnisstand zu

bestimmen, ab dem genügend Erkenntnisse für eine Entscheidung oder zur Durchführung weiterer Studien o. ä. vorliegen. Die Entwicklung und Evaluation einer alternativen, hypothetischen Studienreihenfolge zur Prüfung, ob diese einen schnelleren Erkenntnisgewinn ermöglicht haben würde, ist ebenfalls in der Literatur bisher nicht beschrieben worden. Sie basierte auf der Annahme, dass insbesondere weitere Erkenntnisse für bestimmte paarweise Wirkstoffvergleiche zu einem bestimmen Zeitpunkt von Bedeutung für die weitere Erkenntnisgewinnung generell im Vergleich aller Wirkstoffe miteinander sein könnten. Insbesondere aufgrund der Tatsache, dass dieser Ansatz bisher nicht versucht wurde, sind die hierbei erzielten Ergebnisse zurückhaltend zu werten. Zudem war die Beschränkung auf paarweise Vergleiche insbesondere von ggf. zunächst als kontigent zu wertenden Charakteristika der Studiengrundlagen abhängig: Es lagen nahezu ausschließlich lediglich zwei Wirkstoffe vergleichende, zweiarmige Studien vor. Insbesondere fehlt es an mehrarmigen Studien, die drei oder mehr Wirkstoffe verglichen. Es ist nicht eindeutig, ob der hier gewählte Ansatz lediglich zufällig zu einem (etwas) verbesserten Verlauf der Erkenntnisgewinnung geführt hat. Zukünftig könnte versucht werden, den tatsächlich beobachteten Erkenntnisverlauf basierend auf der faktischen Reihenfolge/Kumulation der Studien mit einer Vielzahl simulierter Verläufe der Erkenntnisgewinnung zu vergleichen. Es könnte dann zumindest bewertet werden, ob es sich um einen eher günstigen oder eher ungünstigen Verlauf handelt. Außerdem wäre es unter Umständen möglich, aus simulierten (sehr) günstigen Verläufen Hinweise auf Kriterien zu gewinnen, die einen solchen Verlauf kennzeichnen und damit auch Anhaltspunkte für die Bewertung des hier vorgeschlagenen Vorgehens geben könnten. Zusammenfassend lässt sich feststellen, dass die Kombination der unterschiedlichen Methoden (kumulative Meta-Analyse, MTC-Meta-Analyse, Informationswertanalyse) nach Erkenntnissen des Autors der vorliegenden Arbeit in ähnlicher Form bisher nicht angewandt wurde. Es erscheint empfehlenswert, weitere Erfahrungen in Bezug auf andere Therapiefelder mit dieser Methode bzw. Methodenkombination zu sammeln, insbesondere vor dem Hintergrund der Ergebnisse, die recht offenkundig Defizite in Bezug auf die Erkenntnisgewinnung zeigen. Es existieren durchaus Hinweise, dass Phänomene wie die hier beobachteten auch aufgrund anderer methodischer Ansätze beobachtet worden sind, insbesondere bezogen auf einen Rückgang des zuvor vermuteten Therapieeffektes (Ioannidis 2008, Pereira & Ioannidis 2011) im weiteren Verlauf von Erkenntnisgewinnungsprozessen. Manche Analysen beziehen sich dabei auch auf das Feld psychischer Erkrankungen, allerdings nicht spezifisch auf die Behandlung von Depressionen (Trikalinos 2004).

IV.2 Fazit und Ausblick

Die Ergebnisse der vorliegenden Analyse zeigen gravierende Defizite im Hinblick auf den Erkenntnisgewinnungsprozess für einen als zentral wichtig erachteten Therapieansatz, nämlich die medikamentöse Therapie, bei einer Erkrankung, nämlich der Depression, deren Relevanz für die gesundheitliche Lage der Bevölkerung und für die vielen Betroffenen offenbar ist. Bei der Erkundung der Ursachen für dieses Defizit kann nicht einfach und ohne weiteres auf fehlende finanzielle Mittel verwiesen werden – schließlich wurde eine Vielzahl von Studien durchgeführt. Da allein für die hier berichteten Studien von über 20,000 Studienteilnehmern und generell sicher von Kosten von einigen Tausend Euro (nach heutigen Preisen, bei hoher Schwankungsbreite) pro Studienteilnehmer ausgegangen werden kann, werden zwei-, eventuell sogar dreistellige Millionenbeträge erreicht.

Die Befunde der vorliegenden Arbeit können, über die spezifisch und exemplarisch untersuchte Indikation und Therapie hinaus, zumindest als Hinweis auf ein generelles Problem verstanden werden. Zwar kann sicherlich nicht von einem Beleg gesprochen werden, dass die Problematik eines hinsichtlich des Erkenntnis-„Ertrages“ fragwürdigen Forschungsgeschehens überall und auf gleiche oder ähnliche Weise besteht. Es existieren andererseits jedoch auch deutliche Hinweise die dafür sprechen. So ist dies zum einen sicherlich auch vor dem Hintergrund des Einflusses generell wohl kaum bestreitbarer Einzelinteressen z. B. von Proponenten einzelner Verfahren (bzw. in diesem Fall, Herstellern von Arzneimitteln, vgl. Ioannidis 2008b, Ioannidis 2009) zu betrachten. Zum anderen sei hierzu noch einmal auf die einleitend bereits genannten Befunde im Hinblick auf die Unsicherheit, dass bei der Planung von Studien und vor deren Durchführung eine hinreichend Klärung des Erkenntnisstandes überhaupt erfolgt verwiesen. Weiter zeigt eine aktuell publizierte Serie von Artikeln eine Vielzahl von Schwachstellen der medizinischen und klinischen Forschung auf (u. a. Chalmers 2014 et al.). Im Übrigen steht mit der in der vorliegenden Arbeit entwickelten Methodik, oder ihren Weiterentwicklungen, auch ein Instrument bereit, um empirisch zu prüfen, ob und in welchem Ausmaß sich die hier erzielten Ergebnisse auch in anderen Therapie- und Erkrankungsgebieten finden lassen.

Die systematische Bewertung von Therapie- und Diagnostikverfahren im Rahmen von Gesundheitstechnologiebewertungen (*Health Technology Assessments*) kann heute vielfach sowohl im Hinblick auf deren politischen Stellenwert als auch in ihrer generellen Methodik als etabliert gelten. Eine wesentliche Schwäche der gesamten Arbeit in diesem Feld stellt jedoch die gegenwärtige, häufig allerdings kaum vermeidbare, Praxis dar, sich eben der Ergebnisse aus vorliegenden Studien bedienen zu müssen – trotz oft wesentlicher Unzulänglichkeiten einer solchen „Evidenzbasis“. Sind vorliegende Studien nur sehr

bedingt geeignet, so sind auch die Wirkungsmöglichkeiten solcher Bewertungen begrenzt, auch wenn viele Versuche unternommen werden (z. B. mit Hilfe von „Modellierungen“) die Defizite der zugrunde liegenden empirischen Primärerkenntnisse zu kompensieren. Zwar sind „systematische Übersichtsarbeit“ (oder „systematischer Review“) und Meta-Analyse, heutzutage als Standardinstrumentarium etabliert und unverzichtbar, aber die Bezugnahme auf den Begriff „Review“ zeigt schon dessen wörtlicher Bedeutung nach: Es handelt sich eben nur um eine rückschauende Zusammenstellung des eben, als kontigent betrachteten, vorhandenen und auffindbaren. Auch bei der vorliegenden Arbeit handelt es sich, im weiteren Sinne, um einen solchen „Review“ bzw. eine erweitere Sekundäranalyse eines solchen. Die vorliegende Arbeit dient allerdings nicht vor allem der Bewertung eines bestimmten, aktuellen Erkenntnisstandes, sondern der exemplarischen Bewertung eines Prozesses und dessen Leistungsfähigkeit im Hinblick auf die spezifische Fragestellung nach dem Erkenntnisverlauf.

Im Unterschied dazu kommt gerade dem Konzept einer prospektiven Herangehensweise im Erkenntnisgewinnungsprozess größte Bedeutung zu. Dem entspricht genau eben die prospektive Studie. Während andere methodische Vorkehrungen in solchen Studien primär auf Verzerrungsfreiheit und Erkenntnissicherheit abstellen, wie z. B. insbesondere Randomisierung und Verblindung, so trägt die prospektive Durchführung darüber hinaus einem zusätzlichen Aspekt Rechnung: Es ist damit nämlich die verbindliche Erwartung verbunden, gerade mit der jeweilig prospektiv geplanten Studie einer bestimmen Unsicherheit abzuhelfen. Insofern ist mit der Durchführung einer solchen Studie dann auch die Verantwortung dafür zu übernehmen, dass die Ergebnisse auch zur Klärung der Frage führen – und nicht zu einer lediglich losen Folge gegebenenfalls mehrerer Studien, ohne ersichtlichen Abschluss im Hinblick auf die zu entscheidende Frage und die hierzu hinreichende Reduktion der Unsicherheit. Hierbei muss es sich nicht notwendig lediglich um eine einzige Studie handeln, da ja auch z. B. zur Zulassung eines Arzneimittels in der Regel mehrere Studien notwendig sind. Entscheidend ist jedoch die konsistente und vorausschauende Planung des gesamten Prozesses. Vorgehensweisen in Bezug auf die Zulassung bzw. den Marktzugang sind hier von besonderem, tendenziell vorbildlichem, Charakter, da ein aktiver Einfluss der Regulierer konstitutiv ist. Dieser ist letztlich bezogen auf die Anwender und Patienten in deren Interesse gehandelt werden soll und impliziert auch einen stärkeren Einfluss auf die Erkenntnisgewinnung, im Sinne der Formulierung von Anforderungen und Nachweisen. Besonders deutlich wird dies auf dem Feld nichtmedikamentöser Verfahren bzw. Medizinprodukten, da hier der Marktzugang im Hinblick auf solche Ergebnisse sehr defizitär ist, sodass eine wesentliche Erkenntnisgewinnung heute, wenn überhaupt, erst später stattfindet (Storz-Pfennig 2013).

Generell ist zu fordern, zumindest dringend zu wünschen, dass zukünftige Prozesse der Erkenntnisgewinnung gemeinsam zwischen Herstellern, Regulierern, Wissenschaft, Anwendern, Betroffenen und Zahlungsträgern erfolgen – bzw. die Anwender, Betroffenen und Finanzierungsträger größeren Einfluss gewinnen. Dies ist eine wesentlich politische Frage und nicht zuletzt ist der Aspekt, wer hier welche Kosten der Erkenntnisgewinnung tragen sollte zu klären. Viele Tendenzen deuten, national wie international (Stafinski 2010, Storz-Pfennig 2013, G-BA 2014) darauf hin, dass sich eine solche Perspektive einer aktiveren und breiteren Beteiligung zunehmend geltend machen wird. Schließlich kann die angemessene Anwendung wissenschaftlicher Erkenntnisse in Situationen mit divergierenden Interessen und Bewertungen besonders dann gelingen wenn man sich, bei noch bestehender substanzieller Unsicherheit, darauf einigt mit welchem Verfahren diese zu beseitigen ist. Dies ist hier letztlich der Sinn einer prospektiven Studie im Kontext möglicher und erstrebter Bedeutung von deren Ergebnissen für das Handeln: Dass man sich *vor* Kenntnis der Ergebnisse darauf einigen kann, dann auch tatsächlich auf deren Grundlage zu entscheiden, wenn sie dann vorliegen. Nur so kann die Bewertung von Gesundheitstechnologien ihren vollen Zweck auch wirklich erfüllen.

V. Zusammenfassung

Heute stehen häufig eine Reihe unterschiedlicher, therapeutischer Verfahren zur Verfügung, die bei der Behandlung spezifischer Erkrankungen eingesetzt werden können bzw. eingesetzt werden. Ihre vergleichende Bewertung im Bezug auf Wirksamkeit, Nutzen und Schaden stellt eine wesentliche Aufgabe dar, die in der Regel auf der Grundlage möglichst erkenntnissicherer Studien (wenn möglich randomisierter Interventionsstudien) erfolgen soll. Häufig wird hierbei, zu Recht, beklagt, dass solche Studien in vielen Gebieten fehlen. Andererseits wird jedoch eine Vielzahl von Studien zumindest in einigen Gebieten durchgeführt, was zu der Frage führt ob die Studien, die tatsächlich durchgeführt werden, auch insbesondere geeignet sind, zu einer vergleichenden Erkenntnis bei verschiedenen, grundsätzlich in Frage kommenden Therapieverfahren zu führen und ob hier nicht ein besserer Einsatz der Forschungsaufwendung möglich wäre?

Diesen Fragen wird exemplarisch am Beispiel antidepressiver Arzneimittelwirkstoffe zur Behandlung depressiver Episoden nachgegangen. Die Analyse basiert auf den in eine systematische Übersichtsarbeit eingeschlossenen 117 randomisierten Studien zu 12 Wirkstoffen, die zwischen 1991 und 2007 publiziert wurden (sowie Ergebnissen unpublizierter Studien). Die zum Zwecke der vorliegenden Analyse entwickelte Methodik kombiniert meta-analytische Verfahren zu direkten und indirekten bzw. gemischten Therapievergleichen (MTC, *mixed treatment comparisons*) und zur Informationswertermittlung (VOI, *value of information*) in einem kumulativen Ansatz. Damit wird der Erkenntnisgewinnungsprozess nachgezeichnet und mit Hilfe der Gegenüberstellung zu einem alternativ konstruierten Erkenntnisgewinnungsprozess vergleichend bewertet. Mit Hilfe von VOI-Analysen die auf den MTC-Meta-Analysen beruhen, wird der Erkenntnisstand für jeden Analysezeitpunkt (Jahr) als Produkt des Ausmaßes des Wirkungsunterschiedes zwischen den Wirkstoffen und des Ausmaßes der Sicherheit, mit der diese Unterschiede tatsächlich erwartet werden können, als erwarteter Nutzen für die Patienten abgeschätzt. Das Ausmaß, in dem die Erkenntnisunsicherheit zu sukzessiven Zeitpunkten sinkt und die Nutzenerwartung sich erhöht, dient als Maßstab zur Bewertung des Erfolges der im Zeitverlauf gewonnenen Erkenntnisse aus Studien. Da hiermit zunächst ein wesentlich beschreibender Zugang eröffnet ist – eine mehr oder minder starke Reduktion der Erkenntnisunsicherheit bzw. Erhöhung der Nutzenerwartung kann verschieden interpretiert werden – wird zudem ein alternativer Verlauf der Erkenntnisgewinnung auf der gleichen Studiengrundlage konzipiert und mit dem faktisch-historischen Verlauf verglichen, um die Frage zu beantworten, ob ein besserer Einsatz der Forschungsressourcen zu einer schnelleren Reduktion der Erkenntnisunsicherheit geführt haben würde.

Die Ergebnisse zu der primären Zielgröße der Analyse (erfolgreich behandelbare Episoden, als Kombination aus Therapieansprechen in der Regel nach 8 Wochen, korrigiert um Therapieabbrüche) zeigen, dass der Verlauf der Erkenntnisgewinnung zu bescheideneren Erwartungen an den Nutzen der untersuchten Therapien Anlass gibt. Zwar hat sich der Informationswert verbleibender Unsicherheit im Zeitverlauf erheblich (im Sinne der Anzahl von 10,000 hypothetisch behandelten Episoden, die von Unsicherheit betroffen sind) von 498/10,000 auf 72/10,000 Episoden verringert. Es ist jedoch zugleich auch das Ausmaß der Unterschiedlichkeit zwischen den Wirkstoffen in späteren im Vergleich zu früheren Jahren ebenfalls geringer – die Reduktion des Informationswertes basiert wesentlich nicht auf einer höheren Sicherheit bei der Wirkstoffauswahl, sondern auf einer geringeren Bedeutung dieser Wahl im Maß der zu erwartenden Wirksamkeitsunterschiede. Dies kann als wesentliches Ergebnis der Erkenntnisgewinnung im Zeitverlauf betrachtet werden. Auch dieses eher "negative" Ergebnis hätte allerdings im Zeitverlauf schneller erzielt werden können, wenn eine am jeweiligen Erkenntnisstand orientierte Durchführung von Studien in einem hypothetisch-alternativen Verlauf der Erkenntnisgewinnung im Zeitverlauf angenommen wird.

Zweifellos ist die Verläßlichkeit der Ergebnisse, an einer exemplarischen Indikation und einer Auswahl von Therapiemöglichkeiten in Bezug auf diese (z. B. wurde die ebenfalls hier bedeutende Psychotherapie nicht berücksichtigt) erzielt, abhängig von einer Reihe von methodischen Unsicherheitsfaktoren. Diese sind bezogen auf die Qualität der zugrundegelegten Studien, der Übersichtsarbeit der diese entnommen wurden sowie Faktoren basierend auf den Elementen der kombinierten methodischen Elementen und deren Kombination zu berücksichtigen. Unter diesen Vorbehalten kann zunächst dennoch gefolgert werden, dass die hier entwickelte, kombinierte Methode auch auf andere Interventionen und Therapiegebieten, auch über Arzneimittel hinaus, angewandt bzw. weiterentwickelt und erprobt werden sollte. Das Potenzial für methodische Weiterentwicklungen bezieht sich insbesondere auf die Frage, an welchem Maßstab die Bewertung eines faktisch erfolgten Erkenntnisverlaufs sich, im Spiegel anderer möglicher Verläufe, orientieren sollte.

Die Ergebnisse der Analyse bekräftigen zumindest die Vermutung, die sich auch in der Literatur aufgrund anderer Ansätze findet, dass das tatsächliche vorgefundene Studiengeschehen alles andere als ideal an den Erkenntnisbedürfnissen der vergleichenden Bewertung orientiert zu sein scheint. In dem Maße, indem solche Befunde als verläßlich bewertet werden können, muß sehr viel ernsthafter als bisher der Versurch unternommen werden, im übergreifenden Interesse aktiven Einfluss auf die Erkenntnisgewinnung im Zeitverlauf auszüben. Schließlich kann die angemessene Anwendung wissenschaftlicher

Erkenntnisse in Situationen mit divergierenden Interessen und Bewertungen besonders dann gelingen wenn man sich, bei noch bestehender substanzieller Unsicherheit, darauf einigt mit welchem Verfahren diese zu beseitigen ist. Wenn dann entsprechend gehandelt bzw. geforscht wird, erhöht sich die reale Chance, dass diesen Ergebnissen entsprechend auch entschieden wird und die Bewertung von therapeutischen und diagnostischen Interventionen schließlich ihren Zweck erfüllt.

VI. Literaturverzeichnis

Ades 2007. Ades AE, Welton N, Lu G. Introduction to Mixed Treatment Comparisons. 2007 www.bris.ac.uk.

Boada 2008. Boada JN, Boada C, García-Sáiz M, García M, Fernández E, Gómez E. Net Efficacy Adjusted for Risk (NEAR): a Simple Procedure for Measuring Risk:benefit Balance. PloS One 3, 10 (2008): e3580

Bollen 2006. Bollen CW, Uiterwaal CS, van Vught AJ, van der Tweel I. Sequential meta-analysis of past clinical trials to determine the use of a new trial. Epidemiology. 2006;17(6):644-649

Briggs 2006. Briggs A, Sculpher M, Claxton K "Decision Making, Uncertainty and the Value of Information." In *Decision Modelling for Health Economic Evaluation*, 165-200. Oxford University Press, 2006

Chalmers 2014. Chalmers I, Bracken MB, Djulbegovic B, Garattini S, Grant J, Gülmezoglu AM, Howells DW, Ioannidis JPA, Sandy O. Research: increasing value, reducing waste 1. How to increase value and reduce waste when research priorities are set. Lancet. 2014;383:156-165

Cipriani 2009a. Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JP, Churchill R, Watanabe N, Nakagawa A, Omori IM, McGuire H, Tansella M, Barbui C. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. Lancet. 2009;373(9665):746-758

Cipriani 2009b. Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JP, Churchill R, Watanabe N, Nakagawa A, Omori IM, McGuire H, Tansella M, Barbui C. "Manga Study. Comparative Efficacy and Acceptability of 12 New Generation Antidepressants: A multiple Treatment Meta-Analysis" (Protocol) http://www.psychiatry.univr.it/docs/Research%20Activities/MTM_Protocol.pdf.

Cipriani 2009c. Cipriani A, Geddes JR, Furukawa TA, Salanti G, Barbui C. Ranking Antidepressants – Authors' Reply. Lancet. 373(9677):1761-1762

Clarke & Chalmers 1998. Clarke M, Chalmers I. Discussion sections in reports of controlled trials published in general medical journals: islands in search of continents? JAMA. 1998;280(3):280-282.

Clarke 2002. Clarke, Mike, Phil Alderson, and Iain Chalmers. "Discussion Sections in Reports of Controlled Trials Published in General Medical Journals." *JAMA: The Journal of the American Medical Association* 287, no. 21 (June 5, 2002): 2799–2801.

Clarke 2007. Clarke M, Hopewell S, Chalmers I. Reports of clinical trials should begin and end with up-to-date systematic reviews of other relevant evidence: a status report. J R Soc Med. 2007 Apr;100(4):187-190

Claxton 2004. Claxton K, Ginnelly L, Sculpher M, Philips Z, Palmer S. A pilot study on the use of decision theory and value of information analysis as part of the NHS Health Technology Assessment programme. Health Technol Assess. 2004;8(31):1-103, iii

Cochrane 1972. Cochrane, A. L. Effectiveness and Efficiency: random reflections on health services. Nuffield Provincial Trust (London). 1972.

Higgins & Green 2011. Higgins JPT, Green S (eds.) Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011] http://www.cochrane-handbook.org/

Deeks 2008. Deeks JJ, Altman DG, Bradburn MJ (2008) "Statistical Methods for Examining Heterogeneity and Combining Results from Several Studies in Meta-Analysis", in: Systematic Reviews in Health Care: Meta-Analysis in Context, Second Edition (eds M. Egger, G. D. Smith and D. G. Altman), BMJ Publishing Group, London, UK Anhänge

Egger 2001. Egger M, Smith GD, Sterne JA. Uses and abuses of meta-analysis. Clin Med. 2001;1(6):478-484

Evans 2011. Evans I, Thornton H, Chalmers I, Glasziou P. Testing Treatments. Better Research for better Healthcare. 2nd edition. Pinter & Martin Ltd. London 2011 http://www.testingtreatments.org/wp-content/uploads/2012/09/TT_2ndEd_English_17oct2011.pdf

Eyding 2010. Eyding D, Lelgemann M, Grouven U, Härter M, Kromp M, Kaiser T, Kerekes MF, Gerken M, Wieseler B. Reboxetine for acute treatment of major depression: systematic review and meta-analysis of published and unpublished placebo and selective serotonin reuptake inhibitor controlled trials. BMJ. 2010 Oct 12;341:c4737

Gartlehner 2009. Gartlehner G, Gaynes BN, Hansen RA, Lohr KN. Ranking antidepressants. Lancet. 2009;373(9677):1761; author reply 1761-2

Gartlehner 2011. Gartlehner G, Hansen RA, Morgan LC, Thaler K, Lux L, Van Noord M, Mager U,Thieda P, Gaynes BN, Wilkins T, Strobelberger M, Lloyd S, Reichenpfader U, Lohr KN. Comparative benefits and harms of second-generation antidepressants fortreating major depressive disorder: an updated meta-analysis. Ann Intern Med. 2011;155(11):772-85

Gartlehner 2012. Gartlehner G, Hansen RA, Morgan LC, Thaler K, Lux L, Van Noord M, Mager U,Thieda P, Gaynes BN, Wilkins T, Strobelberger M, Lloyd S, Reichenpfader U, Lohr KN. „Second-Generation Antidepressants in the Pharmacologic Treatment of Adult Depression: An Update of the 2007 Comparative Effectiveness Review". Rockville (MD): Agency for Healthcare Research and Quality (US); 2011 Dec. Report No.: 12-EHC012-EF., 2012. http://www.ncbi.nlm.nih.gov/books/NBK83442/

G-BA 2014. Erprobung von Untersuchungs- und Behandlungsmethoden nach § 137e SGB V. Gemeinsamer Bundesausschuss http://www.g-ba.de/institution/themenschwerpunkte/erprobungsregelung/)

Geddes 2003. Geddes JR, Carney SM, Davies C, Furukawa TA, Kupfer DJ, Frank E, Goodwin GM. Relapse prevention with antidepressant drug treatment in depressive disorders: a systematic review. Lancet. 2003 Feb 22;361(9358):653-661

Glaeske 2010. Glaeske G, Höffken K, Ludwig WD, Schrappe M, Weißbach L, Wille E. Sicherstellung einer effizienten Arzneimittelversorgung in der Onkologie. Bremen, August 2010. Gutachten im Auftrag des Bundesministeriums für Gesundheit http://www.bmg.bund.de/fileadmin/redaktion/pdf_allgemein/Gutachten_Sicherstellung_einer_effizienten_Arzneimittelversorgung_in_der_Onkologie.pdf)

Goeree & Levin 2006. Goeree R, Levin L. Building bridges between academic research and policy formulation: the PRUFE framework - an integral part of Ontario's evidence-based HTPA process. Pharmacoeconomics. 2006;24(11):1143-56

Ioannidis 2008. Ioannidis JP. Why most discovered true associations are inflated. Epidemiology. 2008;19(5):640-8

Ioannidis 2008b. Ioannidis JP. Effectiveness of antidepressants: an evidence myth constructed from a thousand randomized trials? Philos Ethics Humanit Med. 2008 May 27;3:14

Ioannidis 2009. Ioannidis JP. Ranking antidepressants. Lancet. 2009 May 23;373(9677):1759-60; author reply 1761-2

IOM 2009. Institute of Medicine (IOM). Intitial National Priorities for Comparative Effectiveness Research. Washington DC: The National Academies Press http://www.nap.edu/catalog.php?record_id=12648

Jüni 2004. Jüni P, Nartey L, Reichenbach S, Sterchi R, Dieppe P A, Egger M. Risk of cardiovascular events and rofecoxib: cumulative meta-analysis. Lancet 2004; 364: 2021-2029

Kaymaz 2008. Kaymaz N, van Os J, Loonen AJ, Nolen WA. Evidence that patients with single versus recurrent depressive episodes are differentially sensitive to treatment discontinuation: a meta-analysis of placebo-controlled randomized trials. J Clin Psychiatry. 2008;69(9):1423-36

Kirsch 2008. Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. PLoS Med. 2008 Feb;5(2):e45

Lau 1992. Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. N Engl J Med. 1992;327(4):248-254

Lu & Ades 2004. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med. 2004 Oct 30;23(20):3105-3124

Meltzer 2011. Meltzer DO, Hoomans T, Chung JW, Anirban B. "Minimal Modeling Approaches to Value of Information Analysis for Health Research". Methods Future Research Needs Report. Agency for Healthcare Research and Quality, 2011 http://www.effectivehealthcare.ahrq.gov/ehc/products/197/719/MethodsFRN6_06-28-2011.pdf.

Mills 2011. Mills EJ, Bansback N, Ghement I, Thorlund K, Kelly S, Puhan MA, Wright J.Multiple treatment comparison meta-analyses: a step forward into complexity. Clin Epidemiol. 2011;3:193-202

Muellerleile 2006. Muellerleile P, Mullen B. Sufficiency and stability of evidence for public health interventions using cumulative meta-analysis. Am J Public Health. 2006;96(3):515-522

Mullen 2001. Mullen B, Muellerleihe P, Bryant B. Cumulative Meta-Analysis: A Consideration of Indicators of Sufficiency and Stability. Pers Soc Psychol Bull 2001; 27; 1450-1462

Pereira & Ioannidis 2011. Pereira TV, Ioannidis JP. Statistically significant meta-analyses of clinical trials have modest credibility and inflated effects. J Clin Epidemiol. 2011;64(10):1060-1069

Pogue & Yusuf 1997. Pogue JM, Yusuf S. Cumulating evidence from randomized trials: utilizing sequential monitoring boundaries for cumulative meta-analysis. Control Clin Trials. 1997;18(6):580-593

Robinson & Goodman 2011. Robinson KA, Goodman SN. A systematic examination of the citation of prior research in reports of randomized, controlled trials. Ann Intern Med. 2011;154(1):50-55

Rush 2006. Rush AJ, Trivedi MH, Wisniewski SR, Stewart JW, Nierenberg AA, Thase ME, Ritz L, Biggs MM, Warden D, Luther JF, Shores-Wilson K, Niederehe G, Fava M; STAR*D Study Team. Bupropion-SR, sertraline, or venlafaxine-XR after failure of SSRIs for depression. N Engl J Med. 2006;354(12):1231-1242

Sackett 1996. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ. 1996 Jan 13;312(7023):71-72

Schöttker 2009. Schöttker B, Lühmann D, Boulkhemair D, Raspe H. Indirekte Vergleiche von Therapieverfahren. Schriftenreihe Health Technology Assessment, Bd. 88, 2009, DAHTA@DIMDI.

Scott 2008. Scott NA, Moga C, Harstall C, Magnan J. Using health technology assessment to identify research gaps: an unexploited resource for increasing the value of clinical research. Healthc Policy. 2008 Feb;3(3):e109-127

Song 2009. Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. BMJ. 2009 Apr 3;338:b1147

Song 2011. Song F, Xiong T, Parekh-Bhurke S, Loke YK, Sutton AJ, Eastwood AJ, Holland R, Chen YF, Glenny AM, Deeks JJ, Altman DG. Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study. BMJ. 2011 Aug 16;343:d4909

Stafinski 2010. Stafinski T, McCabe CJ, Menon D. Funding the unfundable: mechanisms for managing uncertainty in decisions on the introduction of new and innovative technologies into healthcare systems. Pharmacoeconomics. 2010;28(2):113-142

Storz-Pfennig 2013. Storz-Pfennig P, Schmedders M, Dettloff M. Trials are needed before new devices are used in routine practice in Europe. BMJ. 2013 Mar 18;346:f1646.

Trikalinos 2004. Trikalinos TA, Churchill R, Ferri M, Leucht S, Tuunainen A, Wahlbeck K, Ioannidis JP; EU-PSI project. Effect sizes in cumulative meta-analyses of mental health randomized trials evolved over time. J Clin Epidemiol. 2004;57(11):1124-1130

Trivedi 2006. Trivedi MH, Fava M, Wisniewski SR, Thase ME, Quitkin F, Warden D, Ritz L, Nierenberg AA, Lebowitz BD, Biggs MM, Luther JF, Shores-Wilson K, Rush AJ; STAR*D Study Team. Medication augmentation after the failure of SSRIs for depression. N Engl J Med. 2006;354(12):1243-1252

Turner 2009. Turner E, Moreno SG, Sutton AJ. Ranking antidepressants. Lancet. 2009;373(9677):1760; author reply 1761-2

Weisbrod 1991. Weisbrod BA. The Health Care Quadrilemma: An Essay on Technological Change, Insurance, Quality of Care, and Cost Containment. Journal of Economic Literature. 1991; 29(2):523-552

Wells 2009. Wells GA, Sultan SA, Chen L, Khan M, Coyle D. “Indirect Evidence: Indirect Treatment Comparisons in Meta-analysis” Ottawa: Canadian Agency for Drugs and Technologies in Health, 2009. http://www.cadth.ca/media/pdf/H0462%5Fitc%5Ftr%5Fe.pdf

VII. Anhänge

VII.1 Programm zur MTC-Analyse

Im folgenden wird das WinBUGS-Programm für die MTC-Analyse der Zielgröße Therapieerfolg (Studien 1991/1993 = erster Analysezeitpunkt) exemplarisch wiedergeben. Die übrigen Programme (jeweils pro Zielgröße, Analysezeitpunkt bzw. alternativen Erkenntnisverlauf) unterscheiden sich nur durch eine abweichende Anzahl einbezogener Studien, Therapien/Wirkstoffen und der entsprechenden Daten.

```
model
{
#all trial arms (N=14, N=7 two-arm trials)
for (i in 1:14) {
       logit(p[i])<-mu[s[i]]+ delta[i]  * (1-equals(t[i],b[i]))
       r[i]~dbin(p[i],n[i])
       delta[i] ~ dnorm(md[i],tau)
       md[i] <- d[t[i]] - d[b[i]] }
#all trials (N=7)
for(j in 1:7)  { mu[j]~dnorm(0,.0001) }
#all treatments (N=6)
d[1]<-0
for (k in 2:6)  {d[k] ~ dnorm(0,.0001) }
sd~dunif(0,2)
tau<-1/pow(sd,2)

#baseline
mA ~ dnorm(0,1)

#absolute efficacy
for (k in 1:6)
{ logit(T[k])<- mA  +d[k] }

#probability pRank (m=rank,k=treatment)  with max k=max m= NT
for (k in 1:6) { rk[k]<- 6+1 - rank(T[],k)}
for (k in 1:6) {
  for (m in 1:6) {
    pRank[m,k]<-equals(rk[k],m)
  }
}
#pairwise lnOR and OR
for (c in 1:(6-1))
{ for (k in (c+1):6)
{ lor[c,k] <- d[k] - d[c]
log(or[c,k]) <- lor[c,k]
}}}

#02 data (s= studyNo, t=treatment, r=events, n=patients, b=baseline therapy/trial
list(
s= c(1,1,2,2,3,3,4,4,5,5,6,6,7,7),
t= c(1,2,1,3,4,6,1,4,1,4,1,4,5,6),
r= c(35,37,26,35,24,24,26,25,27,30,54,53,33,31),
n= c(62,61,56,52,56,64,41,37,45,45,87,89,108,109),
b= c(1,1,1,1,4,4,1,1,1,1,1,1,5,5)
)

#03 intial (vague/informative priors)
list(sd=1, mA=0,
d=c(NA,0,0,0,0,0),
mu=c(0,0,0,0,0,0,0),
delta=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0)
```

VII.2 Anzahl der Studien, Patienten und Wirkstoffe im Zeitverlauf

Tabelle 5 Anzahl einbezogener Studien, Patienten und Wirkstoffe (betreffend der Zielgröße Therapieansprechen)

Jahr	Studien (Jahr) [n]***	Patienten (Jahr) [n]***	Studien kumuliert [n]***	Patienten kumuliert [n]**	Wirkstoffe [n]	Wirkstoffe zusätzlich
1993*	7	912	7	912	6	
1994	3	380	10	1.292	8	VENLA, MILNA
1995	4	648	14	1.940	8	
1996	2	360	16	2.300	8	
1997	9	2.276	25	4.576	10	MIRTA,REBOX
1998	7	1.592	32	6.168	10	
1999	9	1.996	41	8.164	10	
2000	14	2.328	55	10.492	11	ESCIT
2001	4	823	59	11.315	11	
2002	10	2.029	69	13.344	12	DULOX
2003	6	608	75	13.952	12	
2004	6	1.285	81	15.237	12	
2005	9	2.119	90	17.356	12	
2006	9	2.320	99	19.676	12	
2007	5	1.541	104	21.217	12	
2009**	11	3.766	**115**	**24.983**	12	

Tabelle 6 Anzahl einbezogener Studien, Patienten und Wirkstoffe (betreffend der Zielgröße Therapieabbruch)

Jahr	Studien (Jahr) [n]***	Patienten (Jahr) [n]***	Studien kumuliert [n]***	Patienten kumuliert [n]***	Wirkstoffe [n]	Wirkstoffe zusätzlich
1993*	7	912	7	912	6	
1994	3	380	10	1.292	8	VENLA, MILNA
1995	4	648	14	1.940	8	
1996	2	360	16	2.300	8	
1997	8	2.028	24	4.328	10	MIRTA,REBOX
1998	7	1.592	31	5.920	10	
1999	8	1.758	39	7.678	10	
2000	15	2.543	54	10.221	11	ESCIT
2001	4	823	58	11.044	11	
2002	10	2.033	68	13.077	12	DULOX
2003	6	608	74	13.685	12	
2004	5	1.134	79	14.819	12	
2005	8	1.815	87	16.634	12	
2006	10	2.524	97	19.158	12	
2007	4	1.246	101	20.404	12	
2009**	15	4.681	**116**	**25.085**	12	

Tabelle 7 Anzahl einbezogener Studien, Patienten und Wirkstoffe (betreffend der Zielgröße korrigiertes Therapieansprechen)

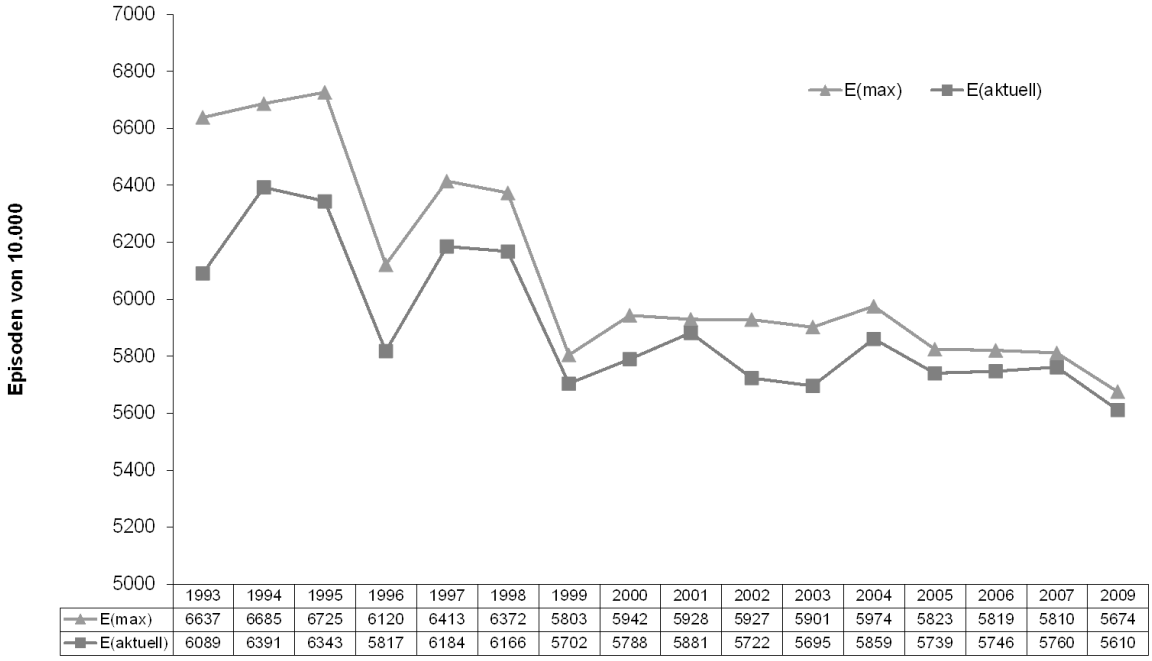
Jahr	Studien (Jahr) [n]***	Patienten (Jahr) [n]***	Studien kumuliert [n]***	Patienten kumuliert [n]***	Wirkstoffe [n]	Wirkstoffe zusätzlich
1993*	7	912	7	912	6	
1994	3	380	10	1.292	8	VENLA, MILNA
1995	4	648	14	1.940	8	
1996	2	360	16	2.300	8	
1997	8	2.028	24	4.328	10	MIRTA,REBOX
1998	7	1.592	31	5.920	10	
1999	8	1.758	39	7.678	10	
2000	14	2.328	53	10.006	11	ESCIT
2001	4	823	57	10.829	11	
2002	10	2.033	67	12.862	12	DULOX
2003	6	608	73	13.470	12	
2004	5	1.134	78	14.604	12	
2005	8	1.820	86	16.424	12	
2006	9	2.320	95	18.744	12	
2007	4	1.246	99	19.990	12	
2009**	11	3.766	110	23.756	12	

*Die Jahre 1991 und 1993 (1992 wurden keine Studien publiziert) wurden zusammengefaßt analysiert, da im Jahr 1991 nur eine einzige Studie publiziert worden war.

**Alle unpublizierten Studien wurden unter dem Jahr 2009 (Publikationsjahr von Cipriani2009a) erfasst.

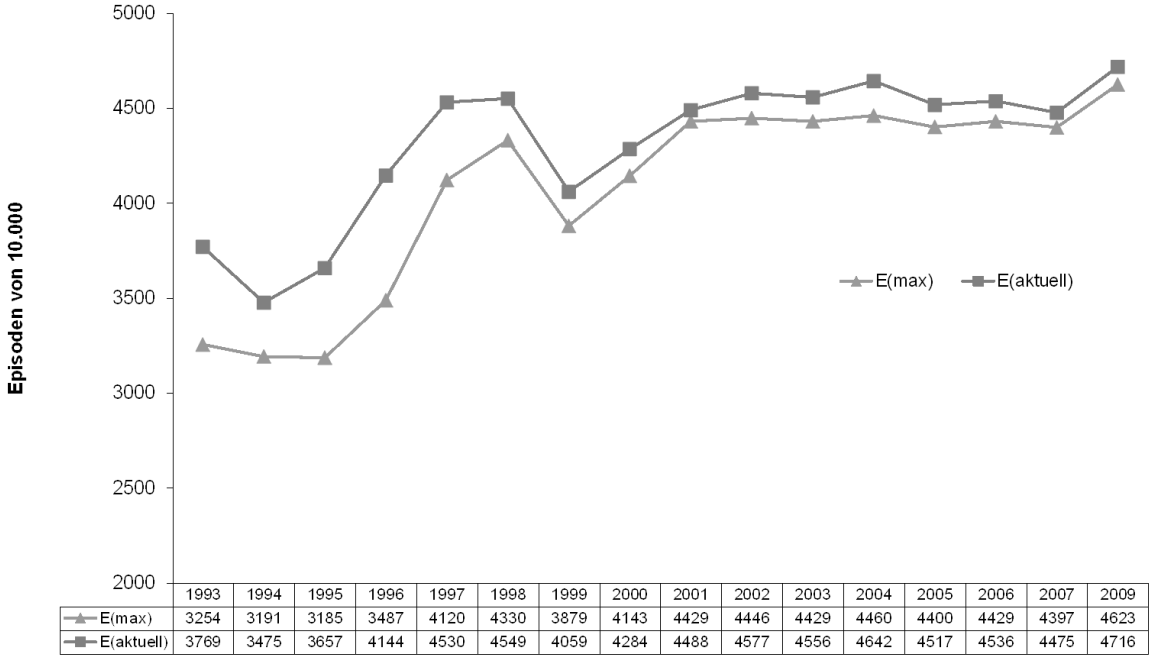
*** Die Anzahl der Studien (und entsprechend die Zahl der Patienten), für die Ergebnisse zur Zielgröße "Therapieansprechen" verfügbar waren, zählt die jeweils zusätzlichen "Studien", die aus der Zerlegung der beiden dreiarmigen Studien resultieren (vgl. Abschnitt II.1.4) mit.

VII.3 Weitere Ergebnisse zur Erkenntnisgewinnung im Zeitverlauf



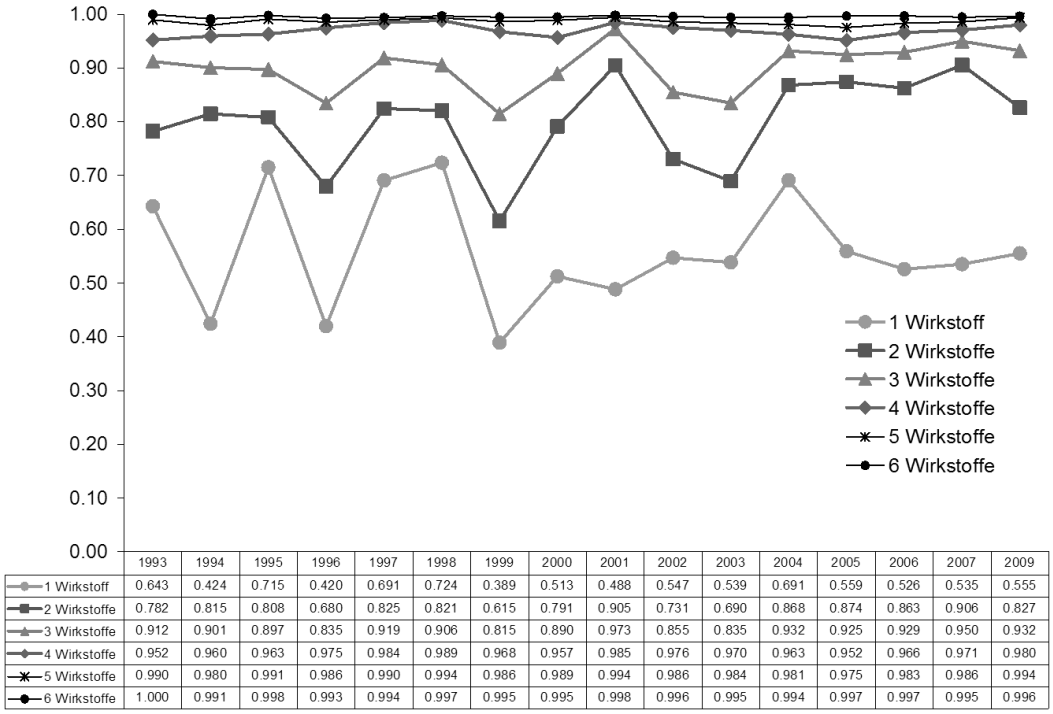
	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2009
E(max)	6637	6685	6725	6120	6413	6372	5803	5942	5928	5927	5901	5974	5823	5819	5810	5674
E(aktuell)	6089	6391	6343	5817	6184	6166	5702	5788	5881	5722	5695	5859	5739	5746	5760	5610

Abbildung 13 Erwartete Erfolgsraten (Therapieansprechen, E_{max}, $E_{aktuell}$) pro 10.000 Episoden



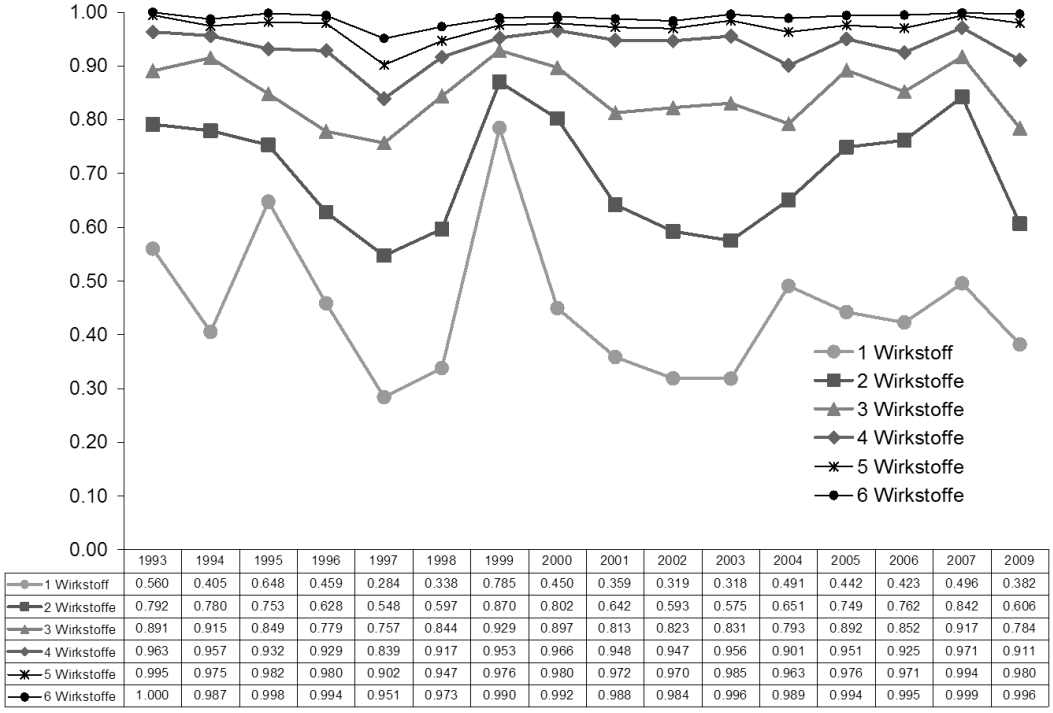
	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2009
E(max)	3254	3191	3185	3487	4120	4330	3879	4143	4429	4446	4429	4460	4400	4429	4397	4623
E(aktuell)	3769	3475	3657	4144	4530	4549	4059	4284	4488	4577	4556	4642	4517	4536	4475	4716

Abbildung 14 Erwartete Erfolgsraten (Therapieabbruch, E_{max}, $E_{aktuell}$) pro 10.000 Episoden



	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2009
1 Wirkstoff	0.643	0.424	0.715	0.420	0.691	0.724	0.389	0.513	0.488	0.547	0.539	0.691	0.559	0.526	0.535	0.555
2 Wirkstoffe	0.782	0.815	0.808	0.680	0.825	0.821	0.615	0.791	0.905	0.731	0.690	0.868	0.874	0.863	0.906	0.827
3 Wirkstoffe	0.912	0.901	0.897	0.835	0.919	0.906	0.815	0.890	0.973	0.855	0.835	0.932	0.925	0.929	0.950	0.932
4 Wirkstoffe	0.952	0.960	0.963	0.975	0.984	0.989	0.968	0.957	0.985	0.976	0.970	0.963	0.952	0.966	0.971	0.980
5 Wirkstoffe	0.990	0.980	0.991	0.986	0.990	0.994	0.986	0.989	0.994	0.986	0.984	0.981	0.975	0.983	0.986	0.994
6 Wirkstoffe	1.000	0.991	0.998	0.993	0.994	0.997	0.995	0.995	0.998	0.996	0.995	0.994	0.997	0.997	0.995	0.996

Abbildung 15 Wahrscheinlichkeit, dass der beste Wirkstoff einer Gruppe von Wirkstoffen bestimmter Größe (1-6 Wirkstoffe) angehört im Zeitverlauf (Zielgröße Therapieansprechen)



	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2009
1 Wirkstoff	0.560	0.405	0.648	0.459	0.284	0.338	0.785	0.450	0.359	0.319	0.318	0.491	0.442	0.423	0.496	0.382
2 Wirkstoffe	0.792	0.780	0.753	0.628	0.548	0.597	0.870	0.802	0.642	0.593	0.575	0.651	0.749	0.762	0.842	0.606
3 Wirkstoffe	0.891	0.915	0.849	0.779	0.757	0.844	0.929	0.897	0.813	0.823	0.831	0.793	0.892	0.852	0.917	0.784
4 Wirkstoffe	0.963	0.957	0.932	0.929	0.839	0.917	0.953	0.966	0.948	0.947	0.956	0.901	0.951	0.925	0.971	0.911
5 Wirkstoffe	0.995	0.975	0.982	0.980	0.902	0.947	0.976	0.980	0.972	0.970	0.985	0.963	0.976	0.971	0.994	0.980
6 Wirkstoffe	1.000	0.987	0.998	0.994	0.951	0.973	0.990	0.992	0.988	0.984	0.996	0.989	0.994	0.995	0.999	0.996

Abbildung 16 Wahrscheinlichkeit, dass der beste Wirkstoff einer Gruppe von Wirkstoffen bestimmter Größe (1-6 Wirkstoffe) angehört im Zeitverlauf (Zielgröße Therapieabbruch)

VII.4 Gegenüberstellung direkter und indirekter/gemischter Vergleiche

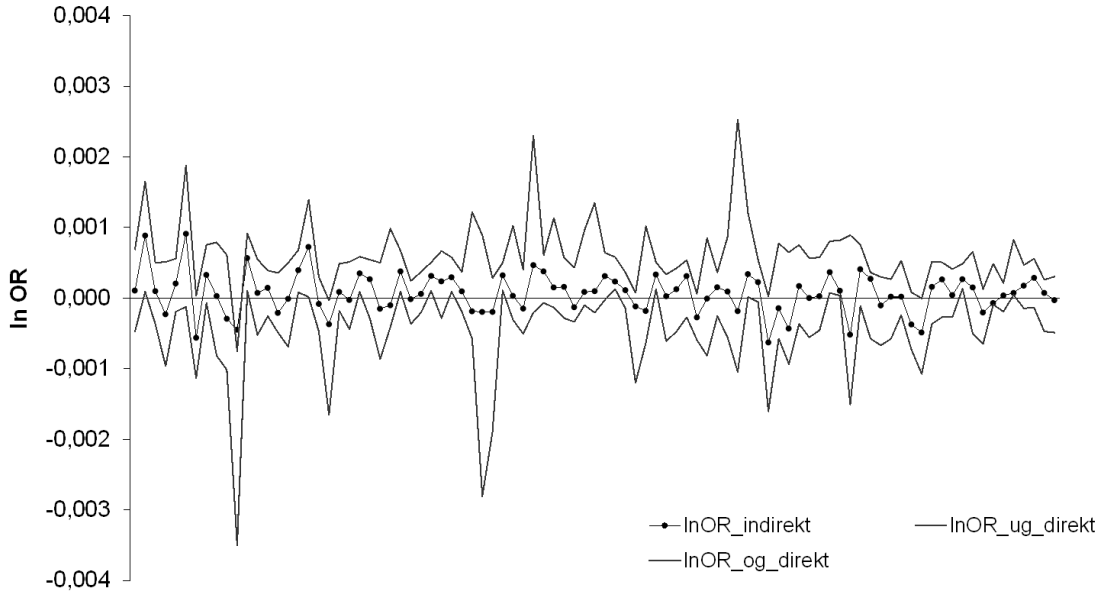


Abbildung 17 Gegenüberstellung gemischter (Punktschätzer) und direkter Vergleiche (Konfidenzintervall) für die Zielgröße Therapieansprechen

VII.5 Ergebnisse direkt vergleichender Meta-Analysen im Zeitverlauf

Hinweis: Es wurden nur Ergebnisse zu Wirkstoffvergleichen aufgenommen, für die zumindest Studienergebnisse zu zwei unterschiedlichen Zeitpunkten zur Verfügung standen, sodass ein Zeitverlauf i. e. S. darstellbar war. Die Studienkennzeichnung entspricht der von Cipriani et. al verwendeten. Die Literaturreferenzen sind in Cipriani et al. (2009a) enthalten. Unpublizierte Studien wurden mit dem „Publikationsjahr“ 2009 gekennzeichnet.

Citalopram (T) - Fluoxetin (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Bougerol1997a	1997	115	173	113	184	1.25	0.81	1.92	0	1.25	0.81	1.92	50.7	50.7
Bougerol1997b	1997	98	158	102	158	0.90	0.57	1.42	4	1.07	0.77	1.47	45.5	96.2
Khanzode2003	2003	3	33	4	34	0.75	0.15	3.64	0	1.05	0.77	1.43	3.8	100.0
Therapieabbruch (Dropout)														
Bougerol1997a	1997	24	173	21	184	1.25	0.67	2.34	0	1.25	0.67	2.34	35.9	35.9
Bougerol1997b	1997	48	158	45	158	1.10	0.68	1.78	0	1.15	0.78	1.69	60.1	95.9
Khanzode2003	2003	3	33	2	34	1.60	0.25	10.25	0	1.17	0.80	1.70	4.1	100.0

Duloxetin (T) - Escitalopram (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Khan2007	2007	66	138	83	140	0.63	0.39	1.01	0	0.63	0.39	1.01	30.3	30.3
Nieremberg2007	2007	113	273	109	274	1.07	0.76	1.50	68	0.84	0.50	1.41	39.0	69.3
Wade2007	2007	81	151	94	144	0.62	0.39	0.98	60	0.77	0.52	1.13	30.7	100.0
Therapieabbruch (Dropout)														
Khan2007	2007	46	138	21	140	2.83	1.58	5.08	0	2.83	1.58	5.08	44.5	44.5
Nieremberg2007	2007	85	273	66	274	1.42	0.98	2.08	73	1.94	0.99	3.78	55.5	100.0

Duloxetin (T) - Paroxetin (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Detke2004	2004	126	188	63	86	0.74	0.42	1.31	0	0.74	0.42	1.31	23.7	23.7
Goldstein2004	2004	81	177	34	87	1.32	0.78	2.22	53	1.00	0.57	1.75	25.6	49.3
Perahia2006	2006	139	196	65	97	1.20	0.71	2.03	16	1.07	0.76	1.50	25.5	74.8
ID#4091	2009	52	175	38	89	0.57	0.33	0.96	54	0.91	0.61	1.35	25.2	100.0
Therapieabbruch (Dropout)														
Detke2004	2004	21	188	10	86	0.96	0.43	2.13	0	0.96	0.43	2.13	15.0	15.0
Goldstein2004	2004	69	177	38	87	0.82	0.49	1.39	0	0.86	0.56	1.33	35.4	50.4
Perahia2006	2006	23	196	11	97	1.04	0.48	2.23	0	0.90	0.62	1.32	16.4	66.8
ID#4091	2009	58	175	31	89	0.93	0.54	1.59	0	0.91	0.67	1.24	33.2	100.0

Escitalopram (T) - Bupropion (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Clayton2006a(WELL_AK130x)	2006	90	138	82	141	1.35	0.83	2.19	0	1.35	0.83	2.19	49.0	49.0
Clayton2006b(WELL_AK130x)	2006	82	149	81	138	0.86	0.54	1.38	41	1.07	0.69	1.67	51.0	100.0
Therapieabbruch (Dropout)														
Clayton2006a(WELL_AK130x)	2006	33	138	32	141	1.07	0.61	1.87	0	1.07	0.61	1.87	31.4	31.4
Clayton2006b(WELL_AK130x)	2006	44	149	39	138	1.06	0.64	1.77	0	1.07	0.73	1.55	37.0	68.4
SCT-MD-35	2009	32	138	34	138	0.92	0.53	1.61	0	1.02	0.75	1.39	31.6	100.0

Escitalopram (T) - Citalopram (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Burke2000	2000	122	252	57	127	1.15	0.75	1.77	0	1.15	0.75	1.77	26.5	26.5
Lepola2001	2001	95	156	79	161	1.62	1.03	2.52	13	1.36	0.97	1.89	24.9	51.4
Colonna2005	2005	104	175	96	182	1.31	0.86	2.00	0	1.34	1.05	1.72	27.4	78.9
Moore2005	2005	105	142	87	152	2.12	1.29	3.47	21	1.48	1.15	1.90	21.1	100.0
Therapieabbruch (Dropout)														
Burke2000	2000	63	252	34	127	0.91	0.56	1.48	0	0.91	0.56	1.48	28.2	28.2
Lepola2001	2001	10	156	9	161	1.16	0.46	2.93	0	0.96	0.62	1.48	11.1	39.3
Colonna2005	2005	25	175	30	182	0.84	0.47	1.50	0	0.92	0.65	1.29	22.8	62.1
Moore2005	2005	10	142	25	152	0.38	0.18	0.83	31	0.78	0.53	1.16	15.0	77.0
SCT-MD-02	2009	33	129	29	128	1.17	0.66	2.08	30	0.86	0.61	1.20	23.0	100.0

Escitalopram (T) - Fluoxetin (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Kasper2005	2005	78	174	61	164	1.37	0.89	2.12	0	1.37	0.89	2.12	63.0	63.0
SCT-MD-16	2009	65	102	65	103	1.03	0.58	1.81	0	1.23	0.87	1.74	37.0	100.0
Therapieabbruch (Dropout)														
Kasper2005	2005	30	174	42	164	0.61	0.36	1.02	0	0.61	0.36	1.02	51.1	51.1
SCT-MD-16	2009	36	102	26	103	1.62	0.88	2.95	83	0.98	0.37	2.56	48.9	100.0

Escitalopram (T) - Paroxetin (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Baldwin2005	2005	112	166	111	159	0.90	0.56	1.43	0	0.90	0.56	1.43	44.3	44.3
Boulenger2006	2006	162	232	144	227	1.33	0.90	1.97	39	1.12	0.76	1.65	55.7	100.0
Therapieabbruch (Dropout)														
Baldwin2005	2005	15	166	14	159	1.03	0.48	2.21	0	1.03	0.48	2.21	33.9	33.9
Boulenger2006	2006	25	232	36	227	0.64	0.37	1.11	0	0.75	0.48	1.17	66.1	100.0

Escitalopram (T) - Sertralin (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Ventura2007	2007	75	107	74	108	1.08	0.60	1.92	0	1.08	0.60	1.92	40.2	40.2
SCT-MD-27	2009	69	136	78	138	0.79	0.49	1.27	0	0.90	0.62	1.29	59.8	100.0
Therapieabbruch (Dropout)														
Ventura2007	2007	19	107	15	108	1.34	0.64	2.80	0	1.34	0.64	2.80	39.9	39.9
SCT-MD-27	2009	28	136	25	138	1.17	0.64	2.14	0	1.24	0.78	1.97	60.1	100.0

Escitalopram (T) - Venlafaxin (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Bielski2004	2004	59	101	47	101	1.61	0.93	2.82	0	1.61	0.93	2.82	49.5	49.5
Montgomery2004	2004	113	148	113	145	0.91	0.53	1.58	51	1.21	0.69	2.11	50.5	100.0
Therapieabbruch (Dropout)														
Bielski2004	2004	29	101	35	101	0.76	0.42	1.38	0	0.76	0.42	1.38	53.8	53.8
Montgomery2004	2004	23	148	21	145	1.09	0.57	2.06	0	0.90	0.58	1.39	46.2	100.0

Fluoxetin (T) - Buproprion (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Feighner1991	1991	35	62	37	61	0.84	0.41	1.72	0	0.84	0.41	1.72	16.3	16.3
Coleman2001	2001	83	154	76	150	1.14	0.73	1.79	0	1.05	0.71	1.53	41.4	57.7
WELL_AK1A4006	2009	88	155	74	158	1.49	0.95	2.33	0	1.21	0.91	1.62	42.3	100.0
Therapieabbruch (Dropout)														
Feighner1991	1991	18	62	16	61	1.15	0.52	2.54	0	1.15	0.52	2.54	14.5	14.5
Coleman2001	2001	57	154	56	150	0.99	0.62	1.57	0	1.03	0.69	1.53	41.8	56.3
WELL_AK1A4006	2009	59	155	62	158	0.95	0.60	1.50	0	0.99	0.73	1.34	43.7	100.0

Fluvoxamin (T) - Fluoxetin (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Rapaport1995	1995	30	51	29	49	0.99	0.44	2.19	0	0.99	0.44	2.19	35.1	35.1
Dalery1998	1998	53	90	54	94	1.06	0.59	1.91	0	1.03	0.64	1.66	64.9	100.0
Therapieabbruch (Dropout)														
Rapaport1995	1995	7	51	8	49	0.82	0.27	2.45	0	0.82	0.27	2.45	27.6	27.6
Dalery1998	1998	24	90	20	94	1.35	0.68	2.66	0	1.17	0.66	2.09	72.4	100.0

Fluvoxamin (T) - Paroxetin (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Annseaau1993	1993	24	64	24	56	0.80	0.38	1.66	0	0.80	0.38	1.66	43.7	43.7
Kiev1997	1997	17	30	16	30	1.14	0.41	3.17	0	0.90	0.50	1.64	22.6	66.4
Kato2006	2006	31	49	37	52	0.70	0.30	1.61	0	0.83	0.51	1.34	33.6	100.0
Therapieabbruch (Dropout)														
Annseaau1993	1993	23	64	16	56	1.40	0.65	3.04	0	1.40	0.65	3.04	46.2	46.2
Kiev1997	1997	11	30	9	30	1.35	0.46	3.97	0	1.38	0.74	2.59	24.5	70.8
Kato2006	2006	8	49	13	52	0.59	0.22	1.57	4	1.08	0.63	1.85	29.2	100.0

Fluvoxamin (T) - Sertralin (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Nemeroff1995	1995	19	49	21	48	0.81	0.36	1.83	0	0.81	0.36	1.83	52.9	52.9
Rossini2002	2002	29	40	28	48	1.88	0.77	4.63	46	1.21	0.53	2.75	47.1	100.0
Therapieabbruch (Dropout)														
Nemeroff1995	1995	21	49	9	48	3.25	1.30	8.15	0	3.25	1.30	8.15	62.7	62.7
Rossini2002	2002	1	40	3	48	0.38	0.04	3.85	65	1.46	0.19	11.16	37.3	100.0

Milnacipram (T) - Fluoxetin (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Annseaau1994	1994	32	97	43	93	0.57	0.32	1.03	0	0.57	0.32	1.03	36.5	36.5
Guelfi1998	1998	109	200	51	100	1.15	0.71	1.86	69	0.83	0.42	1.64	44.9	81.4
Lee2002	2002	15	39	12	31	0.99	0.38	2.61	40	0.87	0.54	1.39	18.6	100.0
Therapieabbruch (Dropout)														
Annseaau1994	1994	23	97	18	93	1.30	0.65	2.60	0	1.30	0.65	2.60	27.5	27.5
Guelfi1998	1998	99	200	50	100	0.98	0.61	1.58	0	1.07	0.72	1.59	57.7	85.3
Lee2002	2002	16	39	15	31	0.74	0.29	1.92	0	1.02	0.71	1.46	14.7	100.0

Mirtazapin (T) - Fluoxetin (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Wheatley1997	1997	39	66	28	67	2.01	1.01	4.01	0	2.01	1.01	4.01	24.3	24.3
Winokur2000	2000	8	9	6	13	9.33	0.89	97.62	34	2.83	0.81	9.96	2.4	26.7
Hong2003	2003	35	66	30	66	1.35	0.68	2.69	25	1.83	1.01	3.31	24.7	51.4
Amini2005	2005	12	18	8	18	2.50	0.65	9.65	0	1.84	1.17	2.88	7.1	58.5
Versiani2005	2005	106	147	104	152	1.19	0.73	1.96	11	1.55	1.07	2.23	41.5	100.0
Therapieabbruch (Dropout)														
Wheatley1997	1997	17	66	21	67	0.76	0.36	1.62	0	0.76	0.36	1.62	41.8	41.8
Winokur2000	2000	1	9	2	13	0.69	0.05	8.96	0	0.75	0.37	1.56	3.6	45.5
Hong2003	2003	30	66	22	66	1.67	0.82	3.37	16	1.12	0.63	2.00	48.1	93.5
Amini2005	2005	2	18	3	18	0.63	0.09	4.28	0	1.09	0.67	1.78	6.5	100.0

Mirtazapin (T) - Paroxetin (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Benkert1999	1999	74	139	66	136	1.21	0.75	1.94	0	1.21	0.75	1.94	38.7	38.7
Schazberg2000	2000	72	128	60	126	1.41	0.86	2.32	0	1.30	0.93	1.83	35.5	74.3
Wade2003	2003	38	99	34	98	1.17	0.66	2.10	0	1.27	0.94	1.70	25.7	100.0
Therapieabbruch (Dropout)														
Benkert1999	1999	30	139	33	136	0.86	0.49	1.51	0	0.86	0.49	1.51	33.5	33.5
Schazberg2000	2000	29	128	39	126	0.65	0.37	1.14	0	0.75	0.50	1.11	33.9	67.4
Wade2003	2003	40	99	38	98	1.07	0.60	1.90	0	0.84	0.61	1.17	32.6	100.0

Mirtazapin (T) - Venlafaxin (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Guelfi2000	2000	48	78	39	79	1.64	0.87	3.10	0	1.64	0.87	3.10	37.6	37.6
Szegedi2005	2005	65	130	52	128	1.46	0.89	2.39	0	1.53	1.03	2.25	62.4	100.0
Therapieabbruch (Dropout)														
Guelfi2000	2000	18	78	28	79	0.55	0.27	1.10	0	0.55	0.27	1.10	35.5	35.5
Szegedi2005	2005	39	130	47	128	0.74	0.44	1.24	0	0.66	0.44	1.01	64.5	100.0

Paroxetin (T) - Fluoxetin (C)														
Studie	**Jahr**	**n(T)**	**N(T)**	**n(C)**	**N(C)**	**OR**	**OR (KI-UG)**	**OR (KI-OG)**	**I-2 [%]**	**kum. OR**	**kum. OR (KI-UG)**	**kum. OR (KI-OG)**	**Gewicht [%]**	**Gewicht - kumuliert [%]**
Therapieansprechen (Response)														
DeWilde1993	1993	25	37	26	41	1.20	0.47	3.07	0	1.20	0.47	3.07	4.2	4.2
Gagiano1993	1993	30	45	27	45	1.33	0.56	3.15	0	1.27	0.67	2.40	4.8	9.0
Tignol1993	1993	53	89	54	87	0.90	0.49	1.65	0	1.06	0.69	1.64	8.5	17.5
Ontiversos1994	1994	41	60	35	62	1.66	0.79	3.49	0	1.19	0.82	1.74	6.2	23.7
Fava1998	1998	32	55	31	54	1.03	0.48	2.21	0	1.16	0.83	1.62	5.9	29.7
Chouinard1999	1999	51	102	53	101	0.91	0.52	1.57	0	1.08	0.81	1.44	9.8	39.5
Fava2000_PF	2000	23	30	26	35	1.14	0.37	3.54	0	1.09	0.82	1.44	2.9	42.4
Fava2002_PF	2002	64	96	57	92	1.23	0.68	2.23	0	1.11	0.86	1.43	8.7	51.1
29060/365	2009	33	68	32	70	1.12	0.57	2.19	0	1.11	0.88	1.41	7.3	58.5
Geretsegger MY 1021/BC	2009	19	54	8	52	2.99	1.17	7.63	0	1.18	0.94	1.48	4.2	62.6
MY-1043/BRL-029060/115	2009	152	284	167	289	0.84	0.60	1.17	0	1.06	0.88	1.28	18.2	80.8
MY-1045/BRL-029060/1	2009	217	357	244	351	0.68	0.50	0.93	26	1.01	0.82	1.23	19.2	100.0
Therapieabbruch (Dropout)														
DeWilde1993	1993	6	37	9	41	0.69	0.22	2.16	0	0.69	0.22	2.16	2.0	2.0
Gagiano1993	1993	8	45	10	45	0.76	0.27	2.14	0	0.73	0.34	1.56	2.4	4.3
Tignol1993	1993	25	89	33	87	0.64	0.34	1.20	0	0.67	0.41	1.10	6.4	10.7
Ontiversos1994	1994	7	60	9	62	0.78	0.27	2.24	0	0.69	0.44	1.08	2.3	13.0
Fava1998	1998	16	55	16	54	0.97	0.43	2.22	0	0.75	0.50	1.10	3.8	16.8
Chouinard1999	1999	40	102	33	101	1.33	0.75	2.36	0	0.90	0.65	1.24	7.7	24.5
Fava2000_PF	2000	13	30	16	35	0.91	0.34	2.42	0	0.90	0.66	1.22	2.7	27.2
Fava2002_PF	2002	27	96	24	92	1.11	0.58	2.11	0	0.93	0.71	1.23	6.2	33.4
2906/421	2009	50	123	45	119	1.13	0.67	1.89	0	0.97	0.76	1.24	9.6	43.0
29060/365	2009	21	68	27	70	0.71	0.35	1.44	0	0.94	0.75	1.18	5.2	48.1
Geretsegger MY 1021/BC	2009	9	54	9	52	0.96	0.35	2.63	0	0.94	0.75	1.18	2.5	50.6
MY-1043/BRL-029060/115	2009	109	284	108	289	1.04	0.74	1.46	0	0.97	0.81	1.17	22.5	73.1
MY-1045/BRL-029060/1	2009	137	357	114	351	1.29	0.95	1.76	0	1.05	0.89	1.23	26.9	100.0

Paroxetin (T) - Sertralin (C)														
Studie	**Jahr**	**n(T)**	**N(T)**	**n(C)**	**N(C)**	**OR**	**OR (KI-UG)**	**OR (KI-OG)**	**I-2 [%]**	**kum. OR**	**kum. OR (KI-UG)**	**kum. OR (KI-OG)**	**Gewicht [%]**	**Gewicht - kumuliert [%]**
Therapieansprechen (Response)														
Zanardi1996	1996	6	22	18	24	0.13	0.03	0.47	0	0.13	0.03	0.47	15.2	15.2
Aberg-Wisted2000	2000	111	177	116	176	0.87	0.56	1.35	87	0.37	0.06	2.42	36.7	51.9
Fava2000_PS	2000	23	30	37	43	0.53	0.16	1.78	74	0.44	0.15	1.32	17.0	68.8
Fava2002_PS	2002	64	96	70	96	0.74	0.40	1.38	61	0.57	0.30	1.07	31.2	100.0
Therapieabbruch (Dropout)														
Zanardi1996	1996	9	22	0	24	34.48	1.86	639.6	0	34.48	1.86	639.6	6.8	6.8
Aberg-Wisted2000	2000	26	177	33	176	0.75	0.43	1.31	86	3.94	0.08	197.9	34.6	41.3
Fava2000_PS	2000	13	30	10	43	2.52	0.92	6.94	80	2.30	0.51	10.43	25.5	66.8
Fava2002_PS	2002	27	96	26	96	1.05	0.56	1.98	70	1.48	0.64	3.39	33.2	100.0

Reboxetin (T) - Citalopram (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Berlanga2004	2004	27	47	33	54	0.86	0.39	1.90	0	0.86	0.39	1.90	32.6	32.6
Langworth2006	2006	83	177	112	173	0.48	0.31	0.74	37	0.58	0.34	0.99	67.4	100.0
Therapieabbruch (Dropout)														
Berlanga2004	2004	8	47	15	54	0.53	0.20	1.40	0	0.53	0.20	1.40	45.5	45.5
Langworth2006	2006	65	177	36	173	2.21	1.37	3.56	85	1.16	0.29	4.63	54.5	100.0

Reboxetin (T) - Fluoxetin (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Bosc1997a	1997	49	126	55	127	0.83	0.50	1.38	0	0.83	0.50	1.38	32.8	32.8
Bosc1997b	1997	46	79	51	89	1.04	0.56	1.92	0	0.91	0.62	1.34	26.3	59.0
Clayton2002	2002	59	150	79	150	0.58	0.37	0.92	17	0.76	0.55	1.06	35.7	94.7
Taner2006	2006	14	22	19	21	0.18	0.03	1.00	40	0.72	0.48	1.08	5.3	100.0
Therapieabbruch (Dropout)														
Bosc1997a	1997	38	126	30	127	1.40	0.80	2.44	0	1.40	0.80	2.44	32.5	32.5
Bosc1997b	1997	20	79	20	89	1.17	0.57	2.38	0	1.30	0.84	2.02	20.1	52.6
Clayton2002	2002	63	150	47	150	1.59	0.99	2.55	0	1.43	1.04	1.97	45.3	98.0
Taner2006	2006	5	22	1	21	5.88	0.62	55.38	0	1.47	1.07	2.02	2.0	100.0

Sertralin (T) - Buproprion (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Kavoussi1997	1997	85	126	82	122	1.01	0.59	1.72	0	1.01	0.59	1.72	32.9	32.9
Coleman1999	1999	67	118	78	122	0.74	0.44	1.25	0	0.86	0.60	1.25	34.4	67.3
Croft1999	1999	79	119	77	120	1.10	0.65	1.88	0	0.93	0.69	1.27	32.7	100.0
Therapieabbruch (Dropout)														
Coleman1999	1999	43	118	27	122	2.02	1.14	3.56	0	2.02	1.14	3.56	49.0	49.0
Croft1999	1999	39	119	36	120	1.14	0.66	1.97	51	1.51	0.86	2.64	51.0	100.0

Sertralin (T) - Fluoxetin (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Aguglia1993	1993	35	52	26	56	2.38	1.09	5.19	0	2.38	1.09	5.19	8.4	8.4
Bennie1995	1995	73	142	63	144	1.36	0.85	2.17	31	1.64	0.98	2.75	23.8	32.2
VanMoffaert1995	1995	49	83	48	82	1.02	0.55	1.90	28	1.41	0.93	2.11	13.4	45.6
Sechter1999	1999	48	118	35	120	1.67	0.97	2.85	3	1.46	1.09	1.95	17.8	63.3
Fava2000_SF	2000	37	43	26	35	2.13	0.68	6.73	0	1.49	1.13	1.97	3.9	67.2
Newhouse2000	2000	85	117	84	119	1.11	0.63	1.95	0	1.41	1.10	1.81	16.1	83.3
Suri2000	2000	9	35	5	18	0.90	0.25	3.24	0	1.39	1.09	1.77	3.1	86.4
Fava2002_SF	2002	70	96	57	92	1.65	0.89	3.06	0	1.42	1.13	1.78	13.6	100.0
Therapieabbruch (Dropout)														
Aguglia1993	1993	17	52	31	56	0.39	0.18	0.86	0	0.39	0.18	0.86	13.9	13.9
Bennie1995	1995	24	142	23	144	1.07	0.57	2.00	74	0.67	0.25	1.78	18.7	32.6
VanMoffaert1995	1995	14	83	16	82	0.84	0.38	1.85	49	0.73	0.40	1.32	13.7	46.3
Fava2000_SF	2000	10	43	16	35	0.36	0.14	0.95	49	0.63	0.37	1.09	10.1	56.4
Newhouse2000	2000	37	117	39	119	0.95	0.55	1.64	42	0.71	0.47	1.09	21.7	78.2
Suri2000	2000	7	35	2	18	2.00	0.37	10.81	38	0.75	0.50	1.14	3.9	82.1
Fava2002_SF	2002	26	96	24	92	1.05	0.55	2.01	32	0.80	0.57	1.14	17.9	100.0

Venlafaxin (T) - Bupropion (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Thase2006	2006	82	177	86	171	0.85	0.56	1.30	0	0.85	0.56	1.30	32.4	32.4
AK130939	2009	127	198	115	204	1.38	0.93	2.07	62	1.09	0.68	1.75	34.4	66.8
WXL101497	2009	120	189	106	188	1.35	0.89	2.03	39	1.17	0.86	1.59	33.2	100.0
Therapieabbruch (Dropout)														
Thase2006	2006	81	177	72	171	1.16	0.76	1.77	0	1.16	0.76	1.77	41.9	41.9
AK130939	2009	46	198	45	204	1.07	0.67	1.71	0	1.12	0.82	1.53	34.5	76.4
WXL101497	2009	25	189	33	188	0.72	0.41	1.26	0	1.01	0.76	1.32	23.6	100.0

Venlafaxin (T) - Fluoxetin (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Clerc1994	1994	24	34	17	34	2.40	0.88	6.51	0	2.40	0.88	6.51	3.2	3.2
Dierick1996	1996	80	153	67	161	1.54	0.98	2.40	0	1.66	1.10	2.49	16.0	19.2
Tylee1997	1997	67	171	58	170	1.24	0.80	1.93	0	1.45	1.08	1.96	16.3	35.5
Costaesilva1998	1998	170	196	153	186	1.41	0.81	2.47	0	1.44	1.11	1.88	10.2	45.7
DiazMaritinez1998	1998	45	70	46	75	1.13	0.58	2.23	0	1.40	1.09	1.79	7.0	52.7
Alves1999	1999	36	40	35	47	3.09	0.91	10.49	0	1.44	1.13	1.83	2.1	54.8
Rudolph1999	1999	54	100	51	103	1.20	0.69	2.08	0	1.40	1.12	1.74	10.5	65.3
Silverstone1999	1999	79	122	75	119	1.08	0.64	1.82	0	1.35	1.10	1.65	11.5	76.8
Tzanakaki2000	2000	36	55	31	54	1.41	0.65	3.05	0	1.35	1.11	1.64	5.3	82.1
DeNayer2002	2002	37	73	29	73	1.56	0.81	3.00	0	1.37	1.13	1.65	7.4	89.4
Nemeroff2007	2007	51	102	45	104	1.31	0.76	2.27	0	1.36	1.14	1.62	10.6	100.0
Therapieabbruch (Dropout)														
Clerc1994	1994	6	34	12	34	0.39	0.13	1.21	0	0.39	0.13	1.21	2.8	2.8
Dierick1996	1996	38	153	40	161	1.00	0.60	1.67	54	0.72	0.30	1.73	13.5	16.3
Tylee1997	1997	47	171	46	170	1.02	0.63	1.65	19	0.91	0.62	1.34	15.6	32.0
Costaesilva1998	1998	29	196	18	186	1.62	0.87	3.03	38	1.03	0.70	1.53	9.1	41.1
DiazMaritinez1998	1998	15	70	20	75	0.75	0.35	1.61	27	0.99	0.71	1.38	6.0	47.1
Alves1999	1999	10	40	9	47	1.41	0.51	3.90	15	1.02	0.76	1.38	3.4	50.5
Rudolph1999	1999	19	100	29	103	0.60	0.31	1.16	26	0.94	0.70	1.27	8.2	58.7
Silverstone1999	1999	37	122	32	119	1.18	0.68	2.07	18	0.98	0.76	1.27	11.3	70.0
Tzanakaki2000	2000	12	55	12	54	0.98	0.39	2.42	7	0.99	0.78	1.24	4.3	74.4
DeNayer2002	2002	29	73	24	73	1.35	0.68	2.65	3	1.02	0.82	1.26	7.8	82.1
Schatzberg2006	2006	36	104	30	100	1.24	0.69	2.22	0	1.04	0.85	1.27	10.3	92.4
Nemeroff2007	2007	24	102	18	104	1.47	0.74	2.91	0	1.07	0.88	1.29	7.6	100.0

Venlafaxin (T) - Sertralin (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Mehtonen2000	2000	49	75	41	72	1.42	0.73	2.77	0	1.42	0.73	2.77	24.5	24.5
Chen2001	2001	31	44	26	45	1.74	0.72	4.19	0	1.53	0.90	2.61	16.0	40.6
Oslin2003	2003	5	27	9	25	0.40	0.11	1.44	47	1.16	0.57	2.36	8.5	49.0
Sir2005	2005	56	84	56	79	0.82	0.42	1.60	37	1.06	0.63	1.77	24.6	73.6
Shelton2006	2006	49	78	45	82	1.39	0.74	2.62	23	1.15	0.78	1.69	26.4	100.0
Therapieabbruch (Dropout)														
Mehtonen2000	2000	16	75	12	72	1.36	0.59	3.11	0	1.36	0.59	3.11	24.7	24.7
Chen2001	2001	1	44	0	45	3.14	0.12	79.13	0	1.43	0.64	3.19	5.8	30.5
Oslin2003	2003	17	27	5	25	6.80	1.94	23.81	55	2.79	0.81	9.60	19.0	49.5
Sir2005	2005	25	79	13	79	2.35	1.10	5.03	33	2.45	1.25	4.80	25.6	75.1
Shelton2006	2006	11	78	19	82	0.54	0.24	1.23	69	1.78	0.75	4.18	24.9	100.0

Citalopram (T) - Sertralin (C)														
Studie	Jahr	n(T)	N(T)	n(C)	N(C)	OR	OR (KI-UG)	OR (KI-OG)	I-2 [%]	kum. OR	kum. OR (KI-UG)	kum. OR (KI-OG)	Gewicht [%]	Gewicht - kumuliert [%]
Therapieansprechen (Response)														
Ekselius1997	1997	136	200	139	200	0.93	0.61	1.42	0	0.93	0.61	1.42	100.0	100.0
Therapieabbruch (Dropout)														
Ekselius1997	1997	36	200	52	200	0.62	0.39	1.01	0	0.62	0.39	1.01	62.5	62.5
Stahl2000	2000	24	107	30	108	0.75	0.40	1.40	0	0.67	0.46	0.98	37.5	100.0

VII.6 Ergebnisse indirekter/gemischter Vergleiche im Zeitverlauf

Tabelle 8 Ergebnisse der kumulativen MTC-Meta-Analyse und der Informationswertermittlung (Zielgröße korrigiertes Therapieansprechen)

korr. Th.-Anspr.												
	p[best]											
Jahr	FLUOX	BUPRO	SERTR	PAROX	CITAL	FLUVO	VENLA	MILNA	MIRTA	REBOX	ESCIT	DULOX
1993	0,005	0,080	0,741	0,037	0,111	0,026						
1994	0,000	0,029	0,466	0,008	0,059	0,009	0,425	0,004				
1995	0,000	0,049	0,064	0,027	0,058	0,002	0,799	0,002				
1996	0,000	0,127	0,386	0,008	0,136	0,006	0,325	0,011				
1997	0,000	0,110	0,157	0,008	0,016	0,003	0,148	0,003	0,552	0,004		
1998	0,000	0,115	0,151	0,002	0,013	0,000	0,050	0,002	0,663	0,003		
1999	0,000	0,706	0,102	0,001	0,052	0,000	0,064	0,003	0,068	0,004		
2000	0,000	0,211	0,024	0,000	0,024	0,000	0,021	0,001	0,315	0,004	0,399	
2001	0,000	0,044	0,013	0,000	0,009	0,000	0,021	0,002	0,277	0,003	0,632	
2002	0,000	0,056	0,010	0,000	0,011	0,000	0,037	0,005	0,062	0,000	0,660	0,158
2003	0,000	0,072	0,013	0,000	0,010	0,000	0,029	0,007	0,042	0,000	0,665	0,161
2004	0,000	0,118	0,025	0,000	0,009	0,001	0,047	0,012	0,075	0,000	0,660	0,054
2005	0,000	0,130	0,035	0,000	0,001	0,001	0,008	0,014	0,147	0,000	0,595	0,069
2006	0,000	0,138	0,015	0,000	0,001	0,000	0,007	0,011	0,136	0,000	0,636	0,055
2007	0,000	0,167	0,021	0,000	0,001	0,000	0,006	0,011	0,140	0,000	0,650	0,005
2009	0,000	0,024	0,099	0,000	0,010	0,002	0,054	0,031	0,360	0,000	0,417	0,002
	T											
	FLUOX	BUPRO	SERTR	PAROX	CITAL	FLUVO	VENLA	MILNA	MIRTA	REBOX	ESCIT	DULOX
1993	0,501	0,532	0,713	0,554	0,513	0,481						
1994	0,499	0,531	0,716	0,563	0,519	0,486	0,706	0,378				
1995	0,501	0,531	0,586	0,565	0,524	0,489	0,711	0,380				
1996	0,500	0,531	0,644	0,519	0,526	0,493	0,624	0,381				
1997	0,500	0,532	0,595	0,519	0,543	0,490	0,586	0,382	0,634	0,460		
1998	0,499	0,532	0,580	0,517	0,536	0,470	0,561	0,453	0,637	0,459		
1999	0,501	0,609	0,573	0,505	0,551	0,469	0,553	0,455	0,536	0,460		
2000	0,500	0,573	0,556	0,511	0,554	0,468	0,549	0,480	0,587	0,460	0,586	
2001	0,499	0,544	0,545	0,509	0,550	0,456	0,550	0,490	0,585	0,460	0,608	
2002	0,501	0,544	0,542	0,514	0,550	0,498	0,551	0,502	0,556	0,426	0,608	0,506
2003	0,500	0,543	0,541	0,511	0,545	0,494	0,545	0,503	0,546	0,425	0,603	0,506
2004	0,498	0,541	0,538	0,508	0,529	0,490	0,545	0,500	0,542	0,438	0,581	0,509
2005	0,501	0,549	0,550	0,522	0,521	0,497	0,540	0,509	0,561	0,440	0,584	0,520
2006	0,501	0,562	0,546	0,521	0,527	0,498	0,541	0,508	0,561	0,407	0,585	0,528
2007	0,500	0,561	0,547	0,518	0,526	0,497	0,540	0,506	0,559	0,406	0,582	0,517
2009	0,501	0,524	0,544	0,503	0,521	0,490	0,539	0,500	0,555	0,406	0,559	0,487
	E[aktuell]	**E[max]**	**V**									
1993	6634	7132	**498**									
1994	6902	7160	**258**									
1995	6787	7112	**325**									
1996	6019	6439	**420**									
1997	6054	6341	**287**									
1998	6101	6373	**272**									
1999	5926	6091	**165**									
2000	5804	5866	**62**									
2001	5955	6078	**123**									
2002	5813	6084	**271**									
2003	5770	6032	**262**									
2004	5653	5810	**157**									
2005	5685	5835	**150**									
2006	5735	5850	**115**									
2007	5733	5823	**90**									
2009	5517	5589	**72**									

Erläuterung (analog für die beiden folgenden Tabellen): Im ersten Segment (p[best]) finden sich, pro Wirkstoff und Jahr, die Wahrscheinlichkeiten, dass es sich bei dem jeweiligen Wirkstoff zum Erkenntnisstand dieses Jahres um den besten Wirkstoff handelt. Im zweiten Segment (T) ist der Anteil der erwarteten, bei Anwendung eines Wirkstoffs erfolgreich behandelbaren Patienten (relativ zum Bezugswirkstoff FLUOX, dem ein Erfolgsrate von 0,5 zugeordnet wird) angegeben. Im dritten Segment ist die Anzahl der erwarteten erfolgreich zu behandelnden Episoden (für eine fiktive Kohorte von 10.000 Patienten) angegeben: E(max) ergibt sich pro Jahr als Produkt von max(p[best]) dieses Zeitpunktes und dem T-Wert dieses Zeitpunktes für den Wirkstoff der dem Wert max(p[best]) zugeordnet ist. E[aktuell] ist die mit den p[best]-Werten gewichtete Summe der T-Werte zum jeweiligen Zeitpunkt. Der Informationswert (V) ergibt sich als Differenz von E[max]-E[aktuell] (s. Abschnitte II.2.2.2 sowie II.2.3 für eine ausführliche Darstellung).

Tabelle 9 Ergebnisse der kumulativen MTC-Meta-Analyse und der Informationswertermittlung (Zielgröße Therapieansprechen)

Th.-Ansprech.												
	p[best]											
Jahr	FLUOX	BUPRO	SERTR	PAROX	CITAL	FLUVO	VENLA	MILNA	MIRTA	REBOX	ESCIT	DULOX
1993	0,010	0,139	0,643	0,038	0,130	0,040						
1994	0,001	0,059	0,391	0,012	0,086	0,020	0,424	0,008				
1995	0,000	0,093	0,066	0,028	0,089	0,006	0,715	0,002				
1996	0,000	0,155	0,260	0,007	0,140	0,011	0,420	0,007				
1997	0,000	0,094	0,065	0,002	0,005	0,004	0,134	0,001	0,691	0,003		
1998	0,000	0,097	0,083	0,001	0,005	0,001	0,085	0,001	0,724	0,003		
1999	0,000	0,389	0,226	0,000	0,019	0,002	0,153	0,002	0,200	0,008		
2000	0,000	0,099	0,032	0,000	0,004	0,000	0,067	0,001	0,513	0,006	0,279	
2001	0,000	0,010	0,011	0,000	0,000	0,000	0,069	0,001	0,417	0,004	0,488	
2002	0,000	0,010	0,010	0,000	0,001	0,000	0,121	0,003	0,183	0,000	0,547	0,125
2003	0,000	0,014	0,011	0,000	0,001	0,001	0,135	0,004	0,151	0,000	0,539	0,145
2004	0,000	0,018	0,013	0,000	0,001	0,000	0,064	0,004	0,177	0,000	0,691	0,031
2005	0,000	0,024	0,021	0,000	0,000	0,000	0,027	0,003	0,315	0,000	0,559	0,051
2006	0,000	0,037	0,014	0,000	0,000	0,000	0,017	0,003	0,337	0,000	0,526	0,067
2007	0,000	0,044	0,015	0,000	0,000	0,000	0,021	0,005	0,370	0,000	0,535	0,009
2009	0,000	0,002	0,048	0,000	0,002	0,001	0,106	0,014	0,555	0,000	0,271	0,001
	T											
	FLUOX	BUPRO	SERTR	PAROX	CITAL	FLUVO	VENLA	MILNA	MIRTA	REBOX	ESCIT	DULOX
1993	0,501	0,536	0,664	0,519	0,493	0,474						
1994	0,501	0,535	0,666	0,541	0,513	0,495	0,669	0,391				
1995	0,499	0,535	0,566	0,542	0,526	0,504	0,673	0,388				
1996	0,500	0,536	0,601	0,509	0,525	0,505	0,612	0,389				
1997	0,499	0,558	0,572	0,505	0,528	0,502	0,579	0,385	0,641	0,481		
1998	0,500	0,556	0,571	0,516	0,526	0,493	0,570	0,468	0,637	0,481		
1999	0,500	0,580	0,576	0,511	0,537	0,494	0,564	0,468	0,558	0,480		
2000	0,500	0,558	0,560	0,518	0,536	0,491	0,565	0,486	0,594	0,483	0,567	
2001	0,500	0,528	0,549	0,516	0,531	0,479	0,565	0,498	0,590	0,482	0,593	
2002	0,500	0,528	0,547	0,521	0,530	0,508	0,567	0,504	0,571	0,442	0,593	0,497
2003	0,500	0,530	0,549	0,523	0,530	0,510	0,566	0,510	0,568	0,436	0,590	0,505
2004	0,501	0,530	0,549	0,522	0,534	0,511	0,565	0,510	0,571	0,444	0,597	0,521
2005	0,500	0,532	0,550	0,528	0,521	0,509	0,555	0,508	0,574	0,442	0,582	0,523
2006	0,500	0,550	0,550	0,529	0,523	0,507	0,554	0,509	0,576	0,422	0,582	0,537
2007	0,501	0,552	0,552	0,530	0,523	0,509	0,555	0,513	0,576	0,421	0,581	0,538
2009	0,501	0,516	0,545	0,507	0,521	0,498	0,551	0,503	0,567	0,421	0,559	0,500
	E[aktuell]	**E[max]**	**V**									
1993	6089	6637	**548**									
1994	6391	6685	**294**									
1995	6343	6725	**382**									
1996	5817	6120	**303**									
1997	6184	6413	**229**									
1998	6166	6372	**206**									
1999	5702	5803	**101**									
2000	5788	5942	**154**									
2001	5881	5928	**47**									
2002	5722	5927	**205**									
2003	5695	5901	**206**									
2004	5859	5974	**115**									
2005	5739	5823	**84**									
2006	5746	5819	**73**									
2007	5760	5810	**50**									
2009	5610	5674	**64**									

Tabelle 10 Ergebnisse der kumulativen MTC-Meta-Analyse und der Informationswertermittlung (Zielgröße Therapieabbruch)

Th.-Abbruch												
	p[best]											
Jahr	FLUOX	BUPRO	SERTR	PAROX	CITAL	FLUVO	VENLA	MILNA	MIRTA	REBOX	ESCIT	DULOX
1993	0,005	0,100	0,560	0,072	0,232	0,031						
1994	0,000	0,041	0,374	0,018	0,136	0,013	0,405	0,012				
1995	0,000	0,084	0,106	0,050	0,095	0,002	0,648	0,016				
1996	0,001	0,151	0,459	0,015	0,169	0,005	0,150	0,051				
1997	0,000	0,210	0,264	0,032	0,063	0,002	0,082	0,049	0,284	0,014		
1998	0,000	0,247	0,259	0,017	0,073	0,000	0,030	0,027	0,338	0,009		
1999	0,000	0,785	0,059	0,002	0,086	0,000	0,014	0,023	0,023	0,008		
2000	0,000	0,450	0,013	0,000	0,069	0,000	0,002	0,013	0,095	0,005	0,353	
2001	0,000	0,284	0,024	0,001	0,136	0,000	0,003	0,016	0,171	0,008	0,359	
2002	0,000	0,230	0,014	0,001	0,124	0,000	0,002	0,023	0,012	0,000	0,319	0,274
2003	0,000	0,257	0,029	0,000	0,125	0,000	0,000	0,012	0,002	0,001	0,318	0,256
2004	0,001	0,491	0,062	0,000	0,109	0,000	0,001	0,026	0,004	0,005	0,141	0,160
2005	0,000	0,442	0,058	0,000	0,025	0,000	0,000	0,018	0,003	0,002	0,307	0,143
2006	0,000	0,339	0,045	0,000	0,073	0,000	0,000	0,024	0,005	0,000	0,423	0,091
2007	0,000	0,346	0,054	0,000	0,075	0,000	0,000	0,023	0,005	0,000	0,496	0,001
2009	0,000	0,224	0,178	0,000	0,127	0,001	0,000	0,069	0,017	0,000	0,382	0,002
	T											
	FLUOX	BUPRO	SERTR	PAROX	CITAL	FLUVO	VENLA	MILNA	MIRTA	REBOX	ESCIT	DULOX
1993	0,501	0,472	0,325	0,422	0,428	0,491						
1994	0,499	0,471	0,319	0,425	0,429	0,494	0,320	0,550				
1995	0,501	0,471	0,409	0,438	0,474	0,543	0,319	0,554				
1996	0,500	0,474	0,349	0,491	0,472	0,541	0,432	0,551				
1997	0,500	0,471	0,412	0,488	0,460	0,554	0,467	0,553	0,445	0,551		
1998	0,500	0,470	0,433	0,496	0,467	0,570	0,482	0,519	0,446	0,555		
1999	0,500	0,388	0,446	0,513	0,455	0,576	0,487	0,518	0,509	0,552		
2000	0,500	0,414	0,463	0,503	0,445	0,573	0,496	0,503	0,454	0,555	0,430	
2001	0,499	0,444	0,470	0,506	0,447	0,573	0,497	0,507	0,456	0,552	0,443	
2002	0,501	0,448	0,477	0,511	0,451	0,543	0,505	0,498	0,498	0,574	0,445	0,478
2003	0,500	0,443	0,467	0,512	0,450	0,545	0,517	0,505	0,512	0,564	0,445	0,481
2004	0,501	0,446	0,472	0,515	0,472	0,552	0,516	0,507	0,515	0,541	0,480	0,487
2005	0,501	0,440	0,463	0,508	0,478	0,549	0,520	0,505	0,504	0,541	0,447	0,479
2006	0,500	0,448	0,470	0,511	0,467	0,538	0,512	0,504	0,503	0,569	0,443	0,491
2007	0,500	0,447	0,467	0,517	0,467	0,540	0,514	0,505	0,504	0,569	0,440	0,529
2009	0,500	0,471	0,471	0,520	0,476	0,540	0,511	0,506	0,506	0,572	0,462	0,536
	E[aktuell]	**E[max]**	**V**									
1993	3769	3254	**515**									
1994	3475	3191	**284**									
1995	3657	3185	**472**									
1996	4144	3487	**657**									
1997	4530	4120	**410**									
1998	4549	4330	**219**									
1999	4059	3879	**180**									
2000	4284	4143	**141**									
2001	4488	4429	**59**									
2002	4577	4446	**131**									
2003	4556	4429	**127**									
2004	4642	4460	**182**									
2005	4517	4400	**117**									
2006	4536	4429	**107**									
2007	4475	4397	**78**									
2009	4716	4623	**93**									

VII.7 Ergebnisse indirekter/gemischter Vergleiche im Zeitverlauf (alternativer Verlauf der Erkenntnisgewinnung)

Tabelle 11 Ergebnisse der kumulativen MTC-Meta-Analyse und der Informationswertermittlung (Zielgröße korrigiertes Therapieansprechen) – alternativer Verlauf der Erkenntnisgewinnung

korr. Th.-Anspr.												
	p[best]											
Jahr	FLUOX	BUPRO	SERTR	PAROX	CITAL	FLUVO	VENLA	MILNA	MIRTA	REBOX	ESCIT	DULOX
1993	0.005	0.080	0.741	0.037	0.111	0.026						
1994	0.000	0.029	0.466	0.008	0.059	0.009	0.425	0.004				
1995	0.000	0.049	0.064	0.027	0.058	0.002	0.799	0.002				
1996	0.000	0.127	0.386	0.008	0.136	0.006	0.325	0.011				
1997	0.000	0.110	0.157	0.008	0.016	0.003	0.148	0.003	0.552	0.004		
1998	0.000	0.115	0.151	0.002	0.013	0.000	0.050	0.002	0.663	0.003		
1999	0.000	0.706	0.102	0.001	0.052	0.000	0.064	0.003	0.068	0.004		
2000	0.000	0.211	0.024	0.000	0.024	0.000	0.021	0.001	0.315	0.004	0.399	
2001	0.000	0.044	0.013	0.000	0.009	0.000	0.021	0.002	0.277	0.003	0.632	
2002	0.000	0.056	0.010	0.000	0.011	0.000	0.037	0.005	0.062	0.000	0.660	0.158
2003	0.000	0.119	0.035	0.001	0.035	0.001	0.059	0.013	0.144	0.000	0.582	0.011
2004	0.000	0.157	0.046	0.001	0.003	0.001	0.077	0.016	0.168	0.000	0.519	0.013
2005	0.000	0.040	0.040	0.000	0.003	0.001	0.051	0.018	0.311	0.000	0.535	0.001
2006	0.000	0.050	0.066	0.000	0.002	0.002	0.043	0.023	0.334	0.000	0.480	0.001
2007	0.000	0.062	0.092	0.000	0.002	0.002	0.057	0.031	0.366	0.001	0.383	0.003
2009	0.000	0.024	0.098	0.000	0.010	0.002	0.055	0.034	0.360	0.000	0.414	0.003
	T											
	FLUOX	BUPRO	SERTR	PAROX	CITAL	FLUVO	VENLA	MILNA	MIRTA	REBOX	ESCIT	DULOX
1993	0.501	0.532	0.713	0.554	0.513	0.481						
1994	0.499	0.531	0.716	0.563	0.519	0.486	0.706	0.378				
1995	0.501	0.531	0.586	0.565	0.524	0.489	0.711	0.380				
1996	0.500	0.531	0.644	0.519	0.526	0.493	0.624	0.381				
1997	0.500	0.532	0.595	0.519	0.543	0.490	0.586	0.382	0.634	0.460		
1998	0.499	0.532	0.580	0.517	0.536	0.470	0.561	0.453	0.637	0.459		
1999	0.501	0.609	0.573	0.505	0.551	0.469	0.553	0.455	0.536	0.460		
2000	0.500	0.573	0.556	0.511	0.554	0.468	0.549	0.480	0.587	0.460	0.586	
2001	0.499	0.544	0.545	0.509	0.550	0.456	0.550	0.490	0.585	0.460	0.608	
2002	0.501	0.544	0.542	0.514	0.550	0.498	0.551	0.502	0.556	0.426	0.608	0.506
2003	0.501	0.555	0.543	0.513	0.542	0.496	0.551	0.503	0.554	0.426	0.580	0.505
2004	0.499	0.555	0.549	0.514	0.522	0.494	0.551	0.503	0.554	0.424	0.576	0.498
2005	0.499	0.542	0.548	0.514	0.522	0.496	0.550	0.504	0.565	0.425	0.575	0.500
2006	0.500	0.540	0.549	0.506	0.514	0.492	0.546	0.502	0.563	0.443	0.570	0.494
2007	0.501	0.537	0.548	0.509	0.510	0.491	0.544	0.502	0.559	0.442	0.561	0.492
2009	0.500	0.524	0.544	0.502	0.521	0.490	0.539	0.501	0.555	0.405	0.558	0.487
	E[aktuell]	**E[max]**	**V**									
1993	6634	7132	**498**									
1994	6902	7160	**258**									
1995	6787	7112	**325**									
1996	6019	6439	**420**									
1997	6054	6341	**287**									
1998	6101	6373	**272**									
1999	5926	6091	**165**									
2000	5804	5866	**62**									
2001	5955	6078	**123**									
2002	5813	6084	**271**									
2003	5672	5804	**132**									
2004	5634	5763	**129**									
2005	5663	5747	**84**									
2006	5617	5699	**82**									
2007	5544	5613	**69**									
2009	5511	5582	**71**									

Erläuterung (analog für die beiden folgenden Tabellen): Im ersten Segment (p[best]) finden sich, pro Wirkstoff und Jahr, die Wahrscheinlichkeiten, dass es sich bei dem jeweiligen Wirkstoff zum Erkenntnisstand dieses Jahres um den besten Wirkstoff handelt. Im zweiten Segment (T) ist der Anteil der erwarteten, bei Anwendung eines Wirkstoffs erfolgreich behandelbaren Patienten (relativ zum Bezugswirkstoff FLUOX, dem ein Erfolgsrate von 0,5 zugeordnet wird) angegeben. Im dritten Segment ist die Anzahl der erwarteten erfolgreich zu behandelnden Episoden (für eine fiktive Kohorte von 10.000 Patienten) angegeben: E(max) ergibt sich pro Jahr als Produkt von max(p[best]) dieses Zeitpunktes und dem T-Wert dieses Zeitpunktes für den Wirkstoff der dem Wert max(p[best]) zugeordnet ist. E[aktuell] ist die mit den p[best]-Werten gewichtete Summe der T-Werte zum jeweiligen Zeitpunkt. Der Informationswert (V) ergibt sich als Differenz von E[max]-E[aktuell] (s. Abschnitte II.2.2.2 sowie II.2.3 für eine ausführliche Darstellung).

Tabelle 12 Ergebnisse der kumulativen MTC-Meta-Analyse und der Informationswertermittlung (Zielgröße Therapieansprechen) – alternativer Verlauf der Erkenntnisgewinnung

Th.-Ansprech.												
	p[best]											
Jahr	FLUOX	BUPRO	SERTR	PAROX	CITAL	FLUVO	VENLA	MILNA	MIRTA	REBOX	ESCIT	DULOX
1993	0.010	0.139	0.643	0.038	0.130	0.040						
1994	0.001	0.059	0.391	0.012	0.086	0.020	0.424	0.008				
1995	0.000	0.093	0.066	0.028	0.089	0.006	0.715	0.002				
1996	0.000	0.155	0.260	0.007	0.140	0.011	0.420	0.007				
1997	0.000	0.094	0.065	0.002	0.005	0.004	0.134	0.001	0.691	0.003		
1998	0.000	0.097	0.083	0.001	0.005	0.001	0.085	0.001	0.724	0.003		
1999	0.000	0.389	0.226	0.000	0.019	0.002	0.153	0.002	0.200	0.008		
2000	0.000	0.099	0.032	0.000	0.004	0.000	0.067	0.001	0.513	0.006	0.279	
2001	0.000	0.010	0.011	0.000	0.000	0.000	0.069	0.001	0.417	0.004	0.488	
2002	0.000	0.010	0.010	0.000	0.001	0.000	0.121	0.003	0.183	0.000	0.547	0.125
2003	0.000	0.021	0.017	0.000	0.003	0.000	0.152	0.004	0.268	0.000	0.436	0.099
2004	0.000	0.031	0.018	0.000	0.000	0.001	0.151	0.006	0.332	0.000	0.365	0.095
2005	0.000	0.005	0.019	0.000	0.000	0.001	0.091	0.009	0.600	0.000	0.268	0.005
2006	0.000	0.009	0.037	0.000	0.000	0.002	0.099	0.013	0.577	0.000	0.258	0.004
2007	0.000	0.012	0.032	0.000	0.000	0.001	0.110	0.015	0.679	0.000	0.148	0.003
2009	0.000	0.002	0.035	0.000	0.001	0.001	0.120	0.015	0.615	0.000	0.207	0.004
	T											
	FLUOX	BUPRO	SERTR	PAROX	CITAL	FLUVO	VENLA	MILNA	MIRTA	REBOX	ESCIT	DULOX
1993	0.501	0.536	0.664	0.519	0.493	0.474						
1994	0.501	0.535	0.666	0.541	0.513	0.495	0.669	0.391				
1995	0.499	0.535	0.566	0.542	0.526	0.504	0.673	0.388				
1996	0.500	0.536	0.601	0.509	0.525	0.505	0.612	0.389				
1997	0.499	0.558	0.572	0.505	0.528	0.502	0.579	0.385	0.641	0.481		
1998	0.500	0.556	0.571	0.516	0.526	0.493	0.570	0.468	0.637	0.481		
1999	0.500	0.580	0.576	0.511	0.537	0.494	0.564	0.468	0.558	0.480		
2000	0.500	0.558	0.560	0.518	0.536	0.491	0.565	0.486	0.594	0.483	0.567	
2001	0.500	0.528	0.549	0.516	0.531	0.479	0.565	0.498	0.590	0.482	0.593	
2002	0.500	0.528	0.547	0.521	0.530	0.508	0.567	0.504	0.571	0.442	0.593	0.497
2003	0.500	0.538	0.544	0.519	0.526	0.506	0.563	0.503	0.568	0.443	0.578	0.546
2004	0.500	0.538	0.548	0.521	0.514	0.505	0.563	0.507	0.569	0.443	0.574	0.542
2005	0.499	0.533	0.549	0.528	0.511	0.510	0.560	0.510	0.580	0.443	0.568	0.522
2006	0.499	0.527	0.545	0.509	0.504	0.501	0.554	0.505	0.572	0.448	0.560	0.509
2007	0.501	0.527	0.545	0.514	0.502	0.502	0.555	0.505	0.576	0.447	0.554	0.509
2009	0.501	0.513	0.542	0.508	0.517	0.497	0.552	0.504	0.569	0.419	0.555	0.506
	E[aktuell]	**E[max]**	**V**									
1993	6089	6637	**548**									
1994	6391	6685	**294**									
1995	6343	6725	**382**									
1996	5817	6120	**303**									
1997	6184	6413	**229**									
1998	6166	6372	**206**									
1999	5702	5803	**101**									
2000	5788	5942	**154**									
2001	5881	5928	**47**									
2002	5722	5927	**205**									
2003	5680	5778	**98**									
2004	5655	5735	**80**									
2005	5731	5802	**71**									
2006	5643	5720	**77**									
2007	5674	5758	**84**									
2009	5619	5694	**75**									

Tabelle 13 Ergebnisse der kumulativen MTC-Meta-Analyse und der Informationswertermittlung (Zielgröße Therapieabbruch)) – alternativer Verlauf der Erkenntnisgewinnung

Th.-Abbruch												
	p[best]											
Jahr	FLUOX	BUPRO	SERTR	PAROX	CITAL	FLUVO	VENLA	MILNA	MIRTA	REBOX	ESCIT	DULOX
1993	0.005	0.100	0.560	0.072	0.232	0.031						
1994	0.000	0.041	0.374	0.018	0.136	0.013	0.405	0.012				
1995	0.000	0.084	0.106	0.050	0.095	0.002	0.648	0.016				
1996	0.001	0.151	0.459	0.015	0.169	0.005	0.150	0.051				
1997	0.000	0.210	0.264	0.032	0.063	0.002	0.082	0.049	0.284	0.014		
1998	0.000	0.247	0.259	0.017	0.073	0.000	0.030	0.027	0.338	0.009		
1999	0.000	0.785	0.059	0.002	0.086	0.000	0.014	0.023	0.023	0.008		
2000	0.000	0.450	0.013	0.000	0.069	0.000	0.002	0.013	0.095	0.005	0.353	
2001	0.000	0.284	0.024	0.001	0.136	0.000	0.003	0.016	0.171	0.008	0.359	
2002	0.000	0.230	0.014	0.001	0.124	0.000	0.002	0.023	0.012	0.000	0.319	0.274
2003	0.000	0.416	0.044	0.001	0.227	0.000	0.003	0.046	0.021	0.001	0.240	0.001
2004	0.000	0.551	0.075	0.002	0.065	0.000	0.005	0.029	0.031	0.001	0.240	0.001
2005	0.000	0.200	0.090	0.000	0.105	0.000	0.003	0.029	0.040	0.001	0.531	0.001
2006	0.000	0.208	0.183	0.000	0.044	0.000	0.002	0.043	0.040	0.006	0.471	0.002
2007	0.001	0.221	0.226	0.000	0.048	0.000	0.001	0.055	0.024	0.007	0.414	0.002
2009	0.001	0.168	0.256	0.000	0.079	0.001	0.001	0.068	0.017	0.000	0.408	0.002
	T											
	FLUOX	BUPRO	SERTR	PAROX	CITAL	FLUVO	VENLA	MILNA	MIRTA	REBOX	ESCIT	DULOX
1993	0.501	0.472	0.325	0.422	0.428	0.491						
1994	0.499	0.471	0.319	0.425	0.429	0.494	0.320	0.550				
1995	0.501	0.471	0.409	0.438	0.474	0.543	0.319	0.554				
1996	0.500	0.474	0.349	0.491	0.472	0.541	0.432	0.551				
1997	0.500	0.471	0.412	0.488	0.460	0.554	0.467	0.553	0.445	0.551		
1998	0.500	0.470	0.433	0.496	0.467	0.570	0.482	0.519	0.446	0.555		
1999	0.500	0.388	0.446	0.513	0.455	0.576	0.487	0.518	0.509	0.552		
2000	0.500	0.414	0.463	0.503	0.445	0.573	0.496	0.503	0.454	0.555	0.430	
2001	0.499	0.444	0.470	0.506	0.447	0.573	0.497	0.507	0.456	0.552	0.443	
2002	0.501	0.448	0.477	0.511	0.451	0.543	0.505	0.498	0.498	0.574	0.445	0.478
2003	0.500	0.445	0.473	0.509	0.458	0.543	0.501	0.497	0.497	0.574	0.455	0.565
2004	0.500	0.441	0.466	0.510	0.475	0.545	0.500	0.503	0.498	0.572	0.453	0.563
2005	0.500	0.460	0.468	0.516	0.471	0.543	0.494	0.506	0.488	0.572	0.445	0.527
2006	0.499	0.462	0.461	0.516	0.483	0.545	0.500	0.505	0.490	0.545	0.450	0.528
2007	0.500	0.468	0.465	0.516	0.489	0.549	0.507	0.505	0.501	0.545	0.458	0.531
2009	0.500	0.474	0.467	0.518	0.484	0.541	0.510	0.506	0.504	0.574	0.461	0.535
	E[aktuell]	**E[max]**	**V**									
1993	3769	3254	**515**									
1994	3475	3191	**284**									
1995	3657	3185	**472**									
1996	4144	3487	**657**									
1997	4530	4120	**410**									
1998	4549	4330	**219**									
1999	4059	3879	**180**									
2000	4284	4143	**141**									
2001	4488	4429	**59**									
2002	4577	4446	**131**									
2003	4556	4454	**102**									
2004	4520	4406	**114**									
2005	4569	4454	**115**									
2006	4609	4501	**108**									
2007	4680	4582	**98**									
2009	4709	4613	**96**									

VII.8 Weitere Ergebnisse zur Erkenntnisgewinnung im Zeitverlauf (alternativer Verlauf der Erkenntnisgewinnung)

Tabelle 14 In die Ermittlung des alternativen Verlaufs der Erkenntnisgewinnung einbezogene Studien mit Publikationsjahr (2003 und später) und Jahr der Verwendung bei der Ermittlung des alternativen Verlaufs der Erkenntnisgewinnung

Studie	Wirkstoff (A)	Wirkstoff (B)	Publikation (Jahr)	Verwendung im alternativer Verlauf (Jahr)
WELL_AK1A4006	FLUOX	BUPRO	2009	2007
29060/365	FLUOX	PAROX	2009	2005
Geretsegger MY 1021/BC	FLUOX	PAROX	2009	2005
MY-1043/BRL-029060/115	FLUOX	PAROX	2009	2005
MY-1045/BRL-029060/1	FLUOX	PAROX	2009	2005
Khanzode2003	FLUOX	CITAL	2003	2005
Nemeroff2007	FLUOX	VENLA	2007	2006
Hong2003	FLUOX	MIRTA	2003	2006
Amini2005	FLUOX	MIRTA	2005	2006
Taner2006	FLUOX	REBOX	2006	2007
SCT-MD-16	FLUOX	ESCIT	2009	2007
Kasper2005	FLUOX	ESCIT	2005	2007
AK130939	BUPRO	VENLA	2009	2004
WXL101497	BUPRO	VENLA	2009	2004
Thase2006	BUPRO	VENLA	2006	2004
Clayton2006a(WELL_AK130x)	BUPRO	ESCIT	2006	2002
Clayton2006b(WELL_AK130x)	BUPRO	ESCIT	2006	2002
Oslin2003	SERTR	VENLA	2003	2005
Sir2005	SERTR	VENLA	2005	2005
Shelton2006	SERTR	VENLA	2006	2005
Yang2003	SERTR	MILNA	2003	2003
Eker2005	SERTR	REBOX	2005	2007
SCT-MD-27	SERTR	ESCIT	2009	2003
Ventura2007	SERTR	ESCIT	2007	2003
29060/785	PAROX	CITAL	2009	2007
Kato2006	PAROX	FLUVO	2006	2007
Wade2003	PAROX	MIRTA	2003	2007
Baldwin2005	PAROX	ESCIT	2005	2006
Boulenger2006	PAROX	ESCIT	2006	2007
ID#4091	PAROX	DULOX	2009	2004
Detke2004	PAROX	DULOX	2004	2004
Goldstein2004	PAROX	DULOX	2004	2004
Perahia2006	PAROX	DULOX	2006	2004
Berlanga2004	CITAL	REBOX	2004	2005
Langworth2006	CITAL	REBOX	2006	2007
Colonna2005	CITAL	ESCIT	2005	2003
Moore2005	CITAL	ESCIT	2005	2003
Szegedi2005	VENLA	MIRTA	2005	2004
Akkaya2003	VENLA	REBOX	2003	2007
Bielski2004	VENLA	ESCIT	2004	2002
Montgomery2004	VENLA	ESCIT	2004	2002
Khan2007	ESCIT	DULOX	2007	2002
Nieremberg2007	ESCIT	DULOX	2007	2002

Die Studienkennzeichnung entspricht der von Cipriani et. al verwendeten. Die Literaturreferenzen sind in Cipriani et al. (2009a) enthalten. Unpublizierte Studien wurden mit dem „Publikationsjahr“ 2009 gekennzeichnet.

VIII. Danksagungen

Ich danke Herrn Prof. Heiner Raspe und Herrn Prof. Martin Scherer, erst von Lübeck, dann von Hamburg aus, für die in mich und meine Arbeit gesetzte Erwartung, dass diese zu einem guten Ende gebracht werden kann und für die allgemeine Unterstützung bei der Anfertigung der Arbeit.

Besonders danke ich Frau Dr. Lühmann für die Unterstützung und Diskussion der Methodik und Zielrichtung der Arbeit sowie Herrn Levente Kriston für die kritische Diskussion insbesondere der Methodik.

Meine Eltern kann ich nicht genug danken für mehr als hier erwähnt werden kann und dafür, dass es für sie außer Frage stand, dass mir die Arbeit gelingen kann, wenn ich es denn will.

Ich danke schließlich ganz besonders meiner Frau, Andrea Pfennig, für so vieles was hier nicht erwähnt zu werden braucht – vor allem aber dafür, dass sie, selbst Ärztin und Wissenschaftlerin, mir klar gemacht hat, dass es nicht ausreicht eine Idee zu haben, große Linien zu skizzieren und Dinge im „Großen und Ganzen“ zu formulieren und es auch nicht ausreicht etwas zu achtzig oder neunzig Prozent oder beinahe fertig zu machen – sondern es vielmehr darauf ankommt, eine wissenschaftliche Arbeit nach den Regeln der wissenschaftlichen Methodik und mit bestem Wissen und Gewissen vollständig fertigzustellen.

IX. Lebenslauf

02.12.1965	Geboren in Düsseldorf (Eltern: Dagmar Storz (geb. Koch), Günter Storz)
1972-1976	Besuch der Grundschule in Kaarst, Gelsenkirchen und Tuttlingen
1976-1985	Besuch des Gymnasiums Tuttlingen und des Gymnasiums Oldenfelde in Hamburg. Abiturprüfung bestanden am 12.06.1985
1985-1994	Studium der Mathematik, Soziologie, Philosophie, Ethnologie und Literaturwissenschaft an der Universität Hamburg
1994-1998	Studium Soziologie und Literaturwissenschaft an der Universität Hannover. Magisterprüfung bestanden am 27.11.1998
1999-2001	Wissenschaftlicher Mitarbeiter Institut für Gesundheits- und Sozialforschung GmbH (IGES GmbH) in Berlin.
2004-2007	Studium Gesundheitswissenschaften/Public Health TU Berlin. Magisterprüfung bestanden am 19.3.2007
2004-2008	Wissenschaftlicher Mitarbeiter IGES-Institut GmbH Berlin
Ab 2009	Referent GKV-Spitzenverband, Berlin
2010	Heirat mit Prof. Dr. Andrea Pfennig