

**Selektive Serotonin- und  
Noradrenalin-  
Wiederaufnahmehemmer  
(SNRI) bei Patienten mit  
Depressionen**

**Dokumentation und Würdigung der  
Stellungnahmen zum Vorbericht**

Auftrag A05-20A  
Version 1.0  
Stand: 16.06.2009

# Impressum

**Herausgeber:**

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

**Thema:**

Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen

**Auftraggeber:**

Gemeinsamer Bundesausschuss

**Datum des Auftrags:**

22.02.2005

**Interne Auftragsnummer:**

A05-20A

**Anschrift des Herausgebers:**

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen  
Dillenburger Str. 27  
51105 Köln

Tel.: +49 221 35685-0

Fax: +49 221 35685-1

Berichte@iqwig.de

www.iqwig.de

# Inhaltsverzeichnis

	Seite
<b>1</b>	<b>Dokumentation des Stellungnahmeprozesses ..... 1</b>
<b>2</b>	<b>Würdigung der Stellungnahmen..... 2</b>
<b>3</b>	<b>Darlegung potenzieller Interessenkonflikte der Stellungnehmenden ..... 3</b>
<b>3.1</b>	<b>Interessenkonflikte von Stellungnehmenden aus Organisationen, Institutionen und Firmen..... 3</b>
<b>3.2</b>	<b>Interessenkonflikte von stellungnehmenden Privatpersonen ..... 5</b>
<b>3.3</b>	<b>Interessenkonflikte von weiteren Teilnehmern der wissenschaftlichen Erörterung (externe Sachverständige, externe Reviewer) ..... 5</b>
<b>4</b>	<b>Dokumentation der wissenschaftlichen Erörterung – Teilnehmerliste, Tagesordnung und Protokoll..... 8</b>
<b>4.1</b>	<b>Teilnehmerliste der wissenschaftlichen Erörterung ..... 8</b>
<b>4.2</b>	<b>Tagesordnung der wissenschaftlichen Erörterung ..... 10</b>
<b>4.3</b>	<b>Protokoll der wissenschaftlichen Erörterung ..... 11</b>
<b>Anhang:</b>	<b>Dokumentation der Stellungnahmen..... 52</b>

## **1    Dokumentation des Stellungnahmeprozesses**

Am 01.12.2008 wurde der Vorbericht in der Version Nr. 1.0 vom 19.11.2008 veröffentlicht und zur Anhörung gestellt. Bis zum 09.01.2009 konnten schriftliche Stellungnahmen eingereicht werden. Insgesamt wurden 9 Stellungnahmen form- und fristgerecht abgegeben. Diese Stellungnahmen sind im Anhang abgebildet.

Unklare Aspekte in den schriftlichen Stellungnahmen wurden in einer wissenschaftlichen Erörterung am 10.03.2009 im IQWiG diskutiert. Das Wortprotokoll der Erörterung befindet sich in Kapitel 4.

Eine Würdigung der in der Anhörung vorgebrachten Aspekte befindet sich im Kapitel „Diskussion“ des Abschlussberichts. Im Abschlussbericht sind darüber hinaus Änderungen, die sich durch die Anhörung ergeben haben, zusammenfassend dargestellt. Der Abschlussbericht ist auf der Website des IQWiG unter [www.iqwig.de](http://www.iqwig.de) veröffentlicht.

## **2    Würdigung der Stellungnahmen**

Die einzelnen Stellungnahmen wurden hinsichtlich valider wissenschaftlicher Argumente für eine Änderung des Vorberichts überprüft. Die wesentlichen Argumente wurden im Kapitel „Diskussion“ des Abschlussberichts gewürdigt. Neben projektspezifischen wissenschaftlichen Aspekten wurden auch übergeordnete Punkte, z. B. zur allgemeinen Institutsmethodik, zur Aufgabe des Instituts, zu rechtlichen Vorgaben für das Institut etc., angesprochen. Auf diese letztgenannten Punkte wurde im Rahmen der projektspezifischen Würdigung der Stellungnahmen nicht weiter eingegangen.

Der Abschlussbericht wurde auf der Website des IQWiG unter [www.iqwig.de](http://www.iqwig.de) veröffentlicht.

### 3 Darlegung potenzieller Interessenkonflikte der Stellungnehmenden

Im Folgenden sind die potenziellen Interessenkonflikte der Stellungnehmenden zusammenfassend dargestellt. Alle Informationen beruhen auf Selbstangabe der einzelnen Personen anhand des „Formblatts zur Darlegung potenzieller Interessenkonflikte“. Das Formblatt ist unter [www.iqwig.de](http://www.iqwig.de) abrufbar. Die in diesem Formblatt aufgeführten Fragen finden sich im Anschluss an diese Zusammenfassung.

#### 3.1 Interessenkonflikte von Stellungnehmenden aus Organisationen, Institutionen und Firmen

Organisation / Institution / Firma	Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6	Frage 7	Frage 8
Arzneimittelkommission der deutschen Ärzteschaft	Bschor, Tom, PD Dr. in Vertretung für Gundert-Remy, Ursula, Prof. Dr.	nein	nein	ja	nein	nein	nein	nein	nein
Arzneimittelkommission der deutschen Ärzteschaft	Gundert-Remy, Ursula, Prof. Dr.	nein	nein	nein	nein	nein	nein	nein	nein
Institut für Arzneimittelsicherheit in der Psychiatrie e. V.	Rüther, Eckart, Prof. Dr.	nein	ja	ja	ja	nein	nein	nein	ja
Boehringer Ingelheim Pharma GmbH & Co. KG	Cairns, Victoria, Dr.	ja	nein	nein	nein	nein	nein	nein	nein
Boehringer Ingelheim Pharma GmbH & Co. KG	Weigmann, Harald, Dr. Dr.	ja	nein	nein	nein	nein	nein	nein	nein
Boehringer Ingelheim Pharma GmbH & Co. KG	Weiser, Martin, Dr.	ja	nein	nein	nein	nein	nein	nein	nein

<b>Organisation / Institution / Firma</b>	<b>Name</b>	<b>Frage 1</b>	<b>Frage 2</b>	<b>Frage 3</b>	<b>Frage 4</b>	<b>Frage 5</b>	<b>Frage 6</b>	<b>Frage 7</b>	<b>Frage 8</b>
Deutsche Gesellschaft für Psychiatrie, Psychotherapie und Nervenheilkunde, Arbeitsgemeinschaft für Neuropsychopharmakologie und Pharmakopsychiatrie	Fritze, Jürgen, Prof. Dr.	nein	ja	ja	nein	nein	nein	nein	ja
Lundbeck GmbH	Flürenbrock, Wibke, Dr.	ja	nein	nein	nein	nein	nein	nein	nein
Lundbeck GmbH	Friede, Michael, Dr.	ja	nein	nein	nein	nein	nein	nein	nein
Lundbeck GmbH	Kessel-Steffen, Markus	ja	ja	nein	nein	ja	nein	ja	nein
Lundbeck GmbH	Pedersen, Anders Gersel, Dr. in Vertretung für Flürenbrock, Wibke, Dr.	ja	ja	nein	nein	nein	ja	nein	nein
Verband Forschender Arzneimittelhersteller e. V.	Rössel, Antje	ja	ja	nein	nein	nein	nein	nein	nein
Verband Forschender Arzneimittelhersteller e. V.	Thole, Henning	ja	ja	ja	nein	nein	nein	ja	ja
Verband Forschender Arzneimittelhersteller e. V.	Wahler, Steffen, Dr.	ja	ja	ja	ja	ja	ja	ja	nein
Wyeth Pharma GmbH	Löschmann, Peter- Andreas, PD Dr.	ja	nein	nein	nein	nein	ja	nein	nein
Wyeth Pharma GmbH	Schmitt, Andreas, PD Dr.	ja	nein	nein	nein	nein	ja	nein	nein

<b>Organisation / Institution / Firma</b>	<b>Name</b>	<b>Frage 1</b>	<b>Frage 2</b>	<b>Frage 3</b>	<b>Frage 4</b>	<b>Frage 5</b>	<b>Frage 6</b>	<b>Frage 7</b>	<b>Frage 8</b>
Wyeth Pharma GmbH	Volmer, Timm, Dr. MPH	ja	nein	nein	nein	nein	ja	nein	nein
Lilly Deutschland GmbH	Hundemer, Hans-Peter, Dr. in Vertretung für Weiser, Martin, Dr.; Boehringer Ingelheim Pharma GmbH	ja	ja	nein	nein	nein	ja	ja	nein

### 3.2 Interessenkonflikte von stellungnehmenden Privatpersonen

<b>Name</b>	<b>Frage 1</b>	<b>Frage 2</b>	<b>Frage 3</b>	<b>Frage 4</b>	<b>Frage 5</b>	<b>Frage 6</b>	<b>Frage 7</b>	<b>Frage 8</b>
Roots, Ivar, Prof. Dr.	nein	ja	ja	nein	nein	nein	nein	nein
Volz, Hans-Peter, Prof. Dr.	nein	ja	ja	nein	ja	nein	nein	nein

### 3.3 Interessenkonflikte von weiteren Teilnehmern der wissenschaftlichen Erörterung (externe Sachverständige, externe Reviewer)

<b>Name</b>	<b>Frage 1</b>	<b>Frage 2</b>	<b>Frage 3</b>	<b>Frage 4</b>	<b>Frage 5</b>	<b>Frage 6</b>	<b>Frage 7</b>	<b>Frage 8</b>
Becker, Thomas, Prof. Dr.	nein	nein	ja	ja	nein	nein	nein	nein
Kösters, Markus	nein	nein	nein	nein	nein	nein	nein	nein
Weinmann, Stefan, Dr. Dr.	nein	nein	ja	nein	nein	nein	nein	nein
Wille, Hans, Dr.	nein	ja	ja	ja	nein	nein	nein	nein



Im „Formblatt zur Darlegung potenzieller Interessenkonflikte“ wurden folgende 8 Fragen gestellt:

*Frage 1:* Sind oder waren Sie innerhalb der letzten 3 Jahre bei einer Person, Institution oder Firma<sup>1</sup> abhängig (angestellt) beschäftigt, die von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut<sup>2</sup> finanziell profitieren könnte?

*Frage 2:* Beraten Sie oder haben Sie innerhalb der letzten 3 Jahre eine Person, Institution oder Firma direkt oder indirekt<sup>3</sup> beraten, die von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut finanziell profitieren könnte?

*Frage 3:* Haben Sie abseits einer Anstellung oder Beratungstätigkeit innerhalb der letzten 3 Jahre im Auftrag einer Person, Institution oder Firma, welche von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut finanziell profitieren könnte, Honorare für Vorträge, Stellungnahmen, Ausrichtung und / oder Teilnahme an Kongressen und Seminaren – auch im Rahmen von Fortbildungsveranstaltungen – oder für (populär-) wissenschaftliche oder sonstige Aussagen oder Artikel erhalten?<sup>4</sup>

*Frage 4:* Haben Sie und/oder die Einrichtung, für die Sie tätig sind, abseits einer Anstellung oder Beratungstätigkeit innerhalb der letzten 3 Jahre von einer Person, Institution oder Firma, welche von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut finanziell profitieren könnte, finanzielle Unterstützung für Forschungsaktivitäten, andere wissenschaftliche Leistungen oder Patentanmeldungen erhalten?

*Frage 5:* Haben Sie und / oder die Einrichtung, bei der Sie angestellt bzw. beschäftigt sind, innerhalb der letzten 3 Jahre sonstige finanzielle oder geldwerte Zuwendungen (z. B. Ausrüstung, Personal, Reisekostenunterstützung ohne wissenschaftliche Gegenleistung) von einer Person, Institution oder Firma erhalten, welche von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut finanziell profitieren könnte?

---

<sup>1</sup> Mit solchen „Personen, Institutionen oder Firmen“ sind im Folgenden alle Einrichtungen gemeint, die direkt oder indirekt einen finanziellen oder geldwerten Vorteil aus dem Ergebnis Ihrer wissenschaftlichen Arbeit für das Institut ziehen könnten. Hierzu gehören z. B. auch medizinische Einrichtungen, die eine zu bewertende medizinische Intervention durchführen und hierdurch Einkünfte erhalten.

<sup>2</sup> Mit „wissenschaftlicher Arbeit für das Institut“ sind im Folgenden alle von Ihnen für das Institut erbrachten oder zu erbringenden Leistungen und / oder an das Institut gerichteten mündlichen und schriftlichen Recherchen, Bewertungen, Berichte und Stellungnahmen gemeint.

<sup>3</sup> „Indirekt“ heißt in diesem Zusammenhang z. B. im Auftrag eines Institutes, das wiederum für eine entsprechende Person, Institution oder Firma tätig wird.

<sup>4</sup> Sofern Sie von einer Person, Institution oder Firma im Verlauf der letzten 3 Jahre mehrfach Honorare erhalten haben reicht es aus, diese für die jeweilige Art der Tätigkeit summiert anzugeben.

*Frage 6:* Besitzen Sie Aktien, Optionsscheine oder sonstige Geschäftsanteile (auch in Fonds) von einer Firma oder Institution, welche von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut finanziell profitieren könnte?

*Frage 7:* Haben Sie persönliche Beziehungen zu einer Person, Firma oder Institution bzw. Mitarbeitern einer Firma oder Institution, welche von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut finanziell profitieren könnte?

*Frage 8:* Gibt es andere bislang nicht dargestellte potenzielle Interessenkonflikte, die in Beziehung zu Ihrer wissenschaftlichen Arbeit für das Institut stehen könnten?

#### 4 Dokumentation der wissenschaftlichen Erörterung – Teilnehmerliste, Tagesordnung und Protokoll

##### 4.1 Teilnehmerliste der wissenschaftlichen Erörterung

Name	Organisation / Institution / Firma
Becker, Thomas, Prof. Dr.	Externer Sachverständiger
Bschor, Tom, PD Dr. in Vertretung für Gundert-Remy, Ursula, Prof. Dr.	Arzneimittelkommission der deutschen Ärzteschaft
Cairns, Victoria, Dr.	Boehringer Ingelheim Pharma GmbH & Co. KG
Eyding, Dirk, Dr.	IQWiG
Friede, Michael, Dr.	Lundbeck GmbH
Grouven, Ulrich, PD Dr.	IQWiG
Hildebrandt, Mandy	IQWiG
Hundemer, Hans-Peter, Dr. in Vertretung für Weiser, Martin, Dr.; Boehringer Ingelheim Pharma GmbH	Lilly Deutschland GmbH
Kaiser, Thomas, Dr.	IQWiG
Kerekes, M. Florina	IQWiG
Kessel-Steffen, Markus	Lundbeck GmbH
Kösters, Markus	Externer Sachverständiger
Löschmann, Peter-Andreas, PD Dr.	Wyeth Pharma GmbH
Pedersen, Anders Gersel, Dr. in Vertretung für Flürenbrock, Wibke, Dr.	Lundbeck GmbH
Roots, Ivar, Prof. Dr.	privat (Charité Berlin, Institut für Klinische Pharmakologie und Toxikologie)
Rössel, Antje	Verband Forschender Arzneimittelhersteller e. V.
Rüther, Eckart, Prof. Dr.	Institut für Arzneimittelsicherheit in der Psychiatrie e. V.
Sawicki, Peter T., Prof. Dr.	IQWiG (Moderation)
Schmitt, Andreas, PD Dr.	Wyeth Pharma GmbH
Schüler, Yvonne-Beatrice, Dr.	IQWiG
Schürmann, Christoph, Dr.	IQWiG

<b>Name</b>	<b>Organisation / Institution / Firma</b>
Thole, Henning (angemeldet, aber nicht erschienen)	Verband Forschender Arzneimittelhersteller e. V.
Weigmann, Harald, Dr. Dr.	Boehringer Ingelheim Pharma GmbH & Co. KG
Weinmann, Stefan, Dr. Dr.	Externer Sachverständiger
Welter, Stefan	Protokollant
Wieseler, Beate, Dr.	IQWiG
Wille, Hans, Dr.	Externer Reviewer

#### 4.2 Tagesordnung der wissenschaftlichen Erörterung

	Begrüßung und Einleitung
<b>TOP 1</b>	Bewertung der Relevanz von Effekten
<b>TOP 2</b>	Stellenwert gepoolter Analysen für die Nutzenbewertung
<b>TOP 3</b>	Methodik zur Beurteilung der Begleitsymptomatik Schmerz
<b>TOP 4</b>	Bewertung von Schadensaspekten
<b>TOP 5</b>	Depressionsskalen-Validierung, Beobachtungszeitraum
<b>TOP 6</b>	Zulassungsstatus von Duloxetin
<b>TOP 7</b>	Validierung des Composite Cognitive Score
<b>TOP 8</b>	Verständnisfrage zum Punkt: Ausschluss von Studien aufgrund der Studiendauer
<b>TOP 9</b>	Verschiedenes / Verabschiedung

### 4.3 Protokoll der wissenschaftlichen Erörterung

- Datum: 10.03.2009, 11:00 bis 12:40 Uhr
- Ort: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG),  
Dillenburger Straße 27, 51105 Köln
- Moderation: Prof. Dr. Peter T. Sawicki

**Moderator Peter T. Sawicki:** Meine Damen und Herren, liebe Kollegen, ich darf Sie ganz herzlich zur Erörterung der Stellungnahmen zum Vorbericht A05-20A begrüßen. Es geht um selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer bei Patienten mit Depressionen.

Ich weise Sie darauf hin, dass wir diese Veranstaltung sowohl stenografisch als auch elektronisch aufzeichnen und daraus ein Wortprotokoll erstellen, das wir zusammen mit dem Abschlussbericht publizieren. Die Sprache ist Deutsch. Sofern jemand nicht damit einverstanden ist, kann er an dieser Veranstaltung nicht teilnehmen.

Für das Gespräch haben wir drei Stunden eingeplant und keine Pausen vorgesehen. Bitte denken Sie daran, ganz kurz Ihren Namen zu sagen, wenn Sie das Wort haben. Denn aufgrund der Größe der Gruppe ist es sonst manchmal schwierig, im Nachhinein eine Aussage der entsprechenden Person zweifelsfrei zuzuordnen, sodass keine Fehler passieren. Haben Sie zum Ablauf noch konkrete Fragen? – Das ist nicht der Fall.

#### **Tagesordnungspunkt 1: Bewertung der Relevanz von Effekten**

**Yvonne-Beatrice Schüler:** Ich möchte einen kurzen Überblick über das Thema geben, bevor wir zu unserer Frage kommen. Bei der Bewertung von Antidepressiva werden zur Erhebung von Behandlungseffekten Fragebögen und verschiedene Skalen eingesetzt. Daraus ergibt sich das Problem der Interpretation der Ergebnisse. Es stellt sich die Frage, welche Änderung der Skalenwerte bei einem einzelnen Patienten als relevanter Effekt angesehen wird und welcher Unterschied zwischen zwei Behandlungsalternativen als relevant gelten kann.

In der internationalen Literatur findet sich hierzu die Empfehlung der sogenannten Minimal Important Difference (MID), mit der nach ihrer Bestimmung eine solche Interpretation möglich ist. Für die in den Studien des Berichts verwendeten Skalen lagen solche MID's jedoch nicht vor. Deswegen haben wir für die Relevanzbewertung von Skalen auf ein ebenfalls international häufig verwendetes Maß zurückgegriffen, nämlich das Cohen's d.

In einigen Stellungnahmen wurde der Umgang mit diesem Cohen's d und der nachfolgenden Relevanzbewertung von Skalen kritisiert. Deshalb möchten wir Sie fragen, was Sie als genaue Alternative zum Cohen's d anbieten und wie Sie die Ergebnisse einer etwaigen solchen Alternative genau interpretieren und bewerten würden. Das würde ich gerne zur Diskussion stellen.

**Moderator Peter T. Sawicki:** Herr Friede.

**Michael Friede:** Die klinische Relevanz, wie Sie es gesagt haben, kann man anhand der individuellen Ebene bewerten, nämlich nach der Response oder der Remission, die Sie bisher außer Acht gelassen haben. Man kann sie anschließend noch diskutieren. Auf der Gruppenebene haben Sie ausgeführt, dass die MID für die verwendeten Skalen nicht definiert wäre. Für die MADRS gibt es eine solche Definition, nämlich die Arbeit von Duru, der zu einer MID von 1,6 kommt. Insofern wäre das eine Möglichkeit, die Relevanz zu bemessen. Man muss dann nicht auf Cohen's d ausweichen, sondern könnte es supportiv verwenden.

**Moderator Peter T. Sawicki:** Frau Wieseler.

**Beate Wieseler:** Ich habe dazu zwei Fragen. Handelt es sich bei dem Wert, der für MADRS angegeben wird, um eine MID für die intraindividuelle Änderung oder für die Gruppenänderung?

**Moderator Peter T. Sawicki:** Herr Friede.

**Michael Friede:** Für die Gruppenänderung.

**Moderator Peter T. Sawicki:** Frau Wieseler.

**Beate Wieseler:** Auf welcher Datenbasis wurde sie festgelegt?

**Moderator Peter T. Sawicki:** Herr Friede.

**Michael Friede:** Das waren über 100 RCTs. Das müsste ich noch einmal nachschauen, oder wir lassen Ihnen die Information zukommen.

**Moderator Peter T. Sawicki:** Frau Wieseler.

**Beate Wieseler:** Haben Sie die Arbeit in Ihrer Stellungnahme zitiert?

(Michael Friede: Ja!)

Dann schauen wir uns das noch einmal an.

Meine zweite Frage zu Ihrer Anmerkung ist folgende: Wir reden nicht nur über die Depressionsskalen, bei denen es in der Tat so ist, dass wir neben der Auswertung der

kontinuierlichen Daten zusätzlich Responderauswertungen aus zwei Responsekriterien haben, nämlich der Remission und der Response, mit in der Regel 50 % Verbesserung. Hier liegt gar nicht unser primäres Problem. Unser Problem liegt in der Bewertung der Begleitsymptomatik, bei der wir auch mit Skalen zur Lebensqualität, zum Schmerz und zur Angst arbeiten. Für diese Skalen sind uns keine MIDs bekannt. Können Sie auch zu diesen Skalen MIDs benennen?

**Moderator Peter T. Sawicki:** Herr Friede.

**Michael Friede:** Nein, nur zu MADRS, bei der es eben diese aktuelle Publikation gibt.

**Moderator Peter T. Sawicki:** Danke schön. – Herr Wille.

**Hans Wille:** Ich möchte zunächst einmal meine Anerkennung für den schönen Bericht aussprechen. Er ist ausgesprochen akribisch und im Detail durchgeführt. Davon war ich beeindruckt.

**Moderator Peter T. Sawicki:** Danke schön.

**Hans Wille:** Ich habe ein Problem mit dem Ausdruck „klinische Relevanz“ für die Größe Cohen's d. Für mich ist es erst einmal nur eine statistische Maßzahl. Für mich wäre es zum Beispiel ein Unterschied, ein Cohen's d von 0,2 bezüglich der Mortalität zu beurteilen, im Vergleich zu einer weniger patientenrelevanten Zielgröße. Ich kann gut verstehen, dass man ein Mittel gesucht hat, mit dem man ein einheitliches Verfahren für alle Zielparameter hat, um die Effektstärken darzustellen. Vielleicht sollte man aber zusätzlich zum Cohen's d jeweils eine Spannbreite der Scores, die man verwendet, und vielleicht auch gleichzeitig die Differenzen der Gruppen darstellen, was zum Teil geschehen ist.

**Moderator Peter T. Sawicki:** Danke schön. – Herr Rüter.

**Eckart Rüter:** Ich möchte es von einer ganz anderen Warte betrachten. Wir haben wirklich versucht, in vielen Gremien und auf verschiedenen Konferenzen und in wissenschaftlichen Gruppen zu diskutieren, was klinische Relevanz in der Psychiatrie anhand von Skalen ist. Ich muss von vornherein gestehen, dass es auch in der Psychiatrie keinen Konsens gibt – nicht, dass Sie sich wundern, dass wir vielleicht diskutieren und nicht wissen, was wir eigentlich wollen.

Ich habe mit der EMEA und der FDA ausführlich diskutiert. Es gibt viele Meinungen, die unterschiedlich interpretiert werden können, zum Beispiel die Frage: Ist eine Hamilton-Differenz von 1, von 5 oder von 10 klinisch relevant? Ist bei der CGI eine Relevanz bei 0,5 oder bei 1 gegeben? Das ist unheimlich schwierig. Je nach Studie, je nach Patientengut und je nach Ausgangswert ist das zu differenzieren. Aus diesem Grunde haben wir eine Notlösung gefunden und gesagt: Lasst uns doch auf NNT zurückgehen. Wenn wir dazu etwas hätten und das anhand der von Ihnen erhobenen Statistik interpretieren könnten, wäre das eine



Möglichkeit, um einen Gesichtspunkt zu haben: Wo fängt es denn überhaupt an, für uns im eigentlichen Sinne klinisch relevant zu sein?

**Moderator Peter T. Sawicki:** Darf ich eine Verständnisfrage stellen: NNT bezüglich welcher Größe?

**Eckart Rütter:** Der Größe, die Sie dann erheben, je nachdem wie viel Differenz Sie zwischen den einzelnen Gruppen gefunden haben. Oder ist das falsch zu verstehen?

**Moderator Peter T. Sawicki:** Das habe ich nicht verstanden. – Herr Weinmann.

**Stefan Weinmann:** Ich möchte noch einmal etwas zur Informierung sagen: Diese Studie von Duru ist ein Versuch, einen minimal klinisch wichtigen Unterschied festzulegen. Es waren drei RCTs, die alle mit Escitalopram durchgeführt worden sind. Man hat die Placebogruppen ausgewählt und zwischen 0 und Woche 1 geschaut, welche Placebopatienten von den Ärzten als nicht verändert geratet werden. Dann hat man versucht, den Standard Error of Measurement als Proxy für diese Minimal Clinical Important Difference herauszufinden. Er berechnet sich aus der Varianz und der Wurzel aus 1 minus Reliabilität.

Das heißt, man hat versucht, sich dieser MCID durch die Reliabilität der Skala anzunähern. Man hat geschaut, wie unterschiedlich die Ärzte bei Patienten raten, die sich eigentlich nicht verändert haben. Dann kam man zu diesem Wert von 1,6 bis 1,9. Man hat es mit einer Beobachtungsstudie kreuzvalidiert: Patienten, bei denen es ein Delta beim antidepressiven Effekt versus Placebo ohne diese MID gibt gegenüber Patienten, bei denen es ein Delta mit MID gibt. Dann hat man gematcht, ob sich diese zwei Kohorten unterscheiden. Das war eben ein auf Verteilung und Reliabilität von Skalen basierender Versuch, einen minimalen klinisch wichtigen Unterschied festzulegen, der nicht auf Patientenbefragung beruht. Das ist aber auch nicht unstrittig.

**Moderator Peter T. Sawicki:** Aber das würde doch nur dann funktionieren, wenn sich Patienten unter Placebo nicht verändern.

**Stefan Weinmann:** Genau, in der ersten Woche.

**Moderator Peter T. Sawicki:** Aber das tun sie doch.

**Stefan Weinmann:** Es basierte auf dem CGI. Der Arzt hat bei Punkt 0 eine CGI-Einschätzung gegeben und bei Punkt 1, also nach einer Woche. Dann hat man die Patienten in der Placebogruppe herausgenommen, die sich nach der Einschätzung des Arztes nicht verändert hatten. Man hat geschaut, wie unterschiedlich reliabel die Skala ist. Das war der Bereich des Standard Error of Measurement. Es gab also keine riesige Studie und keinen Einbezug von Patientenurteilen.

**Moderator Peter T. Sawicki:** Frau Wieseler.

**Beate Wieseler:** Ich möchte einen Punkt von Herrn Weinmann aufgreifen. Internationaler Stand der Diskussion bei der Frage, wie man diese Daten interpretiert, ist, dass es im Grunde genommen zwei Richtungen von Vorschlägen gibt: zum einen die verteilungsbasierten Verfahren wie die Effect-Size und Cohen's d, auf die wir zurückgegriffen haben, weil wir keine anderen Daten haben. Zum anderen gibt es den Strang der sogenannten ankerbasierten Verfahren, bei denen man versucht, die Änderungen, die man auf den Skalen sieht, in Relation zu anderen Ereignissen oder Änderungen zu setzen, die man interpretieren kann. Mit den ankerbasierten Verfahren kann man diese sogenannten MIDs empirisch ermitteln.

Damit möchte ich auf Herrn Rüther zurückkommen: Es geht weniger darum, sich in einem Konsensverfahren darauf zu einigen, was ein klinisch relevanter Unterschied auf der Skala ist. Vielmehr geht es darum, unter Einbezug von in der Regel Patientenurteilen empirisch zu ermitteln, welche Unterschiede auf dieser Skala für die Patienten spürbar sind.

Es ist richtig, was Herr Wille gesagt hat: Wir sollten nicht von klinischer Relevanz sprechen. Das haben wir im Bericht auch ganz bewusst vermieden. Dort sprechen wir von der Relevanz des Effekts. Im Bericht sind Meta-Analysen enthalten und Tausende von Patienten eingegangen. Aufgrund dieser Situation werden sehr kleine Unterschiede statistisch signifikant. Für uns stellt sich die Frage, ob dieser Unterschied im Sinne einer Spürbarkeit für den Patienten relevant ist. Wir sollten uns also in der Tat von dem Begriff „klinisch relevant“ verabschieden, weil dann noch die Aspekte hinzukommen, die Sie genannt haben: Es hängt davon ab, welchen Parameter man betrachtet.

Es geht uns also darum, ein Verfahren zu entwickeln, mit dem wir darstellen können, dass ein Gruppenunterschied nicht nur statistisch signifikant, sondern im Sinne einer Spürbarkeit für den Patienten relevant ist. International diskutiert sind dabei Verfahren, bei denen diese Werte, die Minimal Important Differences, empirisch erhoben werden. Bei diesem Projekt besteht das Problem, dass die empirisch erhobenen Minimal Important Differences für die verwendeten Skalen nicht vorliegen. Deshalb weichen wir auf das verteilungsbasierte Verfahren zur Interpretation aus.

**Moderator Peter T. Sawicki:** Danke schön. – Herr Schmitt, bitte.

**Andreas Schmitt:** Ich möchte direkt daran anknüpfen. Wenn es sehr heterogen und kontrovers diskutiert wird, wie man mit welchen Methoden die klinische Relevanz belegen kann, und der Cohen's d sicherlich kein geeignetes Maß ist, um daran klinische Relevanz festzumachen, stellt sich die Frage: Warum entwickelt man das nicht in einem Scopingprozess für die jeweilige Indikation, um sich beispielsweise den Vor- und Nachteilen der jeweiligen Methoden anzunähern?

Grundsätzlich muss man in diesem Kontext berücksichtigen: Gerade bei Mittelwertveränderungen von Gruppen werden wir immer Schwierigkeiten haben, die klinische Relevanz abzuleiten. Solche Diskussionen sind in den vergangenen zwei Jahren sehr

intensiv geführt worden – getriggert auch durch die eine oder andere Publikation wie etwa von Kirsch. Das hat aber beispielsweise die CHMP dazu motiviert, dem Ganzen zu begegnen und zu sagen, dass Mittelwertveränderungen kein geeignetes Maß sind, gerade bei Antidepressiva eine klinische Relevanz zu bewerten. Kriterien wie Response, Responder-Rates und Remissions-Rates müssen mit eingebracht werden. Da stellt sich tatsächlich die Frage, wie man sich diesem Punkt nähern kann. Absolute Messgrößen wie ARR oder NNTs sind sicherlich eine etablierte Größe. Es gibt aber noch keinen Konsens, wo beispielsweise der Cut-off für ARR oder NNT liegt.

**Moderator Peter T. Sawicki:** Kurz direkt dazu, Frau Wieseler?

**Beate Wieseler:** Ich möchte zunächst zum Scopingprozess kommen, der in einigen Stellungnahmen vorgeschlagen worden ist. Sie haben heute bereits zum zweiten Mal Gelegenheit, zu dieser Frage Stellung zu nehmen. Es gab ein Stellungnahmeverfahren zum Berichtsplan, bei dem Sie aufgerufen waren, zur Methodik des Berichts Stellung zu nehmen. Dazu gehört natürlich auch die Interpretation der in diesem Projekt relevanten Daten. In diesem Stellungnahmeprozess wurden keine substanziellen Anmerkungen zu diesem Thema gemacht.

Mit dem Vorbericht haben Sie unseren Umgang mit diesem Problem und unsere Interpretation der vorliegenden Daten vor sich liegen. Sie können auch jetzt wieder Stellung nehmen und hier und heute Ihre Ideen vorbringen. Ich sehe also nicht, wieso wir von einem dritten Beteiligungsverfahren ein anderes Ergebnis erwarten sollten als von den zwei Beteiligungsprozessen, die wir in diesem Projekt schon hatten.

Ich gebe Ihnen recht, dass eine Analyse von Mittelwertsunterschieden bei der Interpretation dieser Daten schwierig ist. Der Vorschlag lautet deshalb auch international etwa durch die FDA, auf Basis dieser empirisch zu erhebenden MIDs Responderanalysen durchzuführen, sich also gar nicht dem Problem zu stellen, das auch noch niemand gelöst hat, nämlich eine MID für einen Gruppenunterschied zu ermitteln, sondern tatsächlich MIDs für intraindividuelle Änderungen zu ermitteln und damit eine Responderanalyse durchzuführen. Das ist sicherlich ein machbarer und gut zu interpretierender Weg.

Die Cohen's ds zu NNTs weiterzuentwickeln – das hatte auch Herr Rüther gemeint –, ist in einigen Stellungnahmen vorgeschlagen worden. Dazu kann Herr Grouven etwas sagen.

**Moderator Peter T. Sawicki:** Herr Grouven.

**Ulrich Grouven:** Das Problem in unserem Zusammenhang ist, dass wir diese NNTs eben nicht direkt berechnen können, weil wir keine Cut-Points oder MIDs haben. Es gibt einen Ansatz, der auch in einer Stellungnahme angesprochen worden ist, nämlich die NNTs über das Cohen's d herzuleiten. Das ist ein mehrschrittiger Prozess. Der erste Schritt wäre, aus dem

Cohen's d ein Odds Ratio zu berechnen. Das lässt sich machen. Allerdings ist das auch wiederum nur eine approximative Formel, hinter der Verteilungsannahmen stecken.

Dann kommt der eigentlich kritische Schritt, aus dem Odds Ratio wiederum eine Risikodifferenz und daraus eben die NNT zu berechnen. Leider müssen an dieser Stelle zusätzliche Annahmen in den ganzen Prozess eingehen, nämlich die Annahme einer Responderrate für die Kontrollgruppe. Das Problem ist, wie man an diese Responderrate kommt. Wie zuverlässig, wie sicher kann man diese Annahme treffen? Es gibt also noch eine zusätzliche Unsicherheit in dem Prozess.

Am Ende des ganzen Prozesses steht das, was Herr Schmitt gerade anspricht: Selbst wenn wir eine NNT ermittelt haben, müssen wir uns natürlich Gedanken machen, wie die Relevanzschwelle auf der NNT-Skala ist. Der Prozess der Umrechnung auf eine NNT schafft in der Situation, die wir hier vorliegen haben, eigentlich keine größere Klarheit, sondern nach unserer Einschätzung eher noch mehr Unsicherheit.

**Moderator Peter T. Sawicki:** Danke schön. – Sie wollen direkt dazu etwas sagen, Herr Schmitt?

**Andreas Schmitt:** Ich kann das gerade für die Mittelwertveränderung und das entsprechende Ergebnis nachvollziehen. Ich kann es aber noch nicht ganz für die definierten sekundären Parameter Response und Remission nachvollziehen, bei denen man Absolutwerte betrachten kann:

(Zustimmung von Victoria Cairns)

Wie viele Prozent sind es unter Verum versus Placebo? Wie sind die Unterschiede? Dann habe ich doch direkte Differenzen und Absolutgrößen. Mit dem Ergebnis, wie viele Prozent sich verändert haben, kann auch der Kliniker deutlich besser umgehen.

**Moderator Peter T. Sawicki:** Herr Grouven.

**Ulrich Grouven:** Klar, wenn wir Responderanalysen oder die entsprechenden Angaben haben – das hat Frau Wieseler gerade schon gesagt –, wäre das unsere primär erwünschte Analyse. Dann würden wir die Responder direkt auswerten und hätten klare Aussagen. Aber wir sprechen primär über den Fall, in dem wir nur kontinuierliche Skalen und eben nicht die Cut-Points und die Responderdefinition, sondern eben tatsächlich nur den Mittelwertunterschied zur Verfügung haben. Dann kommen wir in diese Interpretationsprobleme.

**Moderator Peter T. Sawicki:** Herr Schmitt.

**Andreas Schmitt:** Ich verstehe es immer noch nicht ganz. Denn in den Publikationen und den Daten sind die entsprechenden Absolutwerte für Response und Remission eingetragen.

Man weiß dann, wie viel Prozent 50 % Reduktionen auf Hamilton oder MADRS aufweisen oder wie viele Patienten beispielsweise eine Remission mit einem HAMD von 7 und kleiner erreicht haben.

**Moderator Peter T. Sawicki:** Ich habe dann noch eine ganze Reihe von Wortmeldungen. Deswegen müssen wir die Rednerliste zu diesem Punkt nun schließen. Sonst kommen wir nicht durch die anderen acht Tagesordnungspunkte. Aber wir müssen diesen Punkt klären; das ist ganz wichtig. – Frau Wieseler.

**Beate Wieseler:** Ich glaube, hier liegt ein Missverständnis vor. Bei den Depressionsskalen haben wir mit Remission und Response im Grunde genommen Responderanalysen und zusätzlich die kontinuierlichen Daten. Bei den Responderanalysen setzen wir keine Relevanzschwelle. Hier sehen wir einen statistisch signifikanten Unterschied als relevant an, weil im Responsekriterium an sich schon eine Relevanzschwelle enthalten ist. Das ist bei der Betrachtung des Übergangs von Cohen's d zu NNT außen vor.

Das Cohen's d berechnen wir ausschließlich in den Fällen, in denen wir kontinuierliche Daten haben. In den Stellungnahmen ist der Vorschlag gemacht worden, das Cohen's d in einem weiteren Schritt in eine NNT umzurechnen. Nur zu dieser Fragestellung hat sich Herr Grouven geäußert. Dieses Verfahren halten wir in der Tat für nicht hilfreich, weil wir denken, dass wir eher eine Unsicherheit als weitere Informationen und Sicherheit hineinbringen.

(Eckart Rüther: Schwierig!)

Aber die Responderanalysen an sich lassen wir so stehen und bewerten sie als solche.

**Moderator Peter T. Sawicki:** Gut, dann fahren wir jetzt mit der Rednerliste fort, die nun allerdings geschlossen ist. – Herr Friede.

**Michael Friede:** Ich möchte eine Anmerkung zu Herrn Weinmann machen. Sie haben recht, dass die Publikation von Duru ein Versuch ist, eine MID zu finden, der evidenzbasiert ist. Wenn Sie auf der anderen Seite großen Wert auf Cohen's d legen, frage ich mich: Wo ist da die Evidenz für deren Schwellen und Cut-offs, die Sie verwenden?

Mein zweiter Punkt ist: Wenn Sie die klinische Relevanz bewerten, bewerten Sie in Ihrem Bericht Gruppen. Dazu ist anzuführen, dass es sich dabei um inhomogene Gruppen handelt, weil sie zum Beispiel die Gesamtheit der SSRIs mit anderen Substanzen vergleichen wie den SNRIs Duloxetine und Venlafaxin. Diese Gruppe ist klinisch und pharmakologisch nicht homogen. Insofern sollte man diesen Versuch nicht machen, sondern einzelne SSRIs untersuchen, um den pharmakologischen und klinischen Besonderheiten Rechnung zu tragen. Daher lautet sicherlich auch eine Frage an das IQWiG zum Bericht zu den SSRIs, ob und wann er durchgeführt wird.

**Moderator Peter T. Sawicki:** Danke schön. – Herr Roots.

**Ivar Roots:** Ich wollte auf eine Limitierung aufmerksam machen, die beim Vergleich solcher Skalen sehr schnell auftreten kann. Wir werden damit konfrontiert, dass die einzelnen Kollektive unter Umständen eine sehr große Varianz aufweisen. Sie hat eine Ursache, die aus meiner Sicht hauptsächlich in der Genetik etwa des Metabolismus der einzelnen Psychopharmaka liegt, sodass es gewissermaßen Ausreißer gibt, die die Varianz natürlich vergrößern, wenn man sie nicht herausnimmt und in einer separaten Gruppe miteinander vergleicht. Man könnte also unter Umständen zu unscharfen, insignifikanten Ergebnissen kommen, wenn der Faktor der Pharmakogenetik nicht berücksichtigt wird.

Das gilt in zwei Situationen: zum einen dann, wenn man verschiedene Medikamente miteinander vergleicht, die diesem genetischen Polymorphismus in einem unterschiedlichen Ausmaß unterliegen. Zum anderen gilt es, wenn man verschiedene Patientengruppen miteinander in Beziehung bringen will, die ebenfalls Unterschiede bei diesen pharmakogenetischen Merkmalen haben. Eine Weiterentwicklung dieser Skalen sollte unbedingt die Pharmakogenetik einschließen. Gerade am Beispiel der SSRIs sieht man ja, wie unterschiedlich die einzelnen Pharmaka sind. Vielleicht kann man darüber noch einmal in einem anderen Tagesordnungspunkt diskutieren.

**Moderator Peter T. Sawicki:** Danke schön. – Frau Cairns.

**Victoria Cairns:** Cohen's d wird als etwas Spürbares vorgeschlagen. Das ist ein Weg, um zu sehen, was ein Patient entdecken kann. Aber spürbar heißt nicht unbedingt „klinisch relevant“. Die Kategorien von Cohen's d für kleine, mittlere und große Effekte sind nur auf dem Gebiet der Verhaltens- und Sozialwissenschaft beobachtet und beschrieben worden. Das hat keine solide Basis.

Genauso könnte man sagen: Wir beobachten in der Psychiatrie NNTs meistens um die 10. Das ist genauso grob geschätzt und eigentlich leichter zu interpretieren, weil man versteht, was eine NNT ist. Aber dann wäre es klarer. Es ist völlig willkürlich und eigentlich deutlich zu sehen: Eine NNT von 10 hat bei Mortalität eine ganz andere Bedeutung als eine NNT bei milden Kopfschmerzen, wenn es andere Nebenwirkungen gibt.

Für die Beurteilung der Relevanz ist eigentlich nicht nur irgendein statistischer Wert von 0,2 wichtig, bei dem man sagt: Über alle Medikamente aller Gebiete aller Patientensorten ist das unser einziger Punkt. – Eigentlich ist es viel komplizierter. Man muss den Gesamtkontext der Krankheit berücksichtigen, die verfügbare Behandlung und das gesamte Nutzenrisikoprofil berücksichtigen. Dabei gibt es nicht nur einen Wert, der überall passt. Auch die EMA schreibt ziemlich deutlich, dass die Effektgröße bei der Beurteilung der klinischen Relevanz eines Gruppenunterschiedes nicht hilft.

**Moderator Peter T. Sawicki:** Danke schön. – Herr Hundemer.

**Hans-Peter Hundemer:** Ich wollte im Prinzip einen vergleichbaren Kommentar abgeben. Man sollte noch zusätzlich bemerken: Wenn man Cohen's d nutzt, muss man die Effektstärke auf das zu untersuchende Kollektiv beziehen. Das ist in diesem Prozess nicht erfolgt. Man muss sich überlegen, ob für das Kollektiv, das man untersuchen will, eine Effektstärke von 0,2 klinisch relevant ist oder eine höhere oder niedrigere. Es gibt Beispiele dafür, dass sich das massiv unterscheiden kann. Das sollte man bedenken.

**Moderator Peter T. Sawicki:** Vielen Dank. – Herr Wille.

**Hans Wille:** Ich habe ein Problem mit der Frage, ob man nur mit den Responderraten weiterkommt. Die Remissionsraten nehme ich aus, weil es sich dabei um eine Sondersituation handelt und sie immer klinisch relevant sind. Wenn man zum Beispiel bei den Gruppen Mittelwerte hat, die relativ nah beieinanderliegen, und einen Cut-Point, der genau dazwischenliegt, kann eine extreme Überschätzung der eigentlichen Effekte auftreten. Insofern glaube ich, dass Responderraten alleine – für welchen Parameter auch immer – nicht reichen. Man muss gleichzeitig auch die Gruppendifferenzen berücksichtigen.

**Moderator Peter T. Sawicki:** Danke schön. – Herr Weigmann.

**Harald Weigmann:** Ich möchte gerne die Lanze für den Scopingworkshop brechen. Sie sagten, er sei nicht nötig. Im Bericht ...

**Moderator Peter T. Sawicki:** Das möchte ich aber nicht, denn das ist nicht der Rahmen dafür. Wir diskutieren nicht über grundsätzliche Verfahrensabläufe zwischen dem Gemeinsamen Bundesausschuss und dem IQWiG. Das müssen wir woanders machen und haben das auch getan. Wir diskutieren über den Vorbericht zu Antidepressiva.

**Harald Weigmann:** Dann möchte ich kurz aus dem Berichtsplan zitieren:

„(...) wird es für den vorliegenden Bericht als nicht sinnvoll erachtet, vorab entsprechende Grenzen zu definieren.“

Daher war uns vorab nicht klar, dass der Cohen's d zur Anwendung kommen wird. Deshalb erachte ich es trotz allem für sinnvoll, sich noch einmal mit Experten darüber abzustimmen, was ein klinisch relevanter oder von Patienten zu spürender Unterschied wäre.

**Moderator Peter T. Sawicki:** Deswegen haben wir das an prominenter Stelle in die Tagesordnung aufgenommen. Es hört mit dieser Diskussion nicht auf. Wir müssen die Erörterung erst auswerten, diskutieren und gegebenenfalls auf bestimmte Experten zurückkommen. Das machen wir sehr wohl, denn es ist natürlich eine wichtige Fragestellung. Danke schön. – Frau Rössel, bitte.

**Antje Rössel:** Wir diskutieren über die einzelnen Aspekte. Ich möchte nur zusammenfassen, dass es derzeit kein akzeptiertes Vorgehen zur Ermittlung der klinischen Relevanz gibt; das

haben Sie auch schon in Ihrem Vorbericht dargelegt. Für uns wäre es natürlich methodisch nachvollziehbarer gewesen, wenn Sie im Berichtsplan erläutert und in Ihrem Papier „Allgemeine Methoden“ Version 3.0, auf die Sie sich beziehen, abgewogen hätten, um was es beim Cohen's d in Abhängigkeit zu den anderen Methoden geht. Denn es gibt noch die ankerbasierten und die verteilungsbasierten Verfahren, die den Cohen's d beinhalten, wie schon Frau Wieseler erwähnt hat. Es gibt aber natürlich auch noch die expertenbasierten Verfahren. Ich halte es doch für wesentlich, das im Rahmen eines Scopingprozesses zu diskutieren.

**Moderator Peter T. Sawicki:** Danke schön. – Dann haben wir den Tagesordnungspunkt 1 abgeschlossen.



## **Tagesordnungspunkt 2: Stellenwert gepoolter Analysen für die Nutzenbewertung**

**Moderator Peter T. Sawicki:** Frau Schüler.

**Yvonne-Beatrice Schüler:** Einige Stellungnahmen nannten uns gepoolte Analysen, deren Ergebnisse in die Nutzenbewertung einfließen oder zumindest ergänzend betrachtet werden sollten. Bei der Sichtung einiger dieser Analysen stellte sich heraus, dass einige auf systematischen Recherchen basierten, andere aber nicht. Internationaler Standard der evidenzbasierten Medizin ist, dass Bewertungen beziehungsweise Meta-Analysen auf Basis einer fragestellungsbezogenen systematischen Recherche durchzuführen sind und gepoolten Analysen unsystematisch zusammengestellter Studien hierbei wegen des hohen Verzerrungspotenzials keine Relevanz zukommt.

Der IQWiG-Bericht selbst ist an sich eine Meta-Analyse auf Basis einer fragestellungsbezogenen systematischen Recherche. Er kann aber natürlich auch von Meta-Analysen gleicher Qualität ergänzt werden. Deswegen würden wir Sie gerne fragen, welche der von Ihnen genannten gepoolten Analysen aus Ihrer Sicht den Bericht adäquat ergänzen und gleichzeitig auf einer fragestellungsbezogenen systematischen Recherche beruhen.

**Moderator Peter T. Sawicki:** Gibt es dazu Wortmeldungen? – Bitte schön, Herr Bschor.

**Tom Bschor:** Ganz neu im Lancet ist ja die Meta-Analyse von Cipriani erschienen. Sie enthält eine systematische Suche. Sie vergleicht Wirksamkeit und Verträglichkeit anhand von Abbruchraten und von UAWs modernerer, neuerer Antidepressiva miteinander. Das käme unserer Fragestellung sicherlich sehr nahe.

**Moderator Peter T. Sawicki:** Frau Wieseler.

**Beate Wieseler:** Das ist richtig; die Arbeit liegt uns auch vor. Wir werden sie dahin gehend überprüfen, ob dort Studien genannt werden, die in unserer Meta-Analyse nicht aufgenommen worden sind. Ansonsten behandelt diese Meta-Analyse Parameter, die wir in unserem Bericht auch betrachten: Wenn ich mich richtig erinnere, geht es um Remission und um die Acceptability – das ist der Studienabbruch, wie Sie richtig erwähnen.

(Tom Bschor: Aufgrund von UAWs!)

Diese beiden Parameter werden auch in unserer Analyse betrachtet. Wir werden diese Meta-Analyse sicherlich auch im Zusammenhang mit unserer Arbeit im Abschlussbericht diskutieren.

**Moderator Peter T. Sawicki:** Danke schön. – Frau Cairns.

**Victoria Cairns:** Bei dieser Studie geht es um Response und Abbrüche wegen Adverse Events und nicht um Remission. Man muss aber ein bisschen aufpassen, weil es einige Schwächen gibt. Sie basiert auf direkten Vergleichen zwischen Präparaten innerhalb der Studie und auf indirekten Vergleichen über Studien. Das ist keine so starke Evidenz wie direkte Vergleiche innerhalb von Studien.

Ob es dann eine Kohärenz in dem ganzen System gibt, ist nicht ganz klar. Es sind viele verschiedene Studien und verschiedene Arten enthalten. Bei solchen Analysen ist es sehr wichtig, dass die Studien eigentlich identisch aufgebaut sind – außer bei der Behandlung. Hier sind placebokontrollierte Studien und nicht placebokontrollierte Studien enthalten, Studien bei schwer kranken Patienten im Krankenhaus und Patienten mit milden bis moderaten Depressionen außerhalb des Krankenhauses sowie Patienten aller Altersgruppen. Es ist sehr gemischt. Man muss bei der Beurteilung dieser Meta-Analyse ein bisschen aufpassen.

**Moderator Peter T. Sawicki:** Danke schön. – Herr Wille.

**Hans Wille:** Nur als Ergänzung: In diese Meta-Analyse sind auch offene Studien eingeflossen.

**Moderator Peter T. Sawicki:** Danke. – Herr Schmitt.

**Andreas Schmitt:** Zu Venlafaxin gibt es eine ganze Reihe von Meta-Analysen. Wir denken aber, dass es wichtig ist, zwei in diesem Kontext zu berücksichtigen: zum einen die Arbeit von Nemeroff mit dem Fokus auf Remission – der Vorteil hier besteht darin, dass sie auf Patient Level Data beruht – und zum anderen die im Januar erschienene Arbeit von Bauer et al., die sich mit dem Thema Venlafaxin auch im Vergleich zu anderen Antidepressiva – nicht nur SSRIs, sondern auch TZAs – sehr umfangreich beschäftigt.

**Moderator Peter T. Sawicki:** Vielen Dank. – Herr Weigmann.

**Harald Weigmann:** Für Duloxetin verfügen wir auch über eine Meta-Analyse im Sinne einer gepoolten Analyse, die sich der Fragestellung der Wirksamkeit in Abhängigkeit vom Schweregrad widmet. Das IQWiG selbst hat eine Methodik gewählt, wenn ich es richtig verstehe, die aus meiner Sicht dazu nur sehr wenig aussagekräftig ist, nämlich einen Basislinienmittelwertvergleich. Ich denke, dass er sicher kaum eine Aussagekraft im Vergleich zu einer gepoolten Analyse hat, in der tatsächlich Patienten mit schwerer Depression selektiert und in einer Meta-Analyse analysiert werden können, wie das in der Publikation von Thase, die wir schon eingereicht haben, durchgeführt wurde und in der sich gegenüber SSRI ein signifikanter Effekt von Duloxetin herausstellt.

**Moderator Peter T. Sawicki:** Danke schön. – Noch einmal Frau Cairns.

**Victoria Cairns:** Dabei war der Gedanke implizit, dass die Auswertungen der Firmen nicht so rigoros sind. Die Firmen haben nur ihre eigenen Daten und können nicht Studien von

anderen auswerten. Denn wenn es um Einzelpatientendaten geht, kann man es nur mit Firmendaten machen. Zumindest Lilly und Boehringer – ich nehme an, auch alle anderen – haben Auswertungen von allen Studien gemacht, die sie zum Zeitpunkt der Meta-Analyse haben. Deswegen kann man sagen: Es ist rigoros. Die Definition dessen, was hineinkommt, ist ziemlich klar: alles, was existiert. Deswegen würde ich sagen, es ist genauso wertvoll wie eine Auswertung von anderen, unabhängigen Leuten, die nehmen, was sie in der Öffentlichkeit finden können.

**Moderator Peter T. Sawicki:** Danke schön. – Herr Friede.

**Michael Friede:** Sie haben nach Poolanalysen gefragt. Beispielsweise gibt es Poolanalysen zwischen Escitalopram und Duloxetin, etwa die von Lam. Zu Venlafaxin und Escitalopram gibt es ebenfalls Poolanalysen, nämlich von Montgomery und Andersen. Zum Vergleich von Escitalopram zu Citalopram gibt es diverse Poolanalysen wie etwa Llorca, Gorman usw. Es gibt schon eine ganze Menge.

Sie sagen, dass bei Poolanalysen die Gefahr eines Bias besteht; das kann ja passieren. Man muss aber sehen, dass bei den Poolanalysen nichts anderes herauskommt als bei den einzelnen RCTs. Wenn ich bei den einzelnen RCTs zu Escitalopram gegenüber Citalopram eine Poolanalyse mache und die beiden Einzelstudien Moore und Yevtushenko beide eine Überlegenheit von Escitalopram gegenüber Citalopram zeigen, kommt in der Poolanalyse nichts anderes heraus. Insofern ist zu der Priorisierung der Wertigkeit dieser Auswertungsmaßnahmen zu sagen, dass die RCTs sicherlich einen sehr hohen Stellenwert haben, aber die Poolanalysen natürlich auch gerechtfertigt und letztlich ein weiteres Puzzlesteinchen zur Beurteilung der Wirksamkeit sind.

**Moderator Peter T. Sawicki:** Es gibt noch einige Wortmeldungen. Ich würde die Rednerliste zu diesem Tagesordnungspunkt erst einmal schließen und die Wortmeldungen abarbeiten. – Herr Kaiser.

**Thomas Kaiser:** Ich habe eine Nachfrage an Herrn Weigmann bezüglich der Analyse, die Sie eben genannt haben: Gibt es zu dieser Analyse ein vorab publiziertes Protokoll, wie die Analyse durchgeführt und welcher Schwellenwert festgelegt werden soll?

**Harald Weigmann:** Das müsste ich nachschauen; das weiß ich nicht auswendig.

**Moderator Peter T. Sawicki:** Frau Cairns.

**Victoria Cairns:** Publiziert ist sie wahrscheinlich nicht. Ein Protokoll gibt es meines Wissens schon, das aber nicht veröffentlicht ist. Es ist aber innerhalb der Firma beschrieben.

**Moderator Peter T. Sawicki:** Herr Kaiser.

**Thomas Kaiser:** Denn man würde schon erwarten, dass eine Rationale gemacht wird und die Sachen publiziert werden – ähnlich, wie Sie auch wünschen, dass wir die Berichtspläne publizieren. Es gibt dazu also kein publiziertes Protokoll, und in der Arbeit ist auch kein Protokoll zitiert?

**Moderator Peter T. Sawicki:** Frau Cairns.

**Victoria Cairns:** Wie ein Studienprotokoll wird es nicht richtig publiziert.

(Thomas Kaiser: Schade!)

**Moderator Peter T. Sawicki:** Herr Hundemer zu einer direkten Erwiderung.

**Hans-Peter Hundemer:** Dass Sie von einer der eben genannten Poolanalysen irgendwo ein zitierfähiges, als Originalarbeit publiziertes Protokoll finden, ist wahrscheinlich nur in Ausnahmefällen passiert, aber nicht regelhaft.

**Moderator Peter T. Sawicki:** Herr Kaiser.

**Thomas Kaiser:** Das ändert nichts daran, dass es sinnvoll ist.

**Moderator Peter T. Sawicki:** Herr Hundemer.

**Hans-Peter Hundemer:** Es hinterfragt aber den Sinn Ihrer Frage. Wenn es im Methodenteil adäquat beschrieben ist, weiß ich nicht, warum Sie eine Extrapublikation fordern.

**Moderator Peter T. Sawicki:** Herr Kaiser.

**Thomas Kaiser:** Nach dem, was Herr Weigmann gerade beschrieben hat, ist eine Untergruppe von Patienten innerhalb der Studien berücksichtigt worden. Wenn ich das richtig verstanden habe, geht es speziell um Patienten mit einem bestimmten Schweregrad. Über die Studienauswahl hinaus besteht dadurch eine große Verzerrungsmöglichkeit, je nachdem welchen Schwellenwert man anlegt. Insofern kann es sein, dass man einen Schwellenwert in Kenntnis der Daten festlegt. Aus diesem Grunde ist insbesondere in solchen Fällen eine Vorabdefinition des Schwellenwerts ohne Kenntnis der Daten sinnvoll. Entsprechend sind auch die Guidelines der EMEA zur Durchführung solcher Analysen. Es ist also nichts Besonderes, was ich anspreche.

**Moderator Peter T. Sawicki:** Frau Cairns.

**Victoria Cairns:** Es wäre dokumentiert.

**Moderator Peter T. Sawicki:** Noch einmal direkt dazu Herr Hundemer.

**Hans-Peter Hundemer:** Es ist bei dieser Analyse im Prinzip genauso gelaufen, wie Sie es auch gemacht haben: Es wurde vorab festgelegt, was eine leichte und was eine schwere Depression ist – sogar auf Basis von Patient Level Data, wie Sie es in Ihrem Bericht gemacht haben. Sie setzen hingegen eine andere Methodik ein, die vorher auch nicht publiziert worden ist und nicht auf Patient Level Data, sondern auf Mittelwertunterschieden zwischen Studien beruht. Dadurch beinhaltet sie ein weitaus verzerrteres Kollektiv als eine Analyse auf Basis der Einzelpatientendaten.

Eine Trennung der interessierten Kollektive sowohl Ihrer als auch unserer Analyse gelingt meiner Meinung nach bei Ihrer Analyse definitiv nicht, weil sich beispielsweise auch in der Gruppe mit leichterer Depression Patienten mit schwerer Depression befinden und umgekehrt, wenn man den Mittelwert zugrunde legt. Bei unserer Art der Analyse ist das definitiv nicht der Fall, weil wir mit Patient Level Data arbeiten. Darauf sollte man hinweisen, wenn man die Wertigkeit der jeweils gefundenen Ergebnisse berücksichtigt.

**Moderator Peter T. Sawicki:** Danke schön. – Zu einer direkten Schlusserwiderung Herr Kaiser.

**Thomas Kaiser:** Ist die Definition des Schweregrads vor Kenntnis der Daten erfolgt, also bevor die Studien beendet und die Datenbanken tatsächlich geschlossen worden sind?

(Zustimmung von Hans-Peter Hundemer)

**Moderator Peter T. Sawicki:** Herr Schmitt.

**Andreas Schmitt:** An dieser Stelle möchte ich nicht versäumen, darauf hinzuweisen, dass vor einer Woche eine gepoolte, auf Patient Level Data beruhende Analyse aller Wyeth-Studien hinsichtlich der Schweregrade als Vollmanuskript veröffentlicht worden ist. Wir hatten das für die Stellungnahme als Abstract zitiert; es wurde dafür noch nicht berücksichtigt. Wir möchten darauf hinweisen, dass es seit einer Woche als Vollpublikation zur Verfügung steht. Es ist ein wichtiger Punkt – das haben Sie in Ihrem Vorbericht formuliert –, dass es gerade bei der Analyse oder Bewertung der Schweregrade natürlich sehr hilfreich wäre, mit Patient Level Data zu operieren.

**Moderator Peter T. Sawicki:** Vielen Dank. – Herr Kessel-Steffen.

**Markus Kessel-Steffen:** Im Vorbericht sind auch indirekte Vergleiche dargestellt, nämlich in Bezug auf die Gruppe der SSRI. Gerade fiel auch das Stichwort der Bedeutung direkter versus indirekter Vergleiche und der Schwächen von indirekten Vergleichen. Von daher sollten im Vorbericht konsequenterweise in Bezug auf die SSRI zunächst die jeweiligen Head-to-Head-Studien berücksichtigt und eine Bewertung innerhalb der SSRI vorgenommen werden, bevor indirekte Vergleiche dargestellt werden. Das heißt, wenn das nicht geschieht, sollte es nur Vergleiche auf Einzelstudienbene geben.

**Moderator Peter T. Sawicki:** Vielen Dank. – Frau Wieseler.

**Beate Wieseler:** Ich habe dazu direkt eine Nachfrage: Was meinen Sie mit der Darstellung indirekter Vergleiche im Bericht?

**Moderator Peter T. Sawicki:** Herr Kessel-Steffen.

**Markus Kessel-Steffen:** Im Vorbericht sind Vergleiche zwischen Duloxetin und Venlafaxin in Bezug zu SSRI dargestellt. Bei Duloxetin gibt es beispielsweise nur drei Wirkstoffe aus der Gruppe der SSRI, für die es direkte Studien gibt. Bei Venlafaxin sind es sicherlich mehr. In unserer Stellungnahme haben wir am Beispiel der Leitsubstanz der SSRI, Citalopram, dargestellt, dass in Bezug zu den von Ihnen als relevant definierten Zielgrößen gemäß der Methodik, die Sie im Vorbericht anwenden, patientenrelevante Unterschiede bestehen. Denen wird nicht Rechnung getragen, wenn man diese heterogene Gruppe der SSRI wie im Vorbericht geschehen darstellt.

**Moderator Peter T. Sawicki:** Frau Wieseler.

**Beate Wieseler:** Zum einen betrachten wir Venlafaxin und Duloxetin zunächst einmal getrennt von- und dann auch im direkten Vergleich zueinander. Aber wir treffen unsere Aussagen nicht über indirekte Vergleiche. Unsere Aussagen zu Venlafaxin im Vergleich zu anderen Präparaten ziehen wir aus direkten Vergleichen dieser Präparate.

Darüber hinaus treffen wir nicht nur eine Aussage zum Vergleich von Venlafaxin versus SSRI, sondern brechen es auf die Wirkstoffebene herunter. Das ist auch eine Anmerkung, die ich zu dem Beitrag von Herrn Friede machen möchte. Wir gehen in dem Bericht so weit, dass wir die SSRI herunterbrechen und alle Einzelwirkstoffe betrachten. Das bilden wir ab. Insofern bilden wir auch den Unterschied zwischen verschiedenen SSRI ab. Die Gegenüberstellung Citalopram versus Escitalopram ist einfach keine Fragestellung des Berichts.

**Moderator Peter T. Sawicki:** Dazu eine Erwiderung, Herr Kessel-Steffen?

**Markus Kessel-Steffen:** In Teilen stimmt das. In Teilen werden aber sehr wohl nicht nur Einzelwirkstoffe, sondern in der Tat Gruppenvergleiche dargestellt, beispielsweise mit den Tabellen, die im Vorbericht mit „Landkarte“ überschrieben sind. Auch im Fließtext wird regelhaft von SSRI gesprochen. Dann seien Sie dort bitte so konsequent und benennen die jeweiligen Einzelwirkstoffe. Denn genau dieser Gruppenvergleich zur Wirkstoffgruppe der SSRI hat in dieser Form nicht stattgefunden.

**Moderator Peter T. Sawicki:** Frau Wieseler.

**Beate Wieseler:** Wir haben im Bericht zwei Dinge getan, nämlich einen Vergleich zum Gesamtpool der SSRI und Vergleiche zu den einzelnen Wirkstoffen vorgenommen. Das

stellen wir so dar. Wir unterscheiden sowohl in der „Landkarte der Beleglage“ als auch im Fließtext zwischen Aussagen zum Testpräparat versus Wirkstoffklasse und Testpräparat versus Einzelsubstanz beziehungsweise einzelner Wirkstoff.

**Moderator Peter T. Sawicki:** Danke schön. Wir haben alles aufgenommen. Ich hoffe, das ist so weit geklärt. Damit können wir Tagesordnungspunkt 2 abschließen.

### **Tagesordnungspunkt 3: Methodik zur Beurteilung der Begleitsymptomatik Schmerz**

**Moderator Peter T. Sawicki:** Frau Schüler.

**Yvonne-Beatrice Schüler:** Im IQWiG-Bericht wurde nicht nur die antidepressive Wirkung von Duloxetin und Venlafaxin untersucht, sondern auch die Wirkung auf verschiedene Einzel- und Begleitsymptome – das klang schon an –, unter anderem die Schmerzsymptomatik.

In der Stellungnahme der Boehringer Ingelheim Pharma GmbH & Co. KG, an die diese Frage hauptsächlich gerichtet ist, wurde ausgeführt, dass in Bezug auf die Bewertung des Baselineschmerzes, also des Schmerzes bei Studienbeginn, ein Wert unter 30 mm auf der visuellen Analogskala als nicht klinisch relevant zu erachten ist. Hierzu würden wir gerne wissen, wodurch diese Aussage begründet ist, da die von Ihnen hierzu zitierte Textstelle, nämlich Collins 1997, lediglich angibt, dass 30 mm die Grenze zu mittelschwerem Schmerz sind.

**Moderator Peter T. Sawicki:** Frau Cairns.

**Victoria Cairns:** So eine genaue Grenze ist fraglich und nicht unser Hauptpunkt. Die sieben Studien, die Sie in einer Meta-Analyse ausgewertet haben, schlossen alle Patienten mit oder ohne Schmerzen an der Baseline ein. Sie hatten zwar alle Depressionen, es war aber offen, ob sie auch Schmerzen hatten. Dabei war zu sehen, dass Duloxetin gegenüber Placebo die Schmerzreduktion signifikant verbessert hat. Man kann erwarten, dass Patienten ohne Schmerzen am Baseline keine Verbesserung erfahren werden. Um nach klinischer Relevanz zu schauen, sollte man eigentlich nur die Patienten anschauen, die Schmerzen an der Baseline haben.

**Moderator Peter T. Sawicki:** Ich glaube, das war nicht die Frage. – Vielleicht können Sie die Frage noch einmal wiederholen.

**Yvonne-Beatrice Schüler:** Es ging letztlich um die Grenze von 30 mm. Sie argumentieren in Ihrer Stellungnahme, dass man bei Werten unterhalb von 30 mm – das gilt für diese sieben Studien in der VAS-Analyse – nicht von einem klinisch relevanten Schmerz sprechen kann und damit der Effekt unterschätzt wird. Das wird mit dem Zitat von Collins belegt. Collins spricht aber nicht von nicht klinischer Relevanz, sondern lediglich davon, dass ein Wert von über 30 mm wahrscheinlich mittelschweren Schmerz darstellt. Man würde also davon ausgehen, dass Werte unter 30 mm eher einem leichten Schmerz entsprechen. Das war letztlich die Frage.

**Moderator Peter T. Sawicki:** Frau Cairns.



**Victoria Cairns:** Aber es ist nicht so relevant, wie hoch wir den Wert setzen, auch wenn wir 2 sagen, dann gibt es in den sieben Studien eine relevante Menge von Patienten, die gar keine oder kaum Schmerzen an der Baseline hatten. Das ist der Punkt: Man muss bei der Beurteilung dieser Ergebnisse aufpassen, ob etwas klinisch relevant ist oder nicht. Denn es gibt Patienten, die gar keine Schmerzen hatten und dadurch gar keine Verbesserung hätten haben können.

**Moderator Peter T. Sawicki:** Herr Rüter.

**Eckart Rüter:** Ich möchte darauf eine klinische Antwort geben: Ein Schmerz unter 30 mm kann nicht besser werden; das ist gemeint. Deswegen sagt man: Wenn der Wert unter 30 mm liegt, wird der Schmerz wahrscheinlich nicht bewertet werden können. Es sollten also nur Patienten mit einem Wert über 30 mm eingeschlossen werden. Das ist der Hintergrund.

**Moderator Peter T. Sawicki:** Herr Hundemer.

**Hans-Peter Hundemer:** Meine Wortmeldung geht in die gleiche Richtung. Ich kenne die zitierte Literaturstelle nicht wörtlich auswendig. Sie sollte zeigen, dass es in der klinischen Übereinkunft bei der Behandlung durch Schmerztherapie eine Grenze gibt, bei der man von klinisch relevantem Schmerz redet, nämlich die 30 mm. Wir sehen gerne noch einmal in die Literatur; wir finden sicherlich auch eine Stelle, die das entsprechend Ihren Wünschen besser belegt.

**Moderator Peter T. Sawicki:** Gut.

**Hans-Peter Hundemer:** Wir möchten aber nochmals betonen, dass das überhaupt nicht unser Punkt ist. Wenn Sie den Endpunkt Schmerz beurteilen und seine Verbesserung durch eine wie auch immer geartete Therapie bewerten wollen, sollte man ein Kollektiv untersuchen, das diese Symptomatik aufweist. In der von Ihnen durchgeführten Analyse wurden Studien herangezogen, in denen Schmerz überhaupt kein Einschlusskriterium war, nämlich die placebokontrollierten Duloxetinstudien bei Patienten mit einer Depression. Ich möchte auf die Argumente verweisen, die Frau Cairns schon aufgeführt hat, und sie nicht wiederholen.

Es sollte aber auch erwähnt werden, dass wir Studien gemacht haben, die genau dieses Zielkollektiv untersuchen und Patienten mit einer Depression einschlossen, die zu Studienbeginn einen relevanten Schmerz aufwiesen. Wenn Sie bei diesen Patienten untersuchen, wie sich der Schmerz bessert, finden Sie einen selbst nach Ihrer Cohen's-d-Definition relevanten Effekt auf die Schmerzbesserung. Ich denke, das ist das adäquate Kollektiv, um Schmerzbesserung in Studien zu untersuchen.

Diese beiden Studien wurden leider durch Sie mit der Begründung ausgeschlossen, dass die Ergebnisse bezüglich der Depression nicht einheitlich waren. Man sollte aber darauf hinweisen, dass primärer Endpunkt in beiden Studien die Messung der Schmerzbesserung

war. Die Besserung der Depression war ein sekundärer Endpunkt, auf den diese Studien nicht gepowert waren. Studien wegen der Nichterfüllung eines sekundären Endpunkts bei signifikantem primärem Endpunkt auszuschließen, ist weder in Ihrem Methodenpapier festgelegt noch entspricht es meiner Meinung nach einem adäquaten wissenschaftlichen Vorgehen.

Unser wesentlicher Punkt ist also, dass die Beurteilung dieses Effekts von Duloxetin in Ihrem Vorbericht nicht adäquat erfolgt und ein Kollektiv beinhaltet, das zur Beurteilung der Fragestellung nicht geeignet ist.

**Moderator Peter T. Sawicki:** Dieser Punkt findet sich auch in Ihrer schriftlichen Stellungnahme?

(Hans-Peter Hundemer: Ja!)

Vielen Dank.

Ich wollte nur darauf hinweisen – das mache ich sonst immer bei der Einleitung –, dass Sie die schriftlichen Stellungnahmen nicht zu wiederholen brauchen; dadurch wird es für die Bewertung nicht mehr. Sie gehen nicht unter. Wir werden sie alle betrachten. Hier sollten wir uns nur darauf beschränken, was zur Klärung bestimmter Fragen beiträgt. – Frau Wieseler.

**Beate Wieseler:** Ich habe eine Reihe von Anmerkungen. Es ist nicht so, dass Studien zu Patienten mit leichtem Schmerz generell als nicht geeignet angesehen werden, eine schmerzlindernde Wirkung zu zeigen. Die EMA beschreibt in den Guidelines zu Medikamenten zum Schmerz sogar explizit, dass für eine breite Indikation eines Schmerzmedikaments auch eine Wirkung bei Patienten mit leichtem Schmerz nachgewiesen werden muss. Es wird also grundsätzlich wohl schon die Möglichkeit gesehen, diese Wirkung bei Patienten mit leichtem Schmerz nachzuweisen.

Es ist richtig, dass wir in unserem Bericht im Grunde genommen zwei verschiedene Studienpools haben, nämlich zum einen den Pool, den Sie ansprechen, den wir analysiert haben. Unser Ergebnis zeigt hier keinen relevanten Gruppenunterschied. Die Frage ist eigentlich, welche Aussage man trifft. Wir treffen damit die Aussage, dass in der Population, die in diesen Studien eingeschlossen worden ist, kein relevanter Effekt bezüglich des Schmerzes nachzuweisen ist. Diese Aussage besteht. Es handelt sich um Patienten mit einer Depression, von denen ein Teil zusätzlich Schmerzen hat. Es gibt kein Einschlusskriterium bezüglich des Schmerzes; das ist richtig. Bei diesen Patienten können Sie keinen Vorteil Ihres Präparats im Vergleich zu Placebo nachweisen. Die Frage lautet, welche Aussage Sie für Ihr Präparat treffen: Geht es darum, dass das Präparat neben der antidepressiven Wirkung auch noch eine günstige Wirkung bezüglich des Schmerzes hat? Oder geht es darum, dass es sich um ein Präparat handelt, das neben der antidepressiven Wirkung auch noch eine günstige Wirkung bezüglich des Schmerzes bei Patienten mit mindestens mittelschwerem Schmerz

hat? Für die Population, die in diesen Studien eingeschlossen war, denke ich, ist diese Aussage richtig.

Den zweiten Studienpool bilden diese beiden Studien, in die Sie Patienten mit einem bestimmten Schweregrad des Schmerzes eingeschlossen haben. Hier haben wir in der Tat ein Kriterium angewendet, um zu entscheiden, ob wir diese Studien anschauen möchten. Wir betrachten hier primär die antidepressive Wirkung von Duloxetin und nicht primär Duloxetin als Schmerzmittel. Zusätzlich betrachten wir gegebenenfalls weitere Wirkungen auf die Begleitsymptomatik.

Nichtsdestotrotz liegt unser Hauptinteresse bei der antidepressiven Wirkung, weil die Patienten deshalb mit diesem Medikament bewertet werden. Wir haben deshalb die Grenze gesetzt, dass in dem Patientenpool, für den wir die Aussage treffen wollen, zunächst einmal die antidepressive Wirkung nachgewiesen sein muss. Denn die Behandlung des Schmerzes ist dem nachgeordnet.

Sie haben die Analyse vorgelegt: Wir können in der Diskussion sicherlich besprechen, wie das Ergebnis in dieser Population im Vergleich zur Population der Patienten, die keine bestimmte Schmerzschwelle überschreiten, aussieht. Allerdings ist es so – das Ergebnis haben Sie selbst berechnet –, dass auch bei diesen Patienten kein relevanter Unterschied bezüglich des Schmerzes vorliegt.

**Victoria Cairns:** Es gibt dabei mehrere Punkte.

**Moderator Peter T. Sawicki:** Sie sind eigentlich noch nicht dran; aber wir nehmen Sie jetzt einmal vorweg.

**Victoria Cairns:** Zum einen haben wir schon über Cohen's d gesprochen. Was wir vorhin gesagt haben, bleibt bestehen. In der Gruppe der sieben Studien, bei denen es auch Patienten ohne Schmerzen gab, ist es wahrscheinlich nicht sinnvoll, mittlere Änderungen anzuschauen. Sinnvoller wäre es, NNTs anzuschauen. Dort würde man eher einen Unterschied in Bezug auf die Proportionen sehen, die Responder bei Duloxetin versus Placebo. Dort könnte man eher etwas bezüglich der klinischen Relevanz erkennen als bei einem Mittelwert und Cohen's d. Wenn wir nur diese Studien, gepoolt in dieser Form, hätten, wäre eine NNT geeigneter, um die klinische Relevanz zu beobachten.

Bei den anderen zwei Studien von Brannan und Brecht ist richtig: In einer Studie gab es bei der Depression keine signifikante Änderung gegenüber Placebo. Aber das kommt auf diesem Gebiet vor. Beide Studien wurden in Ihrer Gesamtauswertung zu Remission und Depression ausgewertet. Wir glauben, dass es immer noch sinnvoll ist, diese beiden Studien zu poolen und in Bezug auf das Merkmal „Schmerz“ auszuwerten.

**Moderator Peter T. Sawicki:** Vielen Dank. – Herr Weigmann.

**Harald Weigmann:** Ich möchte noch eine kurze Anmerkung zur Wirksamkeit bei den verschiedenen Schmerzschweregraden machen. Bei den Studien, die wir untersucht haben, gibt es auch Patienten, die gar keinen Schmerz haben, sodass sie das Ergebnis am Schluss natürlich verwischen, weil sie naturgemäß auch nicht besser werden können. Daher erachte ich es schon als sinnvoll, zumindest einen gewissen unteren Cut-off-Point für die Patienten einzurichten, sodass wenigstens ein minimaler Schmerz vorliegt, der sich bessern kann. Das haben wir letztlich auch in den beiden Studien gemacht.

**Moderator Peter T. Sawicki:** Wenn ich die Frage richtig verstanden habe, ist aber nicht infrage gestellt worden, dass man bei Patienten, die gar keinen Schmerz haben, durch die Therapie keine Schmerzbesserung erreicht. Das kann man ja gar nicht erreichen.

Die Frage ist, warum Sie diesen Grenzwert ansetzen und welchen Grund Sie dafür annehmen. So habe ich Frau Schüler verstanden. Denn es wäre nicht eine Grenze zu keinem Schmerz, wie Sie gesagt haben, Herr Weigmann, sondern zu einem weniger als mittelschweren Schmerz.

**Harald Weigmann:** Es ging um einen moderaten Schmerz. Wir haben uns entschieden, nach Herrn Collins mindestens einen moderaten Schmerz anzunehmen.

**Moderator Peter T. Sawicki:** Direkt dazu, Frau Wieseler.

**Beate Wieseler:** Unsere Aussage lautet auch nicht, dass wir nachgewiesen haben, dass Duloxetin keine schmerzlindernde Wirkung hat. Unsere Aussage lautet vielmehr: Es gibt keinen Nachweis dafür, dass es eine schmerzlindernde Wirkung in dieser Population hat. Darüber hinaus liegen uns die Auswertungen, die Sie vorschlagen, einfach nicht vor. Es gibt in diesen Studien keine nach Schweregrad gestufte Auswertung.

Zudem möchte ich anmerken, dass Sie bei der Interpretation dieser Daten offensichtlich davon ausgehen, dass diese Studien geeignet sind, die Wirkung von Duloxetin bei Schmerzpatienten nachzuweisen. So lese ich es zumindest in Publikationen aus Ihrem Hause. Sie sagen hier, dass diese Studien geeignet sind, die Wirkung von Duloxetin als schmerzlinderndes Medikament nachzuweisen. Insofern wundert es mich, dass Sie jetzt sagen: Diese Studien sind gar nicht für diese Fragestellung geeignet.

**Moderator Peter T. Sawicki:** Frau Cairns.

**Victoria Cairns:** Wenn man nur Mittelwerte betrachtet, sind sie nicht geeignet, um die Relevanz zu beurteilen, weil es eine Mischung mit Patienten ohne Schmerzen gibt. Sie sind aber geeignet, um zu zeigen, dass Duloxetin einen Effekt auf Schmerzen hat, was Sie auch gezeigt haben.

**Moderator Peter T. Sawicki:** Herr Hundemer.

**Hans-Peter Hundemer:** Es sollte noch betont werden, dass der Effekt nicht nicht vorhanden ist, sondern dass diese Studien einen statistisch signifikanten Unterschied zwischen Duloxetin und Placebo gezeigt haben. Es gibt also einen Effekt. In Ihrer Beurteilung ist dieser Effekt nicht relevant. Das ist schon ein wesentlicher Unterschied. Über die Definition von Relevanz oder Nichtrelevanz hatten wir uns schon unterhalten. Darum muss man an dieser Stelle darauf hinweisen. Vor diesem Hintergrund ist es nachvollziehbar, dass die Studien genutzt wurden, um prinzipiell einen Effekt von Duloxetin auch in dieser Population nachzuweisen.

Ich sage aber noch einmal: Die in den genannten beiden Studien von Brecht und Brannan eingesetzte Population ist adäquater. Dort finden Sie einen deutlich stärkeren Effekt, selbst wenn man Ihre Definition zugrunde legt.

**Moderator Peter T. Sawicki:** Danke schön. – Herr Wille.

**Hans Wille:** Ich kann den Schwellenwert von 3, dass die Schmerzen also ab 3 relevant sein sollen, nicht ganz nachvollziehen. Ich kenne es von Schmerztherapiekonzepten und auch von der Entwicklung von Leitlinien zur Schmerztherapie her so, dass es im Grunde genommen ein Grundelement ist, für jeden einzelnen Patienten festzulegen, wo er seine individuelle Schmerzschwelle hat und wo er für sich beispielsweise einen Therapiebedarf sieht. Er kann bei 5 liegen, er kann aber auch bei 2 liegen. Deshalb halte ich den Wert von 3, der häufig angesprochen worden ist, für nur theoretisch. Das hat nichts mit der Praxis zu tun.

**Moderator Peter T. Sawicki:** Wenn ich es richtig verstanden habe, wollten Sie uns die Quelle, die Sie dazu geführt hat, diese Schwelle in den Studien anzunehmen, nachreichen.

**Victoria Cairns:** Diese Schwelle war aber kein großer Punkt. Sie war nur ein Beispiel, um zu zeigen, dass diese Studien Patienten mit und ohne Schmerzen beinhalteten.

**Moderator Peter T. Sawicki:** Es war aber unsere Frage. Deshalb wären wir dafür dankbar. Danke schön. – Herr Rüter.

**Eckart Rüter:** Ich möchte noch eine kurze klinische Bemerkung zu den Schmerzen machen. Ein Schmerzmittel bedeutet bei normalen Schmerzen etwas anderes als bei Schmerzen eines Depressiven. Herr Wille, das ist eine andere Qualität. Deswegen haben wir das in der Klinik unterschiedlich betrachtet. Auf diesen Punkt wollte ich aufmerksam machen. Wenn die EMEA fordert, dass es auch bei niedrigeren Schmerzen besser werden muss, meint man dort echte Schmerzmittel. Aber hier handelt es sich nicht um Schmerzmittel, die man nimmt, wenn man Zahnschmerzen hat.

**Moderator Peter T. Sawicki:** Es ging eben um die Rationale, es so zu machen. Das bekommen wir nachgereicht. Man muss ja nicht alles mitbringen. Es ist durchaus möglich, dass Sie uns das zuschicken. Vielen Dank. – Dann können wir diesen Tagesordnungspunkt abschließen.

#### **Tagesordnungspunkt 4: Bewertung von Schadensaspekten**

**Moderator Peter T. Sawicki:** Frau Schüler, könnten Sie bitte in diesen Punkt einführen?

**Yvonne-Beatrice Schüler:** Neben der Bewertung von positiven Effekten einer Therapie, also dem Nutzen an sich, betrachtet das IQWiG auch potenzielle negative Effekte einer Behandlung, also den potenziellen Schaden einer Therapie. Als Zielgrößen für diese Analyse verwenden wir die Gesamtrate der unerwünschten Ereignisse, die schwerwiegenden unerwünschten Ereignisse und die Therapieabbrüche aufgrund unerwünschter Ereignisse. Darüber hinaus schauen wir uns, soweit möglich, auch spezifische, relevante einzelne unerwünschte Ereignisse an. Deswegen lautet meine Frage an die Kliniker am Tisch: Was sind denn aus Ihrer Sicht die relevanten potenziellen Nebenwirkungen einer Arzneimitteltherapie der Depression?

**Eckart Rüter:** Wie lange darf ich reden?

(Heiterkeit)

**Moderator Peter T. Sawicki:** Fangen Sie mal an; dann werden wir sehen. Ich glaube schon, dass Sie das ausführlich darstellen können.

**Eckart Rüter:** Der wichtigste Punkt ist: Gibt es Todesfälle? Wir haben Sie in unserem AMSP sehr ausführlich untersucht und an Tausenden von Patienten erhoben. Die Bedeutung der einzelnen gegebenen Medikamente in Bezug auf Todesfälle, um eine wahrscheinliche Möglichkeit des Zusammenhangs herzustellen, ist extrem schwierig. Trotzdem ist das der Punkt.

Ein zweiter Punkt sind unsere Organsysteme. Die psychische Seite der Organsysteme ist ein wichtiger Punkt, zum Beispiel das Delir und die epileptische Induktion von Anfällen. Weiter geht es bei Antidepressiva mit neurologischen Erscheinungen wie Tremor, unerwünschten Bewegungen und eigenartigen Gefühlsproblemen der Beweglichkeit. Ein nächster Punkt sind die Leberprobleme.

Bei den SSRIs ist es die Natriumproblematik und die Sexualfunktion. Ihr Punkt vorhin war, was aus klinischer Sicht stört, was gefährlich und wirklich schwierig ist. Das sind andere Nebenwirkungen. Bei den SSRIs ist es die Übelkeit. Es ist also eine Abstufung erforderlich, was für den Patienten wichtig ist und was wichtig ist, wenn wir als Kliniker diese Patienten sehen. Auf diesen Unterschied möchte ich hinweisen.

Ich könnte Ihnen eine Liste von Punkten geben, von denen ich sage: Die sind für uns Kliniker ganz wichtig, so wie ich Ihnen vorhin gesagt habe, was die wichtigsten Punkte sind: Delir, Herz, Leberwerte, Blutbild. Diese Punkte sind für uns wichtig.

**Moderator Peter T. Sawicki:** Herzlichen Dank, Herr Rüter. – Herr Roots.

**Ivar Roots:** Ich glaube, das kann man noch ein bisschen in dem Sinne ergänzen, als man sagt, dass letztlich auch Arzneimittelinteraktionen als Nebenwirkungen in diesen Kontext einzubeziehen sind. Denn Interaktionen sind im Grunde die Hauptursache für Arzneimittelschäden, die wir zu beklagen haben. Selbst bei Todesfällen muss man sagen, dass einerseits die Genetik – darauf habe ich im ersten Teil hingewiesen – und andererseits die Interaktion als Ursache heranzuziehen sind, was von dem behandelnden Arzt in der konkreten Situation häufig gar nicht durchschaut wird.

Ich möchte eine Interaktion besonders hervorheben, die mir aufgrund meiner Forschung ein bisschen am Herzen liegt, nämlich die Einwirkung von Stoffen, die über Cytochrom-P450 2D6 metabolisiert werden, also von vielen Antidepressiva inklusive der heute besprochenen, auf den Metabolismus von Tamoxifen, das Patientinnen über Jahre hinweg bei Brustkrebs gegeben wird. Man muss sagen, dass diejenigen Patientinnen, die bestimmte Antidepressiva erhalten haben, deswegen einen Wirkverlust von Tamoxifen zu beklagen haben. Das ist eine furchtbare Situation, die man bisher nicht so erfasst hat, die man jetzt aber klar sieht. Hier sollte man unbedingt differenzieren, welches Antidepressivum man bei Mammakarzinompatientinnen gibt, die speziell Tamoxifen erhalten haben. – Ich könnte es noch länger ausführen; das würde mir viel Spaß machen. Aber ich glaube, das mag jetzt reichen.

**Moderator Peter T. Sawicki:** Das ist ja auch sehr interessant. Ich habe eine Nachfrage: Das erreicht man nicht durch eine Höherdosierung von Tamoxifen?

**Ivar Roots:** Nein. Sie ist sowieso durch die Verträglichkeit begrenzt. Das ist leider eine sehr ernst zu nehmende Konstellation.

**Moderator Peter T. Sawicki:** Danke schön. – Herr Bschor.

**Tom Bschor:** Ergänzend zu diesem auch aus unserer Sicht sehr wichtigen Interaktionsproblem möchte ich sagen, dass wir hier eine ganz besonders große Diskrepanz zwischen den Studienpatienten und den tatsächlich behandelten Patienten annehmen müssen, weil von den Studien eben komorbide und komedizierte Patienten meistens ausgeschlossen sind, sodass man sich vielleicht überlegen muss, welche zusätzlichen Grundlagen man mit in die Bewertung einfließen lassen kann, als nur die Nebenwirkungsraten in den Studien zu zählen. Denn das Interaktionsproblem fällt in den Studien natürlich fast aus.

**Moderator Peter T. Sawicki:** Das stimmt. Viele leiden darunter – wir auch –, dass die Studien zum Teil in sehr künstlichen Umgebungen durchgeführt werden. Das ist schade. – Bitte schön, Herr Schmitt.

**Andreas Schmitt:** Ich wollte auf den Punkt des differenziellen Nebenwirkungsspektrums und auf die Einschätzung von klinischer Seite eingehen. Grundsätzlich muss man bei Nebenwirkungen das Nebenwirkungsspektrum beachten, das man für den individuellen Patienten einsetzt. Es ist grundsätzlich schwierig, pauschal UEs und SUEs von SSRIs und SNRIs versus Trizyklika aufzuwiegen. Denn Trizyklika haben ein ganz anderes Nebenwirkungsspektrum. Mit anticholinergen Wirkungen ist die Kardiotoxizität dort sicherlich ein größeres Problem als beispielsweise bei den SSRIs und SNRIs. Teilweise kommen da dann direkte Effekte wie Übelkeit zum Tragen, die in der Klinik entsprechend Schwierigkeiten bereiten können, und möglicherweise langfristige Effekte. Aber grundsätzlich muss der Kliniker den individuellen Patienten betrachten und entscheiden, welches Antidepressivum für ihn genau das richtige ist. Man muss auch sagen, dass der klinische Alltag zeigt, dass manche Nebenwirkungen sogar speziell eingesetzt werden, wie etwa der sedierende Effekt von Mirtazapin bei unruhigen Patienten.

**Moderator Peter T. Sawicki:** Danke schön. – Herr Weinmann.

**Stefan Weinmann:** In unserem Bericht geht es nicht um eine Handreichung für die individuelle Auswahl von Antidepressiva, sondern darum, wie das Schadenspotenzial im Verhältnis zum Nutzen zu bewerten ist. Deswegen hat man danach geschaut, welche Nebenwirkungen wie etwa Mortalität, Suizidalität – die haben Sie vergessen –, schwerwiegende UEs und Abbruch wegen UEs so entscheidend sind, dass sie in die Grundüberlegung des Nutzens versus des Schadenspotenzials eingeflossen sind. Wir haben uns dagegen entschieden, einzelne Nebenwirkungen zu betrachten, sondern uns auf die großen beschränkt.

(Eckart Rüter: Todesfälle habe ich genannt!)

Richtig, aber nicht Suizidalität.

**Moderator Peter T. Sawicki:** Danke schön. – Frau Wieseler.

**Beate Wieseler:** Ich möchte zwei Dinge unterscheiden: Wenn man die Nebenwirkungen betrachtet, geht es zum einen um eine unmittelbare Gefährdung des Patienten bei der Suizidalität und damit verbundenen Problemen, wie Sie richtig gesagt haben. Zum anderen gibt es darunter einen Bereich, der eine Belastung für den Patienten darstellt, ihn aber nicht unmittelbar gefährdet.

Ich denke, die Gefährdung des Patienten haben wir abgebildet, indem wir schwerwiegende unerwünschte Ereignisse betrachten und in diesem Projekt explizit Suizidalität zu erfassen versuchen. In Bezug auf die Ereignisse, die eine Belastung für den Patienten darstellen, stellt sich die Frage: Wie äußert sich diese Belastung? Sie wird dann zum Problem, wenn der Patient nicht bereit ist, die Therapie weiterzuführen.



Deshalb schauen wir uns die Therapieabbrüche wegen unerwünschter Ereignisse an. Wir denken, dass wir dadurch eine Art Relevanzschwelle haben und die Ereignisse erfassen, die für den Patienten besonders belastend und relevant sind. Wir sind also der Meinung, dass wir in diesem Parameter eine gewisse Endstrecke all dieser verschiedenen einzelnen unerwünschten Ereignisse zusammenfassen.

In der Regel ist die Datenlage zu den unerwünschten Ereignissen in publizierten Studien viel unzureichender als zu den Nutzenparametern. Häufig haben wir überhaupt nicht die Möglichkeit, einzelne Ereignisse darzustellen. Wir denken aber, dass wir mit dem Parameter „Studienabbruch wegen unerwünschter Ereignisse“, der in der Regel in den Publikationen auch noch einigermaßen gut berichtet ist, tatsächlich die relevanten Ereignisse darstellen und analysieren können.

**Moderator Peter T. Sawicki:** Der Einwand lautete nur, dass es durch die Vorselektion der Patienten gegebenenfalls zu einer Unterschätzung kommen könnte,

(Beate Wieseler: Das kann sein!)

sodass wir den Schaden eher unterschätzen und gegebenenfalls den Nutzen überschätzen. – Direkt dazu, Herr Rüther.

**Eckart Rüther:** Wenn Sie von Abbrüchen sprechen, muss man unterscheiden: Man kann abrechnen, weil es den Patienten stört oder weil es wirklich schwerwiegend für den Patienten ist. Manche Patienten sagen: Ich habe eine leichte Übelkeit, ich muss abrechnen. – Eine andere Möglichkeit für einen Abbruch sind Herzrhythmusstörungen. Das muss man also unterschiedlich bewerten.

**Moderator Peter T. Sawicki:** Das ist klar. Natürlich kann es vom Patienten oder auch vom Arzt ausgehen, weil er es für unverantwortlich hält.

**Eckart Rüther:** Das wird leider Gottes in allen Statistiken immer zusammengemischt. Das ist das Problem.

**Moderator Peter T. Sawicki:** Danke schön. – Herr Weigmann, bitte.

**Harald Weigmann:** Wie wir schon angeführt haben, geht es vielfach darum, ein Medikament gerade individuell für einen Patienten aussuchen zu können. Dazu ist es meiner Meinung nach auch erforderlich, dass man die Datenbasis gerade in Bezug auf Nebenwirkungen so groß wie möglich macht, wie beispielsweise die EMEA, die zur Beurteilung des Nebenwirkungsspektrums auch Studien aus anderen Indikationen heranzieht. Aufgrund dieser größeren Datenbasis kommt sie unter anderem zu dem Ergebnis, dass Duloxetine den SSRIs beim Nebenwirkungsspektrum vergleichbar ist. Eine kleinere Datenbasis führt möglicherweise zu einem verzerrten Ergebnis.

**Moderator Peter T. Sawicki:** Vielen Dank. – Herr Wille.

**Hans Wille:** Mir ist das Thema Suizidalität im Bericht ein bisschen zu kurz gekommen. Dazu gibt es aus den letzten fünf Jahren eine ganze Reihe von Literatur mit gewissen Hinweisen, dass bei jüngeren Patienten unter 25 Jahren und vielleicht gerade bei denjenigen, die auf die Therapie nicht ansprechen und keine positiven Wirkungen zeigen, durchaus eine erhöhte Suizidalität vorliegt. In einem Memorandum der FDA haben gerade die beiden Substanzen Duloxetin und Venlafaxin mit den höchsten Risiken, den höchsten Odds Ratios von 5 oder 6, abgeschnitten. Ich weiß nicht, ob es mit den Daten, die Sie haben, möglich ist, noch Weiteres herauszuziehen und weitere Subanalysen zu machen. Natürlich sind es seltene Ereignisse. Dass sich in den vielen Studien mengenmäßig etwas zeigt, ist wahrscheinlich schwierig.

So weit ich gesehen habe, war Suizidalität andererseits bei allen Studien ein Ausschlussgrund. Zudem ist, glaube ich, lost-to-follow-up in den Studien nicht berücksichtigt beziehungsweise berichtet worden. Das gibt durchaus noch mehr Unsicherheiten, denn es könnte durchaus Patienten betreffen, die das Problem besonders haben.

**Moderator Peter T. Sawicki:** Natürlich; das ist auch so. – Frau Wieseler.

**Beate Wieseler:** Zu Herrn Weigmann möchte ich sagen: So wie ich die Stelle lese, die Sie zitiert haben, bezog sich die Aussage der EMEA zur Vergleichbarkeit von Duloxetin und SSRI auf unerwartete und schwerwiegende unerwünschte Ereignisse. Das spiegelt sich durchaus in unserem Bericht wider. Wir sehen keine Unterschiede zwischen den Substanzen.

Zur Anmerkung von Herrn Wille: Das geben die Daten der Studien leider nicht her. Wir haben im Grunde genommen die Aussagen zu Suizidalität eingeschränkt, weil die Datenbasis das nicht hergibt. Dazu sind aufgrund der Studienanlage und der Häufigkeit ganz andere Fallzahlen notwendig. Es ist aber richtig, dass wir den Punkt in der Diskussion noch einmal stärker adressieren können.

**Moderator Peter T. Sawicki:** Dazu Herr Wille.

**Hans Wille:** Eigentlich kann man überhaupt keine Aussage machen – gerade beim Langzeiteffekt. Das hätte ich mir in dieser Deutlichkeit gewünscht.

**Moderator Peter T. Sawicki:** Vielen Dank. Das nehmen wir auf. – Herr Friede, bitte.

**Michael Friede:** Sie haben eben auf den EMEA-Artikel und auf die Vergleichbarkeit zwischen SNRIs und SSRIs hingewiesen. Hier kommt wieder zum Tragen, dass eine inhomogene Gruppe verschiedener Wirksubstanzen verglichen wird. Wenn Sie das zum Beispiel für den Vergleich von Escitalopram zu Duloxetin oder zu Venlafaxin verwenden, sieht man eindeutig weniger Drop-outs wegen UEs unter Escitalopram im Vergleich zu den einzelnen SNRIs. Insofern gilt es wieder zu berücksichtigen, dass die Wirkstoffe unterschiedlich sind.

**Moderator Peter T. Sawicki:** Vielen Dank. – Bitte schön.

**Peter-Andreas Löschmann:** Ich würde gerne noch allgemein auf einen Punkt hinweisen: Welche Datenbasis steht Ihnen zur Verfügung, wenn Sie randomisierte klinische Studien betrachten und dann die Schadensbewertung vornehmen? Wenn ich es richtig verstanden habe, wird im Methodenpapier ausdrücklich gesagt, dass man eine Kausalität zwischen der Behandlung und den beobachteten Ereignissen herstellen möchte. Nun ist es GCP-Standard, dass Ereignisse innerhalb von klinischen Studien unabhängig vom Zusammenhang gemeldet werden müssen und in die Datenbank eingehen. Ich bezweifle, dass man aus den publizierten Studien aufgrund dieser Datenbasis so einfach eine Kausalbeziehung herstellen kann.

**Moderator Peter T. Sawicki:** Darüber könnten wir jetzt lange diskutieren. Vielleicht machen wir es nur kurz. – Frau Wieseler.

**Beate Wieseler:** Es ist richtig, dass wir unerwünschte Ereignisse berichten. Es ist uns völlig klar, dass sie nach der GCP erhoben werden und erst einmal keine Kausalitätsbewertung dabei ist. Unser Methodenpapier sagt auch, dass die Kausalität im Grunde genommen daraus abgeleitet wird, dass man diese Effekte aus kontrollierten Studien mit einer hinreichend guten Strukturgleichheit ableitet. Damit können die Effekte eigentlich nur auf die Therapie zurückzuführen sein. Wir betrachten nicht das Profil der einzelnen Gruppen, sondern schauen uns die Effekte, also die Gruppenunterschiede an. In einer randomisierten Studie können diese Effekte nur auf den einzigen unterschiedlichen Faktor zwischen den beiden Gruppen zurückzuführen sein, nämlich auf die Behandlung. Das ist unsere Ableitung eines kausalen negativen Effekts.

**Peter-Andreas Löschmann:** Darüber kann man lange diskutieren; wir hören jetzt aber auf.

**Moderator Peter T. Sawicki:** Darüber kann man lange diskutieren; das stimmt. – Herr Weigmann, wollen Sie das diskutieren?

(Harald Weigmann: Nein!)

Gut. Bitte schön.

(Heiterkeit)

**Harald Weigmann:** Ich möchte etwas anderes diskutieren. Wenn ich Sie richtig verstanden habe, sagten Sie, dass die EMEA nur auf schwerwiegende Nebenwirkungen abhebt. Ich muss gestehen, das sehe ich nicht so, wenn ich diesen Satz lese. Darüber kann man sicherlich grammatikalisch streiten; aber für mich lautet der Hauptsatz:

„In conclusion, the safety profile of duloxetine (...) is comparable to that known for SSRIs.“

(Beate Wieseler: Ich denke, es kommt darauf an, was davor steht!)

**Moderator Peter T. Sawicki:** Der Kontext ist immer wichtig. Aber gut: Das nehmen wir auf. Vielen Dank. – Es gibt keine weiteren Wortmeldungen mehr dazu. Ist dieser Tagesordnungspunkt damit auch aus Ihrer Sicht ausreichend diskutiert, Frau Schüler? Sind alle Fragen beantwortet?

(Yvonne-Beatrice Schüler: Ja!)

**Tagesordnungspunkt 5:  
Depressionsskalen-Validierung, Beobachtungszeitraum**

**Moderator Peter T. Sawicki:** Frau Schüler, bitte.

**Yvonne-Beatrice Schüler:** Der Vergleich von Duloxetin und Escitalopram in der Kurzzeitakuttherapie, also in Depressionsendpunkten, erfolgt im Bericht anhand von zwei Studien, nämlich von Khan 2007 und Nierenberg 2007. In der Stellungnahme der Lundbeck GmbH, an die ich diese Frage richten möchte, wird ausgeführt, dass der Einschluss der Studie Nierenberg 2007 nicht gerechtfertigt sei, da die ausgewerteten Depressionsskalen nicht für einen wie in der Studie durchgeführten wöchentlichen Beobachtungszeitraum validiert seien. Daher sei auch die Validität der Ergebnisse nicht gewährleistet. Diese Anmerkung ist uns nicht klar. Deshalb würde ich Sie gerne fragen, was Sie mit der Formulierung „nicht für einen wöchentlichen Beobachtungszeitraum validiert“ meinen?

**Moderator Peter T. Sawicki:** Möchten Sie darauf direkt antworten, Herr Friede?

**Michael Friede:** In der Nierenberg-Arbeit ist bei dem primären Zielkriterium eine Subskala von Hamilton verwendet worden, die Subskala von Maier. Schaut man sich die Ausführungen dazu an, wie häufig diese Subskala verwendet werden kann, stellt sich heraus, dass sie nicht für die wöchentliche Anwendung geeignet ist. Wenn man sie also so häufig anwendet, kann es unserer Meinung nach zu einem Bias kommen.

**Moderator Peter T. Sawicki:** Frau Schüler.

**Yvonne-Beatrice Schüler:** Die Maier-Subskala ist nun nicht das Zielkriterium, das wir erhoben und aus der Studie extrahiert haben, sondern die HAMD-Daten insgesamt. Gibt es dabei aus Ihrer Sicht auch Probleme mit dem wöchentlichen Beobachtungszeitraum?

**Moderator Peter T. Sawicki:** Herr Friede.

**Michael Friede:** Es kann natürlich zu einer Interaktion kommen, weil Sie verschiedene Skalen wie zum Beispiel die Subskala von Maier einfach sehr häufig und die HADS zweimal wöchentlich anwenden. Somit kann es zu einer Interaktion, zu einer Beeinflussung der Patienten kommen. Das führt wiederum zu einem Bias.

**Moderator Peter T. Sawicki:** Frau Wieseler.

**Beate Wieseler:** Ich habe eine Nachfrage dazu. Zum einen betrachten wir den Endpunkt der Studie und schauen uns keinen Verlauf an, bei dem wir wöchentlich die Daten sehen. In den meisten Studien mit Antidepressiva werden die Daten in den ersten vier Wochen wöchentlich erhoben. Danach geschieht das alle zwei Wochen oder in etwas längeren Zeiträumen. Warum gilt dieses Problem nur für die Studie von Nierenberg?

**Moderator Peter T. Sawicki:** Herr Friede.

**Michael Friede:** Das gilt für die Subskala von Maier.

**Beate Wieseler:** Wir verwenden die Subskala von Maier aber nicht.

**Michael Friede:** Ja, aber es wurde eine Reihe von Skalen in einer hohen Frequenz verwendet. Das spiegelt sich in den entsprechenden Ergebnissen wider.

**Beate Wieseler:** Auch in den anderen Studien werden die Skalen in vergleichbarer Frequenz verwendet. Wenn das, was Sie beschreiben, ein Problem wäre, müssten wir wahrscheinlich die Mehrzahl der Studien ausschließen.

**Moderator Peter T. Sawicki:** Möchte jemand in die Diskussion einsteigen? Herr Kessel-Steffen, Herr Pedersen? Herr Weinmann.

**Stefan Weinmann:** In welche Richtung gäbe es einen Bias?

**Anders Gersel Pedersen:** Es gibt immer einen Bias. Wenn Sie mehrere Skalen in den gleichen Studien sehr frequent nutzen, gibt es einen höheren Placeboeffekt, als wenn Sie weniger Instruktionen an die Patienten machen. Das kann man auch in der Khan-Studie sehen, wo man zum Beispiel am Ende keine Separation zwischen Escitalopram und dem Placebo sehen kann. Die meisten anderen Studien haben eine Separation zwischen Placebo und Escitalopram. Bei dieser Skala in diesen Studien gab es keine Separation zwischen Placebo und Escitalopram. Das kann man sehen, wenn man sehr viele Skalen gebraucht, weil man einen großen Placeboeffekt in diesen Studien findet.

**Moderator Peter T. Sawicki:** Frau Wieseler.

**Beate Wieseler:** Ich stelle meine Nachfrage noch einmal: Gilt das nur für die Studie von Nierenberg oder gilt es dann nicht für eine Vielzahl von Studien, die wir betrachtet haben?

**Moderator Peter T. Sawicki:** Herr Pedersen.

**Anders Gersel Pedersen:** Es geht um mehrere Studien. Wenn man mehrere Skalen mit hoher Frequenz gebraucht, hebt man die Placeboeffekte in den Studien.

**Beate Wieseler:** Das ist gegebenenfalls ein Problem. Die Literatur dazu müssten wir uns anschauen. Aber ich denke, das wäre dann auf alle Studien anzuwenden.

Darüber hinaus würde ich das nicht als fehlende Validierung bezeichnen. Die Skala an sich ist dadurch nicht weniger validiert. Es ist gegebenenfalls das Problem eines Einflusses einer häufigen Erhebung.

**Moderator Peter T. Sawicki:** Herr Pedersen.

**Anders Gersel Pedersen:** Nein, es ist immer ein Problem, wenn man die Subskalen als primäre Endpunkte und nachher dann eine ganze Skala für die Kalkulation gebraucht. Es gibt zwei verschiedene Probleme: Das eine ist die Frequenz. Das andere besteht darin, dass Sie eine Subskala als Primärendpunkt und nachher die ganzen Skalen für die Kalkulation gebrauchen.

**Moderator Peter T. Sawicki:** Gut. Jetzt haben wir die Argumente ausgetauscht. Vielleicht haben wir die Frage nicht ganz geklärt, zumindest aber haben wir das Problem beschrieben. Danke schön. – Frau Schüler, haben Sie noch weitere Fragen zu diesem Punkt?

**Yvonne-Beatrice Schüler:** Nein.

**Moderator Peter T. Sawicki:** Damit ist dieser Punkt abgeschlossen.

### **Tagesordnungspunkt 6: Zulassungsstatus von Duloxetin**

**Yvonne-Beatrice Schüler:** In der Stellungnahme der Boehringer Ingelheim Pharma GmbH & Co. KG wurde vorgeschlagen, eine kürzlich beendete Studie zur Rezidivprophylaxe mit Duloxetin in die Bewertung einzuschließen. Hierzu frage ich die Vertreter der Firma Boehringer: Wie sehen Sie den Zulassungsstatus von Duloxetin bezüglich der Therapie in der Rezidivprophylaxe der Depression?

**Moderator Peter T. Sawicki:** Herr Hundemer.

**Hans-Peter Hundemer:** Diese Studie ist als Full Study Report bei der EMEA mit dem Ziel eingereicht worden, die Zulassung zu erreichen. Im Moment haben wir sie nicht.

**Moderator Peter T. Sawicki:** Danke schön, damit ist das geklärt.

**Hans-Peter Hundemer:** Darf ich eine Rückfrage stellen?

**Moderator Peter T. Sawicki:** Selbstverständlich.

**Hans-Peter Hundemer:** Warum ist das relevant für Sie?

**Moderator Peter T. Sawicki:** Weil wir im Rahmen des Zulassungsstatus beurteilen.

**Hans-Peter Hundemer:** In der Bewertung des Vorberichts war kein Bezug auf eine bestehende oder eine nicht bestehende Zulassung erwähnt. Dabei ging es um das Vorliegen oder Nichtvorliegen von Daten.

**Moderator Peter T. Sawicki:** Frau Wieseler.

**Beate Wieseler:** Unsere generelle Methodik ist es, im Rahmen des Zulassungsstatus zu bewerten, weil es um eine Information des G-BA geht, der ebenfalls im Rahmen des Zulassungsstatus arbeitet. Im Vorbericht war uns die Zulassungssituation nicht hinreichend klar. Deshalb haben wir diese Frage an Sie gestellt. Wenn das Präparat zur Rezidivprophylaxe nicht zugelassen ist, können wir es nicht in dieser Indikation bewerten.

**Moderator Peter T. Sawicki:** Herr Wille.

**Hans Wille:** Ist es eigentlich für den Remissionserhalt zugelassen? Wie ich die Zulassung lese, gilt das nur für die akute Therapie. Deswegen habe ich mich gefragt, warum diese drei Studien eigentlich aufgenommen und beurteilt worden sind. Denn Sie liegen nach meiner Lesart außerhalb der zugelassenen Indikation.

**Moderator Peter T. Sawicki:** Interessant. – Frau Schüler.



**Yvonne-Beatrice Schüler:** Dazu möchte ich anmerken, dass wir Rücksprache gehalten haben. Die in den Fachinformationen zu Duloxetin und Venlafaxin festgehaltenen Indikationen, nämlich Episoden einer Major Depression, sind mit dieser sechsmonatigen Rückfallpräventionstherapie abgedeckt. Das ist darin enthalten.

(Hans Wille: Das ist vom Amt so ...)

Ja.

**Moderator Peter T. Sawicki:** Danke schön. Damit haben wir diesen Tagesordnungspunkt abgeschlossen.

**Tagesordnungspunkt 7:  
Validierung des Composite Cognitive Score**

**Yvonne-Beatrice Schüler:** In einigen Stellungnahmen wurde zur Bewertung der Begleitsymptomatik Kognition angeführt, dass der in einer Studie verwendete Composite Cognitive Score als Parameter berücksichtigt werden sollte. Da dieser im Vorbericht aufgrund der fehlenden Validierung ausgeschlossen wurde – wie dort auch beschrieben –, möchten wir fragen, ob Sie uns zu diesem Instrument eine Validierungspublikation oder etwas Ähnliches nennen können?

**Moderator Peter T. Sawicki:** Wer möchte darauf antworten?

**Victoria Cairns:** Nein.

(Heiterkeit)

**Moderator Peter T. Sawicki:** Sie möchten uns das nicht nennen, oder möchten Sie nicht antworten?

**Victoria Cairns:** Nein, leider haben wir hierzu nichts. Dieser Composite Cognitive Score ist eine Kombination von anderen Skalen, die gut bekannt und validiert sind. Aber es stimmt: Wir haben keine formelle Validierung für die Gesamtsumme.

**Moderator Peter T. Sawicki:** Herr Rüter.

**Eckart Rüter:** In der Demenzforschung nehmen wir sehr häufig verschiedene Skalen zusammen, um daraus eine Schlusskala zu berechnen. Sie zu validieren, ist in der Regel nicht üblich, obwohl es wünschenswert wäre; das gebe ich ehrlich zu. Aber das ist hochrangig publiziert, zum Beispiel in den Archives of General Psychiatry. Dabei gibt es auch immer diese Zusammenzählungen. Aber das wird nicht validiert. Es handelt sich um eine übliche Vorgehensweise. Deswegen ist dieser Wert für uns klinisch zu akzeptieren.

**Moderator Peter T. Sawicki:** Danke schön. Die Antwort ist nicht ganz befriedigend, aber es ist halt so.

(Eckart Rüter: Es befriedigt nicht, aber es freut einen!)

**Moderator Peter T. Sawicki:** Es ist, wie es ist. – Herr Hundemer.

**Hans-Peter Hundemer:** Was Frau Cairns sagte, ist richtig. Dieser Summenscore ist als Summenscore nicht validiert. Es gibt jedoch eine deskriptive Validierung, die nicht im Rahmen einer Originalarbeit, sondern im Rahmen eines Posters vorgenommen wurde, das wir Ihnen gern zukommen lassen. Dabei werden die Baselinewerte dieser als Einzelskalen validierten Tests mit normativen Werten der Literatur verglichen, um zu prüfen, ob wir in

dieser Studie Werte finden, die mit den für den jeweiligen Test etablierten Werten vergleichbar sind. Das liegt uns vor; wir können es Ihnen gern nachreichen.

(Eckart Rüther: Das wäre gut!)

Das bezeichnet man als „deskriptive Validierung“, die sich jeweils auf die Einzeltests bezieht.

In Sachen Summenscore möchte ich mich mit Herrn Rüther solidarisch erklären: Dieses Vorgehen ist klinisch üblich, ohne den Summenscore von fünf einzelnen, validierten Tests publikatorisch validieren zu müssen. – Wir können die deskriptive Validierung gern nachreichen.

**Moderator Peter T. Sawicki:** Gern. Wobei Herr Rüther ja gesagt hat, das zu validieren sei wünschenswert. Ein „übliches Vorgehen“ ist nicht das, was er als „wünschenswert“ bezeichnet hat. – Vielen Dank, dann haben wir diesen Tagesordnungspunkt abgeschlossen.

### **Tagesordnungspunkt 8:**

#### **Verständnisfrage zum Punkt: Ausschluss von Studien aufgrund der Studiendauer**

**Yvonne-Beatrice Schüler:** Die AkdÄ, an deren Vertreter ich diese Frage richte, merkt in ihrer Stellungnahme an, dass fünf Studien wegen zu kurzer Studiendauer – das ist das Kriterium „Nicht E6“ – in der Volltextanalyse ausgeschlossen wurden. Weiterhin schreiben Sie, dass weitere 22 Studien aus dem gleichen Grund – ohne Volltextanalyse – ausgeschlossen worden seien, woraus sich die Frage ergebe, wie berechtigt der Ausschluss einer Studie aufgrund der Studiendauer sei.

Unsere Frage lautet, woraus sich die Zahl 22 ergibt? Das konnten wir aus dem Vorbericht weder entnehmen noch erschließen. Denn das Ausschlusskriterium an sich enthält die fünf Studien. Bei den 22 Studien handelt es sich wohl um irgendeinen Subpool.

**Moderator Peter T. Sawicki:** Herr Bschor.

**Tom Bschor:** Ich möchte das Argument zunächst verständlich machen. Uns ging es um Studien, die primär das Ziel hatten, die Verträglichkeit zu erfassen und dabei infrage zu stellen, ob eine bestimmte Studiendauer ein Grund sein kann. Ich möchte es einmal überspitzt darstellen: Bei einer Studie, die fünf Tage lang untersucht und bei der Todesfälle vorkommen, wäre es diskussionswürdig, diese bei der Frage nach dem Schaden aufgrund der kurzen Studiendauer zu berücksichtigen oder auszuschließen. Das ist unser Hauptansatzpunkt.

Die Ausführungen zu den fünf Studien habe ich gerade herausgesucht. Die Zahl von 22 Studien müsste ich recherchieren. Dieser Punkt kam nicht von mir persönlich, sondern von einem anderen AkdÄ-Mitglied. Darauf könnte ich Ihnen in zehn Minuten eine Antwort geben.

**Moderator Peter T. Sawicki:** Das muss nicht sofort passieren. Aber bitte lassen Sie uns die Antwort noch zukommen.

**Tom Bschor:** Dann maile ich Ihnen die Antwort.

**Moderator Peter T. Sawicki:** Danke schön. – Herr Schmitt.

**Andreas Schmitt:** Ich möchte die Studiendauer kurz kommentieren. Denn ich kann nicht ganz nachvollziehen, warum man Studien mit kürzerer Dauer nicht als Evidenz berücksichtigt, weil beispielsweise auch darauf verwiesen wird, dass bei der Zulassung bei der EMEA 6-Wochen-Studien empfohlen würden. Jedoch würde sie die Evidenz von Studien kürzerer Dauer niemals unberücksichtigt lassen. Grundsätzlich würde man damit die Effekte hinsichtlich der Wirksamkeit dann ja eher unterschätzen.

**Moderator Peter T. Sawicki:** Zunächst einmal sind wir keine Zulassungsbehörde. Darüber hinaus haben wir den Aspekt der Studiendauer und des Einschlusses beim vorläufigen Berichtsplan abschließend diskutiert, sodass wir darauf nicht zurückkommen können. Das

wurde diskutiert, beschrieben und im Institut besprochen. Aufgrund der Berücksichtigung aller relevanten Einwürfe haben wir entschieden, so vorzugehen.

**Tagesordnungspunkt 9:  
Verschiedenes**

**Moderator Peter T. Sawicki:** Gibt es etwas, das Ihnen aufgrund der Diskussion wichtig erscheint, das Sie uns noch nicht im mündlichen oder schriftlichen Stellungnahmeverfahren mitgeteilt haben und das sich nicht auf den Scopingworkshop bezieht? – Herr Kessel-Steffen, bitte.

**Markus Kessel-Steffen:** Nutzenbewertung SSRI, Verfahren on hold: Ist das weiterhin so, oder wird es gemäß des Auftrags des Gemeinsamen Bundesausschusses bearbeitet werden?

**Moderator Peter T. Sawicki:** Thomas Kaiser.

**Thomas Kaiser:** Hierzu gibt es ein Informationssystem auf der Website des Institutes; dort werden Sie über einen eventuellen Auftragsbeginn informiert.

**Moderator Peter T. Sawicki:** Herr Kessel-Steffen.

**Markus Kessel-Steffen:** Darf ich daraus schließen, dass es nach wie vor on hold ist, aber der Auftrag seitens des Gemeinsamen Bundesausschusses nach wie vor Bestand hat?

**Moderator Peter T. Sawicki:** Thomas Kaiser.

**Thomas Kaiser:** Der Auftrag hat Bestand. Auf der Internetseite gibt es eine deutliche Trennung, ob sich Aufträge im Bestand befinden oder nicht.

**Moderator Peter T. Sawicki:** Ich bedanke mich ganz herzlich für die sehr konzentrierte, hilfreiche und zielführende Diskussion. Das hat dazu geführt, dass wir vor der geplanten Zeit fertig geworden sind. Wahrscheinlich haben unsere guten Organisationsgeister damit gerechnet, dass wir früher fertig werden, und ein kleines Lunch vorbereitet.

Ich wünsche Ihnen einen trockenen und sicheren Heimweg. Sie werden über die Art der Berücksichtigung Ihrer Stellungnahmen durch die Publikation des Abschlussberichts zu gegebener Zeit wissend werden. – Herzlichen Dank und alles Gute.

(Allgemeiner Beifall)

**Anhang: Dokumentation der Stellungnahmen**

# Inhaltsverzeichnis

	Seite
<b>A 1 Stellungnahmen von Organisationen, Institutionen und Firmen .....</b>	<b>A 2</b>
<b>A 1.1 Arzneimittelkommission der deutschen Ärzteschaft (AkdÄ) .....</b>	<b>A 2</b>
<b>A 1.2 Institut für Arzneimittelsicherheit in der Psychiatrie e. V. (AMSP).....</b>	<b>A 10</b>
<b>A 1.3 Boehringer Ingelheim Pharma GmbH &amp; Co. KG .....</b>	<b>A 19</b>
<b>A 1.4 Deutsche Gesellschaft für Psychiatrie, Psychotherapie und Nervenheilkunde (DGPPN), Arbeitsgemeinschaft für Neuropsychopharmakologie und Pharmakopsychiatrie (AGNP) .....</b>	<b>A 38</b>
<b>A 1.5 Lundbeck GmbH.....</b>	<b>A 43</b>
<b>A 1.6 Verband Forschender Arzneimittelhersteller e. V. (VFA).....</b>	<b>A 81</b>
<b>A 1.7 Wyeth Pharma GmbH .....</b>	<b>A 87</b>
<b>A 2 Stellungnahmen von Privatpersonen.....</b>	<b>A 102</b>
<b>A 2.1 Roots, Ivar, Prof. Dr. med. ....</b>	<b>A 102</b>
<b>A 2.2 Volz, Hans-Peter, Prof. Dr. med. ....</b>	<b>A 126</b>



## **A 1 Stellungnahmen von Organisationen, Institutionen und Firmen**

### **A 1.1 Arzneimittelkommission der deutschen Ärzteschaft (AkdÄ)**

**Autoren:**

Gundert-Remy, Ursula, Prof. Dr. med.

**Adresse:**

Prof. Dr. med. Ursula Gundert-Remy

Arzneimittelkommission der deutschen Ärzteschaft (AkdÄ)

Herbert-Lewin-Platz 1

10623 Berlin

# **Arzneimittelkommission der deutschen Ärzteschaft**

Fachausschuss der Bundesärztekammer



## **Stellungnahme der Arzneimittelkommission der deutschen Ärzteschaft**

**zum Vorbericht A05-20A**

(vorläufige Nutzenbewertung)

**Selektive Serotonin- und Noradrenalin-  
Wiederaufnahmehemmer (SNRI) bei Patienten mit  
Depressionen**

des Instituts für Qualität und Wirtschaftlichkeit  
im Gesundheitswesen (IQWiG)

Version 1.0

Berlin, den 09.01.2009

[www.akdae.de](http://www.akdae.de)

Der Vorbericht des IQWiG legt umfangreich aufgearbeitetes Studienmaterial zur Nutzenbewertung der SNRI Venlafaxin und Duloxetin vor. Große Mühe wurde darauf verwandt, die Studien, die mit zum Teil sehr unterschiedlichen Zielen, Ergebnisparametern und Methoden durchgeführt wurden, vergleichbar zu machen und möglichst einheitliche Parameter aus allen Studien zu extrahieren. Der Bericht liefert eine detaillierte Bewertung der beiden untersuchten Substanzen. Für beide SNRI wird eine eindeutige Wirksamkeitsüberlegenheit gegenüber Placebo festgestellt. Der Vergleich der beiden Pharmaka mit anderen Antidepressiva und die vergleichende Bewertung der Verträglichkeit fällt differenziert aus. Im Folgenden gehen wir auf Aspekte des Vorberichts ein, die aus Sicht der AkdÄ verbesserungswürdig sind.

## 1. Verständlichkeit und Klarheit des Textes

Der Text ist nicht einfach verständlich. Dies resultiert aus verwendeten sprachlichen Formulierungen, von welchen ein Teil durch Ungenauigkeit Verständnisprobleme aufwirft.

Beispiele:

- Seite 10, 4.1.1 Population, 1. Absatz: „Für die Nutzenbewertung wurden Studien mit Patienten mit einer leichten, mittelschweren oder schweren Depression berücksichtigt.“ Es wird fortgefahren im 2. Absatz: „Es wurden auch Patienten mit somatischer bzw. psychiatrischer Komorbidität berücksichtigt.“ (Unterstreichung durch Kommentator). Geht es um Patienten oder um Studien?
- Seite 19, 4.4 Informationssynthese und -analyse, 2. Absatz: „Für Studien mit verschiedenen Wirkstoffen einer Wirkstoffklasse als Vergleichsintervention (z. B. SSRI) war eine Zusammenfassung auf Wirkstoffebene vorgesehen. Bei Hinweisen auf eine durch die verschiedenen Substanzen einer Klasse bedingte Heterogenität erfolgte dann gegebenenfalls eine separate Analyse nach den einzelnen Substanzen.“ Dieser Absatz ist verwirrend, was u. A. an dem wechselnden Gebrauch der Begriffe „Wirkstoffe“, „Wirkstoffklasse“, „Wirkstoffebene“ und „Substanz“ liegt. Sind „Substanz“ und „Wirkstoff“ sowie „Wirkstoffklasse“ und „Wirkstoffebene“ Synonyme? Wenn ja, sollte jeweils nur eine Bezeichnung verwendet werden, wenn nein, müssen die Begriffe definiert werden (gilt insbesondere für die ungebräuchliche „Wirkstoffebene“). Unklar bleibt, woraus sich „Hinweise auf Heterogenität“ ergeben.
- Das Wort „Relevanz“ wird in unterschiedlicher und zum Teil nicht definierter Bedeutung verwendet (z. B. Seite 13 E7 Relevanz der Studie, Seite 20 Schwelle der Relevanz, Seite 609 relevante Menge an Daten).

- In einem anderen Teil der Fälle lässt sich nicht sicher erkennen, welche Information vermittelt werden soll: Beispiel: Seite 614: „Insbesondere in Bezug auf die antidepressive Wirkung beider Substanzen im Vergleich zu aktiven Komparatoren ergibt sich in der Literatur ein vielfältiges Bild.“

**Fazit:** Eine Überarbeitung des Berichts wird angeraten, um Lesbarkeit und Klarheit der Aussagen zu verbessern.

## 2. Konzeptionelle Probleme

### 2.1 Unerwünschte (Arzneimittel)wirkungen/Schadenbetrachtung

- Auf Seite 11 wird als eine der Zielgrößen Unerwünschte Arzneimittelwirkungen aufgeführt. Im Folgenden ist jedoch nicht von unerwünschten Arzneimittelwirkungen die Rede, sondern von unerwünschten Ereignissen, welche per definitionem Ereignisse im Rahmen einer klinischen Prüfung darstellen, ohne dass ein Zusammenhang zu dem verabreichten Arzneimittel hergestellt wurde (siehe auch Tabelle 3 auf Seite 23, hier werden unerwünschte Arzneimittelwirkungen unter Zielgrößen der Nutzenbewertung aufgeführt und mit unerwünschten Ereignissen operationalisiert). Diese Gleichsetzung ist konzeptionell falsch. Ebenfalls findet sich diese Gleichsetzung auf Seite 11, 3. Absatz. Unter Punkt 6.4.5 Unerwünschte Arzneimittelwirkungen wird dieser konzeptionell falsche Ansatz weitergeführt und das Auftreten von schwerwiegenden unerwünschten Ereignissen mit einem Schaden gleichgesetzt, ebenso wie die Gesamtrate an unerwünschten Ereignissen.
- Die jeweiligen Operationalisierungen der Zielgrößen in den Studien wurden hinsichtlich ihrer Angemessenheit, Relevanz und Messsicherheit geprüft. Leider finden sich hierzu in den Diskussionen zu den Studien und Endpunkten zu den unerwünschten (Arzneimittel-) Wirkungen keine Ausführungen.
- Leider fehlt bis auf die Anzahl von Patienten mit Bluthochdruck und sexueller Dysfunktion die Auswertung möglicher weiterer schwerwiegender oder auch nicht schwerwiegender unerwünschter Ereignisse und ihre kausale Zuordnung zum Arzneimittel. Unterschiede in Spektrum, Schweregrad und Häufigkeit von erwünschten Arzneimittelwirkungen sind bei der Auswahl von Antidepressiva aber ein wesentliches Kriterium.
- Im Übrigen wird der Begriff schwerwiegende unerwünschte Ereignisse nicht definiert oder Bezug genommen auf eine bestehende Definition.
- Bei der vergleichenden Risikobewertung von SNRI mit anderen Antidepressiva fehlt ein wesentlicher Aspekt, die Bewertung des pharmakokinetischen Interaktionspotenzials und

ganz generell die der pharmakokinetischen Eigenschaften. Es gibt erhebliche Unterschiede im Interaktionspotenzial der SSRI. Paroxetin, Fluoxetin, sowie Bupropion und Moclobemid als weitere Antidepressiva sind potente Inhibitoren von CYP2D6, Fluvoxamin von CYP1A2, Moclobemid von CYP2C19. Das hat angesichts der Häufigkeit von Multipharmazie im klinischen Alltag hohe klinische Relevanz, betrifft in erhöhtem Maße ältere Patienten und wird immer wieder übersehen. In klinischen Studien spielen diese Interaktionen in der Regel keine Rolle, da eine Behandlung mit interaktionsträchtigen Medikamenten ausgeschlossen wird.

## **2.2 Nutzenbetrachtung**

Auf den Seiten 609 unten bis 614 oben wird die Nutzenbewertung bei Antidepressiva diskutiert. Es fehlen, auch im Hinblick auf die in die Untersuchung einbezogene Studien, eine Betrachtung psychotherapeutischer Behandlungsmöglichkeiten, sowie nicht-medikamentöser Therapieoptionen, die – wie eingangs ausgeführt wird – als mögliche Zusatztherapien zur medikamentösen Therapie angewendet werden können.

## **2.3 Methodisches Vorgehen**

Seite 10, Kriterien für den Einschluss von Studien in die Untersuchung. Bezüglich dieses zentralen Punktes bleiben mehrere wichtige Fragen offen:

- 4.1.2 Intervention und Vergleichsbehandlung, 2. Absatz: „Als Vergleichsinterventionen wurde eine Placebobehandlung oder andere in der ambulanten Versorgung relevante, chemisch definierte Antidepressiva sowie Johanniskraut betrachtet.“ Warum „ambulante Versorgung“, wenn aus der auf Seite 26 aufgeführten Subgruppenanalyse „Versorgungssetting (stationär vs. fachärztlich ambulant vs. nicht fachärztlich ambulant)“ eindeutig hervorgeht, dass auch stationär durchgeführte Studien berücksichtigt wurden?
- Wie ist in diesem Zusammenhang „relevant“ definiert?
- Studien wurden akzeptiert, wenn die Dosierungen den „in Deutschland empfohlenen Dosierungen“ entsprachen. Auf welche „Empfehlungen“ bezieht sich diese Einschätzung. Dieser Punkt scheint besonders relevant, da mehrere doppelblinde Studien aus nicht nachvollziehbaren Gründen ausgeschlossen wurden (s. unten).

## Bewertung der Studien- und Publikationsqualität (Seite 18/19)

- Vier Ausprägungen von möglichen Studienqualitäten werden beschrieben. Eine der Ausprägung ist mit dem Wort „unklar“ benannt. Diese Ausprägung wird nicht definiert und im Weiteren nicht erläutert.
- Auf den Seiten 18/19 wird dargestellt, wie bei einer grob mangelhaften Studienqualität vorgegangen wurde. Das beschriebene Vorgehen ist jedoch nicht nachvollziehbar (Seite 19, 3.-5. Zeile: „...war die Schlussfolgerung damit eventuell herunterzustufen. Bei Einzelstudien in Meta-Analysen musste der Gesamteffekt ohne diese Studie betrachtet und eventuell heruntergestuft werden.“), da unklar ist, auf welche „Stufen“ sich „herunterzustufen“ bezieht.
- Im Übrigen ist nicht erläutert, wie die prozentuale Abweichung von über 30 % („Patienten in Auswertung gegenüber der Randomisierungspopulation“) als Grenzwert für eine grob mangelhafte Studienqualität begründet ist (Seite 18 letzter Absatz, Seite 19 erster Absatz).
- Bei der Subgruppenanalyse wird eine Analyse vermisst, die untersucht, in welchen Ländern mit welchem kulturellen Hintergrund die Studien durchgeführt wurden, um die Übertragbarkeit der Ergebnisse auf die deutsche Versorgungsrealität beurteilen zu können.

## Ausschluss von Studien (Anhang B Seite 675, Volltext) in Verbindung mit 5.1.1 ( Seite 30)

- Sechs Studien wurden in der Volltextanalyse mit der Begründung „keine relevante Vergleichsgruppe“ ausgeschlossen. Bei den Publikationen 3, 5 und 6 lässt sich der genannte Ausschlussgrund nicht nachvollziehen, es handelt sich um randomisierte, doppelblinde Vergleichsstudien gegen Mirtazapin, Imipramin bzw. Bupropion.
- 27 Studien wurden in der Volltextanalyse mit der Begründung „Studien nicht doppelblind, nicht randomisiert (nicht E5)“ ausgeschlossen. Bei Arbeit 23 lässt sich der genannte Ausschlussgrund (nicht-verblindet) nachvollziehen: Die Studie wäre aber für die Diskussion hilfreich gewesen.
- Fünf Studien wurden in der Volltextanalyse wegen zu kurzer Studiendauer (nicht E6) ausgeschlossen. Die Arbeiten 1, 3 und 4 untersuchten bei depressiven Patienten primär Aspekte der Verträglichkeit oder der Lebensqualität. Hier wurden Studien ausgeschlossen, die möglicherweise wichtige Informationen zu klinisch bedeutsamen unerwünschten Wirkungen enthalten. Da weitere 22 Studien aus dem gleichen Grund (ohne Volltextanalyse) ausgeschlossen wurden, ergibt sich die Frage, wie berechtigt der Ausschluss einer Studie aufgrund der Studiendauer ist, wenn der betrachtete Endpunkt nicht die Wirksamkeit, sondern eine andere Zielgröße (z. B. unerwünschte Arzneimittelwirkungen) darstellt.

- Vermisst wird eine Betrachtung von Wirksamkeit und Verträglichkeit der beiden untersuchten Antidepressiva in Abhängigkeit von der Dosierung. Dies scheint insbesondere für Venlafaxin relevant, für das bekannt ist, dass es erst in höherer Dosierung in klinisch relevanter Weise die Noradrenalinaufnahme hemmt (und in niedrigeren Dosierungen also eher als SSRI denn als SNRI wirkt).
- Auch darüber hinaus wäre eine altersdifferenzierte Analyse wünschenswert. Der Nutzen- (und Schadens-)Vergleich zu den trizyklischen Antidepressiva fällt bei älteren Patienten anders als bei jüngeren aus, zum Beispiel in Hinblick auf unerwünschte Ereignisse und die Risiken einer (akzidentellen oder suizidalen) Intoxikation.

## **2.4 Verfügbarkeit relevanter Studien/Daten**

Bei der Diskussion zur Verfügbarkeit relevanter Studien bzw. Daten (6.2, Seite 608) wird in angemessener Weise die unvollständige Überlassung unpublizierter Daten durch die Herstellerfirmen thematisiert. Allerdings bleibt die Schlussfolgerung, es lägen „...keine Hinweise [vor], dass dem IQWiG relevante Mengen an Daten fehlten“ auf dieser Basis nicht nachvollziehbar.

## **3. Zusammenfassende Bewertung**

Das globale Ergebnis des Vorberichts Version 1.0 Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen des IQWiG deckt sich mit der Bewertung der Wirkstoffklasse, die durch die Arzneimittelkommission der deutschen Ärzteschaft u. a. in „Wirkstoff AKTUELL Venlafaxin“ dargestellt wurde. Dennoch weist der Vorbericht Version 1.0 die oben genannten Mängel auf,.

### **Zusätzliche Bemerkung**

Bedauerlicherweise war der Zeitrahmen für die Kommentierung des fast 800 Seiten umfassenden Vorberichts äußerst knapp bemessen. Wegen der an sich jahreszeitlich bedingten Feiertage standen noch weniger Arbeitstage zur Verfügung als im Allgemeinen für die Kommentierung von IQWiG-Vorberichten vorgesehen. Dies bedingte, dass eine Nachrecherche von Studien und damit eine an kritischen Stellen auf der Ebene von Primärdaten durchgeführte Überprüfung nicht möglich war. Im Sinne einer angemessenen Kommentierung durch

weite Fachkreise wäre es wünschenswert, wenn das IQWiG in Zukunft auf eine ausreichende Bearbeitungszeit zur Kommentierung achten würde.



**A 1.2     Institut für Arzneimittelsicherheit in der Psychiatrie e. V. (AMSP)**

**Autoren:**

Rüther, Eckart, Prof. Dr. med.

**Adresse:**

Prof. Dr. Eckart Rüther

Institut für Arzneimittelsicherheit in der Psychiatrie e. V. (AMSP)

Psychiatrische Klinik der LM-Universität

Nussbaumstraße 7

80336 München

AMSP, Prof. Dr. E. Rüter, Psychiatrische Uniklinik, Nussbaumstr.7, 80336 München

**IQWiG**  
**Stellungnahme zum Vorbericht A05-20 A**  
**Prof. Dr. Peter Sawicki**  
**Dillenburgerstr.27**  
**55105 Köln**

Vorsitzender  
 Prof. Dr. Eckart Rüter  
 Psychiatrische Klinik der  
 LM - Universität  
 Nussbaum-Straße 7  
 D-80336 München  
 Telefon: 089 5160 5752  
 Telefax: 089 5160 5562  
 eruethe@gwdg.de

02.01.2009

## **Schriftliche Stellungnahme zum Vorbericht Selektive Serotonin- und Noradrenalin- Wieder aufnamehemmer ( SNRI ) bei Patienten mit Depression (Auftrag A05-20 A)**

Als ehemaliger Direktor der Klinik für Psychiatrie und Psychotherapie der Universität Göttingen erlaube ich mir jetzt in meiner Eigenschaft als Vorsitzender und Begründer des o.g. Instituts AMSP zum o.g. Vorbericht Stellung zu nehmen:

- 1) Der Vorbericht ist sehr sorgfältig gefertigt und zeugt von hoher wissenschaftlicher Kompetenz der Autoren. Leider ist die Auswahl der für die Bewertung ausgewählten Literatur zu sehr an wissenschaftlicher und nicht klinischer Information ausgerichtet.
- 2) Daher erlaube ich mir, auf einige Daten unseres Erfassungssystems AMSP im Folgenden hinzuweisen:
  - 2.1) In Abb. 1 und 2 werden die Anwendungsdaten im Verlauf der letzten Jahre in Psychiatrischen Kliniken ( überwachte Pat etwa 20 00/ Jahr ) dargestellt. Besonders Venlafaxin aber auch neuestens Duloxetin haben in diesen Kliniken trotz erheblichen Kostendrucks eine breite Akzeptanz erfahren. Diese Kliniken sind mit den modernsten Kriterien der Nutzen/ Risiken - Abschätzung und des Qualitätsmanagements durch AMSP vertraut und erhalten durch unser System eine ständige Fortbildung. Eine Akzeptanz von Präparaten ist in diesen Kliniken nur zu einem sehr geringen Teil von Moden abhängig und zeugt von begründeter Anwendungsbereitschaft. Diese beiden Präparate scheinen sich in der Praxis der Kliniken alternativlos durchgesetzt zu haben.
  - 2.2 ) In Abb. 3 -5 werden die UAW Profile der einzelnen AD dargestellt. Es wird deutlich, dass die UAW-Profile der AD sehr voneinander unterschieden sind. Das trifft nicht nur im Vergleich von Venlafaxin mit den anderen AD-Gruppen zu, sondern zeigt sich auch im Vergleich der beiden SNRI untereinander. Dies ist besonders wichtig, da eine der wesentlichen Schlussfolgerungen des Vorberichts ist, dass das Risiko der AD Gruppen und Einzelsubstanzen

in sogn. wissenschaftlichen Studien vergleichbar sind. Dies liegt in der Regel darin, dass die Erforschung der UAW nicht auf gleichmachende Ratingsysteme beruhen darf, sondern zumindest ergänzt werden muss durch so sorgfältige Erhebungen wie in dem AMSP System.

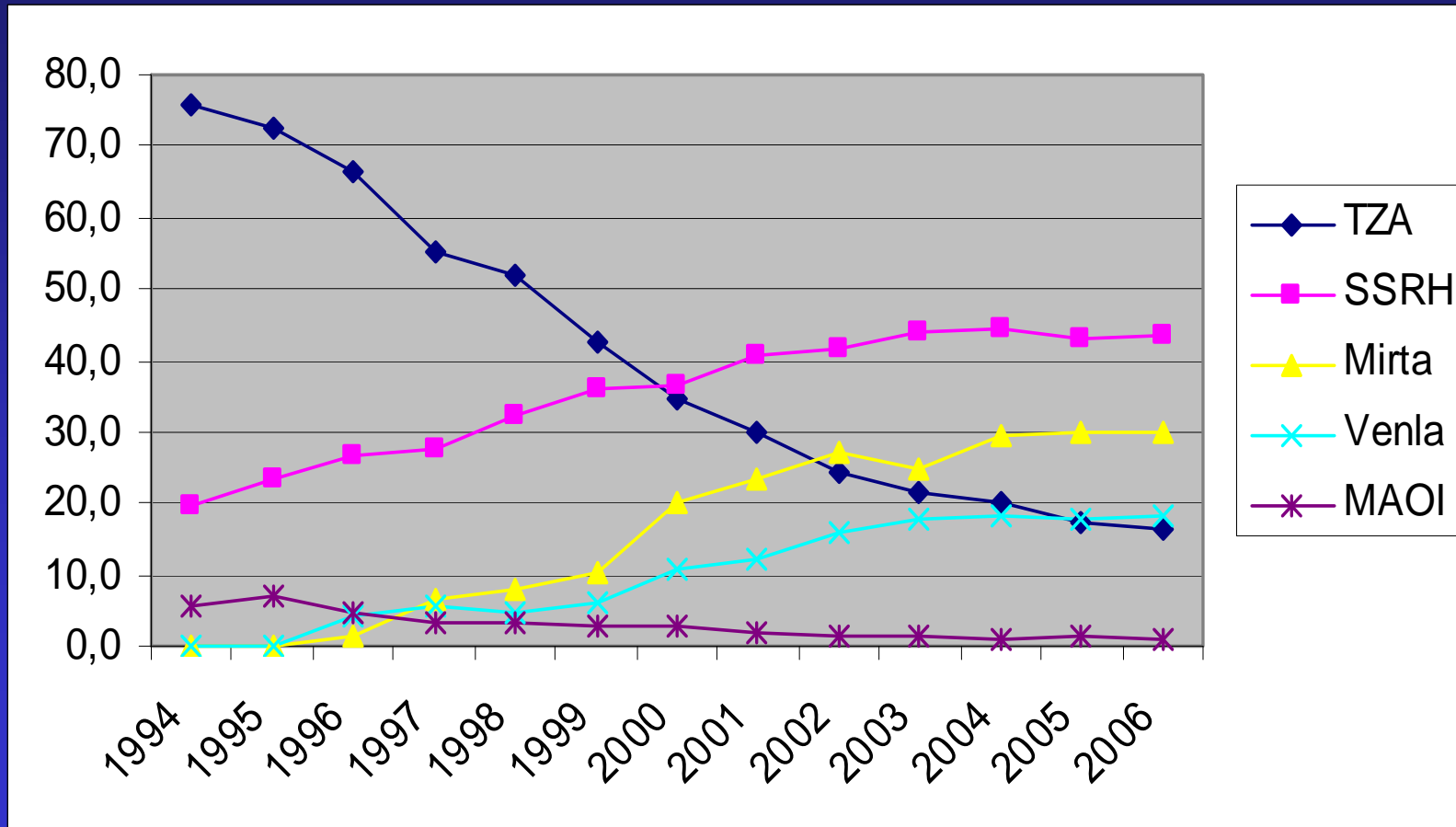
- 3) Als Schlussfolgerung ist vorzuschlagen, dass SNRI eine besondere Substanzgruppe bei vielen PatientInnen alternativlos eine unabdingbare Berechtigung besitzen ,und dass die beiden Substanzen Venlafaxin und Duloxetin jedes für sich als Einzelsubstanz einen eigenen Wert aufzeigen.
- 4) Ich würde mich freuen, wenn diese Argumente im endgültigen Bericht aufgenommen würden. Für die mündliche Anhörung stehe ich gerne zur Verfügung.

Prof Dr. Eckart Rüter

# IQWIG-SNRI Stellungnahme AMSP- Abbildungen 09

A05-20A

## Gebrauch Antidepressiva im Zeitverlauf



# Anwendungsdaten in %

2006

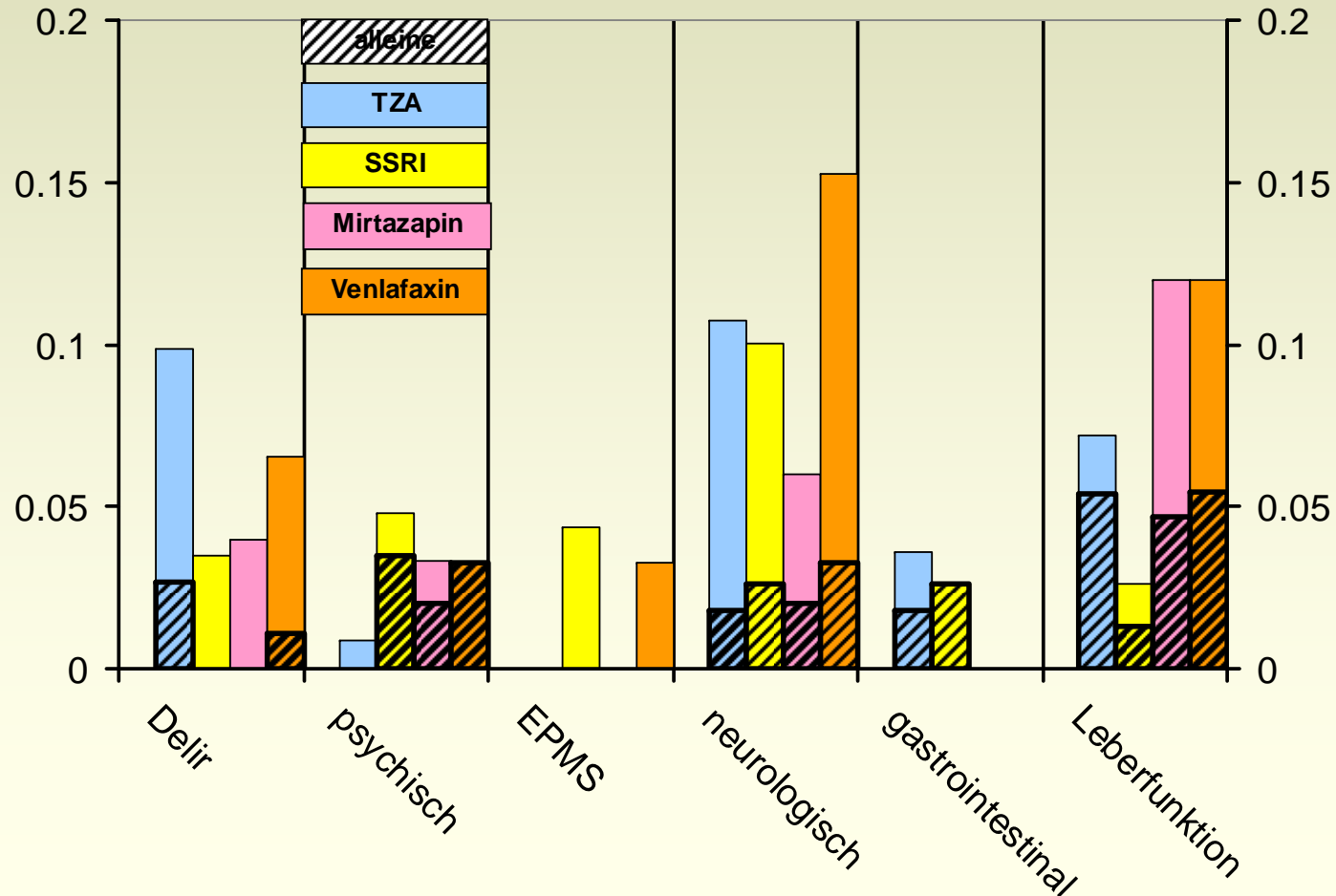
- Mirtazapin 28 %
- Venlafaxin 19 %
- Escitalopram 18 %
- Citalopram 10 %
- Sertralin 8 %
- Trimipramin 6 %
- Duloxetin 5 %

2007

- Mirtazepin 28 %
- Venlafaxin 20 %
- Escitalopram 16 %
- Citalopram 13 %
- Duloxetin 8 %

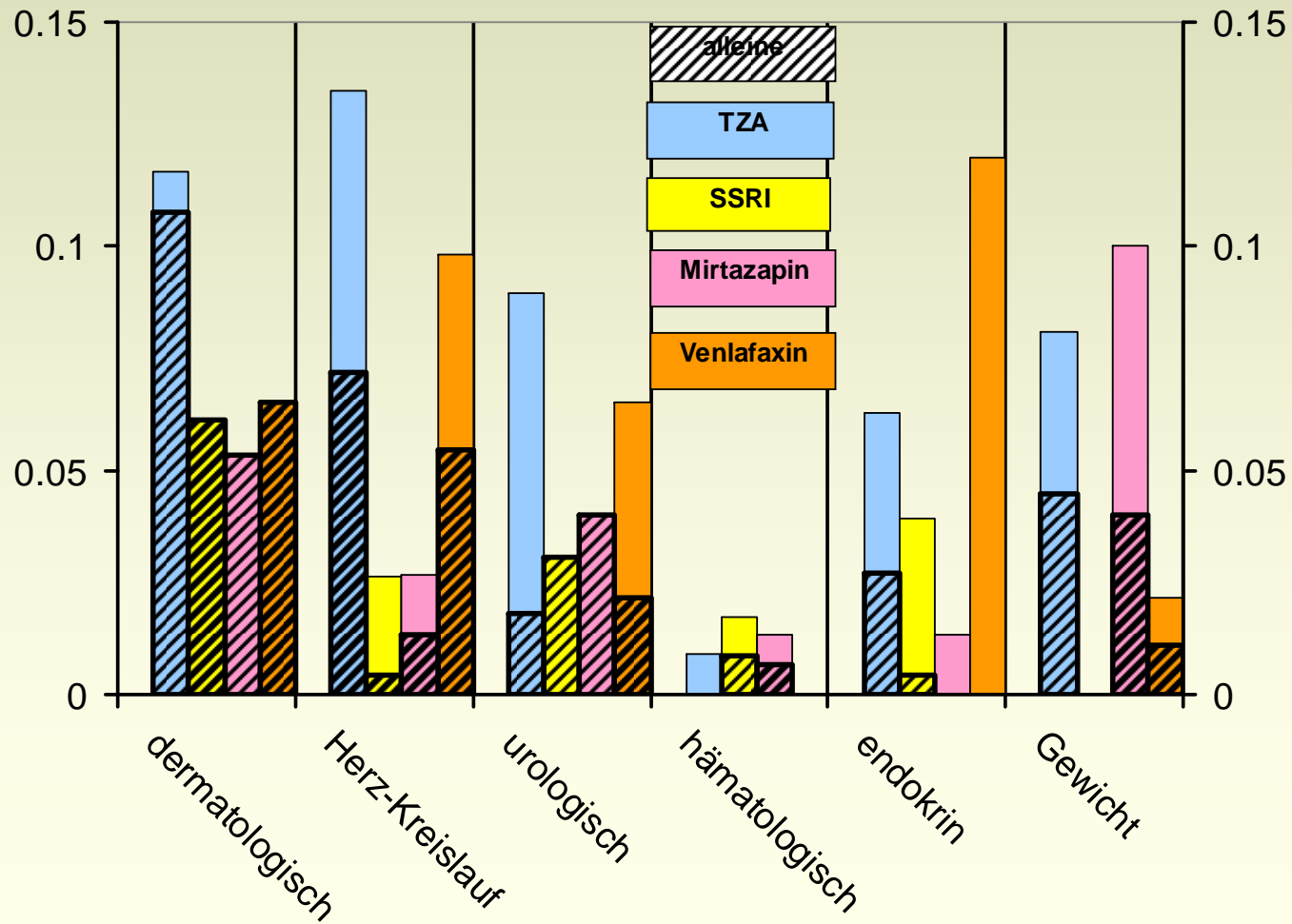
# Schwere UAW – AMSP 2001-2005 (A) Antidepressiva

Anzahl der Fälle in Prozent exponierter Pat. (wahrscheinliche/sichere Fälle)



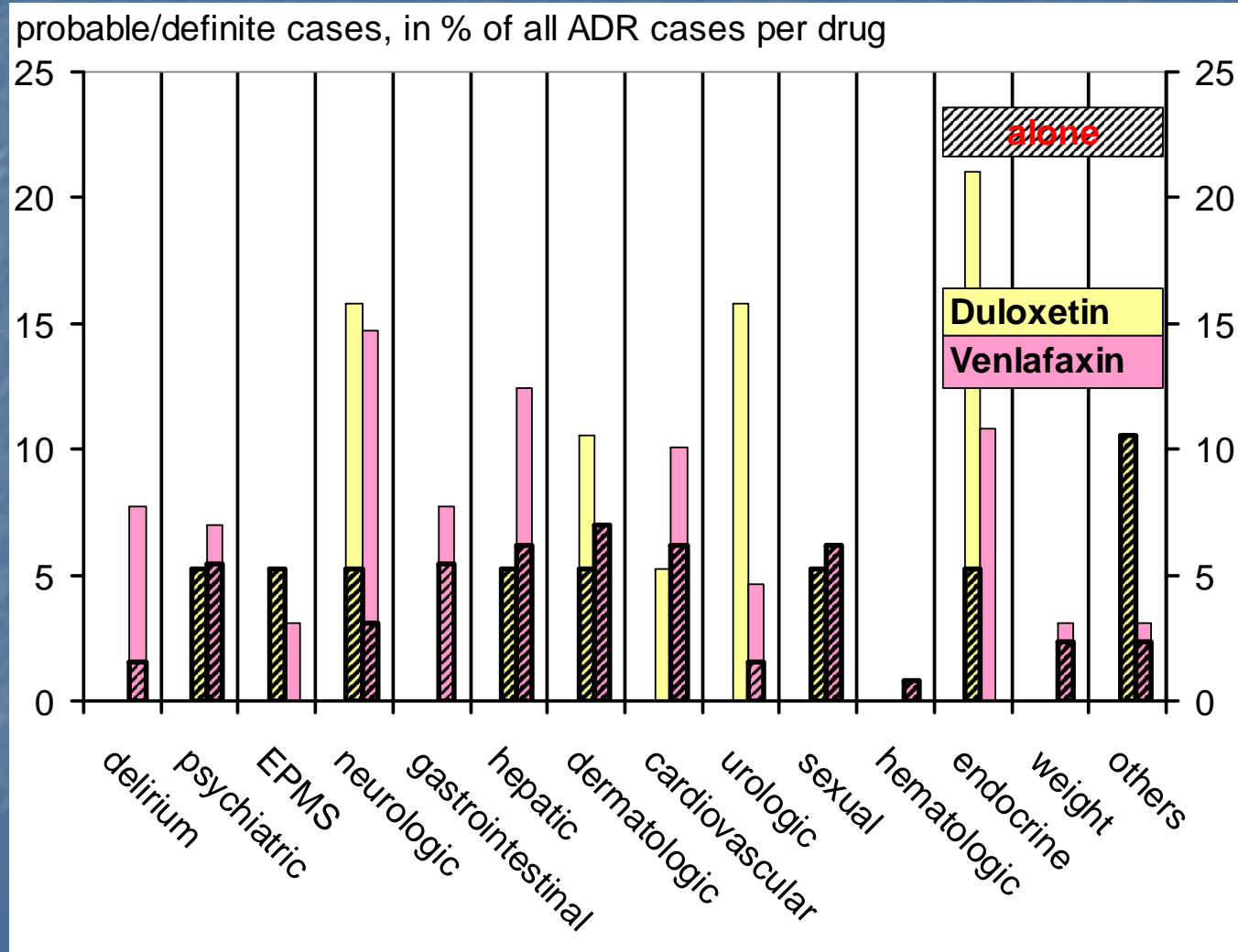
# Schwere UAW – AMSP 2001-2005 (B) Antidepressiva

Anzahl der Fälle pro 10 000 Pat. (wahrscheinliche/sichere Fälle)





# AMSP – AE - Profile



### **A 1.3 Boehringer Ingelheim Pharma GmbH & Co. KG**

**Autoren:**

Weigmann, Harald, Dr. Dr.

Cairns, Victoria, Dr.

Weiser, Martin, Dr.

**Adresse:**

Silke Geier

Boehringer Ingelheim Pharma GmbH & Co. KG

Binger Straße 173

55216 Ingelheim

# Stellungnahme zum Vorbericht der Bewertung A05-20A



## Stellungnahme der Boehringer Ingelheim Pharma GmbH & Co. KG (im Folgenden „Boehringer Ingelheim“) zum Vorbericht der Bewertung A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen“ vom 1. Dezember 2008

### Inhaltsverzeichnis

<b>1. Bewertung und Interpretation der im Vorbericht eingeschlossenen Studien.....</b>	<b>3</b>
1.1 Schmerz im Rahmen der Depression: Meta-Analyse der Placebo-kontrollierten Studien mit VAS (Visuelle Analogskala Schmerz)- und BPI (Brief Pain Inventory)-Daten .....	3
1.2 Schmerz im Rahmen der Depression: Meta-Analyse der Paroxetin-kontrollierten Studien.....	6
1.3 Bewertung der Wirksamkeit von Duloxetin versus SSRIs nach Schweregrad.....	8
1.4 Bewertung der Behandlung älterer Patienten mit Duloxetin .....	9
1.5 Bewertung des Risiko-Profiles von Duloxetin versus SSRIs .....	10
1.6 Kardiovaskuläres Risikoprofil von Duloxetin und Venlafaxin .....	11
1.7 Sexuelle Dysfunktion - Duloxetin versus Paroxetin .....	12
1.8 Problematik des Antidepressiva-Switching.....	12
<b>2. Benennung von zusätzlichen, im Vorbericht 1.0 nicht genannten, relevanten Studien .....</b>	<b>13</b>
<b>3. Anmerkungen zur projektspezifischen Methodik unter jeweiliger Angabe wissenschaftlicher Literatur zur Begründung der Anmerkung insbesondere zur klinischen Relevanz und zum Scoping-Prozess .....</b>	<b>14</b>
<b>4. Fazit .....</b>	<b>15</b>
<b>Literaturverzeichnis.....</b>	<b>16</b>

IQWiG A05-20A Vorbericht:  
Stellungnahme der Boehringer Ingelheim Pharma GmbH & Co. KG  
9. Januar 2009

**Einleitung**

Das IQWiG hat am 1. Dezember 2008 den Vorbericht zum Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahme-Hemmer (SNRI) bei Patienten mit Depressionen“ veröffentlicht.

Boehringer Ingelheim hat sich, ebenso wie andere Unternehmen der deutschen pharmazeutischen Industrie und wie auch der Verband der forschenden Arzneimittelhersteller (VFA), ausführlich schriftlich und mündlich (im Rahmen der Anhörung am 27. November 2007) zu den Berichtsplänen 1.0 und 2.0 geäußert.

Zusätzlich zu den bereits abgegebenen Stellungnahmen von Boehringer Ingelheim zum Berichtsplan 1.0 und 2.0 – die dieser Stellungnahme als Anlage beigefügt sind – kommentiert Boehringer Ingelheim hiermit den Vorbericht. Wir verweisen ergänzend auf die bestehenden, unserer Ansicht nach weiterhin berechtigten Verbesserungsvorschläge an den IQWiG-Berichtsplänen.

## **1. Bewertung und Interpretation der im Vorbericht eingeschlossenen Studien**

### **1.1 Schmerz im Rahmen der Depression: Meta-Analyse der Placebo-kontrollierten Studien mit VAS (Visuelle Analogskala Schmerz)- und BPI (Brief Pain Inventory)-Daten**

In die sieben Placebo-kontrollierten Studien der VAS Meta-Analyse (Vorbericht, Seite 126) wurden depressive Patienten unabhängig vom Vorhandensein von Schmerzen eingeschlossen. Daraus resultiert ein gemischtes Patientenkollektiv mit unterschiedlichen Schmerzgraden, einschließlich Patienten ohne oder mit nur leichten Schmerzen. Dies zeigen auch die VAS Baseline-Mittelwerte, die nur geringen Schmerzgraden entsprachen (siehe Tabelle 1). Dabei sind Werte von  $< 30$  mm auf der VAS als klinisch nicht relevant anzusehen (Collins et al., 1997). Der Median der VAS liegt bei sechs der sieben Studien unter diesen 30 mm, demzufolge leiden mehr als die Hälfte der Patienten in diesen sechs Studien nicht an klinisch relevanten Schmerzen. Daraus folgt, dass hier der schmerzlindernde Effekt von Duloxetin nur eingeschränkt untersucht werden kann und somit die Größe des Effekts tendenziell unterschätzt wird.

Die Meta-Analyse der sieben Placebo-kontrollierten Studien zeigt, dass selbst bei diesem Patientenkollektiv ein statistisch signifikanter schmerzlindernder Behandlungseffekt existiert. Um aber die klinische Relevanz des Effektes adäquat beurteilen zu können, sollten speziell Patienten mit klinisch relevanten Schmerzen zu Beginn der Behandlung berücksichtigt werden.

Tabelle 1: Mittlere und mediane Schmerzen zur Baseline für die sieben Placebo-kontrollierten Studien, die in die Meta-Analyse eingeschlossen wurden

Studie	Duloxetin Patienten Tagesdosis: Baseline-Mittelwert für Schmerz (Median)	Placebo-Patienten Tagesdosis: Baseline-Mittelwert für Schmerz (Median)
Ohne Berücksichtigung von Schmerz zur Baseline (Overall Pain VAS)		
Detke 2002a (HMBHb)	60mg qd: 24,7 (16)	25,6 (19)
Detke 2002b (HMBHa)	60mg qd: 29,2 (24)	27,8 (19,5)
Detke 2004 (HMAYa)	40mg bid: 32,2 (28)	32,0 (28)
	60mg bid: 33,6 (25)	
Goldstein 2004 (HMATb)	20mg bid: 27,0 (18)	24,2 (15)
	40mg bid: 25,6 (18,5)	
HMATa	20mg bid: 28,4 (21)	18,6 (10,5)
	40mg bid: 22,2 (19)	
Perahia 2006b (HMAYb)	40mg bid: 39,4 (37)	34,5 (29)
	60mg bid: 36,2 (31)	
Raskin 2007 (HMBV)	60mg qd: 30,1 (26)	33,5 (27,5)

Erläuterung: qd = quaque die = einmal täglich, bid = bis in die = zweimal täglich

Zwei Placebo-kontrollierte Vergleichsstudien, die Patienten mit zumindest leichten Schmerzen zu Studienbeginn untersuchten, wurden vom IQWiG in seiner Meta-Analyse nicht berücksichtigt (Brannan et al., 2005; Brecht et al., 2007, siehe Tabelle 2). Die Schmerzstärke zur Baseline war in diesen Studien im Mittel deutlich höher im Vergleich zu den oben genannten Studien.

Tabelle 2: Mittlere Schmerzen zur Baseline für die zwei Placebo-kontrollierten Studien, die nicht in die Meta-Analyse des IQWiG eingeschlossen wurden

Studie	Duloxetin-Patienten Baseline-Mittelwerte	Placebo-Patienten Baseline-Mittelwerte
Nur Patienten mit Schmerz zur Baseline (BPI)		
Brannan 2005 (HMCB) [Baseline BPI ≥ 2]	BPI: 4,9 VAS: 49,8	BPI: 4,6 VAS: 46,8
Brecht 2007 (HMDH) [Baseline BPI ≥ 3]	BPI: 5,7 (x10 = 57)	BPI: 5,7 (x10 = 57)

Das IQWiG analysiert diese beiden Studien mit folgender Begründung nicht weiter (Vorbericht, Seite 128):

*"Eine Meta-Analyse beider Studien zur Zielgröße Remission im Vergleich von Duloxetin und Placebo wies eine zu hohe Heterogenität auf, um einen Gesamtschätzer zu ermitteln. Diese Heterogenität war in den gegenläufigen Ergebnissen der beiden Studien begründet. Es erfolgte somit keine weitere Analyse bzw. Darstellung und Diskussion der BPI-Daten zum Pool der beiden Studien, ebenso erfolgte keine Diskussion der Einzelergebnisse."*

Boehringer Ingelheim betrachtet diesen Aspekt aus zwei Gründen differenzierter: Zum einen ist bekannt, dass klinische Studien mit Antidepressiva häufiger keinen statistisch signifikanten therapeutischen Wirksamkeitsunterschied zwischen Verum und Placebo zeigen können (vgl. auch Kapitel 6.3 „Besonderheiten der Nutzenbewertung bei Antidepressiva“, Vorbericht, S. 609), zum anderen waren die Depressions-Remissionsraten für Duloxetin vs. Placebo in den zwei Studien nicht "gegenläufig", sondern ähnlich in HMCB (Brannan et al., 2005: 35% vs. 32%) und besser für Duloxetin in HMDH (Brecht et al., 2007: 53% vs. 29%). Hierbei ist festzustellen, dass in diesen Studien die Remission der depressiven Symptomatik nur ein sekundärer Endpunkt war. Primärer Endpunkt war in beiden Studien die Verbesserung der Schmerzsymptomatik, gemessen durch das Brief Pain Inventory (BPI).

Aus diesen Gründen halten wir die Argumentation des IQWiG nicht für schlüssig. Es ist nicht nachvollziehbar, warum ein in einer Studie negatives Ergebnis in einem sekundären

Endpunkt (Depression) die Analyse des primären Endpunktes beider Studien (Schmerz) nicht erlaubt. Duloxetin ist unseres Wissens auch der einzige SNRI, für den eine Evidenz bezüglich Schmerzreduktion bei depressiven Patienten besteht, basierend auf randomisierten klinischen Studien mit Schmerz im Rahmen der Depression als primärem Endpunkt (Brannan et al., 2005; Brecht et al., 2007). Eine Meta-Analyse dieser beiden Studien zeigt bezüglich des primären Studienendpunktes BPI Average Pain folgendes Ergebnis (random effects Modell):

Schmerz WMD (weighted mean difference; 95% KI): -0,73 (-1,14; -0,33)

Heterogenität:  $I^2 = 4,7\%$

Test für Gesamtbehandlungseffekt:  $p = 0,0003$

Es zeigt sich somit ein statistisch signifikanter Effekt zugunsten von Duloxetin. Weiterhin findet sich bei der Analyse dieses Endpunktes nicht die vom IQWiG bei der Depression kritisierte Heterogenität, sondern vielmehr eine als homogen einzuschätzende Datenbasis ( $I^2 = 4,7\%$ ).

Die Effektstärke, gemessen in Standardabweichungen, liegt für die zwei gepoolten Studien bei -0,29 (95% KI -0,13; -0,46). Dies ist, wie bei dem selektierten Kollektiv zu erwarten, eine deutlich höhere Effektstärke als bei den 7 Studien, die Patienten unabhängig vom Vorhandensein von Schmerzen einschlossen.

## **1.2 Schmerz im Rahmen der Depression: Meta-Analyse der Paroxetin-kontrollierten Studien**

Das oben Gesagte gilt gleichermaßen auch für den Vergleich mit einer aktiven Vergleichssubstanz. Hier sollten zur Beurteilung der klinischen Relevanz ebenfalls Studien berücksichtigt werden, die spezifisch Patienten mit klinisch relevanten Schmerzen einschlossen. Dies war bei den Paroxetin-kontrollierten Studien ebenfalls nicht der Fall.

Für die Meta-Analyse von Duloxetin gegenüber Paroxetin bzgl. Schmerz (Vorbericht, Seite 127) wurden fünf Studien analysiert, in die ebenfalls Patienten unabhängig vom Vorhandensein von Schmerzen eingeschlossen wurden. Die Mediane für VAS Overall-



Pain zur Baseline waren niedrig - sie lagen zwischen 11 und 37 auf einer Skala von 0 bis 100. Wie bei den Placebo-kontrollierten Studien wiesen auch hier die Hälfte der Patienten zur Baseline keine oder nur geringe Schmerzen auf (Tabelle 3).

Tabelle 3: Mittlere und mediane Schmerzen zur Baseline für die 5 Paroxetin-kontrollierten Studien, die in die Meta-Analyse eingeschlossen wurden

Studie	Duloxetin Patienten Tagesdosis: Baseline-Mittelwert für Schmerz (Median)	Placebo-Patienten Tagesdosis: Baseline-Mittelwert für Schmerz (Median)
Ohne Berücksichtigung Schmerz an Baseline (Overall Pain VAS)		
Detke 2004 (HMAYa)	40mg bid: 32,2 (28) 60mg bid: 33,6 (25)	20mg qd: 37,6 (28)
Goldstein 2004 (HMATb)	20mg bid: 27,0 (18) 40mg bid: 25,6 (18,5)	20mg qd: 22,2 (15)
HMATa	20mg bid: 28,4 (21) 40mg bid: 22,2 (19)	20mg qd: 20,4 (11)
Perahia 2006b (HMAYb)	40mg bid: 39,4 (37) 60mg bid: 36,2 (31)	20mg qd: 35,5 (32)
Lee 2007 (HMCV)	60mg qd: 34,7 (29)	20mg qd: 33,6 (30)

Die Auswertungen der vorliegenden Daten durch das IQWiG erscheinen uns unzureichend, um den tatsächlich bestehenden Unterschied zwischen Duloxetin und Paroxetin beurteilen zu können. Es wäre zumindest notwendig, vergleichbar mit den oben geschilderten Studien (Brannan et al., 2005; Brecht et al., 2007), ein selektiertes Patientenkollektiv mit relevanten Schmerzen im Rahmen der Depression zur Baseline auszuwerten.

Duloxetin zeigt versus Placebo eine Effektgröße von 0,29 bezüglich der Reduktion von Schmerzen. Für Paroxetin ist dies für Schmerzen im Rahmen der Depression unseres Wissens nach nicht belegt.

### **1.3 Bewertung der Wirksamkeit von Duloxetin versus SSRIs nach Schweregrad**

Das IQWiG sieht keinen Beleg eines Zusatznutzens von Duloxetin gegenüber den SSRIs in der ambulanten Akuttherapie der Depression bei höherer Erkrankungsschwere. Boehringer Ingelheim teilt diese Ansicht aus statistischen wie inhaltlichen Gründen nicht.

Als Maß für den Schweregrad der Erkrankung der Patienten in den einzelnen Studien wurde in der IQWiG-Bewertung der Mittelwert des HAMD-Gesamtscores zu Studienbeginn verwendet. Die Studien wurden in zwei Gruppen nach niedrigeren bzw. höheren Mittelwerten zur Baseline geclustert (vgl. Vorbericht, Kap. 5.2.3.2). Durch diese Art der Gruppierung nach Mittelwerten gelingt aber keine klare Trennung der beiden Gruppen: Es finden sich in der "niedrig" Gruppe auch Patienten mit schwerer Depression und umgekehrt.

Im Cochrane Handbook for Systematic Reviews of Interventions (Section 18.4.5) wird zu dieser Problematik wie folgt Stellung genommen (Higgins & Green, 2008): *"In conventional analyses using aggregate data from publications, it is usually very difficult to extract sufficient compatible data to undertake meaningful subgroup analyses"*. In den in der Zeitschrift Neuropsychopharmacology im Jahr 2005 erschienenen Consensus Leitlinien schreiben Liebermann und Kollegen (2005): *"Given the loss of information that occurs when the unit of observation is the entire study rather than the individual patient, many experts maintain that pooled analysis of the original patient data is superior to traditional meta-analysis of RCTs ... This potential advantage is most evident when the aim of the research is to investigate interactions between selected patient characteristics and specific treatments."*

Diese Limitierung wird auch vom IQWiG diskutiert: *„Gleichzeitig ist die Aussagekraft dieser Analyse aber begrenzt, da die Unterteilung in die Schweregrad-Subgruppen nicht auf Basis individueller Patientendaten, sondern ausgehend von Mittelwerten der Studienpopulationen vorgenommen wurde.“* (Vorbericht, S.517)

In der IQWiG-Meta-Regression fand sich in der Gruppe mit zu Beginn niedrigerem mittleren HAMD-Wert ein p-Wert von 0,052 für die Remission zum Vorteil von Duloxetin, aber es fand sich kein Vorteil in der Gruppe mit zu Beginn höherem mittleren HAMD-Wert. Dieses Ergebnis steht im Widerspruch zu den Ergebnissen von Thase und Kollegen (2007), einer gepoolten Analyse auf Basis individueller Patientendaten, die zu dem Schluss kommt, dass Duloxetin einen statistisch signifikanten Vorteil gegenüber Fluoxetin und Paroxetin in der Therapie von Patienten mit höherem Schweregrad aufweist. Die IQWiG-Auswertungen mit Studien-Gruppierungen basierend auf Mittelwerten sind daher möglicherweise irreführend.

#### **1.4 Bewertung der Behandlung älterer Patienten mit Duloxetin**

Das IQWiG schreibt zur Subgruppen-Analyse älterer depressiver Patienten (Vorbericht, Kapitel 5.2.3.6): *„Aufgrund der homogenen Datenlage in den allgemeinen Meta-Analysen [...] wurde von einer separaten Analyse zum Alterseffekt abgesehen. Es gab keinen Hinweis für eine maßgebliche Abweichung in Bezug auf die antidepressive Wirkung bei älteren Patienten.“*

Nach unserer Ansicht würdigt die gegenwärtige Formulierung und Darstellung die gute Datenlage zu Duloxetin nicht hinreichend (Raskin et al., 2007; vgl. hierzu auch Vorbericht, Kapitel 5.3.3.6).

Katona und Livingston (2002) sowie Taylor und Doraiswamy (2004) haben einen Vergleich von Antidepressiva auf Basis der Number Needed to Treat (NNT) bei Patienten  $\geq 55$  Jahre vorgenommen. Für den Parameter Ansprechen weist Duloxetin (Raskin et al., 2007) im Vergleich zu Placebo eine NNT von 5,6 bei Patienten  $\geq 65$  Jahre auf. Das ist deutlich günstiger als eine NNT von 8 für die SSRI in der Übersicht von Taylor und Doraiswamy's (2004). Für die älteste Subgruppe (älter als 75 Jahre) liegt die NNT mit einem Wert von 4,5 noch niedriger.

Die Duloxetin-Studie bei Patienten  $\geq 65$  Jahre (Raskin et al, 2007; Wise et al 2007) ist aus einem weiteren Grund erwähnenswert, da sie auch den Aspekt Kognition (Vorbericht, Kapitel 5.2.2.3.3) anhand des „Composite Cognitive Score“, einem zusammengesetzten Instrument aus verschiedenen Skalen, behandelt. Wie auch das IQWiG feststellt, zeigt der

Composite Cognitive Score, der als den primären Endpunkt definiert war, „*einen statistisch signifikanten Unterschied im Mittelwert zum Ende der Studie zugunsten von Duloxetin.*“ (Vorbericht, S. 130). Das IQWiG lässt dieses Ergebnis allerdings unberücksichtigt, da „*es sich in seiner Kombination verschiedener Skalen und deren Aufsummierung nicht um ein validiertes Instrument*“ (Vorbericht, S. 130) handeln würde. Diese Skalen sind ausgewählt worden, um verschiedene Kognitionsaspekte zusammenfassend zu beleuchten: „*they assess aspects of cognition shown previously to be most impaired in patients with depression, specifically verbal learning and memory, attention, executive function and working memory* (Wise et al., 2007. S. 1284).“

Insbesondere weisen wir darauf hin, dass Duloxetin gerade von älteren Patienten gut vertragen wird, was sich in Abbruchraten wegen unerwünschter Ereignisse, die auf Placebo-Niveau (9,7% bzw. 8,7%) liegen, ausdrückt (Raskin et al., 2007).

In der Konsequenz ist Duloxetin bei der Behandlung Älterer wirksam, gut verträglich, und wirkt sich positiv auf die Kognition der Patienten aus.

## **1.5 Bewertung des Risiko-Profiles von Duloxetin versus SSRIs**

In seiner Bewertung kommt das IQWiG zu der Schlussfolgerung, dass für die Endpunkte Gesamtrate unerwünschter Ereignisse und Therapieabbrüche wegen unerwünschter Ereignisse ein Schaden von Duloxetin im Vergleich zu SSRI in der Kurzzeit-Akuttherapie belegt sei.

Im Gegensatz dazu kam die europäische Zulassungsbehörde EMEA in ihrer Bewertung von Duloxetin zu dem Schluss, dass Duloxetin ein den SSRI vergleichbares „Safety-Profil“ aufweist:

„*Overall the safety profile is as expected for a drug with this pharmacological profile. No apparent, unexpected, serious adverse events have been detected in the extensive MDD program) ... In conclusion, the safety profile of duloxetine does not identify any unexpected or serious adverse reaction which would create special concern, and is comparable to that known for SSRIs.*“ (EMEA, 2005: S. 34+35).

Diese Bewertung erfolgte im Rahmen eines umfassenden Zulassungsprozesses auf einer breiten Datenbasis, in die im Gegensatz zum Vorbericht des IQWiG auch Sicherheitsdaten aus Studien in anderen Indikationen und offenen Langzeitstudien eingingen.

Da das IQWiG nur einen kleinen Teil der zur Verfügung stehenden Studien zur Bewertung des Nebenwirkungsprofils heranzieht, ist die Möglichkeit der Verzerrung oder der nur eingeschränkten Wertung des Gesamtbildes gegeben. Wir wiesen auf diesen Punkt bereits in unserer Stellungnahme zum Berichtsplan unter Punkt 2 hin: *„Durch den a-priori Ausschluss von nicht-RCTs wird nur eine begrenzte Sicht vor allem auf den patientenrelevanten Zusatznutzen eines Arzneimittels möglich (..)“*.

Die Wichtigkeit dieses Punktes lässt sich speziell an dem beschriebenen Beispiel des Nutzen-Risiko-Verhältnisses von Duloxetin versus SSRIs zeigen.

Abschnitt 1.5 des IQWiG-Berichtes lässt völlig außer Acht, dass das Verträglichkeitsprofil durch klinische Studien aufgrund einer limitierten Patientenzahl nur unzureichend abgebildet werden kann. Aufgrund der breiteren Datenbasis sind Meldungen aus dem Vertrieb (sog. "Spontanberichte") wichtige Erkenntnisquellen. Auch wenn die Qualität dieser Meldungen oftmals sehr schwankt und keine Vergleichsgruppe existiert, ergeben sich die meisten sicherheitsrelevanten Änderungen der Fach- und Gebrauchsinformation aus solchen Meldungen nach der Zulassung. Die sicherheitsrelevanten Abschnitte der SSRIs und SNRIs weisen unserer Sicht keine medizinisch relevanten Unterschiede von Duloxetin auf. Zur Beurteilung von unerwünschten Ereignissen ist unseres Erachtens eine breitere Datenbasis unabdingbar.

## **1.6 Kardiovaskuläres Risikoprofil von Duloxetin und Venlafaxin**

In einer unlängst erschienenen Studie in der Patienten mit und ohne kardiovaskuläre Erkrankungen unter der Therapie mit Duloxetin untersucht wurden, zeigte sich zwischen den Patienten mit bzw. ohne kardiovaskuläre Erkrankung kein Unterschied (Wernicke et al., 2008). Diese Ergebnisse sprechen für eine gute kardiovaskuläre Verträglichkeit von Duloxetin.

Auch dieses Beispiel zeigt, dass eine Beschränkung auf die Auswertung von unerwünschten Ereignissen nur aus RCTs oft nicht ausreicht um international anerkannte,

elementare Unterschiede in der Verträglichkeit von Substanzen hinreichend würdigen zu können.

In seinem Vorbericht wertete das IQWiG auch Studien zum direkten Vergleich von Duloxetin und Venlafaxin bezüglich kardiovaskulärer Parameter aus, ohne einen klaren Vorteil oder Nachteil einer Substanz entdecken zu können (S. 592). Dies steht im Gegensatz zu Ergebnissen aus klinischen Studien und internationalen Einschätzungen. (Wernicke et al., 2008, NICE 2007)

### **1.7 Sexuelle Dysfunktion - Duloxetin versus Paroxetin**

Das IQWiG schreibt, dass - basierend auf seiner Metaanalyse - der Vergleich von Duloxetin und SSRI im Hinblick auf die sexuelle Dysfunktion keinen statistisch signifikanten Unterschied ergeben habe (Vorbericht, S. 196).

Zumindest für den Vergleich zu Paroxetin existiert eine gepoolte Analyse (mit den vier Paroxetin-Vergleichsstudien, die auch in Abbildung 58 genannt sind), die zu einem dem IQWiG gegenteiligen Ergebnis kommt: „*The incidence of acute treatment-emergent sexual dysfunction was significantly lower among duloxetine-treated patients compared with those receiving paroxetine (p = .015)*“ (Delgado et al., 2005). Dieses Ergebnis sollte vom IQWiG in seiner Bewertung berücksichtigt werden.

### **1.8 Problematik des Antidepressiva-Switching**

Trotz einer großen Anzahl an Therapieoptionen im Bereich der Depression sprechen etwa ein Drittel der Patienten nicht hinreichend auf die erste Therapie an und sind so einem höheren Rückfallrisiko ausgesetzt (Tranter et al., 2002). Zwei Studien (Wohlrreich et al., 2005; Perahia et al., 2008) dokumentieren Duloxetin als eine wertvolle Therapieoption nach Therapieversagen mit anderen Antidepressiva, unabhängig von der Vorbehandlung (keine, Venlafaxin oder SSRI) oder dem Switching-Regime. Der Aspekt der Folgebehandlung bleibt aber vom IQWiG unberücksichtigt. So ist festzustellen, dass eine Nutzenbewertung, die verschiedene Substanzen nur singular betrachtet, der schwierigen Versorgungssituation im psychiatrischen Bereich, insbesondere bei schwer Depressiven mit verschiedenen Therapiekaskaden, nicht gerecht wird (Möller & Maier, 2007).

## **2. Benennung von zusätzlichen, im Vorbericht 1.0 nicht genannten, relevanten Studien**

Im Vorbericht (S. 36) wird die Studie HMDI als nicht abgeschlossene Studie von der Bewertung ausgeschlossen. Entsprechend findet bei der Bewertung der Langzeiteffektivität von Duloxetin nur eine Studie Berücksichtigung (Perahia et al., 2006), in der die Rückfallprävention (Definition nach IQWiG siehe unten) untersucht wurde. Mittlerweile ist die genannte HMDI Studie abgeschlossen. Somit liegen jetzt zusätzliche Daten zur Rezidivprophylaxe (Definition nach IQWiG siehe unten) entsprechend der IQWiG-Definition (Vorbericht, Seite 6) vor. Deshalb reichen wir anbei die erste, bislang publizierte Ergebnispräsentation dieser Studie nach (Perahia et al., 2008a) und stellen bei Bedarf gerne weitere Unterlagen zur Verfügung.

Zusammenfassend war in dieser randomisierten, doppelblinden, 52-wöchigen Studie Duloxetin (60-120 mg/Tag) statistisch signifikant besser als Placebo bezüglich der Zeit bis zum Rezidiv einer Depression (primärer Studienendpunkt).

Wir halten deshalb den Nutzen von Duloxetin auch für die Rezidivprophylaxe als belegt und legen nahe, diese Daten ebenfalls im Vorbericht zu berücksichtigen.

Definitionen der **Rückfall-** und **Rezidivprophylaxe-Studien** nach IQWiG (Auszug Vorbericht, S. 6):

**"Rückfallpräventionsstudien:** *ca. 4 bis 6 Monate doppelblinde Behandlung in Respondern aus der (auch offen und unkontrolliert – wie im Beispiel in der Grafik – durchgeführten) Akutphase nach (Re-)Randomisierung im Anschluss an die Akutphase; Endpunkt: Symptomverschlechterung; bei erstmalig oder rezidivierend erkrankten depressiven Patienten.*

**Studie zur Rezidivprophylaxe:** *mind. 12 Monate doppelblinde Behandlung in Respondern im Anschluss an die Erhaltungsphase; Endpunkt: Symptomverschlechterung; bei erstmalig oder rezidivierend erkrankten depressiven Patienten."*

### **3. Anmerkungen zur projektspezifischen Methodik unter jeweiliger Angabe wissenschaftlicher Literatur zur Begründung der Anmerkung insbesondere zur klinischen Relevanz und zum Scoping-Prozess**

Das IQWiG bewertet die klinische Relevanz der Ergebnisse schwerpunktmäßig unter Bezug auf statistische Maße. Dieser zentrale Punkt wird nur andiskutiert (Vorbericht Seiten 20, 612-613) und wird unseres Erachtens der großen Bedeutung dieses Parameters nicht gerecht. Es bleibt völlig außer Acht, dass in der wissenschaftlichen Literatur weiterhin keine Einigkeit über die beste Methodik zur Bewertung der klinischen Relevanz besteht. Deshalb werden u.a. Scoping Prozesse in nationalen und internationalen Gutachten (Antes et al., 2007, Bekkering & Kleijnen, 2008; NICE, 2008) gefordert, um diese zentrale Frage mit klinisch ausgewiesenen Experten zu diskutieren. Bereits in unserer Stellungnahme zum Berichtsplan 1.0 und 2.0 hatten wir auf die Notwendigkeit solcher Scoping-Prozesse hingewiesen, um eine wissenschaftlich fundierte Diskussion inhaltlich umstrittener, komplexer Parameter die eine Bewertung der patientenrelevanten Wirksamkeit erlaubt, führen zu können.



#### **4. Fazit**

Die Firma Boehringer Ingelheim schätzt am Vorbericht, dass das IQWiG die Wirksamkeit der SNRIs bestätigt hat.

In der Vergangenheit haben wir den Bewertungsprozess des IQWiG zu den SNRIs bereits konstruktiv begleitet. Zusätzlich zu den bereits abgegebenen Stellungnahmen zum Berichtsplan 1.0 und 2.0 möchten wir die folgenden Aspekte noch einmal hervorheben:

- Duloxetin erweist sich nicht nur als wirksam in der Rückfallprophylaxe (Definition siehe 2.), sondern auch in der Rezidivprophylaxe (Definition siehe 2.).
- Duloxetin zeigt in klinischen Studien eine gute Wirksamkeit auf die Schmerzsymptomatik bei Patienten, die zu Beginn der Studien einen relevanten Schmerz aufwiesen. Werden nur Studien ohne dieses Einschlusskriterium zur Beurteilung herangezogen, wird die Wirksamkeit von Duloxetin auf Schmerzen gegenüber SSRI unterschätzt.
- Die Analyse gepoolter Daten aus klinischen Studien zeigt für Duloxetin eine den SSRIs überlegene Wirksamkeit in der Subgruppe der depressiven Patienten mit höherem Schweregrad.
- Die gute Datenlage von Duloxetin in der Behandlung älterer Patienten sollte klarer gewürdigt werden.
- Das Risikoprofil von Duloxetin ist dem der SSRI vergleichbar, bezüglich sexueller Dysfunktionen ist Duloxetin dem Paroxetin überlegen. Duloxetin weist zudem eine gute kardiovaskuläre Verträglichkeit auf.

Abschließend fordern wir erneut einen Scoping-Prozess mit ausgewiesenen klinischen Experten um eine wissenschaftlich fundierte Diskussion inhaltlich umstrittener, komplexer Parameter führen zu können, die eine Bewertung der patientenrelevanten Wirksamkeit erlaubt.

## Literaturverzeichnis

- Antes G, Jöckel KH, Kohlmann T, Raspe H, Wasem J (2007): Kommentierende Synopse der Fachpositionen zur Kosten-Nutzenbewertung von Arzneimitteln – Erstellt im Auftrag des Bundesministeriums für Gesundheit - [http://www.bmg.bund.de/cln\\_110/SharedDocs/Downloads/DE/Neu/Kurzbericht-16-04-08\\_Synopse-Kosten-Nutzen-Arzneimittel,templateId=raw,property=publicationFile.pdf/Kurzbericht-16-04-08\\_Synopse-Kosten-Nutzen-Arzneimittel.pdf](http://www.bmg.bund.de/cln_110/SharedDocs/Downloads/DE/Neu/Kurzbericht-16-04-08_Synopse-Kosten-Nutzen-Arzneimittel,templateId=raw,property=publicationFile.pdf/Kurzbericht-16-04-08_Synopse-Kosten-Nutzen-Arzneimittel.pdf) , (letzter Zugriff: 28.12.2008)
- Brecht S, Courtecuisse C, Debieuvre C, Croenlein J, Desaiyah D, Raskin J (2007): Efficacy and safety of duloxetine 60 mg once daily in the treatment of pain in patients with major depressive disorder and at least moderate pain of unknown etiology: a randomized controlled trial. *J Clin Psychiatry*; 68 (11): 1707-1716
- Bekkering GE, Kleijnen J (2008): Procedures and methods of benefit assessments for medicines in Germany. *Eur J Health Econ*; 9 (Suppl 1): 5-29 bzw. *Dtsch Med Wochenschr.*; 133 (Suppl 7): S225-246
- Brannan SK, Mallinckrodt CH, Brown EB, Wohlreich MM, Watkin JG, Schatzberg AF (2005): Duloxetine 60 mg once-daily in the treatment of painful physical symptoms in patients with major depressive disorder. *J Psychiatr Res.*; 39 (1): 43-53
- Collins SL, Moore RA, McQuay HJ (1997). The visual analogue pain intensity scale what is moderate pain in millimeters? *Pain*; 72 (1-2): 95-97
- Delgado PL, Brannan SK, Mallinckrodt CH, Tran PV, McNamara RK, Wang F, et al. (2005): Sexual functioning assessed in 4 double-blind placebo- and paroxetine-controlled trials of duloxetine for major depressive disorder. *J Clin Psychiatry*; 66 (6): 686-692
- EMA – European Medicines Agencies (2005): European Public Assessment Report: Scientific Discussion Cymbalta, <http://www.emea.europa.eu/humandocs/PDFs/EPAR/cymbalta/19256704en6.pdf>, (letzter Zugriff: 06.01.2009)
- Higgins JPT, Green S (2008): Cochrane Handbook for Systematic Reviews of Interventions: <http://www.cochrane-handbook.org/> (letzter Zugriff: 06.01.2009)
- Katona C, Livingston G (2002): How well do antidepressants work in older people? A systematic review of Number Needed to Treat. *J Affect Disord*; 69 (1-3): 47-52
- Lieberman JA, Greenhouse J, Hamer RM, Krishnan KR, Nemeroff CB, Sheehan DV, et al. (2005): Comparing the effects of antidepressants: Consensus guidelines for evaluating quantitative reviews of antidepressant efficacy. *Neuropsychopharmacology*; 30 (3): 445-460

- Möller HJ, Maier W (2007): Probleme der "evidence based medicine" in der Psychopharmakotherapie. *Nervenarzt*; 8 (9): 1028-1036
- NICE – National Institute for Health and Clinical Excellence (2007): Depression: management of depression in primary and secondary care - NICE guidance, <http://www.nice.org.uk/cg023>, (letzter Zugriff: 06.01.2009)
- NICE – National Institute for Health and Clinical Excellence (2008): Guide to the methods of technology appraisal. <http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf>, (letzter Zugriff: 06.01.2009)
- Perahia DG, Gilaberte I, Wang F, Wiltse CG, Huckins SA, Clemens JW, et al. (2006): Duloxetine in the prevention of relapse of major depressive disorder: double-blind placebo-controlled study. *Br J Psychiatry*; 188 :346-53
- Perahia D, Spann M, Wang F, Walker D, Detke M, Rasmussen K, et al. (2008a): Duloxetine in the Prevention of Depressive Recurrences: A Randomized, Double-Blind, Placebo-Controlled Trial. Poster presented at ACNP; Scottsdale, Arizona, USA; December 7–11
- Perahia DG, Quail D, Desai D, Montejo AL, Schatzberg AF (2008b): Switching to duloxetine in selective serotonin reuptake inhibitor non- and partial-responders: Effects on painful physical symptoms of depression. *J Psychiatr Res*. 2008 Aug 15. [Epub ahead of print]
- Raskin J, Wiltse CG, Siegal A, Sheikh J, Xu J, Dinkel JJ, et al (2007): Efficacy of duloxetine on cognition, depression and pain in elderly patients with major depressive disorder: an 8-week, double-blind, placebo-controlled trial. *Am J Psychiatry*; 164 (6): 900-909
- Taylor WD, Doraiswamy PM (2004): A systematic review of antidepressant placebo-controlled trials for geriatric depression: limitations of current data and directions for the future. *Neuropsychopharmacology*; 29 (12): 2285-2299
- Thase ME, Pritchett YL, Ossanna MJ, Swindle RW, Xu J, Detke MJ (2007): Efficacy of duloxetine and selective serotonin reuptake inhibitors: comparisons as assessed by remission rates in patients with major depressive disorder. *J Clin Psychopharmacol*; 27 (6): 672-676
- Tranter R, O'Donovan C, Chandarana P, Kennedy S (2002): Prevalence and outcome of partial remission in depression. *J Psychiatry Neurosci*; 27 (4): 241-247
- Wernicke JF, Prakash A, Kajdasz DK, Houston J. Safety and tolerability of duloxetine treatment of diabetic peripheral neuropathic pain between patients with and without cardiovascular conditions. *J Diabetes Complications*. 2008 Sep 1. [Epub ahead of print]

9. Januar 2009

Wise TN, Wiltse CG, Iosifescu DV, Sheridan M, Xu JY, Raskin J (2007): The safety and tolerability of duloxetine in depressed elderly patients with and without medical comorbidity. *Int J Clin Pract*; 61 (8): 1283-1293

Wohlschlag, MM, Martinez JM; Mallinckrodt CH et al. (2005): An Open-label Study of Duloxetine for the Treatment of Major Depressive Disorder – Comparison of switching versus initiating treatment approaches . *Journal of Clinical Psychopharmacology* 25 (6) 552 -560

**A 1.4 Deutsche Gesellschaft für Psychiatrie, Psychotherapie und Nervenheilkunde (DGPPN), Arbeitsgemeinschaft für Neuropsychopharmakologie und Pharmakopsychiatrie (AGNP)**

**Autoren:**

Fritze, Jürgen, Prof. Dr. med.

**Adresse:**

Prof. Dr. med. Jürgen Fritze

Deutsche Gesellschaft für Psychiatrie, Psychotherapie und Nervenheilkunde (DGPPN),  
Arbeitsgemeinschaft für Neuropsychopharmakologie und Pharmakopsychiatrie (AGNP)

Asternweg 65

50259 Pulheim



DGPPN Hauptgeschäftsstelle · Reinhardtstraße 14 · 10117 Berlin

Pulheim, 08.01.2009

**Deutsche Gesellschaft  
für Psychiatrie,  
Psychotherapie  
und Nervenheilkunde**

**Stellungnahme zum Vorbericht  
Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer  
(SNRI) bei Patienten mit Depressionen  
Auftrag A05-20A**

**Im Namen der Deutschen Gesellschaft für Psychiatrie, Psycho-  
therapie und Nervenheilkunde (DGPPN) und  
der Arbeitsgemeinschaft für Neuropsychopharmakologie und  
Pharmakopsychiatrie (AGNP)**

IQWiG legt eine beachtliche Leistung, eine sorgfältige Analyse vor.

Dennoch gibt es Anlass zu einigen Fragen. Dabei sollen Grundsatz-  
fragen, die bereits – vergeblich – im Zusammenhang mit dem Be-  
richtsplan bzw. den allgemeinen Methoden gestellt wurden, im we-  
sentlichen nicht erneut im Detail aufgeworfen werden. Die Reihenfol-  
ge der Fragen ist nicht als Rangfolge (im Sinne einer Priorisierung)  
zu verstehen.

1. Mit welcher medizinischen Logik und Berechtigung werden auch die Arzneimittelverträglichkeit und –sicherheit ausschließlich auf der Basis von RCTs bewertet (das steht im krassen Widerspruch zum Verfahren der Zulassungsbehörden nach der Arzneimittelzulassung; die Relevanz dieser Frage ergibt sich exemplarisch auch daraus, dass im gesamten hier berücksichtigten Datenpool zu Duloxetin nur 3 Todesfälle (inkl. einem Suizid) und nur 6 Suizidversuche erfasst wurden)? Entsprechendes gilt für die Analysen zu Venlafaxin.
2. Mit welcher medizinischen Logik und Berechtigung werden Studien, in denen kürzer als 6 Wochen behandelt wurde, ausgeschlossen (4-Wochen-Studien waren bis vor ca. 15 Jahren durchaus – auch von Zulassungsbehörden akzeptierter - Standard; ihr Ausschluß würde für ältere, unverändert zugelassene und verkehrsfähige Antidepressiva bedeuten, dass deren Nutzen unbewertbar wäre, was mit weiterer Verkehrsfähigkeit unvereinbar sein dürfte; außerdem korreliert das Ausmaß der Besserung in den ersten Wochen hoch mit der Besserung nach 6 Wochen)?
3. Mit welcher Berechtigung wurden „Dosisarme geeignet zusammengefasst“, wenn eine Studie mehr als einen Dosisarm zum Wirkstoff enthielt, und was bedeutet dabei „geeignet“ (wenn eine Dosis-Wirkungs-Beziehung besteht, bewirkt das Zusammenfassung grundsätzlich eine „Verdünnung“ des Effektes)?

**Präsident**

Prof. Dr. med. Dr. rer.soc. Frank Schneider, Aachen  
Tel.: 0241/80-89633 Fax: 0241/80-82401  
Email: fschneider@ukaachen.de

**President Elect**

Prof. Dr. med. Peter Falkai, Göttingen  
Tel.: 0551/39-6601 Fax: 0551/39-2798  
Email: pfalkai@gwdg.de

**Past President**

Prof. Dr. med. Wolfgang Gaebel, Düsseldorf  
Tel.: 0211/922-2000 Fax: 0211/922-2020  
Email: wolfgang.gaebel@uni-duesseldorf.de

**Schriftführer**

PD Dr. M. Grözinger, Aachen  
Tel.: 0241/80-88731 Fax: 0241/80-3388731  
Email: schriftfuehrer@dgppn.de

**Kassenführer**

Priv.-Doz. Dr. med. Felix M. Böcker, Naumburg  
Tel.: 03445/72-1800 Fax: 03445/72-1802  
Email: dgppnkasse@t-online.de

**Aus-, Fort- und Weiterbildung**

Prof. Dr. med. Fritz Hohagen, Lübeck  
Tel.: 0451/500-2440 Fax: 0451/500-2603  
Email: fritz.hohagen@psychiatrie.uk-sh.de

**Wissenschaftsförderung**

Prof. Dr. med. Wolfgang Maier, Bonn  
Tel.: 0228/287-15722 Fax: 0228/287-16097  
Email: wolfgang.maier@ukb.uni-bonn.de

**Neurobiologische Forschung**

Prof. Dr. med. Heinrich Sauer, Jena  
Tel.: 03641/935246 Fax: 03641/935280  
Email: heinrich.sauer@med.uni-jena.de

**Psychotherapie und Psychosomatik**

Prof. Dr. med. Sabine C. Herpertz, Rostock  
Tel.: 0381/4949501 Fax: 0381/4949502  
Email: sabine.herpertz@med.uni-rostock.de

**Universitäre Psychiatrie**

Prof. Dr. med. Andreas Heinz, Berlin  
Tel.: 030/450-517001 Fax: 030/450-517921  
Email: andreas.heinz@charite.de

**Stationäre Versorgung, Rehabilitation**

Dr. med. Iris Hauth, Berlin-Weißensee  
Tel.: 030/92790-234 Fax: 030/92790-702  
Email: i.hauth@alexius.de

**Ambulante Versorgung**

Dr. med. Frank Bergmann, Aachen  
Tel.: 0241/36330 Fax: 0241/404972  
Email: bergmann@bvdn-nordrhein.de

**Sozialpsychiatrie**

Prof. Dr. med. Karl H. Beine, Hamm  
Tel.: 02381/18-2525 Fax: 02381/18-2527  
Email: karl-h.beine@marienhospital-hamm.de

**Beauftragte für die Gremien der  
ärztlichen Selbstverwaltung**

Dr. med. Christa Roth-Sackenheim, Andernach  
Tel.: 02632/96400 Fax: 02632/964096  
Email: dgppn@dr-roth-sackenheim.de

**Gesundheitspolitischer Sprecher**

Prof. Dr. med. Jürgen Fritze, Pulheim  
Tel.: 0173/5123250 Fax: 02238/54781  
Email: juergen.fritze@dgn.de

**Hauptgeschäftsführer**

Dr. phil. Thomas Nesselner, Berlin

DGPPN Hauptgeschäftsstelle Berlin  
Reinhardtstraße 14 Tel.: 030-2809 6602  
10117 Berlin Fax: 030-2809 3816  
E-Mail: sekretariat@dgppn.de  
Internet: www.dgppn.de

Hypovereinsbank München (BLZ 700202 70) Konto: 509 511  
VR 26854B, Amtsgericht Berlin-Charlottenburg

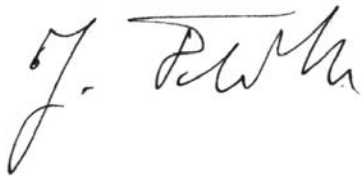
4. Mit welcher Berechtigung wurden – abweichend vom Berichtsplan und anscheinend ohne Amendment zum Berichtsplan – „für die Subgruppen „Vorliegen psychiatrischer und / oder somatischer Komorbidität“ und „spezifische, im Vordergrund der Erkrankung stehende Einzelsymptomatik“ nur solche Studien einer etwaigen Analyse zugeführt, welche explizit Patienten mit relevanten Merkmalen einschlossen“, und auf entsprechende Subgruppenanalysen von Studien ohne entsprechende explizite Ein-schlußkriterien verzichtet?
5. Woraus ergibt sich, die Tagesdosis von 75–225 mg Imipramin bzw. 75–225 mg Imipramin bzw. 225 mg Imipramin bzw. 150–200 mg Imipramin sei überdosiert und entspreche nicht der Zulassung, mit der Konsequenz des Ausschlusses der entsprechenden Studienarme (laut Fachinformation kann bei nicht-geriatrischen Erwachsenen „unter stationären Bedingungen die Tagesdosis auf bis zu 300mg Imipraminhydrochlorid, verteilt auf mehrere Einzelgaben, gesteigert werden“ und ist „die Dosis schrittweise so lange innerhalb des erlaubten Dosisbereichs zu steigern, bis der Patient auf die Behandlung anspricht“, wobei „die mittlere Tagesdosis für erwachsene Patienten im ambulanten Bereich für alle (!) Anwendungsgebiete bei 50 mg bis 150 mg Imipraminhydrochlorid liegt“, so dass eine Tagesdosis von 150 mg nicht als ambulante Höchstdosis zu verstehen ist)?
6. Woraus ergibt sich, die Tagesdosis von 50–100 mg Clomipramin sei „für Ältere überdosiert“ und entspreche nicht der Zulassung, mit der Konsequenz des Ausschlusses der entsprechenden Studienarme (zwar kann die Fachinformation zu Anafranil® oder Clomipramin-ratiopharm® (jeweils Stand Juni 2008) dahingehend verstanden werden, in der Geriatrie sei die Tagesdosis auf 30 bis 50 mg Clomipraminhydrochlorid zu beschränken; in der Fachinformation zu Clomipramin Sandoz® 75 mg (Oktober 2003, unverändert gültig) heißt es aber „die niedrigen Dosen von Clomipraminhydrochlorid, die im Allgemeinen bei älteren Patienten verordnet werden, lassen sich mit Clomipramin Sandoz 75 mg nicht ermöglichen“, und in der Fachinformation zu Clomipramin Sandoz® 10 mg/25 mg „die Dosierung wird schrittweise auf eine optimale Dosierung von 30 – 50 mg täglich (3-mal täglich 1 Tablette zu 10 mg oder 2-mal täglich 1 Tablette zu 25 mg) gesteigert“, und in der Fachinformation zu Hydiphen® 25 (Stand Mai 2008) fehlt ein spezieller Dosierungshinweis zum Einsatz bei älteren Menschen, so dass in der Geriatrie Dosierungen über 50 mg/Tag jedenfalls im Einzelfall mit der Zulassung vereinbar sind, wenn auch die Unschärfen der Formulierungen bedauerlich sind)?
7. Was soll die Formulierung „Somit war der antidepressive Nutzen von Duloxetin gegenüber Placebo im Subpool der Studien mit HAMA-Daten ebenfalls gegeben“ bedeuten (HAMA beansprucht, Angst zu messen, nicht Depressivität; vermutlich eine nur redaktionelle Unschärfe)?
8. Warum wurde die mögliche Abhängigkeit der Schmerzabnahme unter Duloxetin im Vergleich zu Placebo bzw. SSRIs vom initialen Schweregrad des Schmerzes nicht untersucht?
9. Mit welcher medizinischen Logik werden unter dem Begriff „Unerwünschte Arzneimittelwirkungen“ (der ursächliche Zusammenhänge suggeriert) für Duloxetin dann doch Ereignisse berichtet und bewertet, und dies bei Unsicherheit, ob es sich um Patientenzahlen und nicht um Ereigniszahlen handelt?
10. Wie kann – im anzuerkennenden Bemühen und einen mathematischen Auswertungsansatz – die Zusammenfassung unerwünschter Ereignisse zu Duloxetin bzw. seinen Komparatoren unabhängig von der Art des Ereignisses angesichts der Vielfalt der möglichen Ereignisse zu einer medizinisch sinnvollen Bewertung führen? Für Behandlungsabbrüche wegen unerwünschter Ereignisse ist der Ansatz wegen der einheitlichen, klinisch relevanten Konsequenz plausibel, aber sonst?
11. Warum wird – wenn eine bessere Verträglichkeit von Paroxetin bzw. für Behandlungsabbrüche wegen unerwünschter Ereignisse aller geprüften SSRI gegenüber Duloxetin suggeriert wird („Jedoch zeigte der Vergleich von Duloxetin und Paroxetin relativ konsistent niedrigere Ereignisraten für Paroxetin, der Unterschied war aber nicht statistisch signifikant“ bzw. „... erbrachten im Vergleich zu Duloxetin konsistent niedrigere Abbruchraten aufgrund von unerwünschten Ereignissen für die Vergleichs-

substanzen, jedoch ohne statistisch signifikantes Ergebnis“) nicht der mögliche Einfluß der Dosis (z.B. Paroxetin 40% der möglichen Höchstdosis) diskutiert, zumal in der Synopse (Tab. 43) auch nicht-signifikante „Nachteile“ mit einem „D-“ bewertet werden?

12. Wie erklären sich die diskrepanten Befunde „Beleg für einen größeren Schaden von Duloxetin in der ambulanten Akuttherapie der Depression, für beide Zielgrößen Gesamtrate unerwünschter Ereignisse und Abbruchraten aufgrund unerwünschter Ereignisse“ einerseits und „Gesamtraten schwerwiegender unerwünschter Ereignisse ... kein Beleg für einen Schaden von Duloxetin“ andererseits (dieser „Befund“ wirkt kontraintuitiv)?
13. Warum wird der Befund tendenziell geringerer sexueller Dysfunktionen (ASEX-Gesamtscore) unter Duloxetin im Vergleich zu SSRIs und hier insbesondere Paroxetin nicht vor dem Hintergrund möglicher Einflüsse der Dosis (z.B. Paroxetin 40% der möglichen Höchstdosis) diskutiert?
14. Was ist die wissenschaftliche Basis für den Cut-off von 37% des maximal erreichbaren Scores zu Studienbeginn für den Vergleich der Wirkung von Duloxetin im Vergleich zu Placebo bzw. SSRIs bei „leichter“ gegenüber „schwerer“ Depression?
15. Warum wird die Methodik des „Ausschlusses“ einer „einer bestehenden Suizidalität“ bezüglich der Venlafaxin-Studien nicht beschrieben und diskutiert?
16. Woraus rechtfertigt sich die – zu gegenüber der HAMD in den Venlafaxin-Studien abweichenden Ergebnissen führende – Wahl der Cut-Offs für die MADRS-Scores zu Studienbeginn (Scores zwischen 8 und 17 bei leichter Depression, zwischen 18 und 35 bei mittelschwerer Depression und über 35 bei schwerer Depression)? Snaith et al. 1986? In der Literatur werden auch andere Cut-Offs verwendet (z.B. Montgomery SA, Lecrubier Y. Is severe depression a separate indication? ECNP Consensus Meeting September 20, 1996, Amsterdam. European College of Neuropsychopharmacology. Eur Neuropsychopharmacol 1999; 9: 259–64), warum hier gerade diese?
17. Benennen die in Kapitel 5.3.1.2 beschriebenen Mängel solche der Studien (anhand der vorliegenden Studienberichte) oder der Publikationen?
18. Worin hätten die Risiken gelegen, ein „artifizielles Ergebnis“ zu erzielen, wenn die Studien zumindest auch separat zu Venlafaxin IR bzw. mutmaßlich IR bzw. Venlafaxin XR analysiert worden wären (zumindest bezüglich der Verträglichkeit)? Zumal in Anerkenntnis, daß „eine gewisse Verzerrung aufgrund dieses Umstands bei vielen fehlenden Angaben nicht ausgeschlossen werden konnte“?
19. Warum werden die Ergebnisse zu Venlafaxin nicht detaillierter im Kontext anderer Meta-Analysen (insbes. Nemeroff CB et al.: Comprehensive analysis of remission (COMPARE) with venlafaxine versus SSRIs. Biol Psychiatry 2008; 63: 424–434) diskutiert (nur dann könnte wirklich verständlich werden, warum in die IQWiG-Analyse nur 18 (Remission) bzw. 25 (Response) Studien zum Vergleich von Venlafaxin mit SSRIs eingingen, in die z.B. von Nemeroff et al. aber 34 Studien (Remission) als möglicher Grund für die z.T. diskrepanten Ergebnisse)? Aus den Diskussion unter 6.2 wird das nicht transparent.
20. Wie kann – im anzuerkennenden Bemühen und einen mathematischen Auswertungsansatz – die Zusammenfassung unerwünschter Ereignisse zu Venlafaxin bzw. seinen Komparatoren unabhängig von der Art des Ereignisses angesichts der Vielfalt der möglichen Ereignisse zu einer medizinisch sinnvollen Bewertung führen? Stellt nicht gerade das Ergebnis der Meta-Analyse „UEs, Venlafaxin vs. Placebo, Kurzzeit-Akutstudien“ (anscheinend kein Unterschied zwischen Venlafaxin und Placebo, wobei wegen Heterogenität kein Gesamtschätzer berechnet wurde, Abb. 114) bei „besserer“ Verträglichkeit der SSRIs (Abb. 115) die Sinnhaftigkeit dieses Ansatzes in Frage? Müßten danach nicht SSRIs besser verträglich als Placebo sein (eine eher blödsinnige Vorstellung)?
21. Wie kann – im anzuerkennenden Bemühen und einen mathematischen Auswertungsansatz – die Zusammenfassung schwerwiegender unerwünschter Ereignisse zu Venlafaxin bzw. seinen Komparatoren unabhängig von der Art des Ereignisses angesichts der Vielfalt der möglichen Ereignisse zu einer medizinisch sinnvollen Bewertung führen? Stellt nicht gerade das Ergebnis der Meta-Analyse „SUEs, Venlafaxin



- vs. Placebo, Kurzzeit-Akutstudien“ (kein Unterschied zwischen Venlafaxin und Placebo, Abb. 116) bei „gleicher“ Verträglichkeit der SSRIs (Abb. 117) die Sinnhaftigkeit dieses Ansatzes in Frage? Müßten danach nicht SSRIs ebenso verträglich wie Placebo sein (auch dies eine eher blödsinnige Vorstellung)?
22. Wird die Schwäche des Ansatzes, nicht zumindest für die Sicherheitspharmakologie auch Daten aus Nicht-RCTs einzubeziehen, daran (beispielhaft) erkennbar, daß nicht einmal der etablierte Befund einer Blutdruckerhöhung unter Venlafaxin repliziert werden konnte?
  23. Was ist die wissenschaftliche Basis für den Cut-off von 50% des maximal erreichbaren Scores zu Studienbeginn für den Vergleich der Wirkung von Venlafaxin im Vergleich zu SSRIs bei „leichter“ gegenüber „schwerer“ Depression?
  24. Wie erklärt sich, daß der Datenpool von Nemeroff et al. 2008 (s.o.) über 6 Venlafaxin-Studien allein gegen SSRI (3 gegen Paroxetin, 3 gegen Fluoxetin) bei stationären Patienten verfügte, der Datenpool des IQWiG aber über nur 3 (gegen Fluoxetin, Imipramin bzw. Nortriptylin)? Welche Studien fehlen warum?
  25. Welche Altersgrenze wurde mit welcher Begründung den Meta-Analysen getrennt für einen Studienpool älterer und jüngerer Patienten im Vergleich von Venlafaxin und Placebo (deren Ergebnisse Wirkungslosigkeit von Venlafaxin bei älteren suggeriert) zugrunde gelegt ( $\geq 60$  oder 65 Jahre)? Ist aus dem Fehlen einer „relevanten Interaktion“ zwischen Alter und Behandlungseffekt beim Vergleich von Venlafaxin und SSRI bzw. TZA also zu schlußfolgern, daß alle diese Antidepressiva bei älteren Menschen wirkungslos sind? Lassen sich solche Suggestionen auf Basis von nur zwei Studien vertreten?
  26. Worauf gründet die Aussage zu Venlafaxin, „es gab keinen Hinweis für eine maßgebliche Abweichung in Bezug auf die antidepressive Wirkung bei Dosisvariationen“, wenn doch von „einer separaten Analyse der Dosisvariationen abgesehen“ wurde?
  27. Warum erfolgt weder für Duloxetin noch für Venlafaxin eine epikritische Abwägung von Nutzen und Schaden, zum Beispiel in Form von NNT und NNH, wenn sich der Bericht schon derart um Quantifizierung bemüht?



Prof. Dr. med. Jürgen Fritze

## **A 1.5     Lundbeck GmbH**

### **Autoren:**

Flürenbrock, Wibke, Dr.

Friede, Michael, Dr.

Kessel-Steffen, Markus

### **Adresse:**

Lundbeck GmbH

Karnapp 25

21079 Hamburg

# **Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen“ Stellungnahme der Lundbeck GmbH zum Vorbericht (vorläufige Nutzenbewertung), Version 1.0 vom 19. November 2008**

Sehr gerne unterstützt Lundbeck das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) und gibt folgende Stellungnahme zum Vorbericht (Version 1.0) des Auftrags A05-20A.

## **Executive Summary**

Der vorliegende Vorbericht ist sehr umfangreich. Lundbeck stimmt jedoch mit den vorläufigen Schlussfolgerungen nicht überein, dass Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) und Selektive Serotonin-Wiederaufnahmehemmer (SSRI) – insbesondere Escitalopram – äquivalent seien.

Darüber hinaus werden im vorliegenden Vorbericht Vergleiche der SNRI Duloxetin und Venlafaxin mit der Wirkstoffgruppe der SSRI vorgenommen. Voraussetzung für eine solche Vorgehensweise ist, dass die Gleichheit aller SSRI belegt ist. Dies ist nicht der Fall. Der SSRI Escitalopram unterscheidet sich statistisch signifikant und klinisch relevant in Patienten-relevanten Zielgrößen vom SSRI Citalopram.

## **Escitalopram hat ein überlegenes Nutzen-Risiko-Profil im Vergleich zu den SNRI Duloxetin und Venlafaxin**

### **Vergleich Escitalopram vs. Duloxetin**

Im Vorbericht wird ausgeführt, dass Duloxetin und Escitalopram in der Akuttherapie äquivalent seien. Diese Schlussfolgerung basiert auf der Auswahl der eingeschlossenen Studien und der Nichtberücksichtigung der Studie von Wade et al. (2007), die als Langzeit-Studie klassifiziert wird.

Unter Einbezug der vorhandenen Evidenz der drei head-to-head Studien zeigen Meta-Analysen, die gemäß der Vorgehensweise des IQWiG durchgeführt wurden, dass Escitalopram gegenüber Duloxetin überlegen ist. In der Zielgröße Änderungen im MADRS total score im Vergleich zur Baseline ist Escitalopram statistisch signifikant Duloxetin überlegen. Escitalopram ist Duloxetin konsistent numerisch in den Zielgrößen Response, Remission und Änderungen im Vergleich zur Baseline beim HAMD17 score überlegen.

Eine gepoolte Analyse von zwei klinischen Studien (Lam et al., 2008) belegt, dass unter Escitalopram signifikant weniger Patienten ( $P < 0.001$ ) die Studien abgebrochen haben als unter Duloxetin (unerwünschte Ereignisse und sonstige Gründe).

Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0

Darüber hinaus führt die Therapie mit Escitalopram zu signifikant weniger Arbeitsunfähigkeit und signifikant weniger krankheitsbedingter Abwesenheit als die Behandlung mit Duloxetin.

### **Vergleich Escitalopram vs. Venlafaxin XR**

Es liegen zwei head-to-head Studien vor. Beide zeigen Vorteile von Escitalopram gegenüber Venlafaxin XR. Ein aktueller Review (Llorca et al., 2007) kommt zum Ergebnis, dass Escitalopram potentiell wichtige Vorteile gegenüber Venlafaxin XR in den Zielgrößen Zeitdauer bis zur Remission, Verträglichkeit und Absetzsymptome hat.

Auf Basis dieser Evidenz kommen die Autoren zum Ergebnis, dass Escitalopram mindestens genauso wirksam wie Venlafaxin XR in der Behandlung der majoren Depression (MDD) bei besserer Verträglichkeit ist.

### **Escitalopram ist Citalopram überlegen und kann daher nicht in die Wirkstoffgruppe der SSRI einbezogen werden**

#### **Vergleich Escitalopram vs. Citalopram**

Citalopram wurde von den Bundesvertragspartnern als Leitsubstanz der SSRI bestimmt. Daher ist diese Substanz der geeignete Komparator um die Überlegenheit von Escitalopram zu belegen.

Die Überlegenheit von Escitalopram gegenüber Citalopram wurde in vielen RCT und Meta-Analysen bewiesen. In den für das IQWiG relevanten Zielgrößen Response, Remission und Veränderung der Symptomatik zwischen Studienbeginn und –ende ist Escitalopram statistisch signifikant und klinisch relevant Citalopram überlegen.

Die direkten Vergleiche von Escitalopram gegenüber Citalopram zeigen die Unterschiedlichkeit der Substanzen und sind ein Beleg dafür, dass nicht alle SSRI gleich sind. Die pauschale Bildung einer Wirkstoffgruppe SSRI unter Einbezug von Escitalopram, ohne vorherige Bewertung durch das IQWiG, sollte daher unterbleiben.

Die in der Stellungnahme der Lundbeck dargestellte Evidenz und die diskutierten Schlussfolgerungen werden durch unabhängige Health Technology Assessments (HTA) bestätigt. Die schwedische Dental and Pharmaceutical Benefits Agency TLV (Tandvårds- och Läkemedelsförmånsverket) hat im Dezember 2008 einen Review zu Antidepressiva abgeschlossen (Wessling et al., 2008). Die TLV kommt zu folgendem

Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI)  
bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0

Ergebnis: *„Treatment with escitalopram gives the highest probability to reach remission (recovery) in comparison to all other modern antidepressants.“*

Das norwegische Knowledge Centre (Nasjonalt kunnskapssenter for helsetjenesten) veröffentlichte im Jahr 2007 einen Review (Saeterdal et al.; 2007) zu Antidepressiva und kam in der Zusammenfassung zu folgendem Ergebnis: *„Escitalopram was significantly more effective (response and remission rate) than Citalopram.“*

Diese Aussagen bestätigen den unabhängigen HTA-Bericht *“Drug Class Review on Second Generation Antidepressant“* von Gartlehner et al. (2006): *“Overall, effectiveness and efficacy were similar and the majority of trials [55 head-to-head trials compared effectiveness and efficacy of one new AD to another] did not identify substantial differences among drugs. The only exception is the comparison of citalopram to escitalopram. Four fair to good trials indicate consistently that escitalopram has a greater efficacy for the treatment of MDD than citalopram.“*

Aus den dargelegten Gründen soll Escitalopram unter Einbezug der vorhandenen Evidenz ausschließlich als Einzelsubstanz in den Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen“ einbezogen und bewertet werden. Verallgemeinernde Schlussfolgerungen für die Gruppe der SSRI unter Einbezug von Escitalopram sind nicht zulässig.

# 1 Einführung

Der Vorbericht "Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen" (Stand: 19. November 2008) wurde vom IQWiG veröffentlicht und zur Diskussion gestellt. Im Vorbericht werden Aussagen zu SNRI sowie zu anderen Antidepressiva, z.B. dem Selektiven Serotonin-Wiederaufnahmehemmer (SSRI) Escitalopram, getroffen. Hierbei werden direkte Vergleiche zwischen den jeweiligen SNRI und einzelnen Antidepressiva sowie indirekte Vergleiche vorgenommen und Aussagen abgeleitet.

In der Stellungnahme der Lundbeck wird an konkreten Beispielen verdeutlicht, dass zwischen Antidepressiva relevante Unterschiede bestehen. Zunächst werden die Unterschiede des Wirkstoffs Escitalopram unter Bezugnahme auf direkte Vergleichsstudien gegen die SNRI Duloxetin und Venlafaxin XR dargestellt. Im Weiteren wird eine Patienten-relevante Differenzierung von Escitalopram gegen Citalopram auf Basis von direkten Vergleichsstudien belegt. Der Wirkstoff Citalopram wurde exemplarisch gewählt, da die Substanz erneut vom Spitzenverband Bund der Krankenkassen und der Kassenärztlichen Bundesvereinigung (Rahmenvorgaben nach § 84 Abs. 7 SGB V – Arzneimittel – für das Jahr 2009 vom 30. September 2008) als „Leitsubstanz“ der SSRI bestimmt wurde. Die Differenzierung von Escitalopram gegenüber Citalopram beweist, dass trotz pharmakologischer Klassifikation zur Wirkstoffgruppe der SSRI Patienten-relevante Unterschiede zwischen den beiden Wirkstoffen bestehen. Pauschalierte Aussagen über SSRI lassen sich auf Basis des Vorberichts SNRI nicht treffen und sind weder methodisch noch medizinisch zulässig.

In der Stellungnahme der Lundbeck wird aufgezeigt, dass klinisch relevante Unterschiede zwischen Antidepressiva bestehen.

## 2 Methode

### 2.1. SNRI Duloxetin und Venlafaxin XR vs. Escitalopram

#### 2.1.1 Duloxetin vs. Escitalopram

Das IQWiG hat im Vorbericht die 4 Studien Khan et al. (2007), Nierenberg et al. (2007), Wade et al. (2007) und Pigott et al. (2007) (Extension der Studie von Nierenberg et al., 2007) als relevant identifiziert. Die Studien sind in Tabelle 1 aufgelistet. In Abhängigkeit der Studiendauer wurden die Studien entweder als Akuttherapiestudien oder Erhaltungstherapiestudien (Langzeittherapiestudien) unterteilt:

- Akuttherapie: Khan et al. (2007), Nierenberg et al. (2007)
- Erhaltungstherapie (Langzeittherapie): Wade et al. 2007, Pigott et al. 2007

Tab. 1: Übersicht der identifizierten Studien, die in die Meta-Analysen einbezogen wurden

Studie No.	(reference)	Studiendauer, Schweregrad zu Studienbeginn	Design	Setting	Dosierung (mg)	Dosierung (mean / median / mode)*	APTS (n)	Studie beendet (%)
1	(Wade)	24 weeks MADRS $\geq$ 26	Fixed	GP&spec	ESC(20) DUL (60)	20/20/20 60/60/60	143 151	88.1 86.1
2	(Khan)	8 weeks MADRS $\geq$ 26	Flexible	Spec	ESC(10-20) DUL (60)	16.1/20/20 60/60/60	137 133	86.9 69.2
3	(Nierenberg)	8 weeks MADRS $\geq$ 22	Fixed	Spec	ESC(10) DUL (60)	10/10/10 60/60/60	274 273	75.9 68.9
4	(Pigott)	32 weeks MADRS $\geq$ 22	Flexible	Spec	ESC(10-20) DUL (60-120)	15.3/20/20 80.1/60/60	274 273	44.9 38.1

ESC, Escitalopram; DUL, Duloxetine; PBO, Plazebo kontrollierte Studie; Spec, specialist setting; GP, general practice; APTS, all-patients-treated set; \*Daten vom letzten Besuch (LOCF);

Die Studie von Wade et al. (2007) wurde im Vorbericht lediglich bei den Meta-Analysen zur Erhaltungstherapie (Langzeittherapie) berücksichtigt, jedoch nicht in die Meta-Analysen zur Akuttherapie einbezogen. Gemäß dem Studienprotokoll wurden nach 8 Wochen Therapiedauer Untersuchungen durchgeführt, die als sekundärer Endpunkt definiert waren – diese Daten sollen bei den Meta-Analysen zur Akuttherapie berücksichtigt werden.

Gemäß dem Vorgehen des IQWiG wurden Meta-Analysen zur Akuttherapie unter Einbezug der 8 Wochen Daten aus der Studie von Wade et al. (2007) durchgeführt. Ziel bei Meta-Analysen ist, die gesamte verfügbare Evidenz der primären und sekundären Endpunkte einzubeziehen. Folgende Zielgrößen wurden in den Meta-Analysen evaluiert:

- Änderungen im HAMD17 score im Vergleich zur Baseline
- Änderungen im MADRS total score im Vergleich zur Baseline
- Response, definiert als mindestens 50%-ige Reduktion im HAMD17
- Response, definiert als mindestens 50%-ige Reduktion im MADRS total score
- Remission, definiert als Erreichen oder Unterschreiten eines HAMD17 Wertes von 7
- Remission, definiert als Erreichen oder Unterschreiten eines MADRS total score Wertes von 12 (in der Studie von Khan et al. (2007) wurde der cut off mit dem Erreichen oder Unterschreiten eines MADRS total score Wertes von 10 definiert – gemäß der Definition wurden diese Remissionsdaten in die Meta-Analysen einbezogen)
- Studienabbrüche wegen unerwünschter Ereignisse

## 2.1.2 Venlafaxin XR vs. Escitalopram

Das IQWiG hat im Vorbericht die 2 Studien Montgomery et al. (2004) und Bielski et al. (2004) als relevant identifiziert. Die Studien sind in Tabelle 2 aufgelistet.

Tab. 2: Übersicht der identifizierten Studien, die in die Meta-Analysen einbezogen wurden

Studie No.	(reference)	Studiendauer, Schweregrad zu Studienbeginn	Design	Setting	Dosierung (mg)	Dosierung (mean/median/mode) *	APTS (n)	Studie beendet (%)
1	(Montgomery)	8 weeks	Flexible	GP	ESC(10–20)	12.1/10/10	146	85.6
		MADRS $\geq$ 18			VLF (75-150)	95.2/75/75	143	86.7
2	(Bielski)	8 weeks	Fixed	Spec	ESC(20)	20/20/20	98	73.5
		HAMD24 $\geq$ 20			VLF (225)	225/225/225	100	66

ESC, Escitalopram; VLF, Venlafaxin XR; Spec, specialist setting; GP, general practice; APTS, all-patients-treated set; \* Daten vom letzten Besuch (LOCF);

## 2.2 Escitalopram vs Citalopram

Multiple Literaturrecherchen in MEDLINE (1966 - Sep 2008), EMBASE (1998-2008) und Cochrane Collaboration (1980 - Sep 2008) wurden durchgeführt. Vorgegebene Keywords waren escitalopram AND citalopram AND (depression or major depressive disorder or major depressive episode). Weitere Studien in jeglicher Sprache wurden in den Literaturlisten der erhaltenen Artikel gesucht. Unpublizierte Studien wurden in Controlled Trials Database, [www.lundbecktrials.com](http://www.lundbecktrials.com), [www.forestclinicaltrials.com](http://www.forestclinicaltrials.com) und National Institute of Health's Computer Retrieval of Information on Scientific Projects (CRISP) service (1972 to 2005) gesucht. Ergänzend wurden die clinical trial registration sites [www.lundbecktrials.com](http://www.lundbecktrials.com), [www.forestclinicaltrials.com](http://www.forestclinicaltrials.com), [www.clinicaltrials.gov](http://www.clinicaltrials.gov), [www.clinicaltrialresults.org](http://www.clinicaltrialresults.org), [www.ifpma.org/clinicaltrials](http://www.ifpma.org/clinicaltrials) und [www.controlled-trials.com](http://www.controlled-trials.com) durchsucht.

Gemäß der vom IQWiG definierten Ein- und Ausschlusskriterien wurden 6 RCT identifiziert und in die Meta-Analysen einbezogen. Die Studien sind in Tabelle 3 aufgelistet.



Tab. 3: Übersicht der identifizierten Studien, die in die Meta-Analysen einbezogen wurden.

Studie (Referenz)	No.	Studiendauer, Schweregrad zu Studienbeginn	Setting	Dosierung (mg)	Baselinewert MADRS	APTS (FAS)	Studie beendet (%)*
1 (SCT-MD 02)		8 Wochen, MADRS $\geq$ 22	Spec	ESC (10-20) <sup>1</sup> CIT (20-40) <sup>2</sup> PBO	28.7 $\pm$ 4.3 28.3 $\pm$ 5.0 28.8 $\pm$ 5.0	125 (124) 123 (119)	76.8 80.5
2 (Colonna et al., 2005)		24 Wochen MADRS $\geq$ 22	GP	ESC (10) CIT (20)	29.5 $\pm$ 4.3 30.2 $\pm$ 4.7	175 (165) 182 (174)	86.3 84.1
3 (Lepola et al., 2003)		8 Wochen MADRS $\geq$ 22	GP	ESC (10-20) <sup>3</sup> CIT (20-40) <sup>4</sup> PBO	29.0 $\pm$ 4.3 29.2 $\pm$ 4.2 28.7 $\pm$ 4.0	155 (155) 160 (159)	94.2 95.0
4 (Burke et al., 2002)		8 Wochen MADRS $\geq$ 22	Spec	ESC (10) ESC (20) CIT (40) PBO	28.0 $\pm$ 4.9 28.9 $\pm$ 4.6 29.2 $\pm$ 4.5 29.5 $\pm$ 5.0	119 (118) 125 (123) 125 (125)	79.8 75.2 74.4
5 (Moore et al., 2005)		8 Wochen MADRS $\geq$ 30	GP & spec	ESC (20) CIT (40)	36.3 $\pm$ 4.8 35.7 $\pm$ 5.8	142 (138) 152 (142)	93.0 83.6
6 Yevtushenko et al., 2007)		6 Wochen MADRS $\geq$ 25	Spec	ESC (10) CIT (10) CIT (20)	34.8 $\pm$ 3.5 35.4 $\pm$ 3.3 35.7 $\pm$ 3.8	109 (108) 111 (106) 110 (108)	99.1 95.5 98.2

ESC: Escitalopram; CIT: Citalopram; PBO: Plazebo kontrollierte Studien; RCT: Randomisierte kontrollierte Studie; Spec: specialist settings; n/a: not available; APTS: *all-patients-treated set*; FAS: full-analysis set; \* Daten vom letzten Besuch (LOCF); <sup>1</sup>mean dose 17.6mg; <sup>2</sup>mean dose 35.3mg; <sup>3</sup>mean dose 14.0mg; <sup>4</sup>mean dose 24.4mg

Gemäß der Vorgehensweise des IQWiG wurden zwei Evidenzsynthesen durchgeführt:

- Akuttherapie
- Erhaltungstherapie (Langzeittherapie)

Ziel bei Meta-Analysen ist, die gesamte verfügbare Evidenz der primären und sekundären Endpunkte einzubeziehen. Daher wurden in die Meta-Analyse zur Akuttherapie ebenfalls die 8 Wochen Daten aus der Studie von Colonna et al. (2005) einbezogen, obgleich die Studie über 24 Wochen andauerte. Für die Bewertung der Akuttherapie konnten demnach alle 6 identifizierten RCT berücksichtigt werden. Die Aussagen zur Erhaltungstherapie (Langzeittherapie) basieren lediglich auf einer Studie (Colonna et al., 2005).

Folgende Zielgrößen wurden in den Meta-Analysen evaluiert:

- Änderungen im MADRS total score im Vergleich zur Baseline
- Response, definiert als mindestens 50%-ige Reduktion im MADRS total score
- Remission, definiert als Erreichen oder Unterschreiten eines MADRS total score Wertes von 12
- Unerwünschte Ereignisse
- Studienabbrüche wegen unerwünschter Ereignisse

Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0

Wenn in Publikationen oder Studienberichten mit ANCOVA berechnete adjustierte Werte (on centre and baseline MADRS total score) aufgeführt waren, wurden diese für die Analysen der Zielgröße „Änderungen im MADRS total score im Vergleich zur Baseline“ verwendet. Die Zielgrößen Response und Remission wurden mit nicht-adjustierten Daten berechnet. Wurden Standardfehler nur für die Behandlungsunterschiede angegeben, wurden gleiche Werte pro Behandlungsgruppe angenommen.

## 2.3 Meta-Analysen

Die Meta-Analysen wurden jeweils mit gleicher Methodik durchgeführt:

- Meta-Analysen mit zufälligen Effekten (random effect meta-analyses) wurden standardmäßig berechnet. Lag zu große Heterogenität gemäß der Definition im Vorbericht vor ( $p < 0.2$  oder  $I^2 > 50\%$ ), wurden der berechnete und dargestellte Effektschätzer nicht interpretiert.
- Gemäß dem Vorgehen des IQWiG, wurden alle Analysen auf Basis der LOCF Daten (Last Observation Carried Forward; Methode zum Umgang mit fehlenden Daten) durchgeführt. Lagen keine LOCF Daten vor oder war LOCF nicht die primäre Analyseverfahren, wurden MMRM Werte (Mixed Model Repeated Measures) verwendet.

Alle Berechnungen wurden unter Verwendung der folgenden Software durchgeführt: Review Manager (RevMan), Version 5.

### 3 Ergebnisse

#### 3.1 SNRI vs. Escitalopram

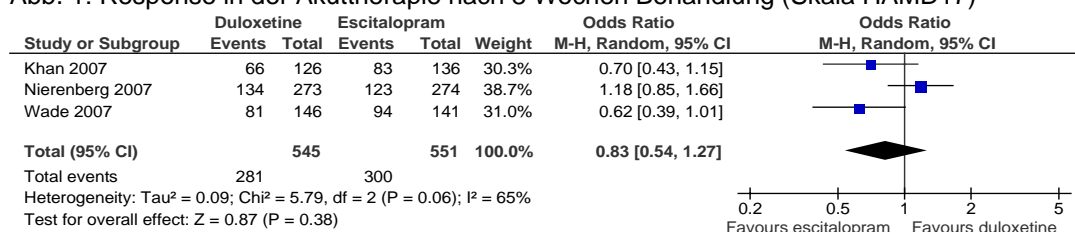
##### 3.1.1 Duloxetin vs. Escitalopram

##### Kurzzeit-Akutstudien

##### Response

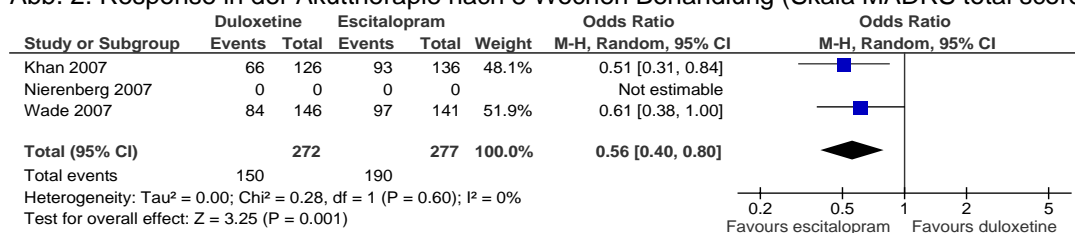
In Abbildung 1 sind die Ergebnisse zur Zielgröße Response für die Akuttherapie dargestellt (Skala HAMD17).

Abb. 1: Response in der Akuttherapie nach 8 Wochen Behandlung (Skala HAMD17)



Zwei Studien zeigten numerische Vorteile für Escitalopram in der Zielgröße Response, erhoben mit der Skala HAMD17. In Abbildung 2 sind die Daten zur Response unter Verwendung der Skala MADRS total score dargestellt.

Abb. 2: Response in der Akuttherapie nach 8 Wochen Behandlung (Skala MADRS total score)

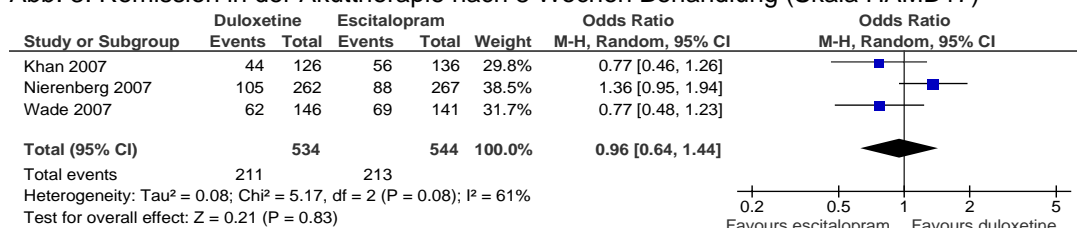


Im MADRS total score ist der Vorteil von Escitalopram sowohl in beiden Einzelstudien als auch in der Meta-Analyse statistisch signifikant.

##### Remission

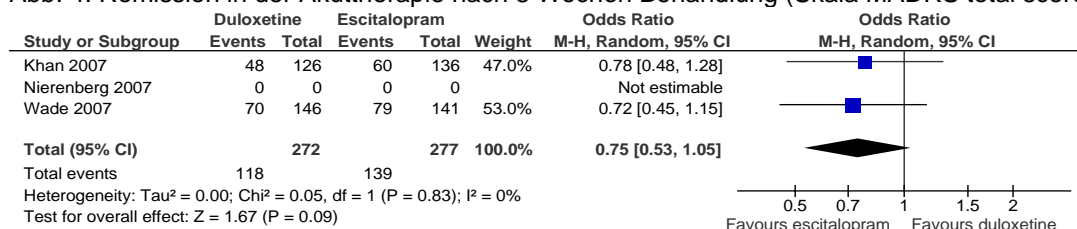
In Abbildung 3 sind die Ergebnisse zur Zielgröße Remission für die Akuttherapie dargestellt (Skala HAMD17).

Abb. 3: Remission in der Akuttherapie nach 8 Wochen Behandlung (Skala HAMD17)



Zwei Studien zeigten numerische Vorteile für Escitalopram in der Zielgröße Remission, erhoben mit der Skala HAMD17. In Abbildung 4 sind die Daten zur Remission unter Verwendung der Skala MADRS total score dargestellt.

Abb. 4: Remission in der Akuttherapie nach 8 Wochen Behandlung (Skala MADRS total score)

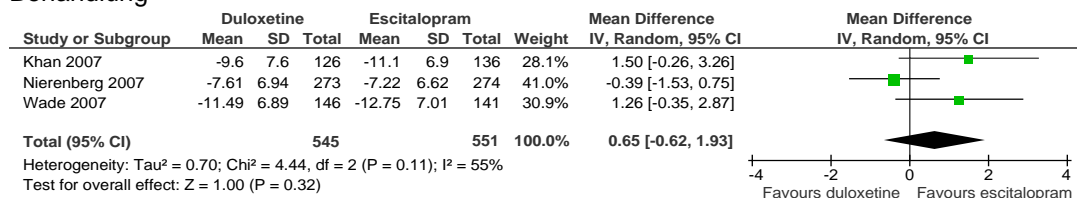


Die numerischen Vorteile von Escitalopram gegenüber Duloxetin werden unter Verwendung der Skala MADRS total score bestätigt.

### Änderungen im HAMD17 und MADRS total score im Vergleich zur Baseline

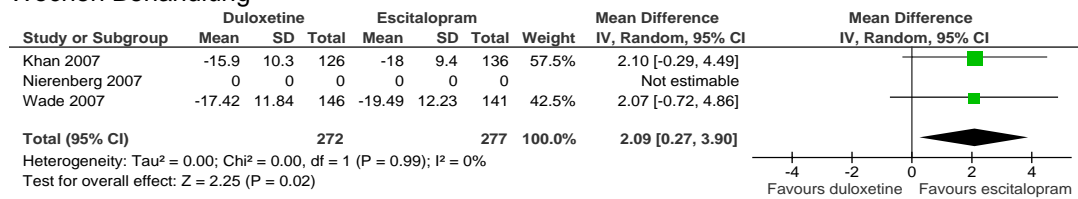
In Abbildung 5 ist die Meta-Analyse der Änderungen im HAMD17 im Vergleich zur Baseline dargestellt.

Abb. 5: Änderungen im HAMD17 im Vergleich zur Baseline in der Akuttherapie nach 8 Wochen Behandlung



Zwei Studien zeigten numerische Vorteile für Escitalopram in der Zielgröße Änderungen im HAMD17. In Abbildung 6 sind die Daten zu Änderungen im Vergleich zur Baseline unter Verwendung der Skala MADRS total score dargestellt.

Abb. 6: Änderungen im MADRS total score im Vergleich zur Baseline in der Akuttherapie nach 8 Wochen Behandlung

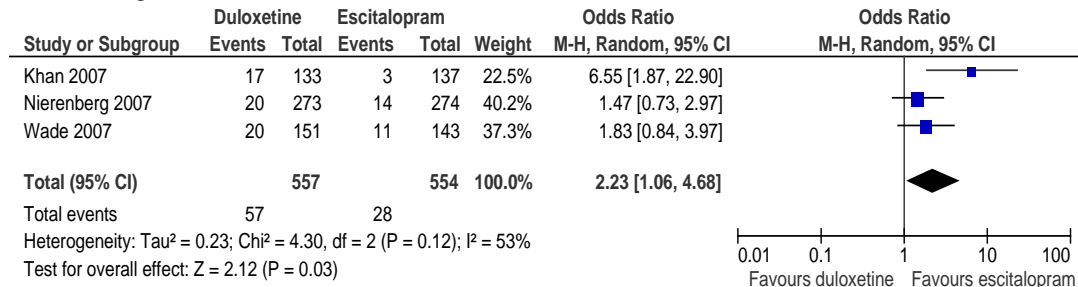


Escitalopram zeigt eine statistisch signifikante Überlegenheit gegenüber Duloxetin im MADRS total score.

### Studienabbruch wegen unerwünschter Ereignisse

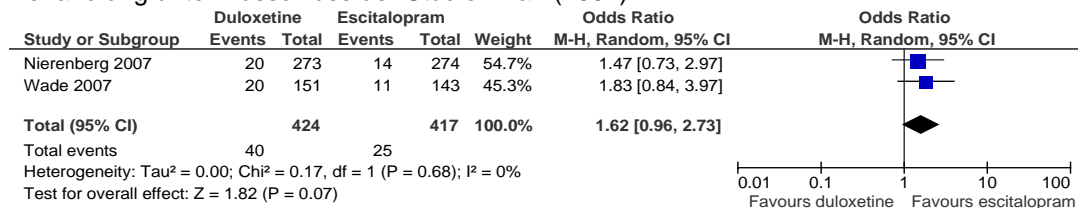
In Abbildung 7 sind die Studienabbrüche in den Akuttherapiestudien wegen unerwünschter Ereignisse dargestellt.

Abb. 7: Studienabbrüche wegen unerwünschter Ereignisse in der Akuttherapie nach 8 Wochen Behandlung



Alle Studien zeigten konsistent weniger Studienabbrüche wegen unerwünschter Ereignisse unter Escitalopram. In der Studie von Khan et al. (2007) war diese Überlegenheit statistisch signifikant. Wegen der Heterogenität wurde eine Meta-Analyse ohne die Studie von Khan et al. (2007) durchgeführt (Abbildung 8).

Abb. 8: Studienabbrüche wegen unerwünschter Ereignisse in der Akuttherapie nach 8 Wochen Behandlung unter Ausschluss der Studie Khan (2007)



Auch die Meta-Analyse unter Ausschluss der statistisch signifikanten Studie Khan et al. (2007) zeigt den numerischen Vorteil von Escitalopram gegenüber Duloxetin in der Zielgröße Studienabbrüche wegen unerwünschter Ereignisse und bestätigt die in Abbildung 7 dargestellte Analyse.

Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen“  
 Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0

## Erhaltungstherapie (Langzeittherapie)

### Arbeitsunfähigkeit und krankheitsbedingte Abwesenheit

Escitalopram hat signifikante Vorteile im Parameter Arbeitsfähigkeit im Vergleich zu Duloxetin. Anhand des Parameters Sheehan Disability Scale (SDS) konnte belegt werden, dass Escitalopram zu signifikant weniger Arbeitsunfähigkeit im Vergleich zu Duloxetin nach zweimonatiger und sechsmonatiger Behandlung führte (Wade et al., 2007).

Wade et al., 2008, zeigten weiterhin, dass die krankheitsbedingte Abwesenheit unter Escitalopram signifikant geringer war als unter Duloxetin. Die durchschnittliche krankheitsbedingte Abwesenheit pro Patient über 24 Wochen Studiendauer betrug unter Escitalopram 30.7 Tage im Vergleich zu 62.2 Tagen bei Duloxetin ( $p < 0.0001$ ) (Wade et al., 2008).

## 3.1.2 Venlafaxin XR vs. Escitalopram

### Kurzzeit-Akutstudien

#### Remission

In Abbildung 9 ist die anhaltende Remission über einen Behandlungszeitraum von 8 Wochen, gemessen mit dem MADRS total score, dargestellt (Montgomery et al., 2004).

Abb. 9: Anhaltende Remission im MADRS total score in der Akuttherapie über 8 Wochen Behandlung

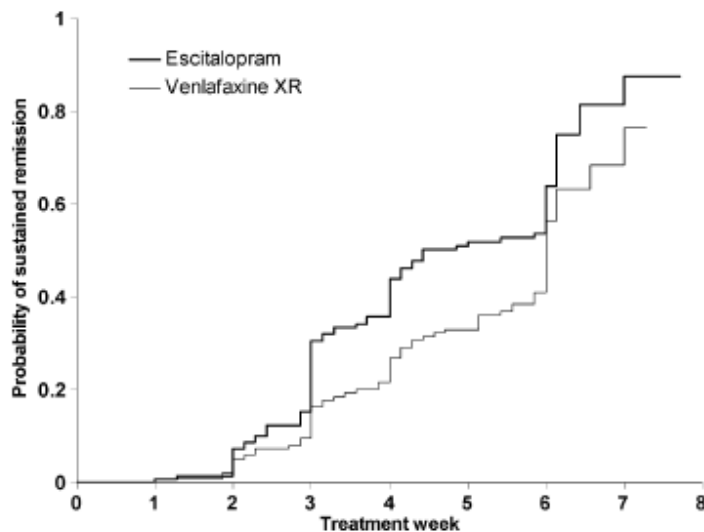


Abbildung 9 (aus Montgomery et al., 2004): Zeitverlauf anhaltende Remission ("time to sustained remission") [Montgomery-Asberg Depression Rating Scale total score  $\leq 12$  maintained until week 8] bei erwachsenen Patienten mit majorer Depression MDD, ambulante Versorgung ( $n = 288$ ; intent-to-treat Population; LOCF-Daten). Patienten erhielten flexible Dosierung von Escitalopram 10–20 mg/day oder Venlafaxin XR 75–150 mg/day in einer 8-wöchigen, doppelblinden, multicentrischen RCT ( $p < 0.01$ )

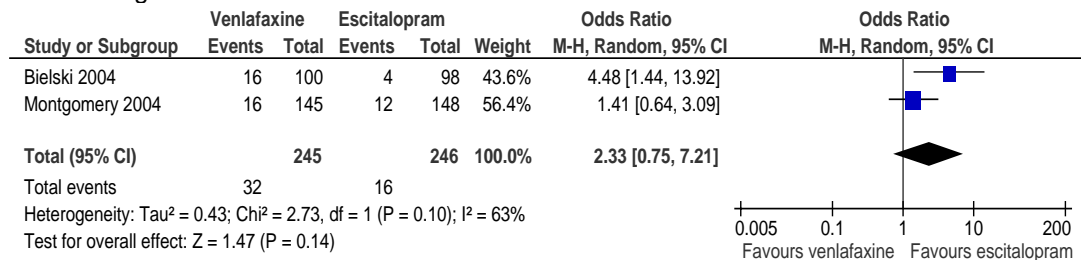
Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0

Die Ergebnisse zeigen, dass unter Escitalopram früher das Kriterium der Remission eintritt und der Effekt anhält als unter Venlafaxin XR.

### Studienabbruch wegen unerwünschter Ereignisse

In Abbildung 10 sind die Studienabbrüche in den Akuttherapiestudien wegen unerwünschter Ereignisse dargestellt.

Abb. 10: Studienabbrüche wegen unerwünschter Ereignisse in der Akuttherapie nach 8 Wochen Behandlung



Beide Studien zeigten konsistent weniger Studienabbrüche wegen unerwünschter Ereignisse unter Escitalopram. Die Überlegenheit von Escitalopram gegenüber Venlafaxin XR in der Studie Bielski et al. (2004) war statistisch signifikant.

In der Studie Montgomery et al. (2004) wurden in der einwöchigen wash-out Periode unter Venlafaxin XR signifikant weniger Absetzsymptome dokumentiert als unter Escitalopram (2.4 vs. 5.0,  $p < 0.001$ ).

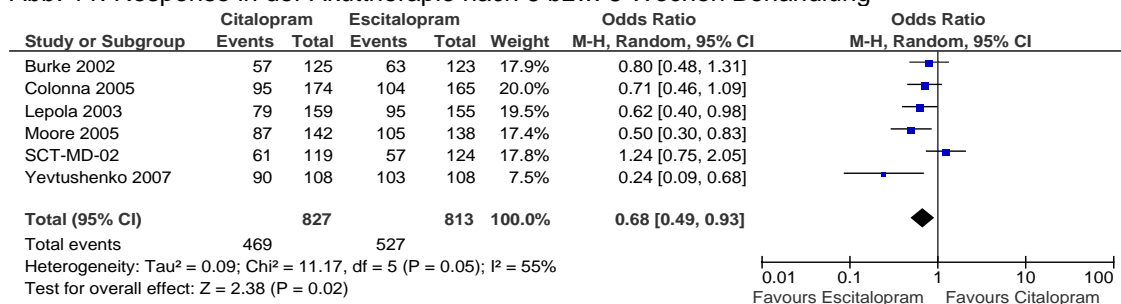
## 3.2 Escitalopram vs. Citalopram

### Kurzzeit-Akutstudien

#### Response

In Abbildung 11 sind die Ergebnisse zur Zielgröße Response für die Akuttherapie dargestellt.

Abb. 11: Response in der Akuttherapie nach 6 bzw. 8 Wochen Behandlung

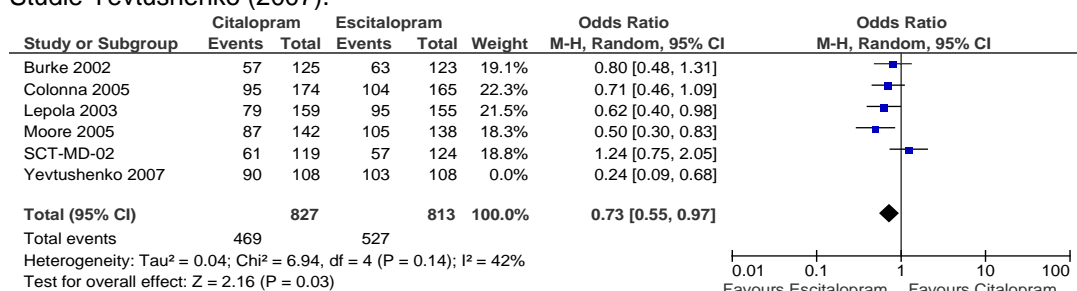


Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0

Fünf Studien zeigten konsistent Effekte zu Gunsten von Escitalopram, wobei die Überlegenheit in den Studien Lepola et al. (2005), Moore et al. (2005) und Yevtushenko et al. (2007) statistisch signifikant ist. Auch die Meta-Analyse aller Studien bestätigt die Signifikanz zu Gunsten von Escitalopram ( $p=0.02$ ).

Die Heterogenität in der Meta-Analyse wird durch die Studie von Yevtushenko et al. (2007) beeinflusst. In Abbildung 12 ist eine Meta-Analyse ohne diese Studie dargestellt, die die vorherigen Ergebnisse bestätigt. Escitalopram ist Citalopram statistisch signifikant überlegen ( $p=0.03$ ).

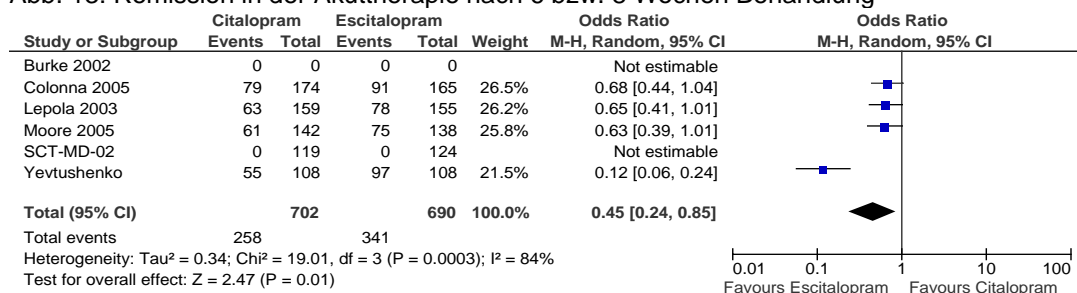
Abb. 12: Response in der Akuttherapie nach 6 bzw. 8 Wochen Behandlung unter Ausschluss der Studie Yevtushenko (2007).



## Remission

In Abbildung 13 sind die Ergebnisse zur Zielgröße Remission für die Akuttherapie dargestellt.

Abb. 13: Remission in der Akuttherapie nach 6 bzw. 8 Wochen Behandlung

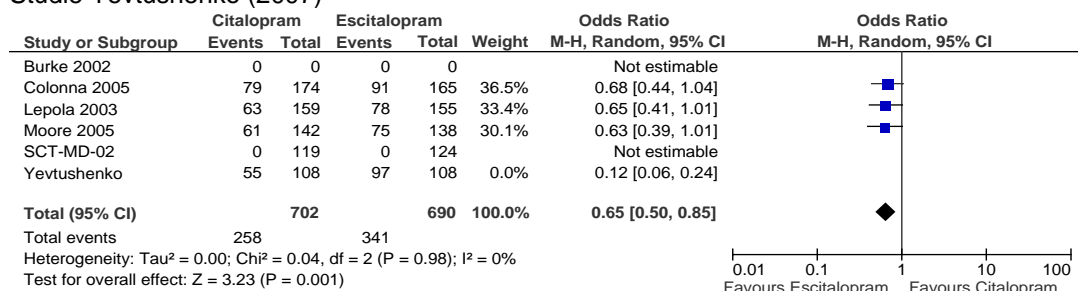


Alle Studien zeigten konsistent Effekte zu Gunsten von Escitalopram, wobei die Überlegenheit in der Studie Yevtushenko et al. (2007) statistisch signifikant ist. Auch die Meta-Analyse aller Studien bestätigt die Signifikanz zu Gunsten von Escitalopram ( $p=0.01$ ).

Die Heterogenität der Meta-Analyse wird durch die Studie Yevtushenko et al. (2007) beeinflusst. In Abbildung 14 ist eine Meta-Analyse ohne diese Studie berechnet. Die Ergebnisse belegen ebenfalls den statistisch signifikanten Effekt zu Gunsten von Escitalopram ( $p=0.001$ ).



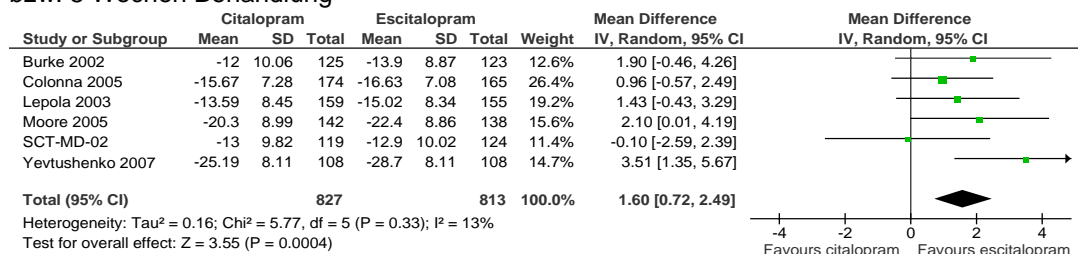
Abb. 14: Remission in der Akuttherapie nach 6 bzw. 8 Wochen Behandlung unter Ausschluss der Studie Yevtushenko (2007)



### Änderungen im MADRS total score im Vergleich zur Baseline

In Abbildung 15 ist die Meta-Analyse der Änderungen im MADRS total score im Vergleich zur Baseline dargestellt.

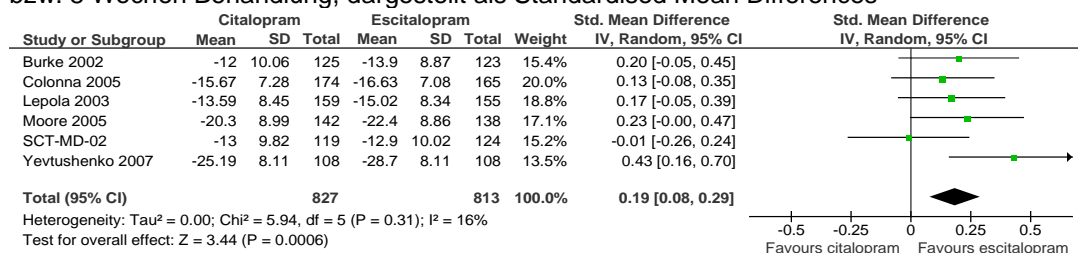
Abb. 15: Änderungen im MADRS total score im Vergleich zur Baseline in der Akuttherapie nach 6 bzw. 8 Wochen Behandlung



5 Studien zeigten konsistent Effekte zu Gunsten von Escitalopram, wobei die Überlegenheit in den Studien Moore et al. (2005) und Yevtushenko et al. (2007) statistisch signifikant ist. Die Meta-Analyse aller Studien belegt den Vorteil von Escitalopram gegenüber Citalopram statistisch signifikant.

In Abbildung 16 sind die Änderungen im MADRS total score im Vergleich zur Baseline als Effektschätzer dargestellt (nach der Methode von Hedge). Die Verwendung von weighted mean differences oder standardised mean differences ergibt geringfügig andere Ergebnisse wegen der unterschiedlichen Kalkulation der Standardfehler bzw. Effektgrößen. Daher unterscheiden sich die Gewichtungen der Studien bei den beiden Analysen (Abbildungen 15 und 16).

Abb. 16: Änderungen im MADRS total score im Vergleich zur Baseline in der Akuttherapie nach 6 bzw. 8 Wochen Behandlung, dargestellt als Standardised Mean Differences



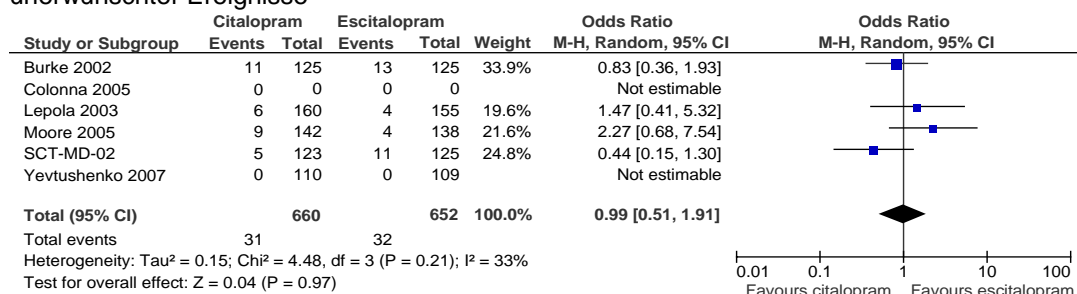
Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0

Wie bereits in Abbildung 15 dargestellt, zeigten 5 Studien konsistent Effekte zu Gunsten von Escitalopram, wobei die Überlegenheit in den Studien Moore et al. (2005) und Yevtushenko et al. (2007) statistisch signifikant ist. Die Meta-Analyse (Abbildung 16) bestätigt den statistisch signifikanten Vorteil von Escitalopram gegenüber Citalopram. Die Effektgröße beträgt 0.19 ( $p=0.0006$ ).

### Studienabbruch wegen unerwünschter Ereignisse

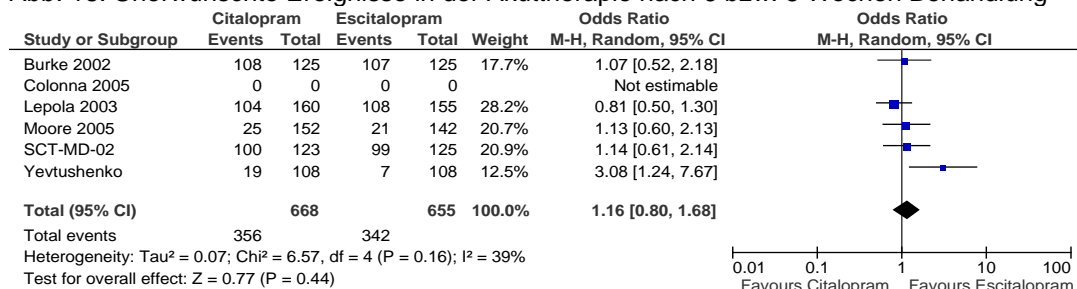
In Abbildung 17 sind die Studienabbrüche in den Akuttherapiestudien wegen unerwünschter Ereignisse dargestellt. In der Studie Yevtushenko et al. (2007) wurden keine Abbrüche wegen unerwünschter Ereignisse im Escitalopram-Arm und Citalopram-Arm berichtet.

Abb. 17: Studienabbrüche in der Akuttherapie nach 6 bzw. 8 Wochen Behandlung wegen unerwünschter Ereignisse



Die Meta-Analyse zeigt keine Unterschiede in Bezug zu Studienabbrüchen wegen unerwünschter Ereignisse zwischen den beiden Substanzen. Die berechnete odds ratio ist annähernd bei 1.

Abb. 18: Unerwünschte Ereignisse in der Akuttherapie nach 6 bzw. 8 Wochen Behandlung



Unerwünschte Ereignisse treten unter Escitalopram weniger häufig ein als unter Citalopram (Abbildung 18). Der Effekt zu Gunsten von Escitalopram ist numerisch.

## **Erhaltungstherapie (Langzeittherapie)**

Die Studie von Colonna et al. (2005) ging über 24 Wochen. Zum Studienende nach 24 Wochen waren die Remissionsraten (76% vs. 71%), die Responderaten (80% vs. 78%) und die Änderungen im MADRS total score (-21.5 vs. -20.5) unter Escitalopram numerisch besser als unter Citalopram. Escitalopram war Citalopram statistisch signifikant im Endpunkt CGI-S score überlegen (ITT, LOCF,  $p < 0.05$ ).

Unter Escitalopram gab es weniger Studienabbrüche wegen unerwünschter Ereignisse als unter Citalopram (6.1% vs. 10.3%). Der Anteil an Patienten mit unerwünschten Ereignissen war unter Escitalopram ebenfalls niedriger als unter Citalopram (62.9% vs. 72%).

## **4 Diskussion**

Hinweis: Wichtige ergänzende Informationen zu den im Folgenden diskutierten Aspekten sind im Appendix aufgeführt.

### **4.1. SNRI Duloxetin und Venlafaxin XR vs. Escitalopram**

#### **4.1.1. Duloxetin vs. Escitalopram**

##### **Kurzzeit-Akutstudien**

Das IQWiG hat die Studien Kahn et al. (2007) und Nierenberg et al. (2007) in die Analyse der Kurzzeit-Akutstudien eingeschlossen. Die Studien Wade et al. (2007) und Pigott et al. (2007) wurden als Langzeit-Akutstudien eingestuft. Die Effektstärke basierend auf den Arbeiten von Khan et al. (2007) und Nierenberg et al. (2007) wurde in dem IQWiG-Vorbericht wegen der Heterogenität nicht kalkuliert.

Das Design der Studie Wade et al. (2007) erlaubt die Einbeziehung der 8 Wochen Daten in die Metaanalyse der Kurzzeit-Akutstudien, da die Erhebung der 8 Wochen Daten als sekundäres Zielkriterium im Studienplan definiert wurde.

Die Inklusion der Publikation von Nierenberg et al. (2007) ist nicht gerechtfertigt, da die ausgewerteten Depressions-Skalen nicht für einen, wie in der Studie durchgeführten, wöchentlichen Beobachtungszeitraum validiert sind. Daher ist auch die Validität der Ergebnisse nicht gewährleistet.

##### **Response**

Unter Einbeziehung der 8 Wochen Daten aus der Studie Wade et al. (2007) zeigt sich eine numerische Überlegenheit von Escitalopram gegenüber Duloxetin (OR 0.83; 95-CI 0.54-1.27;  $I^2 = 65\%$ ).

Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI)  
bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0

Die ausschließliche Berücksichtigung der MADRS-Skala ergibt eine signifikante Überlegenheit von Escitalopram gegenüber Duloxetin (OR 0.56; 95-CI 0.40-0.80;  $I^2=0\%$ ). Hierbei zeigen beide Einzelstudien eine signifikante Überlegenheit von Escitalopram gegenüber Duloxetin.

Die Vergleichsstudien zwischen Escitalopram und Duloxetin zeigen somit im meta-analytischen Ansatz eine numerische Überlegenheit von Escitalopram gegenüber Duloxetin im Zielkriterium Response. Ferner zeigt der Endpunkt Response in 2 voneinander unabhängig durchgeführten RCTs eine numerische Überlegenheit von Escitalopram gegenüber Duloxetin (Khan et al., 2007; Wade et al., 2007). Die Ergebnisse zur Response sind somit sowohl in der Meta-Analyse und den einzelnen RCTs konsistent. Die Resultate zur Response zeigen den Nutzen von Escitalopram im Vergleich zu Duloxetin.

### **Remission**

Im IQWiG-Vorbericht werden die Studien von Kahn et al. (2007) und Nierenberg et al. (2007) eingeschlossen, ohne dass ein klarer Effekt festgestellt werden konnte. Bei Einschluss der 8 Wochen Daten aus der Studie Wade et al. (2007) wurde eine numerische Überlegenheit von Escitalopram gegenüber Duloxetin kalkuliert (OR 0.96; 95-CI 0.64-1.44;  $I^2=61\%$ ).

Die ausschließliche Berücksichtigung der MADRS-Skala ergibt eine numerische Überlegenheit von Escitalopram gegenüber Duloxetin (OR 0.75; 95-CI 0.53-1.05;  $I^2=0\%$ ).

Die Vergleichsstudien zwischen Escitalopram und Duloxetin zeigen somit im metaanalytischen Ansatz eine numerische Überlegenheit von Escitalopram gegenüber Duloxetin im Zielkriterium Remission. Ferner zeigt der Endpunkt Remission in 2 voneinander unabhängig durchgeführten RCTs eine numerische Überlegenheit von Escitalopram gegenüber Duloxetin (Khan et al., 2007; Wade et al., 2007). Die Ergebnisse zur Remission sind somit sowohl in der Meta-Analyse und den einzelnen RCTs konsistent. Die Resultate zur Remission weisen somit auf den größeren Nutzen von Escitalopram im Vergleich zu Duloxetin hin.

### **Depressionsskalen-Gesamtscore**

Die Ergebnisse der IQWiG-Berechnungen zeigen keinen klaren Effekt. Bei Einschluss der 8-Wochen Daten von Wade et al. (2007) in die Stellungnahme von Lundbeck zeigt sich bei der Berechnung der HAMD17 Daten eine numerische Überlegenheit von Escitalopram gegenüber Duloxetin (MD 0.65; 95-CI -0.62-1.93;  $I^2=55\%$ ). Die Berechnung der MADRS Daten zeigt eine signifikante Überlegenheit von Escitalopram gegenüber Duloxetin (MD 2.09; 95-CI 0.27-3.90;  $I^2=0\%$ ).

### **Studienabbruch wegen unerwünschter Ereignisse**

Die Berechnungen des SNRI Vorberichtes ergeben keinen klaren Effekt aufgrund der Heterogenität der beiden eingeschlossenen Studien Kahn et al. (2007) und Nierenberg

Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI)  
bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0

et al. (2007). Werden die 8 Wochen Daten der Arbeit Wade et al. (2007) in die Meta-Analyse eingeschlossen, resultiert eine statistisch signifikant geringere Abbruchrate aufgrund von unerwünschten Ereignissen bei Escitalopram im Vergleich zu Duloxetin.

Die drei unabhängig voneinander durchgeführten RCTs zeigen alle konsistent eine numerisch geringere Rate von Studienabbrüchen wegen unerwünschter Ereignisse bei Escitalopram gegenüber Duloxetin (Khan et al., 2007; Nierenberg et al., 2007; Wade et al., 2007). Dieses Ergebnis wird im meta-analytischen Ansatz bestätigt. Das Verhältnis der Studienabbrüche wegen unerwünschter Ereignisse zeigt einen statistisch signifikanten Unterschied zwischen Escitalopram und Duloxetin (OR 2.23; 95-CI 1.06-4.68;  $I^2= 53\%$ ). Es resultiert eine klinisch relevante bessere Arzneimittelsicherheit für Escitalopram im Vergleich zu Duloxetin.

### **Fazit Duloxetin vs. Escitalopram**

Escitalopram erweist sich in den genannten Zielkriterien im meta-analytischen Ansatz Duloxetin numerisch bzw. statistisch signifikant überlegen. Ferner zeigen 2 von 3 RCTs in diesen Endpunkten einen gleichsinnigen, konsistenten Effekt. Der Effekt weist auf einen Zusatznutzen von Escitalopram gegenüber Duloxetin hin. Darüber hinaus zeigt Escitalopram ein besseres Verträglichkeitsprofil als Duloxetin.

Escitalopram zeigt ein besseres Nutzen-Risiko-Profil als Duloxetin.

## **4.1.2. Venlafaxin XR vs. Escitalopram**

### **Kurzzeit-Akutstudien**

#### **Wirksamkeit**

Die Wirksamkeit (Prä-post-Vergleich des Verlaufs der Depression, Response) von Escitalopram ist mit der von Venlafaxin XR vergleichbar. Escitalopram ist dem Venlafaxin XR in dem Zielparameter Remission numerisch überlegen. Es zeigte sich, dass Patienten unter Escitalopram 4,6 Tage schneller eine anhaltende Response und 6,6 Tage schneller eine anhaltende Remission erreichen. Eine Verzögerung der Response oder Remission in der Akutbehandlung der Depression kann folgendes bedingen:

- zunehmende Inanspruchnahme medizinischer Dienstleistungen
- früheres Wiederauftreten der Depression
- kürzere Intervalle zwischen den Depressionen
- anhaltendes suizidales Risiko
- erhöhtes Risiko für medikamentösen Missbrauch
- erhöhtes Risiko der Entwicklung einer Therapie-resistenten Depression

Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI)  
bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0

Insgesamt werden die Gesamtkosten erhöht und die Gesundheitssysteme häufiger frequentiert.

Patienten unter Escitalopram haben in der 8-wöchigen Akuttherapie statistisch signifikant mehr Tage ohne Depression (30.4 Tage) als unter Venlafaxin (26.2 Tage) ( $p < 0.05$ ). Der Nutzen von Escitalopram gegenüber Venlafaxin XR wurde mit 1.46 kalkuliert. Dies zeigt, dass Patienten unter Escitalopram stärker von der Medikation profitieren als die Patienten, die mit Venlafaxin XR behandelt wurden (Montgomery et al., 2006). Thase et al. (2001) und Thompson et al. (2002) haben einen Wert für den klinisch relevanten Unterschied für Vergleiche von zwei aktiven Wirksubstanzen definiert. Eine Differenz von 2 Punkten auf einer psychometrischen Messskala ist klinisch bedeutsam. Die Differenz zwischen Escitalopram und Venlafaxin XR betrug nach 8-wöchiger Therapie 2.3 Punkte auf der MADRS-Skala (Bielski et al., 2004).

### **Studienabbruch wegen unerwünschter Ereignisse**

In der Stellungnahme von Lundbeck wurden die Studienabbrüche aufgrund von unerwünschten Ereignissen metaanalytisch ausgewertet. Es wurden numerisch mehr Studienabbrüche unter Venlafaxin XR als unter Escitalopram beobachtet (OR 2.33; 95-CI 0.75-7.21;  $I^2 = 63\%$ ).

Die Studie von Bielski et al. (2004) zeigt eine signifikant höhere Abbruchrate aufgrund von unerwünschten Ereignissen unter Venlafaxin XR als unter Escitalopram (OR 4.48; 95-CI 1.44-13.92).

Im Vorbericht zu SNRI wurden die unterschiedlichen Dosierungen, die in diesen beiden Studien angewandt wurden, nicht berücksichtigt. Während bei Bielski et al. (2004) beide Antidepressiva in der Höchstdosierung gegeben wurden (Escitalopram 20 mg/d, Venlafaxin XR 225 mg/d), war in der Studie von Montgomery et al. (2004) nur eine Aufdosierung von bis zu 150 mg/d in der Venlafaxin Gruppe möglich. Es ist bekannt, dass die noradrenerge Wiederaufnahmehemmung von Venlafaxin erst in höheren Dosen anfängt klinisch bedeutsam zu sein. In diesem Zusammenhang wird von Möller ausgeführt, dass bei Venlafaxin "nur bei höheren therapeutischen Dosen eine 50%-ige Hemmung der NA-Wiederaufnahme zu erwarten ist" (Seite 72, dritte Spalte, unten; Möller et al., 2004). Weiter führen die Autoren aus, "Demnach ist im oberen Bereich der Dosierung auch für Venlafaxin mit einem dualen Wirkmechanismus auf der 5-HT- und NA-Transporter zu rechnen" (S73, 1. Spalte, 1. Absatz; Möller et al., 2004).

Die Unterschiede in der Verträglichkeit zwischen Escitalopram und Venlafaxin XR wird auf dieser zusätzlichen Noradrenalin-Wiederaufnahmehemmung basieren. Aufgrund dieser beiden Studien können folgende Aussagen gemacht werden:

Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI)  
bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0

1. Wird Venlafaxin XR  $\leq$  150 mg/d dosiert, wurde bisher kein Unterschied zu Escitalopram (10-20 mg/d) in der Abbruchrate aufgrund von unerwünschten Ereignissen beobachtet.
2. Wird Venlafaxin XR mit der Höchstdosis von 225 mg/d verabreicht, zeigt sich eine signifikant höhere Studienabbruchrate aufgrund von unerwünschten Ereignissen im Vergleich zu der Höchstdosierung von Escitalopram 20 mg/d.

Diese Schlussfolgerung wird auch in der Re-Analyse beider Studien (Montgomery 2006) gezogen.

### **Fazit Venlafaxin XR vs. Escitalopram**

Escitalopram erweist sich im meta-analytischen Ansatz Venlafaxin XR numerisch überlegen. Escitalopram weist klinisch relevante Vorteile in den Zielgrößen Response und Remission auf, da diese Zielparameter unter Escitalopram früher erreicht werden. Der schnellere Wirkeintritt ist von Relevanz für die Therapie depressiver Patienten und ist daher entsprechend zu gewichten.

Die zwei unabhängig voneinander durchgeführten RCTs zeigen konsistent eine numerisch bzw. statistisch signifikant geringere Rate von Studienabbrüchen wegen unerwünschter Ereignisse im Vergleich zum aktiven Komparator Venlafaxin XR (Bielski et al., 2004; Montgomery et al., 2004). Dieses Ergebnis wird im meta-analytischen Ansatz bestätigt (OR 2.33; 95-CI 0.75-7.21;  $I^2 = 63\%$ ). Es resultiert eine klinisch relevante bessere Arzneimittelsicherheit für Escitalopram im Vergleich zu Venlafaxin XR.

Escitalopram zeigt ein besseres Nutzen-Risiko-Verhältnis als Venlafaxin XR.

### **4.2 Escitalopram vs. Citalopram**

Im Vorbericht des IQWiG werden einzelne SSRI den SNRI gegenübergestellt, jedoch Aussagen für die gesamte Wirkstoffgruppe der SSRI abgeleitet. Hierdurch entsteht irrtümlicherweise der Eindruck, dass innerhalb der SSRI keine Unterschiede in Wirksamkeit und Verträglichkeit zwischen den einzelnen Wirkstoffen bestehen.

Der Vergleich des SSRI Escitalopram mit der Leitsubstanz der SSRI – dem Wirkstoff Citalopram – verdeutlicht, dass relevante Nutzenunterschiede bestehen.

## **Kurzzeit-Akutstudien**

### **Response**

Fünf Studien zeigten konsistent Effekte zu Gunsten von Escitalopram, wobei die Überlegenheit in den Studien Lepola et al. (2005), Moore et al. (2005) und Yevtushenko et al. (2007) statistisch signifikant ist. Auch die Meta-Analyse aller Studien bestätigt die Signifikanz zu Gunsten von Escitalopram (OR 0.68; 95-CI 0.49-0.93;  $I^2 = 55\%$ ). Eine Sensitivitätsanalyse ohne die Studie von Yevtushenko et al. (2007) bestätigt die vorherigen Ergebnisse.

Escitalopram ist Citalopram statistisch signifikant überlegen. Gemäß der Definition im Vorbericht SNRI (Version 1.0, S. 613) ist der Beleg für den Zusatznutzen durch Escitalopram gegenüber Citalopram erbracht.

### **Remission**

In die Meta-Analyse konnten vier Studien einbezogen werden. Alle Studien zeigten konsistent Effekte zu Gunsten von Escitalopram, wobei die Überlegenheit in der Studie Yevtushenko et al. (2007) statistisch signifikant ist. Auch die Meta-Analyse aller Studien bestätigt die Signifikanz zu Gunsten von Escitalopram (OR 0.45; 95-CI 0.24-0.85;  $I^2 = 84\%$ ). Eine Sensitivitätsanalyse ohne die Studie von Yevtushenko et al. (2007) bestätigt die vorherigen Ergebnisse.

Auch in der Patienten-relevanten Zielgröße Remission ist der Zusatznutzen von Escitalopram gegenüber Citalopram belegt (Definition: Vorbericht SNRI, Version 1.0, S. 613).

### **Depressionsskalen-Gesamtscore**

Fünf Studien zeigten konsistent Effekte zu Gunsten von Escitalopram, wobei die Überlegenheit in den Studien Moore et al. (2005) und Yevtushenko et al. (2007) statistisch signifikant ist. Die Meta-Analyse aller Studien belegt den statistisch signifikanten Vorteil von Escitalopram gegenüber Citalopram (MD 1.6; 95-CI 0.72-0.29;  $I^2 = 13\%$ ). Die Differenz von 1.6 Punkten auf der MADRS Skala ist klinisch bedeutsam (Duru et al., 2008) und belegt die klinische Relevanz des Zusatznutzen durch Escitalopram.

Die Pool-Analysen von Gorman et al. (2002) und Lepola et al. (2004) bestätigen die statistisch signifikante und klinisch relevante Überlegenheit von Escitalopram gegenüber Citalopram in den für das IQWiG relevanten Zielgrößen Response, Remission und Veränderung der Symptomatik zwischen Studienbeginn und –ende.

### **Sicherheit und Verträglichkeit**

Zwischen Escitalopram und Citalopram zeigten sich keine signifikanten Unterschiede bezüglich der Sicherheit und Verträglichkeit. Escitalopram ist bei überlegener klinischer Wirksamkeit genauso sicher wie Citalopram.

Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI)  
bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0



### **Fazit Escitalopram vs. Citalopram**

Escitalopram ist Citalopram klinisch relevant überlegen und sicher. Der Zusatznutzen ist durch head-to-head Studien belegt und wird durch Meta-Analysen und Pool-Analysen bestätigt.

Zu dieser Bewertung kommen auch unabhängige Health Technology Assessments (HTA). Die schwedische Dental and Pharmaceutical Benefits Agency TLV (Tandvårds- och Läkemedelsförmånsverket) hat im Dezember 2008 einen Review zu Antidepressiva abgeschlossen (Wessling et al., 2008). Die TLV kommt zu folgendem Ergebnis: *"Treatment with escitalopram gives the highest probability to reach remission (recovery) in comparison to all other modern antidepressants."*

Das norwegische Knowledge Centre (Nasjonalt kunnskapssenter for helsetjenesten) veröffentlichte im Jahr 2007 einen Review zu Antidepressiva (Saeterdal et al., 2007) und kam in der Zusammenfassung zu folgendem Ergebnis: *"Escitalopram was significantly more effective (response and remission rate) than Citalopram"*.

Diese Aussagen bestätigen den unabhängigen HTA-Bericht "Drug Class Review on Second Generation Antidepressant" von Gartlehner et al. (2006): *"Overall, effectiveness and efficacy were similar and the majority of trials [55 head-to-head trials compared effectiveness and efficacy of one new AD to another] did not identify substantial differences among drugs. The only exception is the comparison of citalopram to escitalopram. Four fair to good trials indicate consistently that escitalopram has a greater efficacy for the treatment of MDD than citalopram."*

Aus den dargelegten Gründen muss Escitalopram unter Einbezug der vorhandenen Evidenz ausschließlich als Einzelsubstanz in den Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen“ einbezogen und bewertet werden. Verallgemeinernde Schlussfolgerungen für die Gruppe der SSRI unter Einbezug von Escitalopram sind nicht zulässig.

## Zusammenfassung

Lundbeck stimmt nicht mit den vorläufigen Schlussfolgerungen des vorliegenden Vorberichtes überein, dass Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) und Selektive Serotonin-Wiederaufnahmehemmer (SSRI) – insbesondere Escitalopram – äquivalent in Wirksamkeit und Sicherheit seien.

Im Vorbericht zur Nutzenbewertung der SNRI wird fälschlicherweise dargestellt, dass alle SSRI gleich wären. Diese inkorrekte Schlussfolgerung wird durch head-to-head Vergleiche einzelner SSRI mit Duloxetin oder Venlafaxin als Komparator gezogen. Aus dem Vergleich von Wirkstoff A zu Wirkstoff B und Wirkstoff A zu Wirkstoff C ist jedoch keine evidenz-basierte Aussage zum indirekten Verhältnis Wirkstoff B zu Wirkstoff C abzuleiten, wenn direkte Vergleichsstudien Wirkstoff B zu Wirkstoff C vorliegen. Im Vorbericht hingegen wird Äquivalenz von Escitalopram zu den anderen SSRI unterstellt. Diese Schlussfolgerung wäre nur zulässig, wenn eine Bewertung zu SSRI analog der SNRI Bewertung vorliegen würde.

In der Stellungnahme der Lundbeck werden die klinisch relevanten Vorteile von Escitalopram im Vergleich zu Duloxetin und Venlafaxin XR anhand von Wirksamkeits- und Sicherheitsunterschieden zu Gunsten von Escitalopram belegt.

Der direkte Vergleich anhand von head-to-head Studien von Escitalopram mit dem Wirkstoff Citalopram, der von den Bundesvertragspartnern als Leitsubstanz der SSRI bestimmt wurde, belegt den statistisch signifikanten und klinisch relevanten Zusatznutzen von Escitalopram in den Patienten-relevanten Zielgrößen Response, Remission und Veränderung der Symptomatik zwischen Studienbeginn und –ende.

Auf Basis der vorliegenden Evidenz aus Einzel-RCTs und Meta-Analysen, die in der Stellungnahme der Lundbeck dargestellt wird, wird ein konsistenter Zusatznutzen von Escitalopram im Vergleich zu den Komperatoren dokumentiert. Zu dieser Bewertung kommen auch unabhängige Health Technology Assessments (HTA):

- Die schwedische Dental and Pharmaceutical Benefits Agency TLV hat im Dezember 2008 einen Review zu Antidepressiva abgeschlossen (Wessling et al., 2008) und kommt zu folgendem Ergebnis: *„Treatment with escitalopram gives the highest probability to reach remission (recovery) in comparison to all other modern antidepressants.“*
- Das norwegische Knowledge Centre veröffentlichte im Jahr 2007 einen Review zu Antidepressiva (Saeterdal et al., 2007) und kam in der Zusammenfassung zu folgendem Ergebnis: *„Escitalopram was significantly more effective (response and remission rate) than citalopram“.*
- Der HTA-Bericht *“Drug Class Review on Second Generation Antidepressant“* von Gartlehner et al. (2006) belegt: *“Overall, effectiveness and efficacy were similar and the majority of trials [55 head-to-head trials compared effectiveness and efficacy of one new AD to another] did not identify substantial differences among drugs. The only exception is the comparison of citalopram to escitalopram. Four fair to good*

Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI)  
bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0

*trials indicate consistently that escitalopram has a greater efficacy for the treatment of MDD than citalopram."*

Aus den dargelegten Gründen soll Escitalopram unter Einbezug der vorhandenen Evidenz ausschließlich als Einzelsubstanz in den Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen“ einbezogen und bewertet werden. Verallgemeinernde Schlussfolgerungen für die Gruppe der SSRI unter Einbezug von Escitalopram sind nicht zulässig.

Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI)  
bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0

- 25 -

- A 68 -

## Literatur

- 1) Auquier, Pascal et al. "Comparison of escitalopram and citalopram efficacy: A meta-analysis." International Journal of Psychiatry in Clinical Practice 2003, Volume 7, Pages 259-268.
- 2) Bielski, R. J., D. Ventura, and C. C. Chang. "A double-blind comparison of escitalopram and venlafaxine extended release in the treatment of major depressive disorder." Journal of Clinical Psychiatry 65.9 (2004): 1190-96.
- 3) Blier, Pierre. "Pharmacology of Rapid-Onset Antidepressant Treatment Strategies." J Clin Psychiatry 2001;62 (Suppl. 15).
- 4) Burke, William J., Gergel, Ivan, Bose, Anjana. "Fixed-Dose Trial of the Single Isomer SSRI Escitalopram in Depressed Outpatients." J Clin Psychiatry 63:4, April 2002.
- 5) Colonna, L, H. F Andersen, and E. H. Reines. "A randomized, double-blind, 24-week study of escitalopram (10 mg/day) versus citalopram (20 mg/day) in primary care patients with major depressive disorder." Current Medical Research and Opinion 21.10 (2005): 1659-68.
- 6) Colonna L, Menard F, Andersen HF. Escitalopram versus Citalopram: More Efficacious and Well Tolerated in Long-Term Treatment of Moderately Depressed Patients. Poster 2003;16th Congress of the European College of Neuropsychopharmacology (ECNP), Prague, Czech Republic, 20-24 September, 2003 (P.1.097).
- 7) DGPPN (Deutsche Gesellschaft für Psychiatrie, Psychotherapie, und Nervenheilkunde) Praxisleitlinien in Psychiatrie und Psychotherapie, Band 5 Behandlungsleitlinie Affektive Störung, Steinkopff, Darmstadt
- 8) Duru, Gérard; Fantino, Bruno. "The Clinical Relevance of Changes in the Montgomery-Asberg Depression Rating Scale Using the Minimum Clinically Important Difference Approach." Current Medical Research and Opinion Vol. 24, No. 5, 2008, 1329-1335.
- 9) Fachinformation Cipralext (Stand: April 2008)
- 10) Fachinformation Cymbalta (Stand: Juli 2008)
- 11) Gartlehner, G. et al. "Drug Class Review on Second Generation Antidepressants." RTI-UNC Evidence-based Practice Center Cecil G. Sheps Center for Health Services Research University of North Carolina at Chapel Hill, 2006.
- 12) Gorman, J. M., A. Korotzer, and G. Su. "Efficacy Comparison of Escitalopram and Citalopram in the Treatment of Major Depressive Disorder: Pooled Analysis of Placebo-Controlled Trials." CNS Spectrums 7.Suppl. 1 (2002): 40-44.
- 13) Hirschfeld, Robert M.A.; Vornik, Lana A. „Newer Antidepressants: Review of Efficacy and Safety of Escitalopram and Duloxetine.“ J Clin Psychiatry 2004; 65 [Suppl. 4]: 46-52.
- 14) Khan, A., et al. "Double-Blind Comparison of Escitalopram and Duloxetine in the Acute Treatment of Major Depressive Disorder." Clinical Drug Investigation 27.7 (2007): 481-92.

Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0

- 15) Lam, Raymond W., Henning F. Andersen, and Alan G. Wade. "Escitalopram and duloxetine in the treatment of major depressive disorder: a pooled analysis of two trials." International Clinical Psychopharmacology 23.4 (2008): 181-87.
- 16) Leon, Andrew C. "An Ideal Trial to Test Differential Onset of Antidepressant Effect." J Clin Psychiatry 2001;62 (Suppl. 4).
- 17) Lepola, Ulla M., Loft, Henrik, Reines, Elin Heldbo. "Escitalopram (10-20 mg/day) is effective and well tolerated in a placebo-controlled study in depression in primary care." International Clinical Psychopharmacology 2003, 18:211-217.
- 18) Lepola, U., A. Wade, and H. F. Andersen. "Do equivalent doses of escitalopram and citalopram have similar efficacy? A pooled analysis of two positive placebo-controlled studies in major depressive disorder." International Clinical Psychopharmacology 19.3 (2004): 149-55.
- 19) Lieberman, Jeffery A., et al. "Comparing the Effects of Antidepressants: Consensus Guidelines for Evaluating Quantitative Reviews of Antidepressant Efficacy." Neuropsychopharmacology (2005) 30, 445-460.
- 20) Llorca, P. M., et al. "Efficacy of escitalopram in patients with severe depression: a pooled analysis." International Journal of Clinical Practice 59.3 (2005): 268-75.
- 21) Llorca, P.M.; Fernandez, J.-L. "Escitalopram in the treatment of major depressive disorder: clinical efficacy, tolerability and cost-effectiveness vs. venlafaxine extended-release formulation." Int J Clin Pract, April 2007, 61, 4, 702-710.
- 22) Mallinckrodt C., Prakash A., Andorn A. et al. "Duloxetine for the treatment of major depressive disorder: A closer look at efficacy and safety data across the approved dose range" J. Psychiatric Research 2006; 40: 337-348
- 23) Möller E. et al. "Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SSNRI) Antidepressiva mit dualem Wirkmechanismus. Psychopharmakotherapie 2004, 3: 71-75
- 24) Montgomery, Stuart A. "New Developments in the Treatment of Depression." J Clin Psychiatry 1999;60 (Suppl. 14).
- 25) Montgomery SA, Loft H, Sanchez C, Reines EH, Papp M. Escitalopram (S-Enantiomer of Citalopram): Clinical Efficacy and Onset of Action Predicted from a Rat Model. Pharmacol Toxicol 2001; 88(5):282-6
- 26) Montgomery, S. A., A. K. T. Huusom, and J. Bothmer. "A Randomised Study Comparing Escitalopram with Venlafaxine XR in Primary Care Patients with Major Depressive Disorder." Neuropsychobiology 50.1 (2004): 57-64.
- 27) Montgomery, Stuart A., Andersen, Henning F. "Escitalopram versus venlafaxine XR in the treatment of depression." International Clinical Psychopharmacology 2006, 21:297-309.
- 28) Moore, N., H. Verdoux, and D. Fantino. "Escitalopram and citalopram in the treatment of outpatients with major depressive disorder." European Neuropsychopharmacology 15.Suppl. 3 (2005): S426.

Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI)  
bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0

- 29) Mulder, Roger T. et al. "Six Months of Treatment for Depression: Outcome and Predictors of the Course of Illness." Am J Psychiatry 2006; 163:95-100.
- 30) Murdoch, David, Keam, Susan J. "Escitalopram – A Review of its Use in the Management of Major Depressive Disorder." Drugs 2005; 65(16); 2379-2404.
- 31) Nierenberg, A. A., et al. "Duloxetine versus escitalopram and placebo in the treatment of patients with major depressive disorder: onset of antidepressant action, a non-inferiority study." Current Medical Research and Opinion 23.2 (2007): 401-16..
- 32) Pigott, T. A., et al. "Duloxetine versus escitalopram and placebo: an 8-month, double-blind trial in patients with major depressive disorder." Current Medical Research and Opinion 23.6 (2007): 1303-18 (16).
- 33) Rahmenvorgaben nach § 84 Abs. 7 SGB V – Arzneimittel – für das Jahr 2009 vom 30. September 2008.
- 34) Saeterdal, Ingvil et al. "Effekt og sikkerhet for SSRI og andre nyere antidepressive legemidler ved depresjon hos voksne." Nasjonalt kunnskapssenter for helsetjenesten Review 2007.
- 35) Sanchez, Connie, et al. "Escitalopram versus Citalopram: die überraschende Rolle des R-Enantiomers." Psychopharmacology (2004) 174: 163-176.
- 36) Study SCT-MD-02, 2005 "Flexible-Dose Comparison of the Safety and Efficacy of Lu 26-054 (Escitalopram), Citalopram, and Placebo in the Treatment of Major Depressive Disorder  
[http://www.forestclinicaltrials.com/CTR/CTRController/CTRViewPdf?\\_file\\_id=scsr/SCSR\\_SCT-MD-02\\_final.pdf](http://www.forestclinicaltrials.com/CTR/CTRController/CTRViewPdf?_file_id=scsr/SCSR_SCT-MD-02_final.pdf)
- 37) Thase, Michael E. "Treatment of Severe Depression." J Clin Psychiatry 2000;61 (Suppl. 1).
- 38) Thase ME. Methodology to measure onset of action. J Clin Psychiatry 2001; 62(Suppl. 15):18-21.
- 39) Thompson, Chris. "Onset of action of antidepressants: results of different analyses." Hum Psychopharmacol Clin Exp 2002; 17: S27-S32.
- 40) Volz, Hans-Peter. "Chancen und Möglichkeiten der neuen Antidepressiva." Psychopharmakotherapie 2005; Supplement 13:45-50.
- 41) Wade, A., Lemming, Michael, Hedegaard, K. Bang. "Escitalopram 10mg/day is effective and well tolerated in a placebo-controlled study in depression in primary care." International Clinical Psychopharmacology 2002, 17:95-102.
- 42) Wade, Alan, Andersen, Henning Friis. "The onset of effect for escitalopram and its relevance for the clinical management of depression." Current Medical Research and Opinion Vol. 22, No. 11, 2006, 2101-2110.
- 43) Wade, A. Gembert, K. Florea, I. "A comparative study of the efficacy of acute and continuation treatment with escitalopram versus duloxetine in patients with major

Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI)  
bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0

depressive disorder." Current Medical Research and Opinion Vol. 23, No. 7, 2007, 1605-1614.

- 44) Wade, A., K. Gembert, and I. Florea. "Comparative study of the efficacy of acute and continuation treatment with escitalopram versus duloxetine in patients with major depressive disorder." European Psychiatry 23.Suppl. 2 (2008): S268.
- 45) Wessling, Anders; Ramsberg, Joakim. "Genomgången av läkemedel mot depression." Tandvårds- och Läkemedelsförmånsverket Review 2008.
- 46) Yevtushenko, Valery Y., et al. "Efficacy and Tolerability of Escitalopram Versus Citalopram in Major Depressive Disorder: A 6-Week, Multicenter, Prospective, Randomized, Double-Blind, Active-Controlled Study in Adult Outpatients." Clinical Therapeutics 29.11 (2007): 2319-32.

Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI)  
bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0

## Appendix

### A.1 Diskussion zur methodischen Vorgehensweise der Nutzenbewertung

#### A.1.1 Berücksichtigung der gesamten klinischen Evidenz zur Nutzenbewertung

Pool-Analysen sind in die Nutzenbewertung des IQWiG einzuschließen. Entsprechend den Ausführungen des „*Centre For Evidence-Based Medicine*“ bilden randomisierte klinische Studien (RCTs) die höchste Evidenzhierarchie in der Beurteilung von Pharmaka. Außerdem sind systematische Analysen (z.B. Pool-Analysen) bei der Beurteilung der Wirksamkeit und Verträglichkeit berücksichtigt, da auf diese Weise Tendenzen aus den RCTs zu einer zentralen Aussage zusammengefasst werden können und ein etwaiger Bias der RCTs reduziert werden kann (Egger et al., 2001).

Vergleichsstudien, die vom IQWiG für die Nutzenbewertung verwendet werden, sind häufig als „Nicht-Überlegenheitsstudien“ konzipiert worden und werden entsprechend mit geringen Fallzahlen durchgeführt. Diese Analysen sind häufig unterpowerd und daher nicht geeignet einen statistisch signifikanten oder klinisch relevanten Unterschied zu belegen. Aus diesem Grunde stellt die Pool-Analyse eine geeignete Methode zur Evaluierung der Effektivität bei dieser Patientengruppe dar (Lieberman et al., 2005). In der Pool-Analyse wird die Fallzahl erhöht und damit die Power, einen Unterschied zwischen zwei klinisch aktiven Substanzen zu detektieren (Gorman et al., 2002). Außerdem kann das Risiko einer Heterogenität reduziert werden.

Pool-Analysen, die die antidepressive Effektivität von Escitalopram gegen den im Vorbericht aufgeführten Komparatoren Duloxetin und Venlafaxin XR sowie exemplarisch gegenüber Citalopram dokumentieren, sind:

- Escitalopram *versus* Duloxetin (Lam et al., 2008)
- Escitalopram *versus* Venlafaxin XR (Montgomery and Andersen, 2006)
- Escitalopram *versus* Citalopram (Gorman et al., 2002; Lepola et al., 2004; Llorca et al., 2005)

#### Escitalopram vs. Duloxetin – Pool-Analysen

Lam und Mitarbeiter haben zwei vom IQWiG bewertete Vergleichsstudien zwischen Escitalopram und Duloxetin in einer Pool-Analyse re-analysiert (Lam et al. 2008):

- Khan et al. (2007),
- Wade et al. (2007).

Es wurden die Daten zur Effektivität und Verträglichkeit für die Gesamtpopulation mit jenen der schwer depressiven Patienten (initialer MADRS-Score  $\geq 30$  Punkte) verglichen. Der therapeutische Unterschied im primären Zielkriterium der RCTs (Prä-Postvergleich des MADRS-Score) zeigt in der Gesamtpopulation eine statistisch signifikante Überlegenheit zugunsten von Escitalopram gegenüber Duloxetin ( $\Delta 2.6$ ;  $p < 0.01$ ). Für

Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI)  
bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0



die schwer depressiven Patienten war dieser Unterschied stärker ausgeprägt ( $\Delta$  3.7;  $p < 0.01$ ). Auch in der Response- und Remissionsrate wurde die Überlegenheit von Escitalopram gezeigt. Nach 8-wöchiger Therapie wurde unter Escitalopram eine statistisch signifikant höhere Response- und Remissionsrate dokumentiert als unter Duloxetin (Response: 67.1% vs. 53.2%,  $p < 0.001$ ; Remission: 54.3% vs. 44.4%,  $p < 0.05$ ).

### **Escitalopram vs. Venlafaxin XR – Pool-Analysen**

Montgomery und Andersen (2006) haben die Daten der zwei randomisierten klinischen Vergleichsstudien zwischen Escitalopram und Venlafaxin XR reanalysiert:

- Bielski et al. (2004),
- Montgomery et al. (2004).

Die RCTs wurden im Vorbericht des IQWiG für die Nutzenbewertung verwendet. Ziel der Pool-Analyse war, die Wirksamkeit und Verträglichkeit von Escitalopram und Venlafaxin XR hinsichtlich der Lebensqualität und der Nutzen-Risiko-Abwägung zu vergleichen. Außerdem wurden die depressionsfreien Tage unter Escitalopram und Venlafaxin XR kalkuliert. Die Analyse zeigt, dass der therapeutische Unterschied mit dem zu Therapiebeginn dokumentierten Schweregrad der MDD korreliert ist. So wurde in der Gruppe der schwer depressiven Patienten (MADRS  $\geq 30$  Punkte) eine statistisch signifikante stärkere antidepressive Wirksamkeit in Bezug auf das primäre Zielkriterium für Escitalopram gezeigt ( $p < 0.05$ ). In dieser Patientengruppe bewirkte die Behandlung mit Escitalopram (20mg/d) auch eine statistisch signifikant höhere Remissionsrate als unter Venlafaxin XR (47% vs. 29%;  $p < 0.05$ ). Die gepoolten Daten bestätigen den therapeutischen Unterschied zwischen den Wirkstoffen. Das Maß des therapeutischen Unterschiedes steigt mit zunehmendem Schweregrad der Depression. Dieser therapeutische Unterschied ist bei einem MADRS-Punktwert von über 28 Punkten statistisch signifikant und zeigt die klinische Überlegenheit von Escitalopram gegenüber Venlafaxin XR. ( $\Delta$  3.34, 95-CI: 0.24-6.44). Für die schwer depressiven Patienten war dieser Unterschied stärker ausgeprägt ( $\Delta$  3.6, 95-CI: 0.05-7.85).

### **Escitalopram vs. Citalopram – Pool-Analysen**

In der Pool-Analyse von Gorman et al. (2002) wurden drei Einzelstudien (Burke et al., 2002; Wade et al., 2002; Montgomery et al., 2001) eingeschlossen. Die mittlere Reduktion der MADRS-Gesamtpunktzahl unter Escitalopram zeigt bereits in der Woche 1 und der Woche 8 eine statistisch signifikante Überlegenheit gegenüber Citalopram. Dieser statistisch signifikante Unterschied wird als klinisch relevant beurteilt (Hirschfeld und Vomik, 2004).

In der Pool-Analyse von Lepola et al. (2004) wurden Daten der Studien von Burke et al. (2002) und Lepola et al. (2003) ausgewertet. Somit werden äquivalente Dosen (20 bzw. 10-20 mg Escitalopram/d vs. 40 bzw. 20-40 mg Citalopram/d) miteinander verglichen. Die Abnahme im MADRS-Score nach 8 Wochen war in der Escitalopram-Gruppe statistisch signifikant stärker ausgeprägt als unter Citalopram ( $p = 0,026$ ). Die Pool-

Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI)  
bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0

Analyse von Lepola et al. (2004) zeigt auch eine statistisch signifikante Überlegenheit von Escitalopram über Citalopram bezüglich der Responseraten (63,7% vs. 52,6%,  $p=0.021$ ) und Remissionsraten (52,1% vs. 42,8%,  $p=0.036$ ).

Die Pool-Analysen zeigen in den IQWiG-relevanten Parametern Response, Remission und Verlauf der Depressivität eine statistisch signifikante und klinisch relevante Überlegenheit von Escitalopram gegenüber Citalopram. Escitalopram repräsentiert damit einen Wirkstoff, der effektiver ist als die Leitsubstanz Citalopram.

### **A.1.2 Berücksichtigung der unterschiedlichen Dosierungen in der Nutzenbewertung**

Die Dosierungen in den Vergleichsstudien zwischen Duloxetin und Escitalopram werden in dem vorläufigen Vorbericht des IQWiG thematisiert.

- Nierenberg et al. (2007): In dieser RCT wurden depressive Patienten mit Duloxetin (60mg/Tag), Escitalopram (10mg/Tag) bzw. Placebo über einen Zeitraum von 8 Wochen behandelt. Eine Titration der Dosierung wurde nicht durchgeführt. Die Dosierung von Duloxetin entspricht den Empfehlungen des Herstellers (s. Fachinformation Cymbalta; Boehringer Ingelheim, Lilly). Hier wird auch ausgeführt, dass es keinen klinischen Beleg gibt, dass „*Patienten, die nicht auf die empfohlene Anfangsdosis [60mg/Tag] ansprechen, von einer Dosiserhöhung profitieren können*“. Im Gegensatz hierzu wird die Dosierung von Escitalopram mit 10-20mg/Tag beschrieben mit „*Die übliche Dosis beträgt 10mg einmal täglich. Je nach individuellen Ansprechen des Patienten kann die Dosis auf maximal 20mg täglich erhöht werden.*“ (s. Fachinformation Cipralex; Lundbeck). In dieser klinischen Studie werden daher die Höchstdosis (60mg/d Duloxetin) mit der niedrigsten Tagesdosierung (10mg/d Escitalopram) verglichen. Auf dieser Grundlage ist ein Vergleich der beiden Wirkstoffe nicht möglich. Wegen dieser und der unter A1.1 dargelegten Gründe ist die Studie von Nierenberg et al. (2007) aus der Meta-Analyse auszuschließen.
- Khan et al. (2007) und Wade et al. (2007): Das IQWiG führt aus:
  - „*Des Weiteren wurde in 2 Studien eine Dosierung zuungunsten der Zielsubstanz Duloxetin gewählt. Bei Khan 2007 und Wade 2007 wurde Duloxetin jeweils 50% der Maximaldosis dosiert, während der Komparator Escitalopram zwischen 50 und 100% bzw. der Maximaldosis gegeben wurde.*“ (S. 53)
  - „*Zweimal wurde Escitalopram relativ höher dosiert, in diesen Studien war die Dosierung zuungunsten von Duloxetin [Khan et al.2007, Wade et al., 2007].*“

In den Arbeiten von Khan et al. (2007) und Wade et al. (2007) wurde die antidepressive Wirksamkeit von Escitalopram (10-20mg/Tag, 10mg/Tag in den ersten zwei bzw. vier Wochen) mit jener von Duloxetin (60mg/Tag) bei moderater bis schwerer Depression verglichen. Die Dosierung von Duloxetin entsprach den FDA-Vorgaben und entsprach den Empfehlungen in den Gebrauchsinformationen der

Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0

teilnehmenden Länder. Die empfohlene Anfangsdosis von Escitalopram ist 10mg/Tag. Diese Dosis wurde in klinischen Plazebo-kontrollierten Studien als effektiv und verträglich beurteilt (Burke et al., 2002; Wade et al., 2002). Die therapeutische Spanne der Escitalopram-Dosis variiert zwischen 10-20mg/Tag (Lepola et al., 2003). Für Duloxetin wird eine Start- und Erhaltungsdosis von 60mg/Tag beschrieben (Mallinckrodt et al., 2006). Die Tagesdosis von 60mg/Tag Duloxetin wurde in aktuellen Analysen als optimale Dosis ermittelt (Pritchetti et al., 2007). In den Vergleichsstudien von Khan et al. (2007) und Wade et al. (2007) wurden daher adäquate Dosierungen von Escitalopram und Duloxetin verwendet.

Auf dieser Grundlage ist ein Vergleich der beiden Wirkstoffe möglich. Bei einem meta-analytischen Vergleich zwischen Escitalopram und Duloxetin sind daher die Publikationen von Khan et al. (2007) und Wade et al. (2007) zu berücksichtigen und jene von Nierenberg et al. (2007) auszuschließen.

Der Einfluss der Dosierung sollte ebenfalls berücksichtigt werden, wenn das Kriterium „Studienabbruch wegen Unerwünschter Ereignisse“ in den Vergleichsstudien zwischen Venlafaxin XR und Escitalopram analysiert wird.

### **A.1.3 Berücksichtigung des Schweregrades der Depression in der Nutzenbewertung**

Der Schweregrad der Depression ist bei der Nutzenbewertung unzureichend gewürdigt worden. Die Berücksichtigung Patientengruppe mit schwerer Depression ist umso bedeutsamer, als dass Patienten mit schwerer Depression die Zielkriterien Response und Remission seltener erreichen als Patienten mit leichter bis moderater Depression (Thase et al., 2000). Auch das Ansprechen auf eine Plazebogabe ist bei leichter Depression stärker ausgeprägt als bei schweren Formen der Depression (Montgomery et al., 1999). Die Fokussierung auf Behandlungseffekte in der schweren Depression reduziert die Bedeutung der spontanen Remission und des Plazeboeffektes und reflektiert somit den wahren therapeutischen Wert eines Wirkstoffes.

Auswertungen zu Therapieeffekten bei schwer depressiven Patienten resultieren häufig aus Post-hoc Analysen einzelner RCTs oder Vergleichsstudien mit geringen Fallzahlen. Schweregrad-stratifizierte Pool-Analysen wurden für verschiedene Vergleichsstudien mit Escitalopram gegen aktive Komparatoren durchgeführt:

- Escitalopram *versus* Duloxetin (Lam et al., 2008)
- Escitalopram *versus* Venlafaxin XR (Montgomery and Andersen, 2006)

Die relevanten Ergebnisse dieser Pool-Analysen werden im Abschnitt Diskussion bewertet.

Die Analysen belegen, dass Escitalopram den Vergleichssubstanzen Duloxetin, Venlafaxin XR und Citalopram in den IQWiG-relevanten Parametern (Response,

Remission und Verlauf der Depressivität) statistisch signifikant und klinisch relevant überlegen ist.

#### **A.1.4 Wirkeintritt als relevantes Zielkriterium der antidepressiven Therapie**

Das Zielkriterium Wirkeintritt gewinnt aktuell stärker an Bedeutung, wird in der Nutzenbewertung des IQWiG jedoch nicht entsprechend gewürdigt. Die DGPPN führt demgegenüber in der Behandlungsleitlinie für affektive Erkrankungen den Wirkeintritt als relevantes Behandlungsziel auf (DGPPN, 2000). Der Wirkeintritt kann als eine Evaluierungsgröße charakterisiert und entsprechend valide analysiert werden (Volz et al., 2005; Leon et al., 2001). Ein frühes Ansprechen auf eine antidepressive Therapie ist mit dem Verlauf der Depression positiv korreliert (Mulder et al., 2006). Antidepressiva zeigen häufig eine zeitliche Verzögerung im Wirkeintritt, wobei eine klinisch relevante Verbesserung nach wenigstens 2 Wochen beobachtet wird (Blier et al., 2001). In dieser Latenzzeit sind die depressiven Patienten symptomatisch und funktionell eingeschränkt:

- erhöhte Morbidität,
- beeinträchtigte Lebensqualität,
- erhöhtes Suizidrisiko und
- eingeschränkte Arbeitsfähigkeit.

Der Eintritt der antidepressiven Wirksamkeit wird als relevant für das klinische Management der Depression bewertet. Wird ein verzögerter Wirkeintritt beobachtet, führt dies zu wesentlichen Einschränkungen im weiteren Therapieverlauf, wie zum Beispiel:

- Prolongation der depressiven Symptomatik,
- Zunahme des suizidalen Risikos und
- Verlust an Compliance.

Aus diesen Gründen war die Entwicklung neuer innovativer Antidepressiva in den letzten Dekaden fokussiert auf Wirkstoffe mit einem schnellen Eintritt der Wirksamkeit. Daher ist der antidepressive Wirkeintritt in die Nutzenbewertung einzuschließen.

Der Vergleich zwischen Escitalopram und den SNRIs sowie Citalopram hinsichtlich des Wirkeintrittes zeigt:

- Unter Escitalopram wird nach 1-wöchiger Medikation eine statistisch signifikante therapeutische Separation im Verlauf des MADRS-Gesamtscores gegenüber Placebo registriert; unter Citalopram erst nach 6 Wochen (Gorman et al., 2002). Entsprechend wird im CGI-Improvement nach 1 Woche unter der Behandlung mit Escitalopram bzw. nach 4-wöchiger Therapie mit Citalopram ein statistisch signifikanter Unterschied registriert. Die klinische Relevanz des früheren Wirksamkeitseintritts von Escitalopram gegenüber Citalopram wird mittels der MADRS-Remission und MADRS-Response prägnant. In diversen klinischen Studien wird eine frühere Remission unter Escitalopram festgestellt, die zeitliche

Auftrag A05-20A „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI)  
bei Patienten mit Depressionen“  
Stellungnahme der Lundbeck GmbH zum Vorbericht, Version 1.0

- Differenz beträgt 8 Tage (Lepola et al., 2003). Auch die Response wird unter Escitalopram früher beobachtet als unter Citalopram (Llorca et al., 2005).
- Die anhaltende („sustained“) MADRS-Response und MADRS-Remission tritt unter Escitalopram früher und häufiger auf als unter Venlafaxin-XR (Bielski et al., 2004; Murdoch and Keam, 2005).
  - Escitalopram ist Duloxetin im Wirkeintritt überlegen (Wade et al., 2007).

Der etwaige Einfluss des Wirkeintritts auf die wesentlichen Outcome-Parameter einer Akuttherapie wurde für alle publizierten plazebokontrollierten klinischen Studien mit Escitalopram bestätigt (Wade and Andersen, 2006).

Zusammenfassend kann festgestellt werden, dass der Wirkstoff Escitalopram im Vergleich zu anderen Antidepressiva, über einen schnellen Wirkeintritt verfügt. Für Escitalopram wurden im Vergleich zu den aktiven Komparatoren folgende Daten zum Wirkeintritt registriert:

Referenz	Wirkeintritt
<b>Escitalopram vs. Venlafaxin XR</b>	
Montgomery et al., 2004	Die Zeitspanne bis zur Response/Remission ist unter Escitalopram kürzer als unter Venlafaxin XR: <ul style="list-style-type: none"> <li>• MADRS-Response: 4,6 Tage</li> <li>• MADRS-Remission: 6,6 Tage</li> </ul>
<b>Escitalopram vs. Duloxetin</b>	
Wade et al., 2007	Nach 2- wöchiger Therapie wurde der Wirkeintritt unter Escitalopram bei 68.8% der Patienten registriert; gegenüber 54.8% unter Duloxetin.
<b>Escitalopram vs. Citalopram</b>	
Colonna et al., 2003	Unter Escitalopram tritt die Remission (MADRS<12) 13 Tage früher ein als unter Citalopram.
Lepola et al., 2003	Unter Escitalopram tritt die Remission und die Response 8 Tage früher ein als unter Citalopram.
Llorca et al., 2005	Unter Escitalopram tritt die Response 7 Tage früher ein als unter Citalopram.

Der Wirkeintritt von Escitalopram wurde mit verschiedenen Methoden bestimmt. Der Wirkeintritt zeigt in allen Methoden eine gleichsinnige und konsistente Wirklatenz. Die Wirklatenz von Escitalopram umfasst eine kürzere Zeitspanne als bei den entsprechenden Vergleichssubstanzen (Duloxetin, Venlafaxin XR, Citalopram).

## A.2. Escitalopram vs. Citalopram

### A.2.1 Wirksamkeit – Kurzzeit-Akutstudien

Präklinische und klinische Studien belegen konsistent, dass bei äquipotenter Dosierung des Razemates (Citalopram) und des Enantiomers (Escitalopram), Escitalopram eine stärkere pharmakologische Wirkung und klinische Wirksamkeit aufweist (Übersicht: Sanchez et al., 2004; Auquier et al., 2003).

#### Response

Die klinische Relevanz des Ergebnisses kann auch auf der Basis der NNT verdeutlicht werden. In der Arbeit von Moore et al. (2005) wurde für die Response eine NNT=7 dokumentiert; in der Arbeit von Yevtushenko et al. (2007) resultierten entsprechende NNT-Werte von 9.

Die Vergleichsstudien zwischen Escitalopram und Citalopram zeigen somit im meta-analytischen Ansatz eine statistisch signifikante Überlegenheit von Escitalopram gegenüber Citalopram im Zielkriterium Response. Die numerische Überlegenheit von Escitalopram gegenüber Citalopram wird in 5 der 6 RCTs demonstriert, die in die Meta-Analyse eingeschlossen wurden. Ferner zeigt der Endpunkt Response in 3 voneinander unabhängig durchgeführten RCTs eine statistisch signifikante Überlegenheit von Escitalopram gegenüber Citalopram (Lepola et al., 2003; Moore et al., 2005; Yevtushenko et al., 2007). Die Ergebnisse zur Response sind somit sowohl in der Meta-Analyse und den einzelnen RCTs konsistent. Die Resultate zur Response belegen somit den Nutzen und die klinische Relevanz der Überlegenheit von Escitalopram im Vergleich zu Citalopram.

#### Remission

Auch in dem Zielparameter Remission belegt die NNT die klinische Relevanz der Überlegenheit von Escitalopram gegenüber Citalopram. In der Arbeit von Moore et al. (2005) wurde für die Remission eine NNT=9 dokumentiert; in der Arbeit von Yevtushenko et al. (2007) resultierten entsprechende NNT-Werte von 2 (MADRS $\leq$ 12) bzw. NNT=3 (MADRS $\leq$ 10).

Die Vergleichsstudien zwischen Escitalopram und Citalopram zeigen somit im meta-analytischen Ansatz eine statistisch signifikante Überlegenheit von Escitalopram gegenüber Citalopram im Zielkriterium Remission. Ferner zeigt der Endpunkt „Remission“ in einer unabhängig durchgeführten RCT eine statistisch signifikante Überlegenheit von Escitalopram gegenüber Citalopram (Yevtushenko et al., 2007). Die numerische Überlegenheit von Escitalopram gegenüber Citalopram wird in den RCTs registriert, die in die Meta-Analyse eingeschlossen wurden. Die Ergebnisse zur Remission sind somit sowohl in der Meta-Analyse und den einzelnen RCTs konsistent. Die Resultate zur Remission belegen somit den Nutzen und die klinische Relevanz der Überlegenheit von Escitalopram im Vergleich zu Citalopram.

### **Depressionsskalen-Gesamtscore**

Escitalopram erweist sich auch in dem primären Zielkriterium dieser RCTs, im Vergleich zu Citalopram überlegen (MD 1.60; 95-CI 0.72-2.49; p=0,0004). Ferner zeigen 5 von 6 RCTs in diesem Endpunkt eine numerische Überlegenheit von Escitalopram und bestätigen somit den konsistenten Effekt. Auch in zwei unabhängig voneinander durchgeführten RCTs wird eine statistisch signifikante Überlegenheit von Escitalopram gegenüber Citalopram dokumentiert (Moore et al., 2005; Yevtushenko et al., 2007). Der Effekt belegt den Nutzen von Escitalopram gegenüber Citalopram in multipler Art und Weise. Der Effekt ist klinisch relevant.

**A 1.6     Verband Forschender Arzneimittelhersteller e. V. (VFA)**

**Autoren:**

Wahler, Steffen, Dr. med.

Rössel, Antje

Thole, Henning

**Adresse:**

Antje Rössel

Verband Forschender Arzneimittelhersteller e. V. (VFA)

Hausvogteiplatz 13

10117 Berlin



**Stellungnahme des VfA zum Vorbericht „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen (A05-20A)“**

Am 1. Dezember 2008 wurde der Vorbericht „Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen“ vom Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) veröffentlicht und zur Stellungnahme gestellt.

Wir begrüßen die Möglichkeit, mit einer Stellungnahme von Seiten des Verbandes Forschender Arzneimittelhersteller e.V. den Vorbericht kommentieren zu können, sehen jedoch notwendige Ergänzungen am Vorbericht.

**Grundsätzliche Bemerkungen**

Wir teilen die Auffassung des IQWiG, dass Depressionen eine hohe Bedeutung zukommt. Der Vorbericht weist explizit daraufhin, dass Depressionen und ihre Begleitsymptome mit einer massiven Beeinträchtigung des Lebens im Hinblick auf zwischenmenschliche Beziehungen und einer Einschränkung der Arbeitsfähigkeit einhergehen. Auch ist die Relevanz von Depressionen im Hinblick auf die Minderung von Lebensqualität der Betroffenen hinreichend bekannt<sup>1</sup>. Die Anerkennung dieser Volkskrankheit<sup>2</sup> in Verbindung mit den erheblichen Langzeit- und sozialmedizinischen Konsequenzen für Betroffene, deren Angehörige und damit implizit auch der Volkswirtschaft befürworten wir.

Neben den patientenindividuellen Zielen, werden zur Behandlung einer akuten depressiven Episode die Beendigung der depressiven Episode, die anschließende Erhaltungstherapie zur Remissionsstabilisierung und die Rezidivprophylaxe als die Hauptziele bezeichnet, die derzeit mit der akuten antidepressiven Pharmakotherapie erreicht werden können<sup>3</sup>. Die Beurteilung der Wirksamkeit pharmakotherapeutischer Maßnahmen ist vielfach nachgewiesen. Der individuelle Charakter bildet hier die Schlüsselfunktion, da auf kein

---

<sup>1</sup> Gesundheitsberichterstattung des Bundes: Gesundheit in Deutschland. Berlin: 07/2006, S.29.

<sup>2</sup> Rudolf, Bermejo, Schweiger, Hohagen, Härter: Zertifizierte medizinische Fortbildung: Diagnostik depressiver Störungen, In: Dtsch Arztebl 2006; 103(25): A 1754-62.

<sup>3</sup> Arzneimittelkommission der deutschen Ärzteschaft: Therapieempfehlungen der Arzneimittelkommission der deutschen Ärzteschaft. 2. Auflage 2006. Arzneiverordnung in der Praxis, S.11.

Therapieverfahren alle Patienten ansprechen<sup>4</sup> und keine einzelne Medikation bei allen Patienten eine Remission bedingt<sup>5</sup>.

## **Methodische Kritik**

### **Verfahrenstransparenz und Beteiligung im Bewertungsprozess**

Das IQWiG hat für das gesamte Bewertungsverfahren gemäß § 35b Abs. 2 SGB V und § 139a Abs. 5 SGB V eine hohe Verfahrenstransparenz sowie eine angemessene Beteiligung für die maßgeblich zu Beteiligten zu gewährleisten. Bislang finden sich im Vorbericht keine Angaben darüber, welche Expertise für die Erstellung des Berichtsplanes herangezogen worden ist. Diese Vorgehensweise ist intransparent und entspricht damit nicht dem gesetzlichen Auftrag, der die Transparenz und Beteiligung als einen wesentlichen Aspekt des Verfahrensablaufes anerkennt.

Seite 2/5

### **Beschränkung auf RCTs**

Das IQWiG hat den gesetzlichen Auftrag gemäß § 35b Abs. 1 SGB V „[...] auf der Grundlage der in den jeweiligen Fachkreisen anerkannten internationalen Standards der evidenzbasierten Medizin und der Gesundheitsökonomie [...]“ durchzuführen. Für die Bewertung der SNRI zieht das IQWiG nur randomisierte kontrollierte Studien heran. Datenbasis jeglicher Bewertungen müssen gemäß dem Prinzip der bestverfügbaren Evidenz jedoch alle Evidenzgrade berücksichtigen. Die ausschließliche Berücksichtigung auf RCTs widerspricht dem gesetzlichen Auftrag.

## **Inhaltliche Kritik**

### **Unzulässige Interpretation des G-BA Auftrages**

Der gemeinsame Bundesausschuss (G-BA) hat am 22. Februar 2005 das IQWiG mit der „Nutzenbewertung von trizyklischen Antidepressiva, selektiven Serotonin-Wiederaufnahmehemmern und Serotonin-Noradrenalin-Wiederaufnahmehemmern bei Patienten mit Depressionen, auch im Vergleich untereinander“ beauftragt<sup>6</sup>.

<sup>4</sup> Bschor, Mazda: Therapie depressiver Erkrankungen, In: Dtsch Arztebl 2008; 105(45): 782-792.

<sup>5</sup> EMEA 2005: Scientific Discussion; <http://www.emea.europa.eu/humandocs/PDFs/EPAR/cymbalta/19256704en6.pdf> [Zugriff: 21.12.2006]

<sup>6</sup> G-BA: Aufträge des gemeinsamen Bundesausschusses an das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, Beschluss vom 18.01.2005, Beauftragung am 22.02.2005; <http://www.g-ba.de/download/39-261-383/2005-01-18-Auftraege-IQWiG.pdf> [Zugriff: 23.12.2008, 09:43]

Damit hat der G-BA die Evaluation der Wirkstoffe gefordert, die neben den Monoaminoxidasehemmern und den präsynaptischen Rezeptoragonisten die größte pharmakologische Gruppe der Antidepressiva bilden.

Das IQWiG hat diesem Auftrag in Form von Unteraufträgen Folge geleistet, die sich in „Selektive Serotonin- und Noradrenalin-Wiederaufnahme-Hemmer (SNRI) bei Patienten mit Depressionen“ (A05-20A), die „Nutzenbewertung von selektiven Serotonin-Wiederaufnahme-Hemmern (SSRI) bei Patienten mit Depressionen“ (A05-20B) und in die Bewertung neuerer Antidepressiva (A05-20C) aufteilen. Der Auftrag des G-BA fordert neben der Bewertung von SSRI und SNRI zudem die Bewertung von trizyklischen Antidepressiva, die allesamt zur Hemmung der Serotonin- und/oder Noradrenalinwiederaufnahme eingesetzt werden. Das IQWiG lässt diesen Vergleich im Berichtsplan wie auch im Vorbericht nicht zu. Die sachgerechte Umsetzung des G-BA-Auftrages zur Nutzenbewertung von SSRI, SNRI, Trizyklika bleibt damit fraglich.

Seite 3/5

### **Berücksichtigung des Nebenwirkungsprofils**

Die Bedeutung der spezifischen Analyse von Nebenwirkungen stellt einen wesentlichen Aspekt bei der Bewertung von Arzneimitteln dar<sup>7</sup>. So können Nebenwirkungen bei Depressionen in vielfältiger Häufigkeit und Latenz<sup>8</sup> auftreten. Da sich die Antidepressiva in ihrem Nebenwirkungsprofil voneinander unterscheiden, bezeichnen Nebenwirkungen einen bedeutenden Indikator bei der Therapieauswahl. Bei vergleichender Bewertung der Antidepressiva sollte aufgrund dessen spezifisch auf das Nebenwirkungsprofil fokussiert werden. Hier ist insbesondere eine differenzierte Unterscheidung zwischen den einzelnen Nebenwirkungen der Produkte wesentlich.

Eine grundsätzliche Bestimmung der Nebenwirkungen sollte daher immer in einer nachvollziehbaren Darstellung in Abwägung der jeweiligen Substanz mit dem Nebenwirkungs-, Risiko- und Interaktionsprofil erfolgen.

Zudem ist der **Begriff der „unerwünschten Ereignisse“** von Seiten des IQWiG ungeklärt. So beschreibt das IQWiG in seinen Allgemeinen Methoden<sup>9</sup>, **dass bei „unerwünschten Wirkungen“ ein kausaler Zusammenhang mit der Intervention angenommen wird.**

---

<sup>7</sup> Loke, Price, Herxheimer 2008: Adverse effects, In: Higgins, Green: Cochrane Handbook for Systematic Reviews of Interventions, p.433-448.

<sup>8</sup> Arzneimittelkommission der deutschen Ärzteschaft: Therapieempfehlungen der Arzneimittelkommission der deutschen Ärzteschaft. 2. Auflage 2006. Arzneiverordnung in der Praxis, S. 21.

<sup>9</sup> IQWiG 2008: Allgemeine Methoden 3.0, 27.05.2008.

Die Klärung der „unerwünschten Ereignisse“ lässt das IQWiG dementsgegen offen.

### **Interpretation der klinischen Relevanz**

Zur Ermittlung der Minimal Important Difference (MID) können beispielsweise verteilungs-, panelbasierte- oder der ankerbasierte Methoden eingesetzt werden<sup>10</sup>. Das IQWiG nähert sich in seinem Vorbericht der klinischen Relevanz statistisch an und berechnet **relevante Gruppenunterschiede mit Hilfe der Cohen's d Einstufung**<sup>11</sup>, da für keine der eingesetzten Skalen verwendbare MID's vorlagen.

**Seite 4/5**

Die Bestimmung der MID ist subjektiv und daher mit Unsicherheiten behaftet, die bei der Entscheidungsfindung beachtet werden müssen. Die ausschließliche Betrachtung statistisch relevanter Effektgrößen ab 0,2 Standardabweichungen **gemäß der Cohen's d** Klassifizierung kann daher zu irreführenden Schlussfolgerungen führen. Dies legen beispielsweise Patrick et al. 2008 in ihren Ausführungen zur MID im Hinblick auf Patient-reported outcomes deutlich dar: *„If the MID is 0,5 and the mean difference between treatments is 0,4, clinicians may infer that nobody benefits from the intervention. If the mean difference is 0,6, they may conclude that everyone benefits. Both inferences may be misguided“*<sup>12</sup>. Da klinisch relevante Unterschiede sowohl die ärztliche als auch die patientenseitige Perspektive umfassen, sollte die MID im Rahmen eines Scoping-Prozesses bereits unter Beteiligung von Patienten und Klinikern gemeinsam **zum Zeitpunkt der „Konkretisierung der Forschungsfrage und Definition der Zielkriterien“** bestimmt werden. Dadurch kann gewährleistet werden, dass sowohl die klinisch **messbaren, als auch die subjektiv „fühlbaren“** Parameter in einer adäquaten Bewertung münden. Eine ausschließliche statistische Berechnung der klinischen Relevanz ist daher aus Sicht des VFA nicht ausreichend.

### **Fazit**

Der Vorbericht lässt auch nach Empfehlungen durch die Stellungnahme zum Berichtsplan 2.0 methodische und inhaltliche Fragen offen. Die Beschränkung auf RCTs unterminiert den Nutzen von SNRI und widerspricht den gesetzlichen Anforderungen. Die Ausei-

---

<sup>10</sup> Jaeschke, Singer, Guyatt 1989: Measurement of Health Status. Ascertaining the Minimal Clinically Important Difference, In: Controlled Clinical Trials 10, p.407-415.

<sup>11</sup> Cohen 1998: Statistical Power Analysis in the Behavioral Sciences (2nd edition). Hillsdale: Lawrence Erlbaum Associates.

<sup>12</sup> Patrick, Guyatt, Acquadro 2008: Patient-reported outcomes, In: Higgins, Green: Cochrane Handbook for Systematic Reviews of Interventions, p.531-545



Verband Forschender  
Arzneimittelhersteller e.V.

andersetzung mit dem Schaden und der Interpretation klinisch  
relevanter Unterschiede bedingt eine umfangreiche Überarbeitung.

**Seite 5/5**

## **A 1.7 Wyeth Pharma GmbH**

### **Autoren:**

Volmer, Timm, Dr., MPH

Löschmann, Peter-Andreas, PD Dr. med.

Schmitt, Andreas, PD Dr. med.

### **Adresse:**

PD Dr. Andreas Schmitt

Wyeth Pharma GmbH

Wienburgstr. 207

48159 Münster

# **IQWiG Vorbericht: Stellungnahme der Wyeth Pharma GmbH**

## **Zusammenfassung**

Am 1.12.2008 ist der Vorbericht zur Nutzenbewertung der SNRI bei Depression veröffentlicht worden.

Im Vorbericht wird der Nutzen von Venlafaxin in der Behandlung der Akut- und Langzeitbehandlung der Depression bestätigt. Dies ist zweifelsfrei in vielen Studien zur Akutbehandlung sowie in Langzeitstudien zur Rezidivprophylaxe nachgewiesen worden.<sup>1-4</sup> Des Weiteren werden im Vorbericht ein Zusatznutzen in der Akuttherapie gegenüber SSRIs sowie eine bessere Verträglichkeit gegenüber TZAs und dem anderen SNRI Duloxetin festgestellt. Grundsätzlich stimmen wir dieser Einschätzung zu, wobei einige Einzelergebnisse aufgrund der verwendeten Datenbasis nur einen Teil des Nutzens von Venlafaxin widerspiegeln.

Außerhalb Deutschlands finden ähnliche Bewertungen statt, die durchaus zu abweichenden wissenschaftlichen Beurteilungen gelangen können. Die vorläufige IQWiG-Nutzenbewertung ist also nur *ein* wichtiger Beitrag für eine ganzheitliche Bewertung einer Wirkstoffklasse und der dazugehörigen Produkte. Die Bewertung des IQWiG stellt aus Sicht der Wyeth GmbH nur eine Teilbetrachtung dar, da wesentliche Aspekte bei der Nutzenbewertung unberücksichtigt blieben. Die Fokussierung auf die einzelne therapeutische Klasse der SNRI zeigt nur den Ausschnitt eines Gesamtbildes der Pharmakotherapie der Depression.

Eine Depression ist in den meisten Fällen eine chronische Krankheit mit häufigen Wiedererkrankungen<sup>5</sup>, so dass sich die Nutzenbewertung der Antidepressiva durch das IQWiG nicht allein auf die Frage beschränken darf, welches Therapeutikum in der Initialbehandlung besser wirken könnte – vielmehr ist die klinische Praxis häufig von sequenziellen Entscheidungen geprägt: D.h. bei Versagen einer Initialtherapie muss der Arzt individuell entscheiden, wie er die Therapie weiterführt. Eine ganzheitliche Betrachtung bei der Nutzenbewertung spiegelt auch das Interesse des G-BA wider, dessen Ziel es ist, ein möglichst umfassendes Bild zu erhalten und eine qualitativ hochwertige Patientenversorgung sicherzustellen.

Aus Sicht der Wyeth GmbH weist auch die Einzelbetrachtung von Venlafaxin zahlreiche Limitationen auf, insbesondere weil nur ein Teil der verfügbaren Daten vom IQWiG bei der vorläufigen Bewertung berücksichtigt wurde. Einigen Schlussfolgerungen können wir uns daher nicht anschließen.

### **Studienselektion**

Grundsätzlich ist zu betonen, dass die Beschränkung auf doppelblinde randomisierte Studien weder den Prinzipien der evidenzbasierten Medizin noch dem gesetzlichen Auftrag § 35b SGB V, in dem als Grundlage für eine Bewertung die bestmögliche Evidenz gefordert wird, entspricht.

Des Weiteren ist für eine überwiegend meta-analytische Betrachtungsweise von RCTs, wie sie das IQWiG vornimmt, die Einbeziehung aller verfügbaren RCTs notwendig. Aufgrund der besonderen methodischen Kriterien des IQWiG sind nicht alle verfügbaren Studien in diese Nutzenbewertung eingeschlossen worden. Diese Kriterien, die zum Ausschluss wichtiger Studien führte, entsprechen keinen etablierten Standards. Durch dieses Analyseverfahren wird gegenüber anderen umfangreicheren Metaanalysen die Typ-2-Fehlerrate erhöht, was zu weniger präzisen Abschätzungen des klinischen Effektes und zu breiteren Konfidenzintervallen führen kann. Eine Einengung der für die Analyse verwendeten Studien kann demnach zu einem verzerrten Bild der tatsächlichen Evidenz führen.

Wyeth hat dem IQWiG sämtliche verfügbaren Studien, einschließlich der nicht publizierten, zur Verfügung gestellt. Für ein transparentes Vorgehen ist der Einbezug sämtlicher verfügbaren Studien zu fordern. Der Nachweis von zum Teil geringen Effektunterschieden ist bei Ausschluss relevanter verfügbarer Studien somit kein überraschendes Ergebnis. Wir möchten in diesem Zusammenhang insbesondere dem Vorwurf des IQWiG widersprechen, nicht alle Studiendaten zur Verfügung gestellt zu haben. Für die explizit erwähnten älteren Studien (Studien 606, 414, 347) standen keine vollständigen Studienberichte zur Verfügung, die den Anforderungen des IQWiG entsprochen hätten. Dies wurde dem IQWiG entsprechend mitgeteilt.

Für den IQWiG-Vorbericht sind einige Studien zu Venlafaxin für die Bewertung ausgeschlossen worden. Die Gründe hierfür sind zum Teil nicht transparent dargelegt und aus wissenschaftlicher Sicht nicht nachvollziehbar ist.



1. De Nayer<sup>6</sup> et al. 2002: Ist diese Studie nicht berücksichtigt worden, weil laut Einschlusskriterien Patienten mit Depression und begleitender Angst behandelt wurden? Laut Berichtsplan sind Studien mit Patienten mit psychiatrischer Komorbidität jedoch ausdrücklich erlaubt.
2. Li<sup>7</sup> et al. 2004: Bei dieser Studie handelt es sich explizit um eine Studie zum Effekt von Paroxetin und Venlafaxin in der Behandlung der Depression. Der Ausschlussgrund ist nicht erkennbar.
3. Ballus<sup>8</sup> et al. 2000: Ist diese Studie nicht berücksichtigt worden, weil dysthyme Patienten eingeschlossen wurden? Nach den IQWiG Ausschlusskriterien darf der Anteil dysthymen Patienten nicht >20% sein. Der mittlere HAM-D zu Studienbeginn betrug 24 und entspricht den Patientenpopulationen der eingeschlossenen Studien, so dass unseres Erachtens diese Studie einbezogen werden sollte. Ein weiteres Argument ist, dass der Anteil der leichtgradig erkrankten Patienten (nach CGI) bei 15% lag. Es ist aus klinischer Sicht davon auszugehen, dass ausschließlich unter den leichtgradig erkrankten Patienten dysthyme Patienten mit eingeschlossen wurden. Zudem ist ohnehin zu erwarten, dass der Effekt von Venlafaxin durch den Einschluss von dysthymen Patienten eher abgeschwächt wird.
4. Studie 0600-626-NL: Auch hier wurden wie in der Studie von Ballus dysthyme Patienten eingeschlossen. Hierbei betrug der mittlere HAM-D bei Studienbeginn 22 und entsprach somit ebenfalls den Patientenkollektiven der für den IQWiG Vorbericht eingeschlossenen Studien. Ein weiteres Argument ist, dass der Anteil der leichtgradig erkrankten Patienten (nach CGI) bei 6% lag, so dass auch hier der Anteil der dysthymen Patienten deutlich unter 20% lag.
5. Oslin<sup>9</sup> 2003: Ist diese Studie aufgrund des Einschlusses dysthymen Patienten ausgeschlossen worden? Laut Publikation wiesen insgesamt 80,8% aller Patienten (Venlafaxinarm: 85%; Sertralinarms 76%) die Diagnose MDD auf. Der Ausschluss dieser Studie ist daher nicht nachvollziehbar.
6. Studie 0600A-632-NO-SW: In dieser Studie wurden Patienten mit MDD untersucht, so dass der Ausschluss nicht nachvollziehbar ist.
7. Zanardi<sup>10</sup> et al. 2000: Bei dieser Studie führte vermutlich der Anteil der Patienten mit Bipolarer Erkrankung zum Ausschluss für die Bewertung. Nach IQWiG Kriterien ist ein Anteil bipolarer Patienten bis 15% erlaubt. In dieser kleinen Studie, die insgesamt 28 Patienten einschloss, waren somit 2 Patienten ausschlaggebend für den Ausschluss (6/28 statt 4/28)

8. Poirier<sup>11</sup> et al. 1999: Diese Studie wurde offensichtlich aufgrund ihrer Dauer von 4 Wochen ausgeschlossen. Die Rationale dieses Ausschlusskriteriums ist laut IQWiG Methodenpapier in der Zulassungsaufgabe der EMEA begründet, die eine Studiendauer von 6 Wochen empfehlen. Diese Empfehlung begründet unserer Ansicht nach nicht den Ausschluss für eine Bewertung patientenrelevanter Zielkriterien (Response, Remission), die in dieser Studie untersucht wurden. Zudem kann die kurze Studiendauer ohnehin nur zu einer Verringerung des antidepressiven Effektes beitragen.

### **Methodische Bewertung des Nutzens und der klinischen Relevanz**

Das IQWiG nimmt eine Nutzen- Schadenbewertung vornehmlich auf dem Boden von Meta-Analysen vor. Meta-Analysen sind zweifelsfrei eine wichtige und international anerkannte Methode, um anhand mehrerer Studien primäre oder auch sekundäre Zielparameter zu analysieren. Aus wissenschaftlicher Sicht besteht ebenfalls kein Zweifel darüber, dass Meta-Analysen Limitationen aufweisen.<sup>12</sup> Apodiktische Aussagen sind auf dem Boden von Meta-Analysen kaum möglich, so dass deren Ergebnisse eher als Unterstützung einer Aussage verstanden werden. Wie bereits oben beschrieben, ist für die Verwendung von Meta-Analysen v.a. die Berücksichtigung aller relevanten Studien notwendig. Des Weiteren ist die Aussagekraft einer Meta-Analyse bei einer sehr geringen Anzahl von Studien deutlich limitiert. Für die IQWiG Nutzenbewertung einzelner patientenrelevanter Zielparameter sind vielfach nur wenige Studiendaten verfügbar. Auf dem Boden nur weniger Studien sind die apodiktischen Aussagen beispielsweise zur Lebensqualität, sozialem Funktionsniveau, Energie und insbesondere hinsichtlich der Wirksamkeit bei älteren Patienten daher aus wissenschaftlicher Sicht anzuzweifeln.

Des Weiteren sind wir der Ansicht, dass die biometrische Betrachtung bzw. Interpretation anhand des *Cohen`s d* kein geeignetes Maß ist, um die klinische Relevanz von Befunden einzuschätzen. Für die Einschätzung der klinischen Relevanz von Studienergebnissen ist beispielsweise die Zahl der Patienten, die behandelt werden müssen, um bei einem den klinischen Effekt zu erzielen (Number needed to treat - NNT), eine etablierte statistische Kenngröße.<sup>13</sup> Im *Cochrane Handbook for systematic reviews* wird die Bedeutung von absoluten Messgrößen wie ARR oder NNT für Meta-Analysen betont.<sup>14</sup>

Für viele der vom IQWiG analysierten patientenrelevanten Zielparameter gibt es keine allgemein anerkannten Kenngrößen oder „Cut-off-Werte“.

Hier wäre ein entsprechender Scoping-Prozeß unter Einbeziehung von Experten und Patienten sinnvoll. Im Rahmen dieses Scoping-Prozesses sollten bei Verwendung von NNTs u.a. folgende Aspekte berücksichtigt werden: Für die Einschätzung der klinischen Relevanz von NNTs muss insbesondere der Schweregrad der Erkrankung berücksichtigt werden, weshalb in anderen Indikationen wie Sekundärprophylaxe von Schlaganfällen auch NNTs >10 durchaus akzeptiert werden.<sup>15</sup>

Zudem muss berücksichtigt werden, ob der Vergleich gegenüber Plazebo oder einer anderen wirksamen Substanz erfolgt. Bei einem Vergleich zwischen zwei wirksamen Substanzen sind zwangsweise höhere NNTs zu erwarten.

Somit können z.B. NNTs  $\leq 20$  im Vergleich zwischen 2 Antidepressiva als klinisch relevant angesehen werden.<sup>16, 17</sup>

### **Vergleich Venlafaxin und SSRIs**

#### **a) Remission**

In der IQWiG Nutzenbewertung wird im Einklang mit nationalen und internationalen Leitlinien eine Remission als wichtiger patientenrelevanter Parameter beschrieben und u.a. für den Vergleich von Venlafaxin gegenüber SSRIs herangezogen. Dies entspricht nationalen und internationalen Leitlinien, in denen das Erzielen der Remission als wichtigstes therapeutisches Ziel einer antidepressiven Behandlung festgelegt wurde.<sup>18</sup>

Im IQWiG Vorbericht wird kein Zusatznutzen von Venlafaxin gegenüber SSRIs festgestellt. Da bei näherer Betrachtung die Signifikanz nur knapp verfehlt wurde, vermissen wir eine kritische Überprüfung dieser apodiktischen Aussage. Bei einer meta-analytischen Betrachtung von Remissionsraten sind Limitationen zu bedenken, die zu einer Unterschätzung der Therapieeffekte von Venlafaxin beigetragen haben können. Ein wichtiger Aspekt ist, dass eine Remission in praktisch allen klinischen Studien nur ein sekundärer Wirksamkeitsparameter war. Des Weiteren zielten die meisten Einzelstudien nicht auf eine Überlegenheit gegenüber dem Vergleichsmedikament ab, weshalb v.a. durch die häufig sehr kleinen Patientenzahlen (teilweise < 100) selten eine Überlegenheit von Venlafaxin gegenüber SSRI gezeigt wurde.<sup>8, 11, 19</sup> In vielen Einzelstudien wurden Venlafaxin-Dosen bis 150 mg/Tag verwendet, d.h. der mögliche Venlafaxin-Dosisbereich (bis max. 375 mg/Tag)

wurde nicht vollständig ausgeschöpft.<sup>20-23</sup> Diese Aspekte wurden zwar teilweise im Vorbericht erwähnt, unserer Ansicht nach jedoch nicht ausreichend für die Nutzenbewertung berücksichtigt. Des Weiteren bleiben die Unterschiede hinsichtlich der Wirksamkeit zwischen den SSRIs ebenfalls unberücksichtigt.

Im Gegensatz zu dem Ergebnis des IQWiG Vorbericht wurden in mehreren publizierten Metaanalysen konsistent signifikant höhere Remissionsraten unter Venlafaxin im Vergleich zu SSRIs als Substanzklasse berichtet.<sup>24, 25</sup> Dies gilt auch für zwei aktuelle Metaanalysen, die praktisch alle weltweit verfügbaren Studien einschlossen und unabhängig voneinander die bisherigen Vorteile zugunsten von Venlafaxin bestätigten.<sup>16, 17, 24</sup> Meta-Analysen, die nicht alle verfügbare Evidenz berücksichtigen, zeigen zwangsweise geringere Effektunterschiede.<sup>26</sup> Ein knappes Verfehlen der Signifikanz wie im IQWiG Vorbericht spricht gerade bei Meta-Analysen für eine knapp verfehlte Power, so der Einschluss allein einiger der oben beschriebenen ausgeschlossenen Studien zu einem signifikanten Ergebnis zugunsten Venlafaxin hätte führen können. Hierfür spricht auch, dass bei der aktuellen Metaanalysen von Bauer<sup>16, 17</sup> 2007 fünf weitere Studien berücksichtigt wurden, die nicht in der IQWiG Bewertung enthalten sind.

Abgesehen von den oben beschriebenen Studien, die nicht für die IQWiG Bewertung berücksichtigt wurden, ist unklar, warum die Ergebnisse der Studie von Alves 1999 zwar für die Bewertung der Response, nicht aber für die Remission berücksichtigt wurde.

Aufgrund der klinischen Relevanz und pharmako-ökonomischen Bedeutung von Remissionsdaten sind die im Vorbericht unberücksichtigten Studien von Mallick und Shelton bedeutsam.<sup>27, 28</sup> Hier wurden eine größere Anzahl depressionsfreier Tage und höhere Raten für dauerhafte Remission zugunsten Venlafaxin versus SSRIs beobachtet.

#### **b) Behandlungseffekte in Abhängigkeit der Schweregrade**

Im IQWiG Vorbericht wird festgestellt, dass die Datenlage hinsichtlich eines Vorteiles von Venlafaxin gegenüber SSRIs bei schwerer erkrankten Patienten nicht eindeutig sei, so dass lediglich ein Hinweis auf einen Zusatznutzen vorliegt.

Zum einen ist anzumerken, dass das methodische Vorgehen des IQWiGs hinsichtlich der Unterteilung der Schweregrade kein anerkannter Standard ist. In der Literatur werden vornehmlich cut-off-Werte auf Depressionsskalen verwendet.<sup>29</sup>

Im Vorbericht wird angemerkt, dass die Aussagekraft der IQWiG Analyse aufgrund fehlender individueller Patientendaten limitiert ist.

Genau für diese Fragestellung sind Ergebnisse von 2 Meta-Analysen auf dem Boden individueller Patientendaten vor kurzem publiziert

worden.<sup>30,31</sup> Bei diesen Analysen zeigte sich, dass sowohl bei leicht bis mittelschwer, als auch bei schwer erkrankten Patienten eine Venlafaxintherapie mit höheren Remissionsraten gegenüber SSRIs einherging.<sup>30</sup> Des Weiteren zeigte sich, dass die Unterschiede zugunsten Venlafaxin bei sehr schwer erkrankten Patienten sogar noch deutlicher waren. Darüber hinaus zeigte sich, dass eine Venlafaxintherapie zu verbesserter Arbeitsfähigkeitsfähigkeit sowohl bei leicht bis mittelschwer erkrankten Patienten als auch bei sehr schwer erkrankten Patienten führte.<sup>31</sup> Hierbei zeigten wiederum Vorteile von Venlafaxin gegenüber SSRIs, wobei die die Effekte bei den schwer erkrankten Patienten deutlicher ausgeprägt waren. Bei Berücksichtigung dieser Evidenz ergibt sich ein klarer Zusatznutzen für Venlafaxin.

### **c) Verträglichkeit**

Insbesondere bei der Bewertung der Sicherheit und Verträglichkeit von Substanzen ermöglicht die Beschränkung auf doppelblinde RCTs keine ausreichende Einschätzung. Gerade für diese Fragestellung sind Erkenntnisse auf dem Boden von großen Beobachtungsstudien mit Patienten des Praxisalltag und Registerdaten notwendig.

Auf der Basis der Daten von mehreren Metaanalysen ist die Verträglichkeit von Venlafaxin und SSRIs vergleichbar.<sup>16, 17</sup> Die im IQWiG-Vorbericht beschriebene höhere Anzahl an UEs sowie höhere Abbruchraten unter Venlafaxin vs. SSRIs ist gering und nicht als klinisch relevant einzustufen. Es ist zu betonen, dass im IQWiG-Vorbericht (wie auch in vielen Meta-Analysen) keine Unterscheidung zwischen den Darreichungsformen von Venlafaxin (schnell freisetzende Tablette versus Retardkapsel) vorgenommen wurde. Allein im IQWiG Vorbericht sind 38 der 60 für Venlafaxin berücksichtigten Studien mit der schnell freisetzenden Formulierung durchgeführt worden. Dies führt zu einem verzerrten Bild des Verträglichkeitsprofils von Trevilor retard, das im Praxisalltag fast ausschließlich eingesetzt wird. Es ist aus klinischen Studien und aus dem Praxisalltag bekannt, dass die Retardformulierung verträglicher (v.a. weniger Übelkeit) als die schnell freisetzende Formulierung ist.<sup>32</sup>

Beobachtungsstudien, die im Gegensatz zu randomisierten Zulassungsstudien häufig ältere und multimorbide Patienten einschloss und somit eher den Praxisalltag widerspiegeln, bestätigen die Sicherheit und Verträglichkeit von Venlafaxin im Praxisalltag.<sup>33-35</sup>

## **Therapieresistenz**

Die Therapieresistenz ist ein wichtiger, klinisch sehr relevanter Teilaspekt, der zwingend in eine ganzheitliche Bewertung einzelner Substanzen – und damit auch die Entscheidung, was einem Patienten verschrieben wird – einfließen muss.

Aus klinischer Sicht stellen die schwer erkrankten eine wichtige im Praxisalltag häufig anzutreffende Patientengruppe dar, zumal diese häufig schwer zu behandeln sind und ein erhöhtes Risiko für Rezidive und Suizide aufweisen.

Im IQWiG Vorbericht wird festgestellt, dass keine Hinweise auf eine relevante Interaktion zwischen vorbestehender Behandlungsresistenz und Behandlungseffekt vorliegen. Wir können uns dieser Einschätzung nicht anschließen, insbesondere weil vom IQWiG für diese Analyse nur 2 Studien berücksichtigt wurden. Auch in diesem Fall ermöglicht die Beschränkung auf doppelblinde RCTs keine ausreichende Analyse dieser Fragestellung.

So stellt zum Beispiel Venlafaxin bei therapierefraktären Patienten eine wichtige therapeutische Alternative dar. Wichtige klinische Studien, die wegen fehlender Verblindung nicht für den IQWiG Vorbericht berücksichtigt wurden, zeigen konsistent, dass der Einsatz von Venlafaxin bei Therapieversagern zu signifikanter Therapieresponse und Remission führen kann, insbesondere nach Versagen eines SSRI.<sup>36-38</sup> Es ist inzwischen längst klinischer Alltag, dass Venlafaxin insbesondere bei schwer erkrankten Patienten, die häufig nicht ausreichend auf ein anderes Antidepressivum angesprochen haben, eingesetzt wird.<sup>39</sup> GKV-Abrechnungsdaten belegen, dass lediglich 2,2% der Erstverordnungen von Antidepressiva auf Venlafaxin entfallen, 97,8% hingegen auf die anderen therapeutischen Klassen<sup>40</sup> (darunter 68% trizyklische Antidepressiva sowie 20% SSRI). Daten zu schwer erkrankten, therapieresistenten Patienten sind daher ein sehr wichtiger Beitrag für eine ganzheitliche Bewertung eines Antidepressivums.

## **Ältere Patienten**

Im IQWiG Vorbericht wird festgestellt, dass für Venlafaxin kein Nutzen für ältere Patienten besteht. Die Ergebnisse des IQWiG Vorberichtes berücksichtigen nur 2 Studien, so dass eine apodiktische Aussage nicht zulässig ist. Insbesondere spiegelt diese Analyse nicht die vollständige Evidenz zu dieser Fragestellung wider.

Wie für eine Zulassung üblich wurde für Patienten  $\geq 65$  Jahre eine Meta-Analyse auf dem Boden aller eingereichten Wirksamkeitsstudien durchgeführt (*Application for Marketing Authorisation Venlafaxine 25, 37.5, 50 & 75 mg Tablets, Part IVB: Clinical Experience, February 1992*). Dabei zeigte sich für die Wirksamkeit und Sicherheit kein Unterschied zwischen jüngeren oder älteren Patienten. Diese Ergebnisse spiegeln sich entsprechend in der Fachinformation zu Venlafaxin wider.

### **Zulassungsprofile**

Für den therapeutischen Alltag sind die teilweise stark voneinander abweichenden Zulassungsprofilen der verschiedenen Antidepressiva sehr bedeutsam. Venlafaxin verfügt in Deutschland von allen Antidepressiva über das breiteste Indikationsspektrum, wobei allein Venlafaxin sowohl für die Rezidivprophylaxe als auch zur Erhaltungstherapie der Depression zugelassen ist.<sup>41</sup> Die Rezidivprophylaxe ist nach nationalen und internationalen Leitlinien das entscheidende Therapieziel in der Langzeittherapie.<sup>16, 17</sup> Hier existiert für Venlafaxin unter allen Antidepressiva die umfangreichste Evidenz und nur sehr wenige Antidepressiva sind für diese Indikation überhaupt zugelassen.

Des Weiteren ist im Praxisalltag das Indikationsspektrum von Venlafaxin hinsichtlich der verschiedenen Angsterkrankungen ein wichtiger Behandlungsaspekt, weil ein Großteil der Depressionspatienten eine Komorbidität mit Angsterkrankungen aufweist. Hierbei ist hervorzuheben, dass Venlafaxin zu den wenigen Antidepressiva gehört, die zur Behandlung von Depressionen mit Angstsymptomatik zugelassen sind (Fachinformation Trevilor Retard 75/150mg, Stand Mai 2007). Diese wichtigen Aspekte tragen wesentlich zum klinischen Nutzen einer Substanz bei, wurden vom IQWiG in seinem veröffentlichten Vorbericht kaum oder nicht berücksichtigt.

### **Gefahr von Arzneimittelinteraktionen**

Ein weiterer klinisch wichtiger Aspekt, der im IQWiG Vorbericht unberücksichtigt bleibt, ist die Betrachtung der potentiellen Wechselwirkungen der unterschiedlichen Antidepressiva. Venlafaxin wird durch das Cytochrom P450<sub>2D6</sub> (CYP<sub>2D6</sub>) zu ODV metabolisiert; es ist selbst nur ein schwacher Hemmer des CYP<sub>2D6</sub>. Eine Inhibition der Isoenzyme CYP<sub>1A2</sub>, CYP<sub>2C9</sub>,

CYP<sub>3</sub>A<sub>4</sub> und CYP<sub>2</sub>C<sub>19</sub> wurde in vitro nicht beobachtet. Aufgrund dieser Eigenschaften ist die Gefahr von Wechselwirkungen zwischen Venlafaxin mit anderen Arzneimitteln im Vergleich zu vielen anderen Antidepressiva sehr gering. Daher ist der Einsatz von Venlafaxin gerade auch bei älteren und/oder multimorbiden Patienten mit häufiger Komedikation und Komorbidität eine wichtige Behandlungsoption.

### **Schlussfolgerung**

Der IQWiG Vorbericht ist eine umfassende Analyse von vielen wichtigen Daten, basiert jedoch nur auf einen Teil der verfügbaren Evidenz.

Aufgrund der dargelegten Gründe fordern wir, für die Nutzenbewertung u.a. folgende Modifikationen vorzunehmen:

1. Eine Nutzenbewertung der Antidepressiva sollte sämtliche Therapieoptionen berücksichtigen und nicht auf SNRIs beschränkt bleiben.
2. Die Nutzenbewertung sollte auf dem Boden der für die jeweilige Fragestellung verfügbaren bestmöglichen Evidenz erfolgen (z.B. Verträglichkeit, Therapieresistenz)
3. Es sollte eine Re-Analyse der Wirksamkeitsdaten bei Berücksichtigung der aus nicht nachvollziehbaren Gründen ausgeschlossenen Studien erfolgen
4. Die hier gefundenen Ergebnisse zur Wirksamkeit (z.B. Remission) sollten vor dem Hintergrund neuerer, umfangreicher Metaanalysen diskutiert (Diskrepanzen, Limitationen etc.) werden.
5. Für die Bewertung der klinischen Relevanz sollten entsprechende Parameter wie z.B. NNTs im Rahmen eines Scoping-Prozesses ermittelt werden.
6. Für die Bewertung der Schweregrade sollten individuelle Patientendaten berücksichtigt werden, die in Form aktueller publizierter Subanalysen vorliegen.
7. Aufgrund der hohen klinischen Relevanz der Therapieresistenz für den Praxisalltag sollten für diese Fragestellung auch nicht verblindete, jedoch randomisierte Studien berücksichtigt werden.
8. Die Limitationen von meta-analytischen Ergebnissen auf dem Boden nur weniger Studien sollten verdeutlicht werden.
9. Für die Subanalyse bei älteren Patienten sollten entsprechende Analysen, die für die Zulassung vorgenommen werden, berücksichtigt werden.
10. Hinsichtlich der Verträglichkeit sollte eine Unterscheidung zwischen der retardierten und der schnell freisetzenen Formulierung vorgenommen werden
11. Das breite Indikationsspektrum sollte im Vergleich zu Duloxetin, SSRIs und TZAs beschrieben und die Vorteile insbesondere hinsichtlich der häufigen Komorbidität von Depression und Angsterkrankungen diskutiert werden.
12. Die unterschiedlichen pharmakologischen Profile der Antidepressiva sollten im Hinblick auf die Gefahr von Arzneimittelinteraktionen beschrieben und diskutiert werden.



## Literatur

1. Wellington K, Perry CM. Venlafaxine extended-release: a review of its use in the management of major depression. *CNS Drugs* 2001; 15(8):643-669.
2. Thase ME, Entsuah AR, Rudolph RL. Remission rates during treatment with venlafaxine or selective serotonin reuptake inhibitors. *Br J Psychiatry* 2001; 178:234-241.
3. Kocsis JH, Thase ME, Trivedi MH et al. Prevention of recurrent episodes of depression with venlafaxine ER in a 1-year maintenance phase from the PREVENT Study. *J Clin Psychiatry* 2007; 68(7):1014-1023.
4. Keller MB, Trivedi MH, Thase ME et al. The Prevention of Recurrent Episodes of Depression with Venlafaxine for Two Years (PREVENT) Study: Outcomes from the 2-year and combined maintenance phases. *J Clin Psychiatry* 2007; 68(8):1246-1256.
5. Mueller TI, Leon AC, Keller MB et al. Recurrence after recovery from major depressive disorder during 15 years of observational follow-up. *Am J Psychiatry* 1999; 156(7):1000-1006.
6. De NA, Geerts S, Ruelens L et al. Venlafaxine compared with fluoxetine in outpatients with depression and concomitant anxiety. *Int J Neuropsychopharmacol* 2002; 5(2):115-120.
7. Li Y., Zhang X-N., Wu Z.-M. Comparative study on the effects of paroxetine and venlafaxine in treating depression. *Chinese Journal of Clinical Rehabilitation* 2004; 8(21):4174-4175.
8. Ballús C, Quiros G, De Flores T et al. The efficacy and tolerability of venlafaxine and paroxetine in outpatients with depressive disorder of dysthymia. *International Clinical Psychopharmacology* 2000; 15:43-48.
9. Oslin DW, Ten Have TR, Streim JE et al. Probing the safety of medications in the frail elderly: evidence from a randomized clinical trial of sertraline and venlafaxine in depressed nursing home residents. *J Clin Psychiatry* 2003; 64(8):875-882.
10. Zanardi R, Franchini L, Serretti A, Perez J, Smeraldi E. Venlafaxine versus fluvoxamine in the treatment of delusional depression: a pilot double-blind controlled study. *J Clin Psychiatry* 2000; 61(1):26-29.
11. Poirier M-F, Boyer P. Venlafaxine and paroxetine in treatment-resistant depression. Double-blind, randomised-comparison. *British Journal of Psychiatry* 1999; 174:12-16.
12. Moller HJ, Maier W. [Problems of evidence-based medicine in psychopharmacotherapy: problems of evidence grading and of the evidence basis for complex clinical decision making]. *Nervenarzt* 2007; 78(9):1014-1027.
13. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995; 310(6977):452-454.

14. Higgins Julian PT, Green Sally. Cochrane Handbook for Systematic Review of Interventions. Cochrane Collaboration 2008; 5.0.1.
15. Hankey GJ, Warlow CP. Treatment and secondary prevention of stroke: evidence, costs, and effects on individuals and populations. *Lancet* 1999; 354(9188):1457-1463.
16. Bauer M., Tharmanathan P., Volz H.-P., Moeller H.-J., Freemantle N. The effect of venlafaxine compared with other antidepressants and placebo in the treatment of major depression. A meta-analysis. *European Archives of Psychiatry and Clinical Neurosciences*, in press .
17. Meta Analysis of randomised controlled trials of Venlafaxine in the treatment of major depressive disorder [Metaanalyse der randomisierten, kontrollierten Studien zu Venlafaxin in der Behandlung der majoren Depression]. *Der Nervenarzt* 2007; 78 (suppl. 2):115: 2007.
18. Bauer M, Bschor T, Pfennig A et al. World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for Biological Treatment of Unipolar Depressive Disorders in Primary Care. *World J Biol Psychiatry* 2007; 8(2):67-104.
19. Rudolph RL, Feiger AD. A double-blind, randomized, placebo-controlled trial of once-daily venlafaxine extended release (XR) and fluoxetine for the treatment of depression. *J Affect Disord* 1999; 56(2-3):171-181.
20. Costa e Silva. Randomized, double-blind comparison of venlafaxine and fluoxetine in outpatients with major depression. *J Clin Psychiatry* 1998; 59(7):352-357.
21. Mehtonen OP, Sogaard J, Roponen P, Behnke K. Randomized, double-blind comparison of venlafaxine and sertraline in outpatients with major depressive disorder. Venlafaxine 631 Study Group. *J Clin Psychiatry* 2000; 61(2):95-100.
22. Allard P, Gram L, Timdahl K, Behnke K, Hanson M, Sogaard J. Efficacy and tolerability of venlafaxine in geriatric outpatients with major depression: a double-blind, randomised 6-month comparative trial with citalopram. *Int J Geriatr Psychiatry* 2004; 19(12):1123-1130.
23. Montgomery SA, Huusom AK, Bothmer J. A randomised study comparing escitalopram with venlafaxine XR in primary care patients with major depressive disorder. *Neuropsychobiology* 2004; 50(1):57-64.
24. Nemeroff CB, Entsuah R, Benattia I, Demitrack M, Sloan DM, Thase ME. Comprehensive analysis of remission (COMPARE) with venlafaxine versus SSRIs. *Biol Psychiatry* 2008; 63(4):424-434.
25. Smith D, Dempster C, Glanville J, Freemantle N, Anderson I. Efficacy and tolerability of venlafaxine compared with selective serotonin reuptake inhibitors and other antidepressants: a meta-analysis. *Br J Psychiatry* 2002; 180:396-404.
26. Weinmann S, Becker T, Koesters M. Re-evaluation of the efficacy and tolerability of venlafaxine vs SSRI: meta-analysis. *Psychopharmacology (Berl)* 2008; 196(4):511-520.

27. Mallick R, Chen J, Entsuah AR, Schatzberg AF. Depression-free days as a summary measure of the temporal pattern of response and remission in the treatment of major depression: a comparison of venlafaxine, selective serotonin reuptake inhibitors, and placebo. *J Clin Psychiatry* 2003; 64(3):321-330.
28. Shelton C, Entsuah R, Padmanabhan SK, Vinall PE. Venlafaxine XR demonstrates higher rates of sustained remission compared to fluoxetine, paroxetine or placebo. *Int Clin Psychopharmacol* 2005; 20(4):233-238.
29. Montgomery SA, Lecrubier Y. Is severe depression a separate indication? ECNP Consensus Meeting September 20, 1996, Amsterdam. European College of Neuropsychopharmacology. *Eur Neuropsychopharmacol* 1999; 9(3):259-264.
30. Schmitt A, Jiang Q, Loeschmann PA, Ahmed S, Wyeth Pharma GmbH, Wyeth Pharmaceuticals. Differential effects of venlafaxine compared to selective serotonin reuptake inhibitors (SSRIs) in the treatment of MDD according to baseline severity. *European Psychiatry* 2008; 23:262-263.
31. Schmitt A.B., Volz H.-P., Wiedemann K, Fritze J, Loeschmann P.-A. Therapieeffekte durch Venlafaxine auf die Arbeitsfähigkeit im Vergleich zu Serotonin-Wiederaufnahmehemmern bei majorer Depression verschiedener Schweregrade. *Gesundheitsökonomie & Qualitätsmanagement* 2008; 13:269-275.
32. Entsuah R, Gao B. Global Benefit-risk Evaluation of Antidepressant Action: Comparison of Pooled Data for Venlafaxine, SSRIs, and Placebo. *CNS Spectr* 2002; 7(12):882-888.
33. Linden M, Ludewig K, Munz T, Dierkes W. Dosage Finding and Outcome of Venlafaxine Treatment in Psychiatric Outpatients and Inpatients: Results of a Drug Utilization Observation Study. *Pharmacopsychiatry* 2003; 36:197-205.
34. Baca E, Roca M, Garcia-Calvo C, Prieto R. Venlafaxine extended-release in patients older than 80 years with depressive syndrome. *Int J Geriatr Psychiatry* 2006; 21(4):337-343.
35. Zimmer B, Kant R, Zeiler D, Brilmyer M. Antidepressant efficacy and cardiovascular safety of venlafaxine in young vs old patients with comorbid medical disorders. *Int J Psychiatry Med* 1997; 27(4):353-364.
36. Thase ME, Shelton RC, Khan A. Treatment with venlafaxine extended release after SSRI nonresponse or intolerance: a randomized comparison of standard- and higher-dosing strategies. *J Clin Psychopharmacol* 2006; 26(3):250-258.
37. Baldomero EB, Ubago JG, Cercos CL, Ruiloba JV, Calvo CG, Lopez RP. Venlafaxine extended release versus conventional antidepressants in the remission of depressive disorders after previous antidepressant failure: ARGOS study. *Depress Anxiety* 2005; 22(2):68-76.
38. Rush AJ, Trivedi MH, Wisniewski SR et al. Bupropion-SR, sertraline, or venlafaxine-XR after failure of SSRIs for depression. *N Engl J Med* 2006; 354(12):1231-1242.

39. Mines D, Hill D, Yu H, Novelli L. Prevalence of risk factors for suicide in patients prescribed venlafaxine, fluoxetine, and citalopram. *Pharmacoepidemiol Drug Saf* 2005; 14(6):367-372.
40. Gothe H, Hörer A, Volmer T, Runge C, Glaeske G, Häussler B. Co-Medikation bei Patienten mit Antidepressiva-Therapie unter Alltagsbedingungen - eine Analyse von Routineabrechnungsdaten einer deutschen gesetzlichen Krankenversicherung. IGES Institut GmbH 2007.
41. Therapieempfehlungen der Arzneimittelkommission der deutschen Ärzteschaft (Depression). *Arzneimittelverordnung in der Praxis (AVP)* 2006;(2):1-41.

## **A 2 Stellungnahmen von Privatpersonen**

### **A 2.1 Roots, Ivar, Prof. Dr. med.**

**Adresse:**

Prof. Dr. med. Ivar Roots

Institut für Klinische Pharmakologie

Charité

Charitéplatz 1

10117 Berlin

Stellungnahme zum Auftrag **A05-20A**  
„Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI)  
zur Behandlung der Depression“  
Stellungnahme zum Vorbericht 1.0

**Vergleichende Pharmakokinetik der SSRIs  
(Citalopram, Escitalopram, Fluoxetin, Fluvoxamin,  
Paroxetin, Sertralin)**

Prof. Dr. med. Ivar Roots  
Facharzt für Klinische Pharmakologie  
Facharzt für Pharmakologie und Toxikologie

Direktor des  
Instituts für Klinische Pharmakologie und Toxikologie  
Charité - Universitätsmedizin Berlin  
Charitéplatz 1  
10117 Berlin  
Tel.: 030-450-525112, Fax: 030-450-525932  
E-Mail: [ivar.roots@charite.de](mailto:ivar.roots@charite.de)

## Einleitung

In den letzten 20 Jahren wurden bei der Behandlung der Depression die klassischen tricyclischen Antidepressiva (TCA) wie Imipramin und Amitriptylin zunehmend von den moderneren selektiven Serotonin-Wiederaufnahmehemmern (SSRI) verdrängt. Global gesehen ist dies durch die insgesamt bessere Verträglichkeit dieser Gruppe, d.h. das geringere Auftreten von Nebenwirkungen und toxischer Erscheinungen, begünstigt worden. Darüber hinaus weisen einige Substanzen unter den SSRIs auch ein deutlich kleineres Interaktionspotential mit anderen Medikamenten auf, die – mit gleicher Indikation oder zur Behandlung einer Komorbidität – erforderlich sind. Ein weiterer, zunehmend als wichtig angesehener Vorteil einiger der SSRIs gegenüber den TCAs liegt in der weniger starken Abhängigkeit ihrer Pharmakokinetik und damit ihrer Wirkstärke von pharmakogenetischen Besonderheiten im Metabolismus.

Die Unterscheidungsmerkmale zwischen TCAs und SSRIs – Unterschiede in Pharmakokinetik, Metabolisierung, Pharmakogenetik, Wechselwirkung – spielen auch innerhalb dieser beiden großen Antidepressivagruppen eine wichtige Rolle. Für die seit langem klinisch eingesetzten TCAs ist dies ausführlich in der Literatur dargestellt. Für die Gruppe der SSRIs liegt ebenfalls eine umfangreiche Literatur vor, die im Folgenden analysiert werden soll.

Ziel der vorliegenden Stellungnahme ist es zu prüfen, ob die sechs SSRIs (Citalopram, Escitalopram, Fluoxetin, Fluvoxamin, Paroxetin, Sertralin), für deren Wirkung ein sehr ähnlicher Mechanismus angenommen wird, sich nicht doch aufgrund der oben genannten Kriterien in einem Maße unterscheiden, dass dies auch klinisch relevant werden könnte. Wenn das der Fall ist, würde das bedeuten, dass die Substanzen individuell betrachtet werden müssten und beim klinischen Einsatz nicht ohne Bedacht ausgetauscht werden dürften.

Es soll hier nicht die klinische Wirksamkeit der SSRIs beim Patienten betrachtet werden. Vielmehr steht der Vergleich der Verträglichkeit, Sicherheit und ärztlichen Handhabbarkeit der SSRIs beim individuellen Patienten anhand klinisch-pharmakologischer Kriterien im Vordergrund.

## Kurze Darstellung der SSRIs

Zum besseren Verständnis der späteren Diskussion werden hier für die einzelnen SSRIs wichtige Charakteristika zur Pharmakokinetik und zum Metabolismus aufgeführt.

### Citalopram

Citalopram ist ein Razemat, das S-Enantiomer ist pharmakologisch weitaus aktiver als das R-Enantiomer. Darüber hinaus konnte gezeigt werden, dass R-Citalopram die Aktivität des S-Enantiomers behindert (Sanchez et al., 2004).

Resorption, Verteilung, Elimination: Citalopram wird bei oraler Verabreichung gut resorbiert. Das Ausmaß der Resorption ist unabhängig von der Nahrungsaufnahme. Die Substanz unterliegt einem First-Pass-Metabolismus, der jedoch kleiner ist als bei anderen SSRIs. Die orale Bioverfügbarkeit beträgt ca. 80%. Maximale Plasmaspiegel werden nach 2-4 Std. erreicht. Dosiserhöhung führt zu einer proportionalen Erhöhung des Plasmaspiegels. Das Verteilungsvolumen beträgt ca. 15 l/kg. Nach ca. 1 Woche sind die Plasmaspiegel bei oraler Gabe im Gleichgewicht. Die Ausscheidung erfolgt zu 85% über die Leber und zu 15% über die Niere (El-Armouche et al., 2003).

Metabolismus: Citalopram wird überwiegend N-demethyliert, wobei N-Desmethylcitalopram entsteht, das dann einer weiteren N-Demethylierung zum Desmethylcitalopram unterliegt. An der ersten Reaktion sind CYP2C19, CYP2D6 und CYP3A4 etwa mit gleichem Anteil beteiligt, die zweite erfolgt über CYP2D6 (Sindrup et al., 1993; Oleson & Linnet, 1999). Die Eliminationshalbwertszeit im Steady-State beträgt für Citalopram ca. 36 Std., die der N-demethylierten Metaboliten ist ca. 2-3 mal so lang. Sie sind allerdings für die therapeutische Wirkung nur von geringer Bedeutung (Brøsen & Naranjo, 2001).

Interaktionen: Citalopram hemmt CYP2D6 schwach und CYP1A2, CYP2C19 und CYP3A4 ebenfalls nur schwach. Aus dieser Konstellation ergibt sich theoretisch nur ein geringes Interaktionspotential (Spina et al., 2008).

### Escitalopram

Escitalopram ist das pharmakodynamisch wirksame S-Isomer von Citalopram. Der Patient ist also befreit von der Belastung durch R-Citalopram. Untersuchungen zeigten, dass R-Citalopram möglicherweise die Besetzung von S-Citalopram am Serotonin-Reuptaketransporter (SERT) im Mittelhirn abschwächt (Klein et al., 2007).



Resorption, Verteilung, Elimination: Escitalopram wird unabhängig von der Nahrungsaufnahme nahezu vollständig resorbiert. Die orale Bioverfügbarkeit liegt bei ca. 80%. Maximale Blutspiegel werden nach etwa 3-4 Std. erreicht. Die Beziehung zwischen Dosis und Plasmaspiegel ist linear. Das Verteilungsvolumen liegt bei 10-25 l/kg. Der Steady-State wird nach ca. 7-10 Tagen erreicht (Rao 2007).

Metabolismus: Escitalopram wird hauptsächlich in der Leber verstoffwechselt. Es wird unter Beteiligung der Enzyme CYP2C19, CYP3A4 und CYP2D6 zu S-Desmethylcitalopram umgesetzt. Ein weiterer Demethylierungsschritt erfolgt über CYP2D6. Die Produkte beider Reaktionen haben eine gewisse, wenn auch klinisch nicht bedeutende pharmakologische Aktivität. Die Plasmahalbwertszeit von Escitalopram beträgt ca. 30 Std. (Rao 2007).

Interaktionen: Escitalopram ist ein schwacher Inhibitor von CYP2D6, die anderen CYP-Enzyme werden kaum in ihrer Aktivität beeinflusst (Spina et al., 2008).

## **Fluoxetin**

Fluoxetin ist ein Razemat. Das S-Enantiomer hemmt die Serotoninaufnahme ca. 1,5fach stärker als das R-Enantiomer. Bei der Verstoffwechslung entsteht der pharmakologisch aktive Metabolit Norfluoxetin, dessen S-Enantiomer 20fach stärker wirksam ist als das R-Enantiomer. Norfluoxetin ist ein maßgeblicher Träger der antidepressiven Wirkung (Liu et al., 2001).

Resorption, Verteilung, Elimination: Fluoxetin wird bei oraler Gabe fast vollständig resorbiert, da es jedoch einem First-Pass-Metabolismus unterliegt, beträgt die orale Bioverfügbarkeit weniger als 90%. Spitzenkonzentrationen werden im Blut relativ spät nach 6-8 Stunden erreicht. Das Verteilungsvolumen ist mit 14-45 l/kg das höchste in der Gruppe der SSRI. Ein Steady-State tritt erst nach 1-22 Monaten ein; diese sehr breite Spanne dürfte insbesondere dem genetischen Polymorphismus des für den Abbau maßgeblichen Enzyms CYP2D6 geschuldet sein. Zwischen der Dosis und der Blutkonzentration besteht eine nichtlineare Beziehung (Hiemke & Härtter, 2000).

Metabolismus: Fluoxetin wird N-demethyliert, wodurch der aktive Metabolit Norfluoxetin entsteht. Das wichtigste Enzym für diese Reaktion ist CYP2D6, wobei jedoch auch CYP2C9, CYP2C19 und CYP3A4 einen Beitrag leisten. Die Halbwertszeit von Fluoxetin beträgt 1-4 Tage, die von Norfluoxetin 7-15 Tage. Im Steady-State ist die Norfluoxetinkonzentration im Blut höher als die von Fluoxetin (Hiemke & Härtter, 2000).

Interaktionen: Sowohl die Muttersubstanz als auch der Metabolit Norfluoxetin sind potente Inhibitoren von CYP2D6, in mäßigem Ausmaß werden auch CYP2C19 und CYP3A4 gehemmt. Damit bestehen unter der Therapie mit Fluoxetin mannigfaltige Möglichkeiten für Arzneimittelinteraktionen, die aufgrund der langen Halbwertszeit von Fluoxetin und Norfluoxetin auch nach Absetzen noch wochenlang bestehen bleiben können (Spina et al., 2008).

### **Fluvoxamin**

Resorption, Verteilung, Elimination: Fluvoxamin wird zwar zu über 90% resorbiert, aufgrund des ausgeprägten First-Pass-Metabolismus in der Leber beträgt die Bioverfügbarkeit aber nur wenig mehr als 50%. Bei einer Erhöhung der Dosis verlängert sich die Halbwertszeit, Fluvoxamin folgt somit einer Sättigungskinetik. Das Verteilungsvolumen liegt bei 25 l/kg. Steady-State-Konzentrationen sind innerhalb einer Woche erreicht (Hiemke & Härter, 2000).

Metabolismus: Die Biotransformation von Fluvoxamin erfolgt in der Leber unter Mitwirkung von CYP1A2 und CYP2D6. Die Plasmahalbwertszeit von Fluvoxamin beträgt 8-28 Std. (Hiemke & Härter, 2000).

Interaktionen: Fluvoxamin beeinflusst die Aktivität einer Reihe von CYP-Enzymen. Es ist ein starker Inhibitor von CYP1A2 und CYP2C19, hemmt jedoch auch CYP2C9 und CYP3A4, wenngleich mäßig. Seine Wirkung auf CYP2D6 ist nur gering. Die breite Beeinflussung der CYP-Aktivität führt zu einer Fülle von Interaktionsmöglichkeiten (Spina et al., 2008).

### **Paroxetin**

Resorption, Verteilung, Elimination: Trotz fast vollständiger oraler Resorption liegt die orale Bioverfügbarkeit einer Einzeldosis von Paroxetin aufgrund eines ausgeprägten First-Pass-Metabolismus in der Leber unter 50%. Dieser First-Pass-Metabolismus ist sättigbar, so dass die orale Bioverfügbarkeit bei Mehrfachgabe steigt. Die Halbwertszeit von Paroxetin ist abhängig von Dosis und Anwendungsdauer, bei therapeutischer Dosierung liegt sie bei 16-21 Stunden. Steady-State-Plasmakonzentrationen werden nach 7-14 Tagen erreicht. Die Plasma-Eiweiß-Bindung liegt um 95% (Bourin et al., 2001; Hiemke & Härter, 2000).

Metabolismus: Paroxetin wird in der Leber zu pharmakologisch inaktiven Metaboliten abgebaut. An diesem Prozeß sind die Enzyme CYP2D6 und CYP3A4 beteiligt. Die Umsetzung über CYP2D6 unterliegt der Sättigung, was die nichtlineare Kinetik von Paroxetin erklärt (Bourin et al., 2001; DeVane, 2003).

Interaktionen: Paroxetin ist ein potenter Hemmstoff von CYP2D6. Interaktionen sind daher zu erwarten, wenn der Patient weitere Substrate von CYP2D6 einnimmt. Die Hemmung dauert so lange an, wie ausreichende Paroxetinkonzentrationen vorhanden sind (3-7 Tage) (Spina et al., 2008).

### **Sertralin**

Resorption, Verteilung, Elimination: Sertralin wird fast vollständig resorbiert, maximale Plasmaspiegel werden nach 4-8 Std. erreicht. Die Beziehung zwischen Dosis und Plasmaspiegel ist linear. Das Verteilungsvolumen beträgt mehr als 20 l/kg. Die Plasma-Eiweiß-Bindung liegt bei 98%. Die Elimination ist bei Frauen und älteren Patienten langsamer, Grund hierfür sind möglicherweise Unterschiede im Verteilungsvolumen oder im Metabolismus (DeVane et al., 2002)

Metabolismus: Sertralin wird in der Leber verstoffwechselt. Die Halbwertszeit beträgt ca. 26 Std. Der wichtigste Schritt ist die N-Demethylierung. N-Desmethylsertralin hat 5-10% der pharmakologischen Aktivität der Muttersubstanz, aber eine dreimal so lange Halbwertszeit (60-100 Std.). An der N-Demethylierung sind die Enzyme CYP2C9, CYP2C19, CYP2D6 und CYP3A4 beteiligt (Spina et al., 2008; Hiemke & Härtter, 2000).

Interaktionen: Sertralin ist ein mittelstarker Inhibitor von CYP2D6, das Ausmaß der Hemmung könnte dosisabhängig sein und unter Umständen klinische Bedeutung erlangen. Andere CYP-Enzyme werden in ihrer Aktivität allenfalls mäßiggradig beeinflusst (Spina et al., 2008).

## **Deskriptive Pharmakokinetik der SSRI**

Tab. 1 faßt die wichtigen pharmakokinetischen Parameter der SSRI zusammen. Bei allen betrachteten SSRI mit Ausnahme von Paroxetin, besitzen auch die primären Metaboliten pharmakodynamische Wirksamkeit. Der Beitrag zur antidepressiven Wirkung dürfte jedoch meist gering sein, mit Ausnahme des Norfluoxetin, dessen Kinetik separat aufgeführt ist

**Tab. 1** Pharmakokinetik der SSRI (nach Hiemke & Härtter, 2000; ergänzt, \*Sögaard et al., 2005)

Substanz	Tagesdosis (mg)	T ½	Zeit bis Steady-State	Verteilungsvolumen (l/kg)	Lineare Kinetik
Citalopram	10-60	36 h	6-10 Tage	14-16	ja
Escitalopram*	10-20	29 h	5 Tage	18	ja
Fluoxetin Norfluoxetin	20-80	1-4 Tage 7-15 Tage	> 4 Wochen	20-45	nein
Fluvoxamin	50-300	15 h	10 Tage	5	nein
Paroxetin	20-50	20 h	7-14 Tage	3-12	nein
Sertralin	50-150	26 h	5-7 Tage	20	ja

Alle sechs SSRI zeichnen sich durch eine vergleichsweise lange Halbwertszeit aus, die – mit Ausnahme von Fluoxetin – bei einem Tag liegt. Dementsprechend stellt sich das Fließgleichgewicht bei Mehrfachdosierung auch erst nach frühestens einer Woche ein. Hiervon unterscheidet sich Fluoxetin signifikant mit einer deutlich längeren Halbwertszeit von 1-4 Tagen und einem aktiven Metaboliten (Norfluoxetin), der klinisch bedeutsam ist und eine Halbwertszeit von 1-2 Wochen aufweist. Dementsprechend dauert es länger als einen Monat, bis sich ein Fließgleichgewicht eingestellt hat.

Das Vorliegen einer nichtlinearen Kinetik bei Fluoxetin, Fluvoxamin und Paroxetin könnte notwendige Dosisveränderungen erschweren und evtl. durch unerwartet hohe Konzentrationsanstiege leichter toxische Effekte hervorrufen. Die Dosislinearität von Citalopram, Escitalopram und Sertralin dürfte als Vorteil zu werten sein.

## Elimination der SSRI

Alle hier betrachteten SSRI werden ganz überwiegend durch Metabolismus eliminiert, die renale Elimination spielt keine Rolle. Die dabei beteiligten Enzyme wurden oben bei der Besprechung der einzelnen Substanzen genannt und sind in Tab. 2 zusammengefaßt. Es sind im Wesentlichen nur Cytochrom-P450-Enzyme involviert. Die Betrachtung der Enzyme ist in diesem Kontext wichtig, da sich hieraus Informationen für Arzneimittelwechselwirkungen und für die individuelle Varianz der Aktivität beim einzelnen Patienten ergeben.

**Tab. 2**      Verstoffwechslung von SSRI über Cytochrom-P450-Enzyme  
(zusammengestellt gemäß der Einschätzung nach Spina et al., 2008)

Substanz	Metabolisierung über				
	CYP1A2	CYP2C9	CYP2C19	CYP3A4	CYP2D6
Citalopram			+	+	+
Escitalopram			+	+	+
Fluoxetin		+	+	+	+
Fluvoxamin	+				+
Paroxetin				+	++
Sertralin		+	+	+	+

**Kurze Charakterisierung der beim Stoffwechsel von SSRI involvierten CYP-Enzyme**

CYP2D6: Dieses Cytochrom-P450-Enzym ist am Metabolismus von ca. 25% der Arzneimittel beteiligt. Wie Tab. 2 zeigt, gehören hierzu auch alle sechs SSRI. Von klinischer Bedeutung ist, dass auch tricyclische Antidepressiva und Neuroleptika fast ausnahmslos über CYP2D6 maßgeblich metabolisiert werden. CYP2D6 ist nicht induzierbar (im Gegensatz zu z.B. CYP3A4 und CYP1A2). In der Bevölkerung ist CYP2D6 in seiner Aktivität hochgradig variabel. Dies beruht auf dem Vorliegen mehrerer genetischer Varianten. Man unterscheidet die folgenden Phänotypen der CYP2D6-Aktivität, die heute durch einfache genotypische Messungen erfasst werden können:

- *Schneller Metabolisierer* (extensive metaboliser, EM; ca. 75% der deutschen Bevölkerung)
  - es liegen beim betreffenden Individuum 2 aktive Allel vor.
- *Intermediärer Metabolisierer* (IM; ca. 15% der deutschen Bevölkerung) – heterozygoter Zustand mit einem aktiven und einem defekten Allel oder Kombination zweier Allele mit jeweils verminderter Aktivität.
- *Langsamer Metabolisierer* (poor metabolizer, PM; ca. 7% der deutschen Bevölkerung)
  - beide Allele sind defekt und der Träger hat kein CYP2D6-Enzym.
- *Ultraschneller Metabolisierer* (UM; ca. 3% der deutschen Bevölkerung) – hier liegt eine Genduplikation von zwei aktiven Allelen vor, die mit einem weiteren aktiven Allel gepaart sind. Die Träger dieses Genotyps weisen eine weit überdurchschnittliche CYP2D6-Aktivität auf.

Die spezifischen Allelfrequenzen sind in den verschiedenen ethnischen Gruppen weltweit sehr unterschiedlich.

CYP2C19: Die Zahl der Substrate von CYP2C19 ist begrenzt, dieses CYP-Enzym spielt jedoch bei der Verstoffwechslung etlicher Psychopharmaka eine wichtige Rolle. Auch für CYP2C19 gibt es in der Bevölkerung verschiedene mutierte Allele, einschließlich eines defekten Allels. Für die CYP2C19-Aktivität finden sich in der deutschen Bevölkerung die folgenden Phänotypen: PM (ca. 3%), IM (ca. 25 %), UM (aufgrund einer aktivitätssteigernden Mutation, CYP2C19\*17, ca. 5 %). Die Aktivität von CYP2C19 kann durch Gabe enzyminduzierender Pharmaka gesteigert werden.

CYP1A2: Die Zahl der Substrate, soweit sie Arzneimittel sind, ist klein. CYP1A2 kann stark induziert werden, vor allem durch polycyclische Kohlenwasserstoffe (z.B. im Zigarettenrauch). In der Bevölkerung kommen Mutanten vor, die eine höhere Induzierbarkeit aufweisen.

CYP2C9: Dieses Enzym ist für die anteilige Metabolisierung einiger Psychopharmaka von Wichtigkeit. Große Bedeutung hat CYP2C9 u.a. für nichtsteroidale Antirheumatika und etliche blutzuckersenkende Substanzen. In der Bevölkerung kommen Mutanten mit verminderter Aktivität vor. Bei ca. 1% sind diese klinisch relevant.

CYP3A4: Dieses Enzym metabolisiert ca. die Hälfte unserer Arzneimittel. Es ist sehr stark induzierbar. Viele der CYP3A4-Substrate werden auch über CYP3A5 metabolisiert, was in der Vergangenheit nicht immer differenziert wurde. Das letztere Enzym weist genetisch bedingte Aktivitätsunterschiede in der Bevölkerung auf.

## **Differenzierung der SSRI nach pharmakogenetischen Kriterien in ihrem Metabolismus**

Wie oben dargestellt, sind die am Metabolismus der sechs SSRI beteiligten Enzyme durch genetische Variation von Person zu Person sehr unterschiedlich aktiv. Zum Erreichen vergleichbarer Plasmaspiegel sind daher unterschiedliche Dosen je nach Genotyp des Patienten erforderlich. Da zur Zeit der betreffende Genotyp nur selten getestet wird, ist davon auszugehen, dass die Exposition des individuellen Patienten gegenüber den einzelnen SSRIs sehr unterschiedlich ist. Bei einem Teil der Patienten muß man daher mit individueller Überdosierung und möglichen toxischen Wirkungen rechnen, bei anderen Patienten mit

ungenügender Wirkung. Will man diese Unterschiede ausgleichen, ergibt sich – je nach Genotyp – ein unterschiedlicher Dosisbedarf (Kirchheiner et al., 2001; 2004)

### **Einfluß des CYP2D6-Polymorphismus auf die Dosis**

Auch wenn alle sechs SSRI über CYP2D6 metabolisiert werden, das – wie oben dargestellt – sich durch einen hochgradigen Polymorphismus in der Bevölkerung auszeichnet, so hat dies nur für Fluoxetin und Paroxetin klinische Konsequenzen im Sinne einer notwendigen Dosisanpassung. Im Einzelnen:

Citalopram: Nach unseren Berechnungen (Kirchheiner et al., 2004) ist keine Dosisanpassung bei unterschiedlichen CYP2D6-Genotypen erforderlich (Sindrup et al., 1993).

Escitalopram: Es liegen keine Berechnungen in der Literatur vor. Die Konstellation gleicht jedoch der von Citalopram, damit ist eine Dosisanpassung bei den CYP2D6-Varianten nicht erforderlich.

Fluoxetin: PMs sollten ca. die halbe Dosis der EMs erhalten (Fjordside et al., 1999; Hamelin et al., 1996).

Fluvoxamin: Dosisreduktion bei PM auf ca. die Hälfte empfohlen (Carrillo et al., 1996; Spigset et al. 1997).

Paroxetin: Dosisreduktion bei PM auf ca. 66% der Normdosis (Sindrup et al., 1992).

Sertralin: Keine Dosisanpassung erforderlich (Hamelin et al., 1996).

### **Einfluß des CYP2C19-Polymorphismus auf die SSRI-Dosis**

Citalopram: Die Dosis müsste beim PM auf 2/3 reduziert werden (Sindrup et al., 1993). Der Patient mit UM-Genotyp müsste nach den Untersuchungen von Rudberg et al. (2008) mit Escitalopram einen Zuschlag von ca. 30% auf die Normaldosis erhalten.

Escitalopram: Ähnlich wie Citalopram (Rudberg et al., 2008).

Fluoxetin: Bei Berücksichtigung auch des Desmethylmetaboliten (Norfluoxetin) ist eine Dosisanpassung nicht erforderlich (Liu et al., 2001).

Fluvoxamin: wird nicht über CYP2C19 verstoffwechselt (Jan et al., 2002).

Paroxetin: wird nicht über CYP2C19 verstoffwechselt.

Sertralin: keine oder nur leichte (25%) Dosisverminderung bei PM empfohlen (Wang et al., 2001).

## **Fazit**

Diese Zusammenstellung zeigt, dass Citalopram, Escitalopram und Sertralin unabhängig vom CYP2D6-Polymorphismus dosiert werden können, wohingegen Standarddosen von Fluoxetin, Fluvoxamin und Paroxetin bei PM zu hohe Spiegel hervorrufen könnten. Bei CYP2C19 ist vor allem Citalopram und Escitalopram vom Polymorphismus betroffen, die anderen SSRI nicht. Dies dürfte vor allem für die Anwendung bei fernöstlichen Völkern (wie Chinesen und Japanern) von Bedeutung sein, da dort die PMs von CYP2C19 ca. 20% der Bevölkerung ausmachen, bei der mitteleuropäischen Bevölkerung sind es nur 3%.

## **Klinisch wichtige Interaktionen der SSRI aufgrund einer Konkurrenz an metabolisierenden Enzymen**

Zwei Medikamente, die über dasselbe Enzym metabolisiert werden, können sich dabei gegenseitig hemmen. Dasjenige Substrat, das eine affinere Bindung an das Enzym eingeht, wird das kompetierende Substrat stärker hemmen, als es selbst durch das andere Substrat gehemmt wird. Daher kann das Ausmaß einer kompetitiven Hemmung nicht ohne weiteres vorhergesagt werden und muß experimentell in vitro oder am Patienten ermittelt werden. Ferner gibt es Beispiele, dass ein Medikament an ein Enzym durchaus hochaffin bindet und es hemmt, dabei aber selbst nicht metabolisiert wird.

Tab. 3 gibt eine graduierte Abschätzung der Hemmwirkung spezifischer CYP-Enzyme durch die sechs SSRI wieder. Man erkennt, dass Escitalopram, Citalopram und in abgeschwächtem Maße auch Sertralin keine oder nur eine geringe Hemmwirkung auf die aufgeführten CYP-Enzyme ausüben. Demgegenüber zeichnen sich Fluoxetin, Fluvoxamin und Paroxetin durch ein hohes Interaktionspotential aus. Im Falle von Fluvoxamin erstreckt sich die inhibitorische Wirkung auf alle fünf CYP-Enzyme. Hierin liegt ein gravierendes Unterscheidungskriterium zwischen den älteren und neueren SSRI-Präparaten. Es liegen hierzu eine große Zahl von klinischen Studien vor, von denen einige im Folgenden zitiert werden, und die auch in zahlreichen Reviews bewertet wurden.



**Tab. 3** Einfluß von SSRI auf die Aktivität verschiedener Cytochrom-450-Enzyme (Spina et al. 2008)

Substanz	Hemmung von				
	CYP1A2	CYP2C9	CYP2C19	CYP3A4	CYP2D6
Citalopram	(+)		(+)	(+)	+
Escitalopram					+
Fluoxetin	+	++	+(+)	+(+)	+++
Fluvoxamin	+++	++	+++	++	+
Paroxetin	+	+	+	+	+++
Sertralin	+	+	+	+	+(+)

### Citalopram

Dieser SSRI wird in der vorliegenden vergleichenden Literatur stets als nur schwach mit anderen Pharmaka interagierend eingestuft (Spina et al., 2008; Hall et al., 2003; Kirchheiner et al., 2001; 2004; Jeppesen et al., 1996; Hiemke & Härtter, 2000). Dennoch finden sich einzelne Berichte über signifikante Interaktionen. In einer Studie mit gesunden Probanden wurden unter Gabe von Citalopram relevante Plasmaspiegelanstiege von Desipramin (Gram et al., 1993) beobachtet. Dieser Effekt dürfte durch Hemmung von CYP2D6 zustande kommen. Im Lichte der neueren Literatur könnte auch eine Hemmung des Auswärtstransporters P-Glykoprotein (MDR1; ABCB1) dafür verantwortlich sein (Weiss et al., 2003). Jedoch ist die Hemmwirkung von Citalopram nur schwach im Vergleich zu den anderen SSRI.

Hervorzuheben sei, dass Citalopram offensichtlich keine Wirkverminderung der Tamoxifentherapie bei Brustkrebs hervorruft (Lash et al., 2008). Dies bedeutet, dass offensichtlich Citalopram den letzten Schritt zur Bildung des aktiven Tamoxifenmetaboliten Endoxifen nicht relevant inhibiert. Dieser Schritt verläuft über CYP2D6.

### Escitalopram

Ähnlich wie bei Citalopram ist auch bei dessen S-Enantiomer nur mit wenigen Interaktionen zu rechnen, gemessen an den drei älteren SSRI (Fluvoxamin, Fluoxetin, Paroxetin). Auch bei Escitalopram liegen dennoch einige Interaktionsstudien vor. Escitalopram führte bei gesunden Probanden zu Plasmaspiegelerhöhungen von Desipramin (Rao, 2007) und Metoprolol (Preskorn et al., 2007). Ebenfalls in Studien mit gesunden Probanden zeigte sich, dass Gabe von Cimetidin oder Omeprazol, beides Inhibitoren von CYP2C19, zu signifikanten Erhöhungen der Plasmaspiegel von Escitalopram führen (Malling et al., 2005), was durch die wichtige Rolle von CYP2C19 im Escitaloprammetabolismus zu erklären ist. Die bei

Citalopram erwähnte Untersuchung von Lash et al. (2008) schließt auch Escitalopram ein, das somit ebenfalls die Tamoxifenaktivierung über CYP2D6 nicht signifikant hemmt.

## **Fluoxetin**

In der vergleichenden Literatur (s.o.) wird Fluoxetin neben Fluvoxamin und Paroxetin den stark interagierenden SSRI zugewiesen. Vor allem sind Pharmaka, die über CYP2D6 und CYP2C9 metabolisiert werden, betroffen.

Nach Verabreichung von Fluoxetin zusammen mit tricyclischen Antidepressiva wurden bei gesunden Freiwilligen TCA-Plasmaspiegelanstiege auf das 2-4fache gemessen, die mit Symptomen einer Überdosierung einhergingen. Ursache für die Plasmaspiegelanstiege ist die hemmende Wirkung von Fluoxetin und Norfluoxetin auf CYP2D6 (Aranow et al., 1989; Westermeyer, 1991; Bergstrom et al., 1992; Preskorn et al., 1994).

Andere Antipsychotika, bei denen klinisch relevante Plasmaspiegelanstiege beobachtet wurden, sind Haloperidol (Avenoso et al., 1997), Clozapin (Centorrino et al., 1994; 1996; Spina et al., 1998) und Risperidon (Spina et al., 2002; Bondolfi et al., 2002). Auch hier ist die Erklärung die Hemmung von CYP2D6, ggf. zusätzlich die von CYP3A4 und CYP2C19.

Bei Gabe zusammen Citalopram stiegen die Plasmaspiegel von R- und S-Citalopram um 50-100% an (Bondolfi et al., 2000).

Fluoxetin in höheren Dosen scheint den Abbau von Carbamazepin so zu beeinträchtigen, dass toxische Plasmakonzentrationen mit entsprechenden klinischen Symptomen beobachtet wurden (Pearson, 1990; Grimsley et al., 1991). Dies wird auf die Hemmung von CYP3A4 zurückgeführt. Eine Anhebung der Plasmaspiegel in supratherapeutische Bereiche wurde auch für Phenytoin beobachtet (Jalil, 1992; Woods et al., 1994); hier ist die Erklärung die Hemmung des Phenytoinmetabolismus über CYP2C9.

Gleichzeitige Gabe von Fluoxetin und Warfarin kann durch Hemmung der Verstoffwechselung des Warfarin-S-Enantiomers über CYP2C9 zu einem unerwartet starken Effekt auf die Blutgerinnung und zur Blutungsgefahr führen (Woolfrey et al., 1993; Dent & Orrock, 1997; Duncan et al., 1998).

β-Blocker werden teilweise ebenfalls über CYP2D6 verstoffwechselt. Fälle von Bradykardie bei Koadministration von Fluoxetin und Metoprolol (Walley et al., 1993) bzw. von schweren Überleitungsstörungen bei Gabe zusammen mit Propranolol (Drake et al., 1994) wurden berichtet.

## **Fluvoxamin**

Fluvoxamin sticht durch seine Hemmwirkung auf CYP1A2, CYP2C9, CYP2C19 und CYP3A4 hervor. Im Vergleich dazu wird CYP2D6 weniger stark gehemmt, jedoch gibt es auch hierzu zahlreiche Interaktionsberichte. Die breite Hemmwirkung auf mehrere CYP-Enzyme hat zur Folge, dass alternative Metabolisierungswege verschlossen sind bei solchen Pharmaka, die über mehrere CYP-Enzyme metabolisiert werden.

Fluvoxamin wirkt sich durch Hemmung von CYP2C19, CYP1A2 und CYP3A4 negativ auf Demethylierungsreaktionen beim Abbau von TCAs aus. Entsprechend wurden Plasmaspiegelerhöhungen in den toxischen Bereich für Amitriptylin, Imipramin, Clomipramin und Mirtazapin im Rahmen klinischer Studien beschrieben (Bertschy et al., 1991; Spina et al., 1993; Härter et al., 1993; Anttila et al., 2001).

Auch die Plasmaspiegel der Psychopharmaka Haloperidol (Daniel et al., 1994), Clozapin (Hiemke et al., 1992; Jerling et al., 1994; Szegedi et al., 1999; Fabrazzo et al., 2000; Shader & Greenblatt, 1998; Olesen & Linnet, 2000), Olanzapin (Weigmann et al., 2001; de Jong et al., 2001; Hiemke et al., 2002) und Quetiapin (Castberg et al., 2007) steigen bei gleichzeitiger Gabe von Fluvoxamin in klinisch relevantem Ausmaß an.

Gleichzeitige Medikation mit Citalopram und Fluvoxamin führte zu einer Plasmaspiegelerhöhung von S-Citalopram auf das ca. Zweieinhalbfache und von R-Citalopram auf das ca. Doppelte (Bondolfi et al., 1996).

Bei Patienten unter Methadonsubstitution erhöht Fluvoxamin die Methadonplasmaspiegel um 30-50% (Eap et al., 1997).

Fluvoxamin kann auch den Abbau von Theophyllin über CYP1A2 beeinträchtigen und zu Intoxikationszeichen führen (Sperber, 1991; de Vane et al., 1997).

Auch eine klinisch bedeutende Erhöhung der Warfarinwirkung auf die Blutgerinnung durch Hemmung von CYP2C9 ist möglich (Benfield and Ward, 1986; Yap & Low, 1999).

Ein fünffacher Anstieg der Propranolol-Plasmakonzentration bei gleichzeitiger Gabe von Fluvoxamin wurde in einer Studie mit gesunden Probanden beobachtet (Benfield & Ward, 1986).

## **Paroxetin**

Paroxetin zeigt eine intensive Hemmung von CYP2D6-Reaktionen, was sich bei einer Vielzahl komedizierter Arzneien negativ bemerkbar macht.

Die hemmende Wirkung von Paroxetin auf die CYP2D6-Aktivität wirkt sich negativ auf den Abbau anderer Psychopharmaka aus, deren Biotransformation maßgeblich von

diesem Enzym abhängt. Folge sind Plasmaspiegelanstiege über den therapeutischen Bereich, die zu Nebenwirkungen bzw. toxischen Wirkungen führen. Berichte über das Auftreten klinisch relevanter Interaktionen liegen für Desipramin (Brøsen et al., 1993; Alderman et al., 1997), Perphenazin (Ozdemir et al., 1997), Risperidon (Spina et al., 2001; Saito et al., 2005) und Atomoxetin (Belle et al., 2002) vor.

Tramadol wird unter Mitwirkung von CYP2D6 durch O-Demethylierung zu seinem aktiven Metaboliten verstoffwechselt, der für die schmerzlindernde Wirkung verantwortlich ist. Diese Reaktion wird unter Paroxetin nachweisbar verlangsamt, was zu einer Wirkungsabschwächung von Tramadol führt (Laugesen et al., 2005).

Paroxetin erwies sich als starker Inhibitor der über CYP2D6 erfolgenden Aktivierung von Tamoxifen (Stearns et al., 2003). Jin et al. (2005) konnten dies bestätigen. Goetz et al. (2007) zeigten, dass bei tamoxifenbehandelten Brustkrebspatientinnen niedrigere Endoxifenspiegel mit einer erheblich verschlechterten Prognose einhergehen.

### **Sertralin**

Dieser SSRI gehört nach Einschätzung der oben genannten Literatur zu den schwächer interagierenden neueren Antidepressiva. Es werden alle in Tab. 3 aufgeführten CYP-Enzyme schwach gehemmt, darunter CYP2D6 schwach bis mäßiggradig. Bei der Interpretation der Interaktionsergebnisse müsste darüber hinaus auch geprüft werden, ob nicht ein beobachteter Effekt durch Hemmung von P-Glykoprotein durch Sertralin zustande kommt. In einem der Testmodelle von Weiss et al. (2003) wies Sertralin eine besonders starke Hemmwirkung auf.

Sertralin in hoher therapeutischer Dosierung führt über die dosisabhängige Hemmung der CYP2D6-Aktivität zu klinisch relevanten Plasmaspiegelerhöhungen von Desipramin (Kurtz et al., 1997), Nortriptylin (Solai et al., 1997) und Risperidon (Spina et al., 2004).

Bei Patienten mit Epilepsie, die Lamotrigin erhielten, sind Überdosierungserscheinungen beobachtet worden, sobald Sertralin dazugegeben wurde. Man vermutet, dass die beiden Stoffe um die Glucuronidierung konkurrieren (Kaufman & Gerner, 1998).

Komedikation mit Sertralin führte zu um ca. 40% verminderten Endoxifenspiegeln bei tamoxifenbehandelten Brustkrebspatientinnen. Die Hemmung war nur halb so groß wie bei Paroxetin (Jin et al., 2005), dürfte aber die Tamoxifenwirkung deutlich einschränken.

## Zusammenfassung wichtiger Differenzen der SSRI

Anhand der angewandten Kriterien Pharmakokinetik, Metabolismus, Pharmakogenetik und Interaktionen zeigen sich teils markante Unterschiede zwischen den einzelnen SSRI.

Pharmakokinetik: Fluoxetin und insbesondere sein wirksamer Metabolit Norfluoxetin weisen eine besonders lange Halbwertszeit (Norfluoxetin 1-2 Wochen) auf, entsprechend dauert die Einstellung eines Steady-State über einen Monat. Dieses Medikament ist beim Auftreten von Nebenwirkungen oder Interaktionen ausgesprochen schlecht steuerbar. Die übrigen SSRI differieren bezüglich Pharmakokinetik nicht signifikant voneinander.

Metabolismus: Alle SSRI werden durch CYP-Enzyme eliminiert. CYP2D6 ist bei allen SSRI involviert, darüber hinaus sind aber stets noch 1-3 weitere CYP-Enzyme beteiligt. Citalopram und Escitalopram werden vorwiegend über CYP2C19 verstoffwechselt. Bei Fluvoxamin steht die Metabolisierung über CYP1A2 im Vordergrund. Die Art der Stoffwechselwege kann per se nicht als Vor- oder Nachteil gewertet werden.

Pharmakogenetik: Die großen interindividuellen pharmakogenetisch bedingten Aktivitätsunterschiede zwischen fehlender Aktivität (Defizienz) und deutlich überhöhter Aktivität bei CYP2D6 und CYP2C19 müssen durch Dosisanpassung kompensiert werden. Wenn dies unterbleibt, droht eine Überdosierung mit toxischen Effekten oder eine Unterdosierung mit fehlender Wirkung. Dementsprechend ist bei Fluoxetin, Fluvoxamin und Paroxetin eine Dosisanpassung gemäß dem CYP2D6-Status vorzunehmen. Bei Citalopram, Escitalopram und Sertralin ist dies nicht erforderlich. Die drei neueren SSRI (Citalopram, Escitalopram, Sertralin) bedürfen bezüglich des CYP2C19-Polymorphismus, wenn überhaupt, nur einer leichten Dosisanpassung.

Interaktionen: Hier ergibt sich eine deutliche Zweiteilung der SSRI-Gruppe in die älteren Substanzen (Fluoxetin, Fluvoxamin, Paroxetin), die mit einem großen Interaktionspotential belastet sind, und die Gruppe der neueren (Citalopram, Escitalopram, Sertralin), die wesentlich schwächer mit anderen Pharmaka interagieren.

Die Nachteile, die sich hier aus Pharmakogenetik oder Arzneimittelinteraktionen ergeben, lassen sich schwer durch klinische Studien statistisch belegen, da die betreffenden besonderen Merkmale im untersuchten Kollektiv selten sind. Daraus darf man aber nicht schließen, dass sie für den einzelnen Patienten ohne Relevanz wären.

## **Resümee**

Die hier angewandten Kriterien belegen relevante klinisch-pharmakologische Unterschiede zwischen den einzelnen SSRI. Diese sind daher beim einzelnen Patienten nur sehr bedingt austauschbar. Die Kriterien sind Patienten-relevant und daher bei der Nutzenbewertung des IQWiG zu berücksichtigen, indem die Gruppe der SSRIs differenziert dargestellt wird.

Berlin, den 7. Januar 2009



---

Prof. Dr. I. Roots

## Literatur

- Alderman J, Preskorn SH, Greenblatt DJ *et al.*: Desipramine pharmacokinetics when coadministered with paroxetine or sertraline in extensive metabolizers. *J Clin Psychopharmacol* 1997; 17: 284–291.
- Anttila AK, Rasanen L, Leinonen EV: Fluvoxamine augmentation increases serum mirtazapine concentrations three- to fourfold. *Ann Pharmacother* 2001; 35: 1221–1223.
- Aranow AB, JI Hudson JI, Pope Jr HG *et al.*: Elevated antidepressant plasma levels after addition of fluoxetine. *Am J Psychiatry* 1989; 146: 911–913.
- Avenoso A, Spinà E, Campo G *et al.*: Interaction between fluoxetine and haloperidol: Pharmacokinetic and clinical implications. *Pharmacol Res* 1997; 35: 335–339.
- Belle DJ, Ernest CS, Sauer JM *et al.*: Effect of potent CYP2D6 inhibition by paroxetine on atomoxetine pharmacokinetics. *J Clin Pharmacol* 2002; 42: 1219–1227.
- Benfield P and Ward A: Fluvoxamine: A review of its pharmacodynamic and pharmacokinetic properties, and therapeutic efficacy in depressive illness. *Drugs* 1986; 32: 313–334.
- Bergstrom RF, Peyton AL, Lemberger L: Quantification and mechanism of the fluoxetine and tricyclic antidepressant interaction. *Clin Pharmacol Ther* 1992; 51: 239–248.
- Bertschy G, Vandel S, Vandel B *et al.*: Fluvoxamine-tricyclic antidepressant interaction. An accidental finding. *Eur J Clin Pharmacol* 1991; 40: 119–120.
- Bondolfi G, Chautems C, Rochat B *et al.*: Non-response to citalopram in depressive patients: Pharmacokinetic and clinical consequences of a fluvoxamine augmentation. *Psychopharmacology* 1996; 128: 421–425.
- Bondolfi G, Eap CB, Bertschy G *et al.*: The effect of fluoxetine on the pharmacokinetics and safety of risperidone in psychotic patients. *Pharmacopsychiatry* 2002; 35: 50–56.
- Bondolfi G, Lissner C, Kosel M *et al.*: Fluoxetine augmentation in citalopram non-responders: Pharmacokinetic and clinical consequences. *Int J Neuropsychopharmacol* 2000; 3: 55–60.
- Bourin M, Chue P, Guillon Y: Paroxetine: a review. *CNS Drug Rev.* 2001;7:25-47.
- Brøsen K, Hansen JG, Nielsen KK *et al.*: Inhibition by paroxetine of desipramine metabolism in extensive but not in poor metabolizers of sparteine. *Eur J Clin Pharmacol* 1993; 44: 349–355.
- Brøsen K, Naranjo CA: Review of pharmacokinetic and pharmacodynamic interaction studies with citalopram. *Eur Neuropsychopharmacol.* 2001; 11: 275-83.
- Carrillo JA, Dahl ML, Svensson JO, Alm C, Rodriguez I, Bertilsson L. Disposition of fluvoxamine in humans is determined by the polymorphic CYP2D6 and also by the CYP1A2 activity. *Clin Pharmacol Ther* 1996; 60: 183–190.
- Castberg I, Skogvoll E, Spigset O: Quetiapine and drug interactions: Evidence from a routine therapeutic drug monitoring service. *J Clin Psychiatry* 2007; 68: 1540–1545.
- Centorrino F, Baldessarini RJ, Frankenburg FR *et al.*: Serum levels of clozapine and norclozapine in patients treated with selective serotonin reuptake inhibitors. *Am J Psychiatry* 1996; 153: 820–822.

- Centorrino F, Baldessarini RJ, Kando J *et al.*: Serum concentrations of clozapine and its major metabolites: Effects of cotreatment with fluoxetine or valproate. *Am J Psychiatry* 1994; 151: 123–125.
- Daniel DG, Randolph C, Jaskiw G *et al.*: Coadministration of fluvoxamine increases serum concentrations of haloperidol. *J Clin Psychopharmacol* 1994; 14: 340–343.
- de Jong J, Hoogenboom B, van Troostwijk LD, de Haan L: Interaction of olanzapine with fluvoxamine. *Psychopharmacology* 2001; 155: 219–220.
- Dent LA, Orrock MW: Warfarin-fluoxetine and diazepam-fluoxetine interaction. *Pharmacotherapy* 1997; 17: 170–172.
- DeVane CL: Pharmacokinetics, drug interactions, and tolerability of paroxetine and paroxetine CR. *Psychopharmacol Bull.* 2003; 37 Suppl 1: 29-41.
- DeVane CL, Liston HL, Markowitz JS: Clinical pharmacokinetics of sertraline. *Clin Pharmacokinet.* 2002; 41: 1247-66.
- DeVane CL, Markowitz JS, Hardesty SJ *et al.*: Fluvoxamine-induced theophylline toxicity. *Am J Psychiatry* 1997; 154: 1317–1318.
- Drake WM, Gordon GD: Heart block in a patient on propranolol and fluoxetine. *Lancet* 1994; 343: 425–426.
- Duncan D, Sayal K, McConnell H, Taylor D: Antidepressant interactions with warfarin. *Int Clin Psychopharmacol* 1998; 13: 87–94.
- Eap CB, Bertschy G, Powell K, Baumann P: Fluvoxamine and fluoxetine do not interact in the same way with the metabolism of the enantiomers of methadone. *J Clin Psychopharmacol* 1997; 17: 113–117.
- El-Armouche A, Zolk O, Eschenhagen T: Citalopram. *Dtsch Med Wochenschr.* 2003; 128: 2253-6.
- Fabrazzo M, La Pia S, Monteleone P *et al.*: Fluvoxamine increases plasma and urinary levels of clozapine and its major metabolites in a time- and dose-dependent manner. *J Clin Psychopharmacol* 2000; 20: 708–710.
- Fjordside L, Jeppesen U, Eap CB, Powell K, Baumann P, Brøsen K: The stereoselective metabolism of fluoxetine in poor and extensive metabolizers of sparteine. *Pharmacogenetics* 1999; 9: 55–60.
- Goetz MP, Knox SK, Suman VJ, Rae JM, Safgren SL, Ames MM, Visscher DW, Reynolds C, Couch FJ, Lingle WL, Weinshilboum RM, Fritcher EG, Nibbe AM, Desta Z, Nguyen A, Flockhart DA, Perez EA, Ingle JN: The impact of cytochrome P450 2D6 metabolism in women receiving adjuvant tamoxifen. *Breast Cancer Res Treat.* 2007; 101: 113-21.
- Gram LF, Hansen MG, Sindrup SH *et al.*: Citalopram: Interaction studies with levomepromazine, imipramine, and lithium. *Ther Drug Monit* 1993; 15: 18–24.
- Grimsley SR, Jann MW, Carter JG *et al.*: Increased carbamazepine plasma concentrations after fluoxetine coadministration. *Clin Pharmacol Ther* 1991; 50: 10–15.
- Hall J, Naranjo CA, Sproule BA, Herrmann N. Pharmacokinetic and pharmacodynamic evaluation of the inhibition of alprazolam by citalopram and fluoxetine. *J Clin Psychopharmacol.* 2003; 23: 349-57.
- Hamelin BA, Turgeon J, Vallee F, Belanger PM, Paquet F, LeBel M: The disposition of fluoxetine but not sertraline is altered in poor metabolizers of debrisoquin. *Clin Pharmacol Ther* 1996; 60: 512–521.



- Härtter S, Wetzel H, Hammes E, Hiemke C: Inhibition of antidepressant demethylation and hydroxylation by fluvoxamine in depressed patients. *Psychopharmacology* 1993; 110: 303–308.
- Hiemke C, Härtter S: Pharmacokinetics of selective serotonin reuptake inhibitors. *Pharmacol Ther.* 2000; 85:11-28.
- Hiemke C, Peled A, Jabarin M *et al.*: Fluvoxamine augmentation of olanzapine in chronic schizophrenia: Pharmacokinetic interactions and clinical effects. *J Clin Psychopharmacol* 2002; 22: 502–506.
- Hiemke C, Weigmann H, Härtter S *et al.*: Elevated levels of clozapine in serum after addition of fluvoxamine. *J Clin Psychopharmacol* 1994; 14: 279–281.
- Jalil P: Toxic reaction following the combined administration of fluoxetine and phenytoin: Two case reports. *J Neurol Neurosurg Psychiatry* 1992; 55: 412–413.
- Jan MW, ZumBrunnen TL, Kazmi YR, VanDenBerg CM, Desai HD, Weidler DJ *et al.*: Pharmacokinetics of fluvoxamine in relation to CYP2C19 phenotype and genotype. *Drug Metabol Drug Interact* 2002; 19: 1–11.
- Jeppesen U, Gram LF, Vistisen K, Loft S, Poulsen HE, Brøsen K: Dose-dependent inhibition of CYP1A2, CYP2C19 and CYP2D6 by citalopram, fluoxetine, fluvoxamine and paroxetine. *Eur J Clin Pharmacol.* 1996; 51: 73-8.
- Jerling M, Lindström L, Bondesson U, Bertilsson L: Fluvoxamine inhibition and carbamazepine induction of the metabolism of clozapine: Evidence from a therapeutic drug monitoring service. *Ther Drug Monit* 1994; 16: 368–374.
- Jin Y, Desta Z, Stearns V, Ward B, Ho H, Lee KH, Skaar T, Storniolo AM, Li L, Araba A, Blanchard R, Nguyen A, Ullmer L, Hayden J, Lemler S, Weinshilboum RM, Rae JM, Hayes DF, Flockhart DA: CYP2D6 genotype, antidepressant use, and tamoxifen metabolism during adjuvant breast cancer treatment. *J Natl Cancer Inst.* 2005; 97: 30-9.
- Kaufman KR and Gerner R: Lamotrigine toxicity secondary to sertraline. *Seizure* 1998; 7: 163–165.
- Kirchheiner J, Brøsen K, Dahl ML, Gram L, Kasper S, Roots I, Sjöqvist F, Spina E, Brockmüller J: CYP2D6 and CYP2C19 genotype-based dose recommendations for antidepressants: A first step towards subpopulation specific dosages. *Acta Psych Scand.* 2001; 104: 173-192.
- Kirchheiner J, Nickchen K, Bauer M, Wong ML, Licinio J, Roots I, Brockmüller J: Pharmacogenetics of antidepressants and antipsychotics: the contribution of allelic variations to the phenotype of drug response. *Mol Psychiatry* 2004; 9: 442-473
- Klein N, Sacher J, Geiss-Granadia T, Mossaheb N, Attarbaschi T, Lanzenberger R, Spindelegger C, Holik A, Asenbaum S, Dudczak R, Tauscher J, Kasper S: Higher serotonin transporter occupancy after multiple dose administration of escitalopram compared to citalopram: an [123I]ADAM SPECT study. *Psychopharmacology (Berl).* 2007; 191: 333-9.
- Kurtz DL, Bergstrom RF, Goldberg MJ, Cerimele BJ: The effect of sertraline on the pharmacokinetics of desipramine and imipramine. *Clin Pharmacol Ther* 1997; 62: 145–156.
- Lash TL, Pedersen L, Cronin-Fenton D, Ahern TP, Rosenberg CL, Lunetta KL, Silliman RA, Hamilton-Dutoit S, Garne JP, Ewertz M, Sørensen HT: Tamoxifen's protection against breast cancer recurrence is not reduced by concurrent use of the SSRI citalopram. *Br J Cancer.* 2008; 99: 616-21.

- Laugesen S, Enggaard TP, Pedersen RS *et al.*: Paroxetine, a cytochrome P450 2D6 inhibitor, diminishes the stereoselective O-demethylation and reduces the hypoalgesic effect of tramadol. *Clin Pharmacol Ther* 2005; 77: 312–323.
- Liu ZQ, Cheng ZN, Huang SL, Chen XP, Ou-Yang DS, Jiang CH *et al.*: Effect of the CYP2C19 oxidation polymorphism on fluoxetine metabolism in Chinese healthy subjects. *Br J Clin Pharmacol* 2001; 52: 96–99.
- Malling D, Poulsen MN, Sjøgaard B: The effect of cimetidine or omeprazole on the pharmacokinetics of escitalopram in healthy subjects. *Br J Clin Pharmacol* 2005; 60: 287–290.
- Olesen OV & Linnet K: Fluvoxamine-clozapine drug interaction: Inhibition in vitro of five cytochrome P450 isoforms involved in clozapine metabolism. *J Clin Psychopharmacol* 2000; 20: 35–42.
- Olesen O & Linnet K: Studies on the stereoselective metabolism of citalopram by human liver microsomes and cDNA-expressed cytochrome P450 enzymes. *Pharmacology* 1999;59:298–309.
- Ozdemir V, Naranjo CA, Herrmann N *et al.*: Paroxetine potentiates the central nervous system side effects of perphenazine: Contribution of cytochrome P4502D6 inhibition in vivo. *Clin Pharmacol Ther* 1997; 62: 334–347.
- Pearson HJ: Interaction of fluoxetine with carbamazepine. *J Clin Psychiatry* 1990; 51: 126.
- Preskorn SH, Alderman J, Chung M *et al.*: Pharmacokinetics of desipramine coadministered with sertraline or fluoxetine. *J Clin Psychopharmacol* 1994; 14: 90–98.
- Preskorn SH, Greenblatt DJ, Flockhart D *et al.*: Comparison of duloxetine, escitalopram, and sertraline effects on cytochrome P450 2D6 function in healthy volunteers. *J Clin Psychopharmacol* 2007; 27: 28–34.
- Rao N: The clinical pharmacokinetics of escitalopram, *Clin Pharmacokinet* 2007; 46: 281–290.
- Rudberg I, Mohebi B, Hermann M, Refsum H, Molden E: Impact of the ultrarapid CYP2C19\*17 allele on serum concentration of escitalopram in psychiatric patients. *Clin Pharmacol Ther.* 2008; 83: 322-7.
- Saito M, Yasui-Furukori N, Nakagami T *et al.*: Dose-dependent interaction of paroxetine with risperidone in schizophrenic patients. *J Clin Psychopharmacol* 2005; 25: 527–532.
- Sánchez C, Bøgesø KP, Ebert B, Reines EH, Braestrup C: Escitalopram versus citalopram: die überraschende Rolle des R-Enantiomers. *Psychopharmacology (Berl)*. 2004; 174: 163-76.
- Sanchez C & Hyttel J: Comparison of the effects of anti-depressants and their metabolites on reuptake of biogenic amines and on receptor binding. *Cell Mol Neurobiol* 1999; 19: 467–489.
- Shader RI and Greenblatt DJ: Clozapine and fluvoxamine, a curious complexity. *J Clin Psychopharmacol* 1998; 18: 101–102.
- Sindrup SH, Brøsen K, Gram LF, Hallas J, Skjelbo E, Allen A *et al.*: The relationship between paroxetine and the sparteine oxidation polymorphism. *Clin Pharmacol Ther* 1992; 51: 278–287.
- Sindrup SH, Brøsen K, Hansen MG, Aaes Jorgensen T, Overo KF, Gram LF: Pharmacokinetics of citalopram in relation to the sparteine and the mephenytoin oxidation polymorphisms. *Ther Drug Monit* 1993; 15: 11–17.

- Søgaard B, Mengel H, Rao N, Larsen F: The pharmacokinetics of escitalopram after oral and intravenous administration of single and multiple doses to healthy subjects. *J Clin Pharmacol*. 2005; 45: 1400-6.
- Solai LK, Mulsant BH, Pollock BG *et al.*: Effect of sertraline on plasma nortriptyline levels in depressed elderly. *J Clin Psychiatry* 1997; 58: 440–443.
- Sperber AD: Toxic interaction between fluvoxamine and sustained release theophylline in an 11-year-old boy. *Drug Safety* 1991; 6: 460–462.
- Spigset O, Granberg K, Hagg S, Norstrom A, Dahlqvist R: Relationship between fluvoxamine pharmacokinetics and CYP2D6/CYP2C19 phenotype polymorphisms. *Eur J Clin Pharmacol* 1997; 52: 129–133.
- Spina E, Avenoso A, Facciola G *et al.*: Effect of fluoxetine on the plasma concentrations of clozapine and its major metabolites in patients with schizophrenia. *Int Clin Psychopharmacol* 1998; 13: 141–145.
- Spina E, Avenoso A, Facciola G *et al.*: Plasma concentrations of risperidone and 9-hydroxyrisperidone during combined treatment with paroxetine. *Ther Drug Monit* 2001; 23: 223–227.
- Spina E, Avenoso A, Scordo MG *et al.*: Inhibition of risperidone metabolism by fluoxetine in patients with schizophrenia: A clinically relevant pharmacokinetic drug interaction. *J Clin Psychopharmacol* 2002; 22: 419–423.
- Spina E, D'Arrigo C, Migliardi G *et al.*: Plasma risperidone concentrations during combined treatment with sertraline. *Ther Drug Monit* 2004; 26: 386–390.
- Spina E, Pollicino AM, Avenoso A *et al.*: Effect of fluvoxamine on the pharmacokinetics of imipramine and desipramine in healthy subjects. *Ther Drug Monit* 1993; 15: 243–246.
- Spina E, Santoro V, D'Arrigo C: Clinically relevant pharmacokinetic drug interactions with second-generation antidepressants: an update. *Clin Ther*. 2008;30:1206-27.
- Stahl SM, Gergel I, Li D: Escitalopram in the treatment of panic disorder: a randomized, double-blind, placebo-controlled trial. *J Clin Psychiatry*. 2003; 64: 1322-7.
- Stearns V, Johnson MD, Rae JM, Morocho A, Novielli A, Bhargava P, Hayes DF, Desta Z, Flockhart DA: Active tamoxifen metabolite plasma concentrations after coadministration of tamoxifen and the selective serotonin reuptake inhibitor paroxetine. *J Natl Cancer Inst*. 2003; 95:1758-64.
- Szegedi A, Anghelescu I, Wiesner J *et al.*: Addition of low-dose of fluvoxamine to low-dose clozapine monotherapy in schizophrenia: Drug monitoring and tolerability data from a prospective clinical trial. *Pharmacopsychiatry* 1999; 32: 148–153.
- Walley T, Pirmohamed M, Proudlove C, Maxwell D: Interaction of metoprolol and fluoxetine. *Lancet* 1993; 341: 967–968.
- Wang JH, Liu ZQ, Wang W, Chen XP, Shu Y, He N *et al.*: Pharmacokinetics of sertraline in relation to genetic polymorphism of CYP2C19. *Clin Pharmacol Ther* 2001; 70: 42–47.
- Weigmann H, Gerek S, Zeisig A *et al.*: Fluvoxamine but not sertraline inhibits the metabolism of olanzapine: Evidence from a therapeutic drug monitoring service. *Ther Drug Monit* 2001; 23: 410–413.

Weiss J, Dormann SM, Martin-Facklam M, Kerpen CJ, Ketabi-Kiyanvash N, Haefeli WE: Inhibition of P-glycoprotein by newer antidepressants. *J Pharmacol Exp Ther.* 2003; 305:197-204.

Westermeyer J. Fluoxetine-induced tricyclic toxicity: Extent and duration. *J Clin Pharmacol* 1991; 31: 388–392.

Woods DJ, Coulter DM, Pillans P: Interaction of phenytoin and fluoxetine [2]. *NZ Med J* 1994; 107: 19.

Woolfrey S, Gammack NS, Dewar MS, Brown PJ: Fluoxetine-warfarin interaction. *BMJ* 1993; 307: 241.

Yap KB and Low ST: Interaction of fluvoxamine with warfarin in an elderly woman. *Singapore Med J* 1999; 40: 480–482.

**A 2.2 Volz, Hans-Peter, Prof. Dr. med.**

**Adresse:**

Prof. Dr. med. Hans-Peter Volz

Krankenhaus für Psychiatrie, Psychotherapie und Psychosomatische Medizin

Schloss Werneck

Balthasar-Neumann-Platz 1

97440 Werneck

## **Einleitende Bemerkung**

Als psychopharmakologisch interessierter Kliniker werde ich mich in der folgenden Stellungnahme in erster Linie auf meiner Ansicht nach klinisch wichtige Aspekte des Vorberichts zu den SNRIs konzentrieren.

### **4.4 Informationssynthese und –analyse**

Im Bericht wird hier sinngemäß ausgeführt, dass für die Einschätzung der Relevanz (= Spürbarkeit des Effekts für den Patienten oder behandelnden Arzt) eines Gruppenunterschiedes bezüglich stetiger Skalen die „Minimal important Differences“ (MID) für einen Gruppenunterschied berücksichtigt werden sollten. Da für keine der eingesetzten Skalen eine verwendbare MID vorlag, wurde eine Berechnung von Cohen`s d durchgeführt, als Relevanzgrenze wurde 0,2 angenommen, da nach einer Interpretation von Cohen`s d eine Effektgröße von 0,2 bis 0,5 Standardabweichungen als kleiner Effekt angesehen wird. Bei der Interpretation der Effektstärken wurde das Konfidenzintervall zu dieser Relevanzgröße in Beziehung gesetzt. Lag das Konfidenzintervall vollständig oberhalb der Relevanzgrenze wurde von einer relevanten Effektgröße ausgegangen und ein Nutzenbeleg oder- hinweis attestiert, lag es teilweise oder komplett unterhalb dieser Grenze, blieb unklar, ob die Effektstärke eine so relevante Größenordnung erreichte, dass von einem Nutzen gesprochen werden konnte.

Hier muss bemängelt werden, dass das IQWiG hinter seinen eigenen Zielvorgaben zurückbleibt. Angestrebt war, einen für den Patienten oder Arzt spürbaren Effekt zu definieren, z.B. auch umschrieben mit „klinischer Relevanz“. Ein artifizielles Maß wie das Miteinander-In-Beziehung-Setzen zweier statistischer Maße, um dann (hilfsweise) zu beurteilen, ob es sich um einen (für den Patienten und/oder Arzt) spürbaren Effekt handelt, kann in diesem Zusammenhang nicht befriedigen.

Es bleibt unklar, wieso nicht z.B. auf number-needed-to-treat (NNT) zurückgegriffen wurde, als ein anschaulicher und von großer klinischer Wichtigkeit geprägter Parameter.

#### **5.2.3.6 – Duloxetin, Alter**

Dort wird abschließend Folgendes angeführt: „Es gab keinen Hinweis für eine maßgebliche Abweichung in Bezug auf die antidepressive Wirkung bei älteren Patienten.“

Mir ist die Bedeutung dieses Satzes nicht klar. Heißt dies, dass Duloxetin bei älteren Patienten ebenso effektiv war wie bei normal-alten Patienten? Wenn dies gemeint ist, handelte es sich ja um einen klaren Vorteil der Substanz, da die Wirksamkeit von Antidepressiva bei älteren Patienten z.T. als abgeschwächt betrachtet wird, was ja mutmaßlich auch der Grund ist, warum hier Subanalysen durchgeführt wurden.

Gerade in diesem Punkt ist die Datenlage zu Duloxetin sehr gut, wie die Studien von Raskin et al. (2007) und Wise et al. (2007) demonstrieren.

#### **5.2.5 Zusammenfassung der Ergebnisse Duloxetin – Remission und Response**

Hier stellt sich, ähnlich wie unter 5.3.2 für Venlafaxin aufgeführt, die Frage nach den in die entsprechenden Meta-Analysen eingegangenen Studien, da auch Meta-Analysen existieren, die eine Überlegenheit von Duloxetin über die SSRI zeigen:

- Die Meta-Analyse von Cookson et al. (2006) zeigte eine Überlegenheit von Duloxetin sowohl gegenüber Placebo als auch (tendenziell) gegenüber den SSRI, auch unter Verwendung des relevanten Merkmals NNT.
- Mallinckrodt et al. (2007) zeigten eine Überlegenheit in der Gesamtwirksamkeit von Duloxetin versus SSRIs (vor allem verursacht durch einen Wirksamkeitsunterschied bei den leichteren Depressionsformen) sowie eine überlegene Wirksamkeit auf HAMD-Einzelitems, die mit Aktivität zusammenhängen.
- Thase et al. (2007) konnten im Gesamtvergleich versus SSRIs in Bezug auf die Remissionsrate keinen Unterschied feststellen, hier zeigten sich aber Vorteile für Duloxetin in der Gruppe der schwerer erkrankten depressiven Patienten.

Zumindestens sollte in Fällen, wo die IQWiG-eigene Meta-Analyse von anderen Meta-Analysen abweicht, vom Institut diskutiert werden, warum es zu solchen diskrepanten Ergebnissen im Vergleich zu publizierten Analysen gelangt (dies gilt analog für die Ergebnisse zur Remission für den Vergleich Venlafaxin – SSRIs). Neben der mutmaßlich unterschiedlichen Auswahl der in die Meta-Analysen einbezogenen Studien könnte dies auch auf den vom IQWiG zu Grunde gelegten Mittelwertvergleich beruhen, bei der Verwendung von Rohdaten werden in der Regel validere Ergebnisse, wie in den hierzu publizierten Meta-Analysen, erzielt.

#### **5.2.5 Zusammenfassung der Ergebnisse Duloxetin – Kognition**

Dass Daten zur Kognition lediglich in einer Studie an älteren Patienten erhoben wurden, ist unrichtig. In den Studien von Raskin et al. (2007) und Wise et al. (2007) war der primäre Endpunkt ein sog. 'composite score', der u.a. auch kognitive Parameter umfasste. Insofern müssten diese Studien zumindestens hier erwähnt und gewürdigt werden, diese Studien können durchaus im Sinne eines besonderen Wertes von Duloxetin in diesem speziellen Symptombereich bei älteren depressiven Patienten gewertet werden.

#### **5.2.5 Zusammenfassung der Ergebnisse Duloxetin - Schmerz**

Hier wird zwar über einen Vorteil von Duloxetin vs. Placebo berichtet, die Relevanz dieses Effektes sei aber unklar. Eine Studie (es muss sich wohl um die Studie von Brannan et al., 2005, handeln) sei nicht berücksichtigt worden, da in dieser Studie der antidepressive Nutzen nicht hätte gezeigt werden können. In dieser Studie war der primäre Zielparameter die Schmerzsymptomatik, es wurden depressive Patienten, die durch eine hohe Schmerzintensität gekennzeichnet waren, eingeschlossen. In dieser Studie war nie primär intendiert, den antidepressiven Nutzen (Effektivität) zu zeigen, hierfür war die Studie auch nicht gepowert. Gerade diese Studie, die bestens geeignet ist, den Effekt von Duloxetin auf ausgeprägte Schmerzen im Bereich der Depression zu zeigen, wird ausgeschlossen, da in einem sekundären Zielparameter (Depressivität) kein Unterschied zu Placebo gezeigt werden konnte. Dies ist m.E. nicht nachvollziehbar.

Auch die Studie von Brecht et al. (2007) ist leider in dieser Meta-Analyse unberücksichtigt geblieben, obwohl auch diese Studie gut geeignet erscheint, den Effekt von Duloxetin auf Schmerzen zu zeigen.

### 5.3.2 Ergebnisse zu Therapiezielen - Venlafaxin

Hier verfehlte Venlafaxin im Vergleich zu den SSRI knapp die statistische Signifikanz für die Remission. In zahlreichen Meta-Analysen war aber gerade auch für die Remission ein statistisch signifikanter Vorteil von Venlafaxin vs. SSRIs gezeigt worden, auch in einer Meta-Analyse, in der ich Ko-Autor bin (Bauer et al., 2008). Am wahrscheinlichsten ist das vom IQWiG gefundene negative Ergebnis auf den Ausschluss von Studien, die in anderen Meta-Analysen zu Mindestens zum Teil berücksichtigt wurden, zurückzuführen.

Bei Durchsicht der Daten scheinen folgende Studien (im Vergleich zu der Meta-Analyse von Bauer et al., s.o.) nicht berücksichtigt worden zu sein (chronologisch geordnet):

- Poirier et al. 1999
- Ballús et al., 2000
- Zanardi et al., 2000
- De Nayer et al., 2002
- Oslin et al, 2003
- Li et al., 2004
- Unpublished, 2007

Bei der Diskussion zum Berichtsplan habe ich mehrfach darauf hingewiesen, welche Gefahren damit verbunden sind, auf Evidenzen zu verzichten. Selbst wenn das IQWiG auf seiner Studienselektion insistiert, besteht doch kein Grund, nicht auch noch eine zweite Meta-Analyse unter Einschluss auch der ausgeschlossenen Studien (s.u.) durchzuführen und dann beide Ergebnisse, die wahrscheinlich diskrepanz sind, zu diskutieren. So wie jetzt vorgegangen wurde (d.h. Ausschluss von Studien nach arbiträren, nicht allgemein verbindlichen Ausschlusskriterien) wird sich das Argument, dass so das in diesem Punkt negative Ergebnis für Venlafaxin generiert wurde, nicht ausräumen lassen.

Hier ist ein weiterer Punkt, der analog auch auf die Ergebnisse zu Duloxetine zutrifft, anzumerken: Bei den Meta-Analysen des IQWiG werden die SSRI als eine einheitliche Klasse gewertet. Allerdings gibt es deutliche Hinweise darauf, dass Escitalopram effektiver als die anderen SSRI sein könnte, zu Mindestens aber effektiver als Citalopram (Gorman et al., 2002; Lepola et al., 2004; Moore et al., 2005). Um somit den Effekt von Venlafaxin und Duloxetine gegenüber den SSRI umfassend zu beschreiben, würde sich eine zusätzliche Meta-Analyse ohne die Studien zu Escitalopram empfehlen.

### 6.4.7 Offene Forschungsfragen

Hier wird u.a. dargelegt, dass bei Duloxetine Vergleiche zu trizyklischen Antidepressiva fehlten (für Venlafaxin, das wesentlich früher entwickelt wurde, liegen solche Vergleiche vor). Es bleibt offen, was aus solchen Vergleichen für zusätzliche Schlussfolgerungen gezogen werden sollten, da die Wirksamkeit der Substanz eindeutig belegt ist und allenfalls zu erwarten wäre, dass die Verträglichkeitsnachteile der Trizyklika zum wiederholten Male gezeigt würden. Warum sollten dann Patienten im Rahmen



einer klinischen Studie dem erhöhten Risiko, das eine Trizyklika-Gabe mit sich bringt, ausgesetzt werden?

Zu beiden SNRI, also auch zu Venlafaxin, wird bemerkt, dass die Evidenz zu Langzeittherapie noch optimierbar sei. In diesem Zusammenhang sind die Maßstäbe, die hier zu Grunde gelegt werden, von Interesse. Welches Antidepressivum hat denn (sowohl qualitativ als auch quantitativ) umfänglichere Langzeitdaten als Venlafaxin? Natürlich können immer noch mehr Studien gefordert werden, allerdings sollten gewisse vernünftige Maßstäbe zugrunde gelegt werden.

Der abschließende Satz dieses Kapitels, dass für die Auswahl der Antidepressiva auf der Basis der aktuellen Datenlage die differenzielle Wirksamkeit auf Einzel- und Begleitsymptome der Depression keine ausreichende Begründung sei, ist in der allgemeinen Formulierung nicht haltbar, da lediglich die SNRI Venlafaxin und Duloxetin untersucht wurden.

In diese Wertung geht nicht ein, dass z.B. für Venlafaxin breite Studienprogramme mit positiven Ergebnisse für eine Reihe von Angsterkrankungen vorliegen, für Duloxetin für Schmerzsyndrome. Dies ist auch ein Teil eines differenziellen Wirkspektrums, das bei der Auswahl eines Antidepressivums für einen individuellen Patienten Berücksichtigung finden sollte.

## Literatur

- Ballús C, Quiros G, De Flores T, de la Torre J, Palao D, Rojo L, Gutiérrez M, Casais L, Riesgo Y. The efficacy and tolerability of venlafaxine and paroxetine in outpatients with depressive disorder or dysthymia. *Int Clin Psychopharmacol* 2005; 15: 43-48
- Bauer M, Tharmanathan P, Volz HP, Möller HJ, Freemantle N. The effect of venlafaxine compared with other antidepressants and placebo in the treatment of major depression. A meta-analysis. *Eur Arch Psychiatry Clin Neurosci* 2008, in press
- Brannan SK, Mallinckrodt CH, Brown EB, Wohlreich MM, Watkin JG, Schatzberg AF. Duloxetine 60 mg once-daily in the treatment of painful physical symptoms in patients with major depressive disorder. *J Psychiatr Res* 2005; 39: 43-53
- Brecht S, Courtecuisse C, Debieuvre C, Croenlein J, Desaiyah D, Raskin J, Petit C, Demyttenaere K. Efficacy and safety of duloxetine 60 mg once daily in the treatment of pain in patients with major depressive disorder and at least moderate pain of unknown etiology: A randomised controlled trial. *J Clin Psychiatry* 2007; 68: 1707-1716
- Cookson J, Gilaberte I, Desaiyah D, Kajdasz DK. Treatment benefits of duloxetine in major depressive disorder as assessed by number needed to treat. *Int Clin Psychopharmacol* 2006; 21: 267-273
- De Nayer A, Geerts S, Ruelens L, Schittecatte M, De Bleeker E, Van Eeckhoutte I, Evrard JL, Linkowski P, Fossion P, Leyman S, Mignon A. Venlafaxine compared with fluoxetine in outpatients with depression and concomitant anxiety. *J Neuropsychopharmacol* 2002; 5: 115-120
- Gorman JM, Korotzer A, Su G. Efficacy comparison of escitalopram and citalopram in the treatment of major depressive disorder: Pooled analysis of placebo-controlled trials. *CNS Spectrum* 2002; 7 (suppl 1): 40-44

- Lepola U, Wade A, Andersen HF. Do equivalent doses of escitalopram and citalopram have similar efficacy? A pooled analysis of two positive placebo-controlled studies in major depressive disorder. *Int Clin Psychopharmacol* 2004; 19: 149-155
- Li Y, Zhang X-N, Wu Z-M. Comparative study on the effects of paroxetine and venlafaxine in treating depression. [Chinese]. *Zhongguo Linchuang Kangfu* 2004, 8: 4174-4175
- Mallinckrodt CH, Prakash A, Houston JP, Swindle R, Detke MJ, Fava M. Differential antidepressant symptom efficacy: Placebo-controlled comparisons of duloxetine and SSRIs (fluoxetine, paroxetine, escitalopram). *Neuropsychobiology* 2007; 56: 73-85
- Moore N, Verdoux H, Fantino B. Prospective, multicentre, randomized, double-blind study of the efficacy of escitalopram versus citalopram in outpatient treatment of major depressive disorder. *Int Clin Psychopharmacol* 2005; 20: 131-137
- Oslin DW, Ten Have TR, Streim JE, Datto CJ, Weintraub D, DiFilippo S; Katz IR. Probing the safety of medications in the frail elderly: evidence from a randomized clinical trial of sertraline and venlafaxine in depressed nursing home residents. *J Clin Psychiatry* 2003; 64: 875-882
- Poirier M-F, Boyer P. Venlafaxine and paroxetine in treatment-resistant depression. Double-blind, randomised comparison. *Brit J Psychiatry* 1999; 174: 12-16
- Raskin J, Wiltse CG, Siegal A, Sheikh J, Xu J, Dinkel JJ, Rotz BT, Mohs RC. Efficacy of duloxetine on cognition, depression, and pain in elderly patients with major depressive disorder: An 8-week, double-blind, placebo-controlled trial. *Am J Psychiatry* 2007; 164: 900-909
- Thase ME, Pritchett YL, Ossanna MJ, Swindle RW, Xu J, Detke MJ. Efficacy of duloxetine and selective serotonin reuptake inhibitors. *J Clin Psychopharmacol* 2007; 27: 672-676
- Unpublished Report-Wyeth-Ayerst. Randomised double-blind parallel-group comparison of the efficacy and safety of venlafaxine versus fluoxetine in the treatment of moderately depressed outpatients (Protocol 600A-626-NL). 2007
- Wise TN, Wiltse CG, Iosifescu DV, Sheridan M, Xu JY, Raskin J. The safety and tolerability of duloxetine in depressed elderly patients with and without medical comorbidity. *Int J Clin Pract* 2007; 61: 1283-1293
- Zanardi R, Franchini L, Serretti A, Perez J, Smeraldi E. Venlafaxine versus fluvoxamine in the treatment of delusional depression: A pilot double-blind controlled study. *J Clin Psychiatry* 2000; 61: 26-29