

# Existieren Schwellenwerte für klinische Relevanz?

## Exemplarische Darstellung von Wirksamkeitsunterschieden bei Escitalopram im Vergleich mit anderen Antidepressiva

Rebecca Schennach, München, Michael Riedel, Rottweil/München, Florian Seemüller, Sandra Dehning und Hans-Jürgen Möller, München

Für die Beurteilung der klinischen Relevanz wissenschaftlicher Ergebnisse werden verschiedene Kriterien herangezogen, die im Folgenden dargestellt werden. Die Aussagekraft dieser Kriterien wird am Beispiel von Untersuchungsergebnissen aus Escitalopram-Studien exemplarisch evaluiert. Die Vor- und Nachteile der Darstellung von Mittelwertsunterschieden in Psychopathologie-Ratingskalen, sowie dem Erfassen von Response und Remission und der Number needed to treat werden anhand von Einzelstudien und Metaanalysen diskutiert. Für den Kliniker scheinen besonders die Darstellung von Response und Remission von klinischer Relevanz zu sein, für Statistiker und Mathematiker sind Mittelwertsvergleiche hinsichtlich der klinischen Relevanz ein härteres und belastbareres Kriterium. Aus wissenschaftlicher Sicht sollten alle hier dargestellten und diskutierten Kriterien gleichwertig und gleichzeitig berücksichtigt werden, wenn die klinische Relevanz von Studienergebnissen eingestuft werden soll.

**Schlüsselwörter:** Klinische Relevanz, Mittelwertsunterschiede, Response, Remission, Number needed to treat

*Psychopharmakotherapie* 2012;19:10–7.

Im Zentrum medizinischer Forschung und Entwicklung stehen die Durchführung klinischer Studien und die Beurteilung unterschiedlicher Behandlungsmöglichkeiten. Die Beurteilung von Ergebnissen klinischer Studien stellt jedoch seit jeher eine große Herausforderung dar. Dies ist besonders in der psychiatrischen Forschung relevant, da hier biologische Marker zur Evaluierung eines möglichen Behandlungserfolgs gänzlich fehlen [15]. Daher kommen in psychiatrischen Studien überwiegend etablierte Psychopathologie-Ratingskalen zum Einsatz, um den psychopathologischen Ausgangszustand des Patienten und den Therapieerfolg zu messen. Verschiedene statistische Vorgehensweisen und Definitionen für einen Therapieerfolg sind seither entwickelt worden, wie zum Beispiel der Vergleich von Mittelwertsunterschieden oder auch das Erfassen von Response und Remission. Jedoch bergen alle diese statistischen Größen auch Nachteile

und etwaige Studienergebnisse müssen mit Bedacht gewertet werden [9, 35]. Denn statistische Gruppenvergleiche sind von unterschiedlichen Einflussvariablen abhängig und durch deren Auswahl beeinflussbar. So hängt das Ergebnis zum einen von der Größe der Stichprobe, der Größe des zu untersuchenden Unterschieds, aber auch von der Variabilität der Daten ab. Je größer die Stichprobe ist, desto kleiner kann der Unterschied zwischen den zu untersuchenden Gruppen sein, bei dem mit ausreichender Sicherheit von einem statistisch signifikanten Unterschied zwischen den Gruppen gesprochen werden kann. Folglich kann in Untersuchungen an großen Stichproben und erst recht in Metaanalysen mit umfangreichen Patientenzahlen selbst ein sehr geringer Unterschied statistische Signifikanz erreichen, so dass sich die Frage nach der klinischen Relevanz solcher Ergebnisse stellt [19]. Basierend auf einer Publikation von Montgomery & Möller [24]

sollen daher im Folgenden Kriterien zur Beurteilung der klinischen Relevanz am Beispiel der in mehreren Studien beschriebenen statistisch signifikanten Überlegenheit von Escitalopram im Vergleich mit anderen Antidepressiva exemplarisch angewandt werden. Die zugrunde liegenden Studien mussten in Englisch verfasst und in Medline gelistet sein. Folgende Stichworte wurden bei der Literatursuche berücksichtigt: major depressive disorder, escitalopram, comparative trial, randomized double-blind, efficacy. Es wurden sowohl Ein-

---

*Dr. Rebecca Schennach, Dr. Florian Seemüller, Dr. Sandra Dehning, Prof. Dr. Hans-Jürgen Möller, Klinik für Psychiatrie und Psychotherapie der Ludwig-Maximilians-Universität München, Nussbaumstraße 7, 80336 München, E-Mail: Rebecca.Schennach@med.uni.muenchen.de Prof. Dr. Michael Riedel, Vinzenz von Paul Hospital gGmbH, Klinik für Psychiatrie und Psychosomatik, Schwenninger Straße 55, 78628 Rottweil, und Klinik für Psychiatrie und Psychotherapie der Ludwig-Maximilians-Universität München*

zelstudien als auch Metaanalysen in diese Beurteilung mit aufgenommen.

## Ansätze zur Bewertung klinischer Relevanz

### Mittelwertsunterschiede in Depressionsskalen

Eines der ältesten und am besten etablierten Vorgehen zur Evaluation von Gruppenunterschieden ist der Vergleich von Mittelwerten in Depressionsskalen. Hierbei können ein allgemeiner Mittelwertsunterschied und ein gewichteter Mittelwertsunterschied (= Effektstärke) differenziert werden. Im Unterschied zu allgemeinen Mittelwertsunterschieden wird bei gewichteten Mittelwertsunterschieden eine Normierung durch die Variabilität der Daten vorgenommen. Bei der Bewertung der klinischen Relevanz von Mittelwertsunterschieden wird in den meisten klinischen Studien, die zu einer Zulassung eines neuen Antidepressivums führen, ein Mittelwertsunterschied von 2 Punkten auf der Hamilton-Depressionsskala (HAMD) zwischen Verum und Placebo zugrunde gelegt [18]. Dieser Unterschied wird von den Zulassungsbehörden als ausreichend anerkannt. Es gibt aber auch strengere Forderungen, so gilt nach den Leitlinien des britischen National Institute for Health and Clinical Excellence (NICE) zum Beispiel erst ein Mittelwertsunterschied von 3 Punkten auf der Hamilton-Depressionsskala als klinisch relevant [26], wobei dieser Grenzwert willkürlich festgelegt und nicht durch empirische Studien oder Expertenmeinungen unterstützt ist.

Die entwickelten Kriterien zur Beurteilung klinischer Relevanz beziehen sich überwiegend auf den Vergleich eines Arzneistoffs mit Placebo. Für den Vergleich zweier aktiver Behandlungsverfahren wäre es möglicherweise ein zu strenges Vorgehen, wenn man ein Kriterium anlegen würde, das für den Vergleich zwischen einem Arzneistoff und Placebo entwickelt wurde. Angesichts der Tatsache, dass Antidepressiva im Mittel einen Unterschied von 2 Punkten im Vergleich zu Placebo aufzeigen, würde dies nämlich bedeuten, dass ein

Behandlungsverfahren eine doppelt so große Effektstärke aufweisen müsste, um anderen Vergleichsmitteln klinisch überlegen zu sein: 2 Punkte Unterschied zu Placebo und ein weiterer Unterschied um 2 Punkte im Vergleich zu dem anderen Medikament. Um Unterschiede im Mittel zwischen Medikamenten zu untersuchen, wäre es demnach sinnvoller, wenn eine Überlegenheit eines Medikaments mit einem Mehrnutzen von 50 % nachgewiesen werden könnte. Daher wird ein mittlerer Unterschied von 1 Punkt beim Vergleich zweier Behandlungsgruppen als klinisch relevant angesehen.

### Unterschiede in den Response-Raten

Responder-Analysen sind in Europa wahrscheinlich das am besten etablierte Vorgehen, um klinische Relevanz zu ermitteln [8]. Eine psychopathologische Verbesserung um mindestens 50 %, gemessen mit gängigen Depressions-Skalen, ist das übliche Response-Kriterium. Eine andere Definition für Response, die auch gelegentlich angewandt wird, basiert auf der Clinical-Global-Impression-Skala, auf welcher ein Wert von 1 oder 2 (Zustand ist sehr viel besser/Zustand ist viel besser) am Ende der Studie Response angibt.

Eine Metaanalyse der Ergebnisse von Antidepressiva-Zulassungsstudien, die die Behandlung mit seit 1994 zugelassenen, selektiven Serotonin-Wiederaufnahmehemmern und Serotonin-Noradrenalin-Wiederaufnahmehemmern mit einer Placebo-Behandlung vergleicht [18], ergab einen mittleren Unterschied von 16 Prozentpunkten bei einer Streuung von 5 bis 20 Prozentpunkten für die einzelnen Antidepressiva. Nach Thase et al. gilt grundsätzlich ein Unterschied zwischen einem Antidepressivum und Placebo von 10 bis 20 Prozentpunkten als klinisch relevant [30]. In verschiedenen Zulassungsstudien anderer Medikamente werden ebenso Unterschiede zwischen einer medikamentösen Behandlung und Placebo von 10 Prozentpunkten oder mehr als ausreichend angesehen, um die klinische Relevanz eines Unterschieds zu Placebo zu begründen.

### Unterschiede in den Remissionsraten

Remission ist ein weiteres international anerkanntes Zielkriterium zur Beurteilung patientenrelevanter Therapieeffekte [8, 21]. Remission ist definiert als die Verringerung der depressiven Symptomatik unter einen bestimmten Schwellenwert auf einer Depressions-Skala, der sich nicht wesentlich von dem Durchschnittswert der normalen Bevölkerung unterscheidet. Remission gilt als primäres Behandlungsziel und stellt ein hartes Kriterium zur Erfassung der klinischen Relevanz dar, da Remission in Kurzzeitstudien wesentlich seltener auftritt als Response. Zudem haben viele Studien keine ausreichende Fallzahl, um diesbezüglich Wirksamkeitsunterschiede im Sinne von statistischer Signifikanz zu belegen. Daher gelten grundsätzlich statistisch signifikante Unterschiede in den Remissionsraten als klinisch relevant.

### Number needed to treat

Einen guten Anhaltspunkt bezüglich klinischer Relevanz bekommt man auch über die „Number needed to treat“ (NNT, Anzahl der notwendigen Behandlungen). Die NNT ist ein anerkannter Parameter zur Beschreibung klinischer Effekte zwischen zwei Behandlungsarmen. Je kleiner der Wert, desto höher die therapeutische Wirksamkeit im Vergleich zu einem anderen Verfahren (z. B. Placebo, z. B. Standardpräparat). Die Umrechnung der Responderaten in NNT von neueren Antidepressiva im Vergleich zu Placebo liegt im Mittel bei etwa 7. Das heißt, dass 7 Patienten behandelt werden müssen, um einen Responder mehr im Vergleich zu Placebo zu erzielen. Als geeigneter Grenzwert, um die klinische Relevanz eines Unterschieds zu Placebo bzw. eines Vergleichspräparats darzustellen, gilt eine NNT von 10 oder weniger.

### Klinische Relevanz in Einzelstudien mit Escitalopram

Die genannten Ansätze zur Bewertung klinischer Relevanz sollen im Folgenden am Beispiel von Escitalopram näher betrachtet werden, zunächst als

Tab. 1. Einzelstudien zu Escitalopram (ESC) im Vergleich mit anderen Antidepressiva

	Montgomery et al. (2004) [23]	Bielski et al. (2004) [2]	Moore et al. (2005) [25]	Colonna et al. (2005) [7]	Yevtushenko et al. (2007) [36]	Boulenger et al. (2006) [3]	Khan et al. (2007) [11]	Wade et al. (2007) [33]	Nierenberg et al. (2007) [27]	Clayton et al. (2006) [6]	Ventura et al. (2007) [32]	
	ESC/ Venlafaxin	ESC/ Venlafaxin	ESC/ Citalopram	ESC/ Citalopram	ESC/ Citalopram	ESC/ Paroxetin	ESC/ Duloxetin	ESC/ Duloxetin	ESC/ Duloxetin	ESC/ Duloxetin	ESC/ Bupropion <sup>3</sup>	ESC/ Sertralin
MADRS-Verbesserung vs. Baseline	-20,74 -20,41	-15,9 -13,6	-22,4 -20,3	-16,6 -15,6	28,7 25,2***	23,2 21,2**	18,0 15,9*	19,5 17,4*	7,22 7,61	-13,6 -13,2	-19,1 -18,4	
Response	77,4% <sup>1</sup> 79,6%	58,8% <sup>1</sup> 48,0%	76,1% 61,3%**	63% 55%*	95,4% 83,3%***	82% 76,7%	68% 50%*	68,8% 57,5%*	33% 40,1%	65% 62%*	75% 70%	
Remission	69,9% <sup>2</sup> 69,7%	50,5% <sup>2</sup> 41,8%	56,1% 43,6%*	55% 45%	89,9% 50,9%***	75% 66,8*	44% 38%	56% 47,9%	32% 37%	45% 43%*	58% <sup>4</sup> 58%	
NNT												
• Response			7		9		9	6				
• Remission			8		3 (2,6)	9						

<sup>1</sup>: Response-Definition: MADRS-Verbesserung um  $\geq 50\%$  von Baseline bis zum Endpunkt

<sup>2</sup>: Remission-Definition: MADRS-Punktwert von  $\leq 12$  Punkten bei Entlassung

<sup>3</sup>: Punktverbesserung, Response und Remission nach HAMD

<sup>4</sup>: Remission-Definition: MADRS-Punktwert von  $\leq 10$  Punkten bei Entlassung

MADRS: Montgomery-Åsberg Depression Rating Scale; NNT: Number needed to treat; Vergleich mit Escitalopram: \*  $p < 0,05$ ; \*\*  $p < 0,01$ ; \*\*\*  $p < 0,001$

Prüfung der klinischen Relevanz von Wirksamkeitsunterschieden in Einzelstudien mit Escitalopram (Tab. 1).

### Unterschiede in Mittelwertanalysen

In allen hier aufgeführten Studien wurde der Wirksamkeitsunterschied zwischen Escitalopram und anderen Antidepressiva anhand der Verbesserung der depressiven Symptomatik in der Montgomery-Åsberg-Depressionsskala (MADRS) oder der Hamilton-Depressionsskala (HAMD) gemessen. Aus mathematisch-statistischer Sicht stellt die Veränderung von Mittelwertsunterschieden eines der härtesten und belastbarsten Kriterien zur Beurteilung klinisch relevanter Unterschiede dar, auch wenn aus klinischer Sicht ein Unterschied von 2 oder 3 Punkten zwischen Escitalopram und anderen Antidepressiva zunächst wenig aussagekräftig scheinen mag. Bei dem Vergleich zwischen Escitalopram und *Citalopram* zeigen die meisten der aufgelisteten Studien einen Unterschied in der Veränderung der Mittelwerte von mehr als 2 Punkten, was darauf hindeutet, dass Escitalopram den entsprechenden Vergleichssubstanzen überlegen ist. In der Untersuchung von Yevtushenko et al. [36] konnte in der LOCF (Last observa-

tion forward)-Analyse ein signifikanter Behandlungserfolg im Sinne eines klinisch relevanten Unterschieds von Escitalopram im Vergleich zu Citalopram von sogar 3,5 Punkten auf der MADRS nach einer 6-wöchigen Behandlung gezeigt werden. Zusätzlich zu der gesamten Studienpopulation wurden in dieser Vergleichsstudie auch die Patienten mit einer sehr schweren Depression ausgewertet. Im Vergleich zu der gesamten Studienpopulation wurde bei dieser Subgruppe eine noch stärkere Wirksamkeit von Escitalopram im Vergleich zu Citalopram belegt [36]. Ein grundsätzlich signifikanter und klinisch relevanter Vorteil für Escitalopram zeigte sich hingegen nicht bei Vergleichsstudien mit Bupropion, Venlafaxin oder Sertralin. Bei der Beurteilung der klinischen Relevanz von Studienergebnissen sollte auch immer die *statistische Berechnungs- und Auswertestrategie* berücksichtigt werden. Denn auch das entsprechende Analyseverfahren kann zu deutlichen Unterschieden in den Mittelwertsveränderungen und somit zu einer Veränderung der schlussgefolgerten klinischen Relevanz der Ergebnisse führen. So wurde zum Beispiel bei zwei Studien, in denen Escitalopram mit *Duloxetin* verglichen wurde, für die jeweilige einzel-

ne Studie kein Vorteil von Escitalopram, bezogen auf eine 24-wöchige Behandlung, mittels LOCF-Analyse nachgewiesen (1,7 Punkte auf der MADRS). Anhand einer zusätzlich durchgeführten Observed-Case-Analyse konnte aber ein signifikanter Behandlungsvorteil von 2,2 Punkten auf der MADRS für die 24-wöchige Behandlung gezeigt werden [33].

Weiterhin sollte bei der Beurteilung klinischer Relevanz auch die *Studiendauer* bedacht werden. Denn der Grad des Unterschieds zwischen zwei Behandlungsarmen zeigt sich möglicherweise erst nach längerer Behandlungsdauer, was Unterschiede in der klinischen Relevanz nach sich ziehen kann. In einer Studie von Boulenger et al. [3] zeigte der Vergleich zwischen Escitalopram 20 mg und 40 mg *Paroxetin* zum Beispiel einen signifikanten Vorteil für Escitalopram mit einem Unterschied von 2,4 Punkten auf der MADRS in der LOCF-Analyse und einem Beobachtungszeitraum von 24 Wochen. Wurden nur 12 Wochen der Studie untersucht, zeigte sich zwar immer noch ein signifikanter Behandlungsvorteil zugunsten von Escitalopram, die Differenz betrug aber nur noch 2 Punkte auf der MADRS.

**Responderanalysen**

Ähnlich wie die Ergebnisse der Mittelwertsunterschiede wurden auch die Responderaten der unterschiedlichen Behandlungsarme in den einzelnen Studien explizit dargestellt. Besonders aus klinisch-therapeutischer Sicht stellen die Responderaten ein relevantes Kriterium dar. Denn das Ansprechen auf eine Behandlung und das Erreichen einer vordefinierten psychopathologischen Mindestverbesserung ist grundsätzlich mit einem günstigeren Erkrankungsverlauf assoziiert. Die meisten der hier aufgeführten Studien konnten das geforderte Kriterium für einen klinisch relevanten Unterschied in der Responderate (mindestens 10 Prozentpunkte Unterschied) erreichen. Ähnlich wie bei der Beurteilung der Ergebnisse der Mittelwertsunterschiede zeigten die meisten Vergleichsstudien zwischen Escitalopram und Citalopram einen Vorteil für Escitalopram. In der Studie von Moore et al. [25] konnte zum Beispiel gezeigt werden, dass es signifikant mehr Responder während der Behandlung mit Escitalopram (76,1%) als mit Citalopram (61,3%) gab. Der Unterschied von 14,8 Prozentpunkten zwischen Escitalopram und Citalopram liegt höher als das oben erwähnte Kriterium von 10 Prozentpunkten und erreicht fast den Mittelwert der Plazebo-Verum-Differenz in der Responderdifferenz von 16 Prozentpunkten in der Metaanalyse von Melander [18]. Eine ähnliche Differenz konnte in der kürzlich publizierten Arbeit von Yevtushenko et al. gezeigt werden [36]. Unter der Behandlung mit Escitalopram ergaben sich 95,4% Responder im Vergleich zu einer Responderate von 83,3% unter der Behandlung mit Citalopram. Die Differenz von 12,1 Prozentpunkten erfüllt erneut den 10%-Grenzwert.

Bei einer Studie, in der die Wirksamkeit von Escitalopram bei 143 Patienten im Vergleich zu einer Behandlung mit Duloxetin bei 151 Patienten untersucht wurde [33], fand sich für Escitalopram in der achten Behandlungswoche eine Responderate von 68,8% im Vergleich zu 57,5% unter Duloxetin. Dieser Unterschied war statistisch si-

gnifikant ( $p < 0,05$ ) und ist wegen der mehr als 10%igen Responder-Differenz als klinisch relevant zu werten. Auch in der Studie von Khan et al. konnten anhand derselben Responsedefinition signifikant mehr Responder in der Escitalopram-Gruppe im Vergleich zur Duloxetin-Gruppe nachgewiesen werden (68% für Escitalopram versus 50% für Duloxetin) [11]. Die Differenz der Responderrate von 18 Prozentpunkten liegt auch wieder weit über dem geforderten 10%-Grenzwert. Hingegen zeigten Vergleichsstudien zwischen Escitalopram und Sertralin sowie Paroxetin keinen klinisch relevanten Vorteil für Escitalopram; der Unterschied der Responderaten erfüllte in diesen Studien nicht das erwähnte 10%-Differenzkriterium für klinische Relevanz. Im Vergleich zu Venlafaxin konnte nur in der Untersuchung von Bielski et al. ein klinisch relevanter Vorteil für Escitalopram nachgewiesen werden.

**Remitteranalysen**

Neben den Ergebnissen zu Responderaten wurden in allen aufgelisteten Einzelstudien auch Angaben zu den Remissionshäufigkeiten der Patienten in den unterschiedlichen Behandlungsarmen gemacht. Remission, dem traditionellen Verständnis nach ein Zustand ohne Krankheitssymptome, ist für die Kliniker unter den Lesern von besonderem Interesse, da dieses Kriterium eine Aussage darüber trifft, ob der Patient im weitesten Sinne krank oder gesund ist [28]. Wenn also eine Therapie einen größeren Anteil an Patienten bei Entlassung in einen „gesunden“ Zustand versetzt als eine Vergleichsbehandlung, dann ist dies von äußerster klinischer Relevanz. Im Gegensatz zur Response wird aber keine Aussage darüber gemacht, inwieweit sich der Patient verbessert hat.

Grundsätzlich zeigt sich in der Übersichtstabelle, dass die meisten Studien, die bereits einen relevanten Unterschied bei den Responderaten fanden, auch einen relevanten Unterschied bei dem Vergleich der Remissionshäufigkeiten zeigen konnten. Bei der Beurteilung der Remissionsraten gilt es noch zu beden-

ken, dass hier das psychopathologische Ausgangsniveau des Patienten ein deutlicher Einflussfaktor ist. So können vor allem leicht oder moderat erkrankte Patienten am Ende einer Therapie einen vordefinierten Grenzwert für Remission erreichen, da sie wahrscheinlicher einen „gesunden“ Zustand zum Endpunkt der Studie erreichen werden, wenn sie von vornherein nicht schwer krank waren. Dies gilt es bei der klinischen Interpretation solcher Ergebnisse zu bedenken. Vor dem Hintergrund, dass statistisch signifikante Unterschiede in den Remissionsraten grundsätzlich als klinisch relevant zu werten sind, kann für die Beispielstudien gesagt werden, dass Escitalopram einen klinisch relevanten Vorteil zu haben scheint. So fanden sich beispielsweise in der Studie von Moore et al. [25] und auch in der Untersuchung von Yevtushenko et al. [36] signifikant mehr Remitter in der jeweiligen Escitalopram-Gruppe (Escitalopram 56,1%/Citalopram 43,6% bei Moore et al. und Escitalopram 89,9%/Citalopram 50,9% bei Yevtushenko et al.). Beide genannten Differenzen der Remissionsraten erfüllen wegen der grundsätzlichen Bedeutung von Remission und wegen der Signifikanz der Unterschiede das Kriterium für klinische Relevanz.

Bei dem Vergleich von Escitalopram und Duloxetin hinsichtlich der Remissionsraten wird ein weiteres grundsätzliches Problem der Outcome-Forschung deutlich, nämlich die ausgeprägte Beeinflussbarkeit der Ergebnisse durch die entsprechend angelegte Definition und auch wieder der Einfluss der Stichprobengröße. Denn in den Untersuchungen von Khan et al. und Wade et al. zeigten sich keine signifikanten Unterschiede in der Remissionsrate von Escitalopram im Vergleich zu Duloxetin [11, 33], wobei in den beiden Studien unterschiedliche Definitionen für Remission anlegt wurden: in der Studie von Khan et al. [11] ein MADRS-Wert  $\leq 10$  Punkte, bei Wade et al. [33] ein MADRS-Wert  $\leq 12$  Punkte. Hingegen zeigte sich bei einer gepoolten Analyse dieser beiden Einzelstudien und einer einheitlichen Definition für Remission (MADRS-Punktwert  $\leq 12$ ) ein

signifikanter Vorteil für Escitalopram im Vergleich zu Duloxetin von 9,9 Prozentpunkten (54,3 % im Vergleich zu 44,4 %) [13]. Wenn Remission als ein MADRS-Wert  $\leq 10$  Punkte definiert war, zeigte sich ein Unterschied von 8,8 Prozentpunkten (47,5 % versus 38,7 %). In einer Subgruppenuntersuchung der gepoolten Daten von Patienten, die an einer schweren Depression litten (MADRS-Wert  $\geq 30$  Punkte), betrug der signifikante Vorteil für Escitalopram verglichen mit Duloxetin 14,3 Prozentpunkte (51,7 % zu 37,4 %), wenn Remission als ein MADRS-Punktwert  $\leq 12$  definiert war [13]. Bei der Analyse mit einer weiteren Definition von Remission (MADRS-Wert  $\leq 10$  Punkte) betrug die Differenz 13,9 Prozentpunkte (47,2 % zu 33,3 %). Alle genannten Differenzen der gepoolten Daten erfüllen das Kriterium für klinische Relevanz, die Ergebnisse der Einzelstudien hingegen nicht.

#### Number needed to treat

Nur wenige Studien enthalten Angaben zu der Number needed to treat. In der Untersuchung von Moore et al. [25] beträgt die auf Responderraten basierende NNT 7 für Escitalopram im Vergleich zu Citalopram. Dieser Unterschied ist kleiner als 10 und somit ausreichend, um klinische Relevanz zu belegen. In der kürzlich durchgeführten Studie von Yevtushenko et al. [36] betrug die auf den Responderraten beruhende NNT 9. Dieser Wert entspricht ebenfalls den oben dargestellten Kriterien für klinisch relevante Unterschiede. Auch die auf den Remissionsraten basierenden NNT-Analysen dieser beiden Vergleichsstudien [25, 36] zeigen eine Überlegenheit von Escitalopram mit NNT-Werten zwischen 3 und 8. Im Praxisalltag bedeutet dies, dass von 100 behandelten depressiven Patienten unter Escitalopram bis zu 34 Patienten mehr remittieren als unter Citalopram, was die außerordentliche therapeutische Bedeutung der Überlegenheit von Escitalopram verdeutlicht. In zwei Vergleichsuntersuchungen mit Duloxetin betrug die auf Response basierende NNT einmal 9 und einmal 6 [11, 33]. In der Studie

von Boulenger et al., die 232 Patienten mit Escitalopram und 227 Patienten mit Paroxetin über 24 Wochen untersuchten, lag die NNT bezogen auf Remission am Ende einer Langzeitbehandlung bei 9 [3]. Auch dieser Wert belegt die klinische Relevanz der Ergebnisse gemäß den dargestellten Kriterien.

#### Escitalopram in meta-analytischen Untersuchungen

Häufig werden in Metaanalysen alle aufgeführten Kriterien zur Beurteilung klinischer Relevanz dargestellt, wobei der Fokus auf den Mittelwertsunterschieden und Response liegt. Deutlich weniger Studien berichten von Remissionsraten oder auch der NNT. Das gilt auch für Metaanalysen mit Escitalopram (**Tab. 2**).

In einer kürzlich veröffentlichten Cochrane-Analyse zu Escitalopram [5] von sechs Studien mit insgesamt 1 823 Patienten zeigte sich ein statistisch signifikanter Wirksamkeitsvorteil für Escitalopram in den Responderraten (61 % für Escitalopram gegenüber 54 % für Citalopram (Odds-Ratio [OR]: 0,67; 95%-Konfidenzintervall [95%-KI]: 0,50–0,89;  $p=0,006$ ). In den aufgeführten Werten bedeutet ein  $OR < 1$  eine Überlegenheit von Escitalopram gegenüber der Vergleichssubstanz. Bezüglich des Unterschieds zwischen den Responderraten von Escitalopram und Citalopram ist das 10%-Kriterium nicht erreicht, bei der Differenz von 7 Prozentpunkten kann die klinische Relevanz des Unterschieds aber noch vermutet werden. Auch in den Remitteranalysen fand sich ein statistisch signifikanter und nach den oben aufgeführten Kriterien klinisch relevanter Vorteil für Escitalopram (OR: 0,57; 95%-KI: 0,36–0,90;  $p=0,02$ ) [5].

In der jüngsten Metaanalyse von Montgomery et al. [22] ergaben sich ebenfalls signifikante Wirksamkeitsvorteile für Escitalopram im Vergleich zu Citalopram. Die Wirksamkeit der beiden Antidepressiva wurde anhand des Behandlungsunterschieds in der MADRS bei Woche 8 gemessen. Zudem wurden Response (50%-Verbesserung von

Aufnahme zu Endpunkt) und Remission ( $< 12$  Punkte auf der MADRS) evaluiert. Acht randomisierte kontrollierte Studien (RCTs) und eine naturalistische Studie mit zusammen 2 009 Patienten (Escitalopram: 995 Patienten; Citalopram: 1 014 Patienten) wurden in dieser Analyse berücksichtigt. Escitalopram war signifikant wirksamer als Citalopram mit einer mittleren Differenz von 1,7 Punkten auf der MADRS bei Woche 8 (95%-KI: 0,8–2,6;  $p=0,0002$ ), wobei vier der randomisierten Studien die MADRS verwendeten, zwei Studien die MADRS und HAMD und zwei Studien nur die HAMD-Skala. In den Responderanalysen zeigte sich ein Unterschied von 8,4 Prozentpunkten (95%-KI: 4,4–12,3) (8 RCTs) und bei den Remitteranalysen ein Unterschied von 17,6 Prozentpunkten (95%-KI: 12,1–23,1) bei 4 der untersuchten RCTs. Die NNT entsprach 11,9 für Response ( $p < 0,0001$ ) und 5,7 für Remission ( $p < 0,0001$ ). Das Odds-Ratio lag zugunsten von Escitalopram bei 1,44 für Response ( $p < 0,0003$ ) und bei 1,86 für Remission ( $p < 0,0001$ ). Aus all dem ist zu schließen, dass die in dieser Metaanalyse gefundene, statistisch signifikant überlegene Wirksamkeit von Escitalopram im Vergleich zu Citalopram auch klinisch relevant ist [22].

Interessanterweise konnten Cipriani et al. in ihrer metaanalytischen Untersuchung mit 12 Antidepressiva der neuen Generation bezüglich Wirksamkeit und Verträglichkeit keinen signifikanten Wirksamkeitsvorteil für Escitalopram gegenüber Citalopram hinsichtlich der untersuchten Responderraten nachweisen [4]. Auch in der Metaanalyse von Trkulja, in der sieben randomisierte kontrollierte Vergleichsstudien von Escitalopram und Citalopram untersucht wurden, zeigte sich keine klinisch relevante Überlegenheit von Escitalopram [31]. In der Arbeit von Trkulja werden zwar dieselben Kriterien (Mittelwertsunterschiede, Responderanteile, NNT) zur Evaluation klinischer Relevanz angelegt wie in der vorliegenden Arbeit, die Hintergründe zu der Wahl der Grenzwerte für die entsprechenden Kriterien bleiben aber unklar. So beurteilte

Tab. 2. Gepoolte Analysen und Metaanalysen zu Escitalopram (ESC) im Vergleich mit anderen Antidepressiva

	Auquier et al. (2003) [1]	Lepola et al. (2004) [14]	Llorca et al. (2005) [16]	Lam et al. (2006) [12]	Saeterdahl et al. (2007) [29]	Kennedy et al. (2009) [10]	Ramsberg (2008) [34]	Trkulja (2010) [31]	Lam et al. (2008) [13]	Cipriani et al. (2009) [5] (Cochrane-A.)	Cipriani et al. (2009) [4]	Montgomery et al. (2011) [22]
	ESC/ Citalopram	ESC/ Citalopram	ESC/ Citalopram	ESC/ Citalopram	ESC/ Citalopram	ESC/ Citalopram	ESC/ Citalopram	ESC/ Citalopram	ESC/ Duloxetine	ESC/ a. Citalopram b. Duloxetine c. Paroxetin	ESC/ a. Citalopram b. Duloxetine c. Paroxetin	ESC/ Citalopram
MADRS Verbesserung vs. Baseline	1,77*	3,6**	6*						18,9% 16,5%			
• SMD (95%-KI)								-0,131 (-0,235; -0,027)	a. -0,17 (-0,30; -0,04)** b. -0,10 (-0,30; 0,09) c. -0,05 (-0,36; 0,26)			
• MD (95%-KI)	1,2 (0,09; 1,95)*					1,2 (0,3; 2,1)***						1,7 (0,8; 2,6)***
Response	55,5% 50,8%**	56% 41%**	58% 39%**			59,7% 52,3%**			67,1% 53,2%**			72,3% 63,9%**
• OR (95%-KI)					0,70 (0,52; 0,93)*					a. 0,67 (0,50; 0,89)** b. 0,72 (0,43; 1,20) c. 0,89 (0,61; 1,32)	a. 0,84 (0,70; 1,01) b. 0,75 (0,60; 0,93)* c. 1,30 (1,10; 1,53)*	
Remission	51,6% 44,3%*					47,8% 42,7%*			54,3% 44,4%*			61,6% 44,0%**
• OR (95%-KI)					0,70 (0,54; 0,91)**		0,659 (0,509; 0,854)*			a. 0,57 (0,36; 0,90)* b. 0,90 (0,62; 1,29) c. 0,87 (0,45; 1,68)		
NNT (95%-KI)												
• Response								14 (7; 11)				11,9 (8,1; 22,8)***
• Remission												5,7 (4,3; 8,3)***

MADRS: Montgomery-Åsberg Depression Rating Scale; SMD: standardisierte MD; MD: Mittelwertdifferenz; OR: Odds-Ratio; NNT: Number needed to treat; \*p&lt;0,05; \*\*p&lt;0,01; \*\*\*p&lt;0,001

Trkulja zum Beispiel einen Mittelwertsunterschied von 2 Punkten als nicht ausreichend, um einen Vorteil von Escitalopram über Citalopram festzumachen [31]. Dies sollte, entsprechend den hier vorgestellten und diskutierten Kriterien zur Beurteilung von klinischer Relevanz, kritisch diskutiert und hinterfragt werden und zeigt noch einmal die Notwendigkeit der Entwicklung eines einheitlichen Verständnisses bei der Einschätzung von Therapieergebnissen. Hinsichtlich des Vergleichs zwischen Escitalopram und anderen Antidepressiva ist die Datenlage ebenfalls widersprüchlich. Bei einer metaanalytischen Untersuchung von 12 modernen Antidepressiva zeigte sich ein signifikanter Wirksamkeitsvorteil von Escitalopram gegenüber *Duloxetine* in den Responderaten [4]. Hingegen war Escitalopram in der kürzlich veröffentlichten Cochrane-Analyse von Cipriani et al. in drei untersuchten Studien an 1 120 Patienten bezüglich der Anzahl der Responder einer Behandlung mit Duloxetine über 6 bis 12 Wochen nicht signifikant überlegen (OR: 0,72; 95%-KI: 0,43–1,20; p=0,21) [5]. Auch in Follow-up-Untersuchungen bis zu 24 Wochen konnte kein signifikanter Unterschied gefunden werden (OR: 0,72; 95%-KI: 0,42–1,25; p=0,25). Hinsichtlich der weiteren Outcome-Kriterien Remission und Mittelwertveränderungen in der MADRS fand sich ebenfalls kein sicherer Hinweis auf eine Überlegenheit von Escitalopram (OR: 0,90; 95%-KI: 0,62–1,29; p=0,56; SMD -0,10; 95%-KI: -0,30 bis 0,09; p=0,28).

Auch bezüglich des Vergleichs zwischen Escitalopram und *Paroxetin* konnte in der von Cipriani et al. [5] durchgeführten Cochrane-Analyse von zwei Studien mit 784 Patienten keine Überlegenheit für Escitalopram hinsichtlich Responder- und Remitteranalysen sowie hinsichtlich der Veränderungen der MADRS-Mittelwerte in der akuten Behandlungsphase gefunden werden [5].

## Diskussion

In der vorliegenden Arbeit wurden anhand eines Vergleichs von Escitalopram

mit anderen Antidepressiva verschiedene Kriterien zur Beurteilung klinischer Relevanz vorgestellt. In den beiden Übersichtstabellen wird schnell deutlich, wie unterschiedlich einzelne wissenschaftliche Untersuchungen ihre Ergebnisse präsentieren und dass dabei häufig nicht alle Kriterien zur Beurteilung klinischer Relevanz dargestellt werden.

In allen zusammengetragenen Studien finden sich Daten zu Veränderungen und Unterschieden in den Mittelwertverläufen der angewandten Depressionsskalen – Angaben, die besonders auch im Zulassungsprozess eines Medikaments von großer Bedeutung sind. Aus klinischer Sicht stellt sich jedoch die Frage, was ein Unterschied von beispielsweise 2 Punkten auf der MADRS zwischen Escitalopram und Citalopram bedeutet. „Klinischere“ Kriterien für die Beurteilung von klinischer Relevanz sind Response und Remission, da grundsätzlich auch beide Outcome-Bereiche mit einem günstigeren Krankheitsverlauf, weniger Rückfällen und einer besseren Gesamtfunktionalität der Patienten in Verbindung gebracht werden. Die Number needed to treat beruht meist auf Responder- oder Remitteranalysen mit ähnlicher klinischer Relevanz. Im Vergleich von Response und Remission gilt Remission, also das Erzielen eines Zustands von Symptombefreiheit bzw. dem Vorhandensein von nur noch minimalen psychopathologischen Symptomen, als strengeres und schwieriger zu erreichendes Kriterium. Im Gegensatz zu Response wird bei Remission zwar nicht der Grad der Verbesserung angegeben, dafür aber der Endzustand, und somit fast direkt der Therapieerfolg beurteilt. Daher scheinen Unterschiede in Remissionsraten zwischen verschiedenen Behandlungsregimen von großer Bedeutung und vor allem von hoher klinischer Relevanz zu sein.

Interessanterweise fand sich in fast allen Metaanalysen und gepoolten Analysen ein statistisch signifikanter Vorteil für Escitalopram bezüglich der Mittelwertveränderung sowie der Response und Remission. Metaanalysen stellen eine Zusammenfassung verschiedener

Einzelstudien dar, wobei jedes Einzelergebnis in eine Effektgröße transferiert wird und diese Größen anschließend zu einem mittleren Effekt zusammengefasst werden [20]. Dabei werden die unterschiedlichen Ergebnisse nach ihrer Stichprobengröße gewichtet. Bei sehr heterogenen Effektgrößen wird üblicherweise zusätzlich der Einfluss einer Moderatorvariablen auf diese Ergebnisse untersucht. Die Cochrane Collaboration geht daher davon aus, dass Metaanalysen das höchste Maß an Evidenz bieten [17].

Allerdings weisen sie auch eklatante methodische Schwächen auf, die dieser Einschätzung widersprechen: So kann eine Metaanalyse nur so gut sein, wie es die Güte der Einzelergebnisse zulässt. Mit anderen Worten: Qualitativ minderwertige Einzelanalysen führen auch zu qualitativ minderwertigen Metaanalysen. Überdies werden oft unterschiedlich operationalisierte Einzelergebnisse zusammengefasst, auch wenn sie vielleicht durch ihre verschiedenartigen Definitionen schwer miteinander vergleichbar sind. Ein großes Problem, dem zwar methodisch häufig versucht wird Rechnung zu tragen, das aber dennoch einen unbestreitbar negativen Einfluss auf die Güte einer Metaanalyse hat, ist der Publikationsbias: Negative Studienergebnisse werden aus unterschiedlichsten Gründen häufig nicht publiziert, wodurch sich beim Blick auf die veröffentlichten Ergebnisse oft ein beschönigendes Bild des wahren Effekts ergibt. In den in **Tabelle 2** dargestellten Metaanalysen fällt auf, dass die meisten Arbeiten einen signifikanten Vorteil für Escitalopram im Vergleich zu den anderen Antidepressiva fanden, während sich diese Ergebnisse in den in **Tabelle 1** aufgeführten Einzelstudien nicht immer so zeigten. Das Zusammentragen verschiedener Datensätze erhöht die Fallzahl und dadurch können auch sehr kleine, klinisch eigentlich unbedeutende Unterschiede statistisch signifikant werden, was ein weiterer Nachteil von metaanalytischen Untersuchungen ist und bei der Diskussion der Ergebnisse bedacht und berücksichtigt werden sollte.

Zusammenfassend kann festgehalten werden, dass für Kliniker die Ergebnisse zu Response- und Remissionsraten von besonderer Bedeutung sind, um die klinische Relevanz wissenschaftlicher Studien zu beurteilen. Für Statistiker und Mathematiker stellen Mittelwertsunterschiede ein noch härteres und belastbareres Kriterium dar, unter anderem weil keine vordefinierten Grenzwerte wie bei Response oder Remission benötigt werden, wodurch wie bereits diskutiert die Ergebnisse beeinflusst werden können. Aus wissenschaftlicher Sicht ist es empfehlenswert, alle hier dargestellten Kriterien gleichwertig zu berücksichtigen und mit den jeweils ausgeführten Vor- und Nachteilen zu beurteilen. Vor diesem Hintergrund sollte eine valide Aussage über die klinische Relevanz von Studienergebnissen möglich sein.

#### Interessenkonflikte

Prof. Möller und Prof. Riedel geben die folgenden Interessenskonflikte an:

Von den folgenden Firmen hat Prof. Möller Forschungsgelder erhalten, ist Mitglied des Advisory Boards oder erhält Honorare für Vorträge: AstraZeneca, Bristol-Myers Squibb, Eisai, Eli Lilly, GlaxoSmithKline, Janssen-Cilag, Lundbeck, Merck, Novartis, Organon, Pfizer, Sanofi-Aventis, Schering-Plough, Schwabe, Sepracor, Servier, Wyeth.

Von den folgenden Firmen hat Prof. Riedel Forschungsgelder erhalten, ist Mitglied des Advisory Boards oder erhält Honorare für Vorträge: AstraZeneca, Pfizer, Otsuka Pharma, Janssen-Cilag.

Alle anderen Autoren haben keine Interessenskonflikte.

#### How to evaluate the clinical relevance of scientific research results – an example using escitalopram efficacy studies

Aim of this article is to provide an overview of criteria evaluating the clinical relevance of scientific research results. Statistically significant results of clinical studies are often biased by the sample size or variability of data. Therefore, it is important to validate the clinical relevance of such results. Exemplarily, using studies on escitalopram compared to other antidepressants, we will discuss different criteria (changes of mean differences of psychopathology rating scales, response, remission, number needed to treat) which can help clinicians to rank and critically discuss results of clinical studies as well as meta-analyses. Also in this context, pros and cons of single studies and meta-analyses will be discussed in order to provide tools helping clinicians to choose the best treatment available for the patient. For the treating clinician the consideration of response and remission bears greatest

clinical relevance, in contrast, for statisticians the changes in mean differences offer greatest relevance and validity. From a scientific point of view, all of the here discussed criteria should be concurrently considered in order to reliably conclude on the clinical relevance of research results.

**Key words:** Clinical significance, mean difference, response, remission, number needed to treat

## Literatur

- Auquier P, Robital S, Llorca P, Rive B. Comparison of escitalopram and citalopram efficacy: A meta-analysis. *Int J Psychiatry Clin Pract* 2003;7:7259–68.
- Bielski RJ, Ventura D, Chang CC. A double-blind comparison of escitalopram and venlafaxine extended release in the treatment of major depressive disorder. *J Clin Psychiatry* 2004;65:1190–6.
- Boulenger JP, et al. A comparative study of the efficacy of long-term treatment with escitalopram and paroxetine in severely depressed patients. *Curr Med Res Opin* 2006;22:1331–41.
- Cipriani A, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet* 2009;373:746–58.
- Cipriani A, et al. Escitalopram versus other antidepressive agents for depression. *Cochrane Database Syst Rev* 2009;2.
- Clayton AH, Croft HA, Horrigan JP, Wightman DS, et al. Bupropion extended release compared with escitalopram: effects on sexual functioning and antidepressant efficacy in 2 randomized, double-blind, placebo-controlled studies. *J Clin Psychiatry* 2006;67:736–46.
- Colonna L, Andersen HF, Reines EH. 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. *Curr Med Res Opin* 2005;21:1659–68.
- Committee for Proprietary Medicinal Products. Note for Guidance on Clinical Investigation of Medicinal Products in the Treatment of Depression. The European Agency for the Evaluation of Medicinal Products Evaluation of Medicines for Human Use. CPMP/EWP 2002:518–97.
- Guyatt GH, Deyo RA, Charlson M, Levine MN, et al. Responsiveness and validity in health status measurement: a clarification. *J Clin Epidemiol* 1989;42:403–8.
- Kennedy SH, Andersen HF, Thase ME. Escitalopram in the treatment of major depressive disorder: a meta-analysis. *Curr Med Res Opin* 2009;25:161–75.
- Khan A, et al. Double-blind comparison of escitalopram and duloxetine in the acute treatment of major depressive disorder. *Clin Drug Investig* 2007;27:481–92.
- Lam RW, et al. The influence of baseline severity on efficacy of escitalopram and citalopram in the treatment of major depressive disorder: an extended analysis. *Pharmacopsychiatry* 2006;39:180–4.
- Lam RW. Escitalopram and duloxetine in the treatment of major depressive disorder: a pooled analysis of two trials. *Int Clin Psychopharmacol* 2008;23:181–7.
- 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;19:149–55.
- Leucht S, Kane JM. Measurement-based psychiatry: definitions of response, remission, stability, and relapse in schizophrenia. *J Clin Psychiatry* 2006;67:1813–4.
- Llorca PM, Azorin JM, Despiegel N, Verpillat P. Efficacy of escitalopram in patients with severe depression: a pooled analysis. *Int J Clin Pract* 2005;59:268–75.
- Maier W, Möller HJ. Metaanalysen. Methode zur Evidenzmaximierung von Therapiestudien? *Nervenarzt* 2007;78:1028–36.
- Melander H, Salmonson T, Abadie E, van Zwieten-Boot B. A regulatory apologia – a review of placebo-controlled studies in regulatory submissions of new-generation antidepressants. *Eur J Neuropsychopharmacol* 2008;18:623–7.
- Möller HJ. Antidepressants: controversies about their efficacy in depression, their effect on suicidality and their place in a complex psychiatric treatment approach. *World J Biol Psychiatry* 2009;10:180–95.
- Möller HJ. Isn't the efficacy of antidepressants clinically relevant? A critical comment on the results of the metaanalysis by Kirsch et al. 2008. *Eur Arch Psychiatry Clin Neurosci* 2008;258:451–5.
- Möller HJ. Outcomes in major depressive disorder: The evolving concept of remission and its implications for treatment. *World J Biol Psychiatry* 2008;9:102–14.
- Montgomery S, Hansen T, Kasper S. Efficacy of escitalopram compared to citalopram: a meta-analysis. *Int J Neuropsychopharmacol* 2011;14:261–8.
- Montgomery SA, Huusom AK, Bothmer J. A randomised study comparing escitalopram with venlafaxine XR in primary care patients with major depressive disorder. *Neuropsychobiology* 2004;50:57–64.
- Montgomery SA, Möller HJ. Is the significant superiority of escitalopram compared with other antidepressants clinically relevant? *Int Clin Psychopharmacol* 2009;24:111–8.
- 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;20:131–7.
- National Institute for Clinical Excellence. Depression: management of depression in primary and secondary care. Clinical practice guideline no. 23. National Institute for Clinical Excellence, London, 2004.
- Nierenberg AA, Greist JH, Mallinckrodt CH, Prakash 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. *Curr Med Res Opin* 2007;23:401–16.
- Nierenberg AA, Wright EC. Evolution of remission as the new standard in the treatment of depression. *J Clin Psychiatry* 1999;60(Suppl 22):7–11.
- Sæterdal I. Efficacy and safety for SSRIs and other recent antidepressants for treatment of depression in adults. Norwegian Knowledge Centre for Health Services, 2007.
- Thase ME. Antidepressant effects: The suit may be small, but the fabric is real. *Prevention & Treatment* 2002;5:article 33.
- Trkulja V. Is escitalopram really relevantly superior to citalopram in treatment of major depressive disorder? A meta-analysis of head-to-head randomized trials. *Croat Med J* 2010;51:61–73.
- Ventura D, Armstrong EP, Skrepnek GH, Haim EM. Escitalopram versus sertraline in the treatment of major depressive disorder: a randomized clinical trial. *Curr Med Res Opin* 2007;23:245–50.
- 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 depressive disorder. *Curr Med Res Opin* 2007;23:1605–14.
- Wessling A, Ramsberg J. The review of antidepressants. The Dental and Pharmaceutical Benefits Agency, 2008.
- Yeaton WH, Sechrest L. Critical dimensions in the choice and maintenance of successful treatments: strength, integrity, and effectiveness. *J Consult Clin Psychol* 1981;49:156–67.
- Yevtushenko VY, 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. *Clin Ther* 2007;29:2319–32.