

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

**Im Namen der Deutschen Gesellschaft für Psychiatrie, Psychotherapie 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 Grundsatzfragen, die bereits – vergeblich – im Zusammenhang mit dem Berichtsplan bzw. den allgemeinen Methoden gestellt wurden, im wesentlichen nicht erneut im Detail aufgeworfen werden. Die Reihenfolge 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)?
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 Einschlusskriterien 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 Einzeldosen, 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 Vergleichssubstanzen, 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