

Wirksamkeit von Antidepressiva in der stationären Depressions-Therapie

Ergebnisse einer kontrollierten Studie mit Ratingskalen-Vergleichen

Cai Christian Heimsoth, München, Guido Distler, Wasserburg, und Gerd Laux, Soyen/München

Die Frage, ob die in Studien beschriebenen Wirksamkeitsraten von Antidepressiva auch in der Routinepraxis erzielt werden können und ob sich Wirksamkeitsunterschiede zwischen den einzelnen Wirkstoffgruppen nachweisen lassen, ist von hoher klinischer Relevanz. Zur Evaluation des Therapieverlaufs sind unterschiedliche Beurteilungsskalen im Gebrauch, wobei die Angaben zur Korrelation und die Handhabung von Cut-off-Werten teilweise differieren. Ziel der vorliegenden klinischen Studie war es, diesbezügliche Erkenntnisse zu erlangen. Ein Sample von 125 stationär behandelten depressiven Patienten wurde randomisiert mit verschiedenen Antidepressiva behandelt, der Therapieverlauf wurde anhand sieben unterschiedlicher Beurteilungsskalen zu fünf Messzeitpunkten evaluiert. Folgende Skalen kamen zum Einsatz: Jeweils zwei Versionen der Clinical Global Impressions Scale (CGI-S, CGI-I) und der Hamilton Depression Scale (HAMD-17, HAMD-7), die Montgomery Åsberg Depression Rating Scale (MADRS), die Befindlichkeits-Skala (Bf-S) und das Beck-Depressions-Inventar (BDI-II). Die Ratingskalen-Werte verringerten sich im Therapieverlauf kontinuierlich und signifikant. Der Cut-off-Wert ≤ 12 der Skala MADRS zeigte die höchste Übereinstimmung mit dem Cut-off-Wert ≤ 7 der Skala HAMD-17. Die kalkulierten Response- und Remissionsraten lagen zwischen 67 % und 75,9 % bzw. 58,9 % und 70,6 %. In den initialen drei Wochen der Behandlung konnten keine signifikanten Unterschiede zwischen den einzelnen Antidepressiva-Gruppen nachgewiesen werden. Anhand unterschiedlicher Response- und Remissionsraten deuteten sich allerdings tendenzielle Wirksamkeitsunterschiede zwischen den Antidepressiva-Wirkstoffgruppen an, die hier aufgrund der geringen Gruppenstärken nur rein deskriptiv angemerkt werden. Die Daten der Studie zeigen, dass sich in der Routinepraxis hohe Wirksamkeitsraten mit schnell einsetzenden Therapieeffekten nachweisen lassen. Verschiedene Fremdbeurteilungsskalen bewerten den Schweregrad einer Depression sehr ähnlich, die erfassten Symptomfortschritte scheinen aber nicht mit der subjektiven Befindlichkeit der Patienten zu korrelieren. Für zukünftige Untersuchungen wird anhand der vorliegenden Daten der Remissionsgrenzwert ≤ 12 der Skala MADRS empfohlen.

Schlüsselwörter: Antidepressiva, MADRS-Remissions-Cut-off-Wert, Korrelation von Beurteilungsskalen, Wirksamkeitsunterschiede von Antidepressiva

Psychopharmakotherapie 2020;27:2–11.



Effectiveness of antidepressants in inpatient treatment – Results from a controlled clinical study with comparisons of different rating scales

Of great importance is the question of whether the described efficacy rates of antidepressants can also be achieved in routine practice and whether differences between substances can be demonstrated. Various evaluation scales are in use with different cut-off values, aim of this clinical study was to gain further insights into this. A sample of 125 inpatients was randomly assigned to different antidepressants, the course of the therapy was evaluated on 7 different assessment scales. The following scales were used: Two versions of the Clinical Global Impression Scale (CGI-S, CGI-I) and the Hamilton Depression Scale (HAMD-17, HAMD-7), the Montgomery Asberg Depression Scale (MADRS), the emotional Scale (Bf-S) and the Beck Depression Inventory (BDI-II). The scores of all assessment scales decreased significantly, the calculated response and remission rates were between 67% and 75.9% and 58.9% and 70.6%. The cut-off value ≤ 12 of the MADRS scale showed the highest agreement with the cut-off value ≤ 7 of the HAMD-17 scale. The data reveal high effectiveness rates in routine practice, no significant differences could be detected between antidepressant drug groups in the initial three-week therapy phase. Different scoring scales assess the severity of depression in a very similar way, but the recorded ratings seem not to correlate with the subjective mood of the patient. For future studies, the cut-off value ≤ 12 of the MADRS scale is recommended on the basis of this data.

Key words: Antidepressants, MADRS remission cut-off value, correlation of evaluation scales, effectiveness different antidepressants

Dr. med. dent. Cai Christian Heimsoth, Maximilianstr. 2a, 80539 München, E-Mail: info@palais-dental.de

Guido Distler, kbo-Inn-Salzach-Klinikum gemeinnützige GmbH, 83512 Wasserburg am Inn

Prof. Dr. med. Dipl.-Psych. Gerd Laux, Institut für Psychologische Medizin (IPM) GbR, Nussbaumstraße 9, 83564 Soyen, E-Mail: ipm@ipm-laux.de

Das Krankheitsbild der Depression zählt gemessen am Indikator YLD (Years lived with disability) zu den häufigsten Krankheiten überhaupt [34]. Nach der Studie DEGS1-MH (Studie zur Gesundheit Erwachsener in Deutschland und ihr Zusatzmodul Psychische Gesundheit) liegt die 12-Monats-Prävalenz für unipolare Depressionen bei 7,7% (Frauen: 10,6%, Männer: 4,8%) [11].

Zur Abschätzung der aktuellen psychopathologischen Symptomatik eines Patienten werden standardisierte Beurteilungsverfahren eingesetzt. Selbstbeurteilungsskalen gelten dabei als zeitökonomischer und unabhängig von untersucherbedingten Verzerrungen, allerdings fallen Fehleinschätzungen durch eine gestörte Selbstwahrnehmung der Patienten stärker ins Gewicht. Fremdbeurteilungsskalen verringern diesen Fehler. Auch werden leichte Gesamtverhaltensfortschritte eher registriert [19].

Entscheidend für die Bewertung des Therapieerfolgs ist, ob ein Patient am Ende einer Behandlung Remissionswerte aufweist. Aufgrund der unterschiedlichen Definitionen von Remissionsgrenzwerten ist ein Vergleich verschiedener Studien nur eingeschränkt möglich [26]. Der HAMD-17-Remissionsgrenzwert ≤ 7 gilt weitgehend als akzeptiert [5, 38]. Bei der Fremdbeurteilungsskala MADRS besteht hingegen kein Konsens über einen optimalen Remissionsgrenzwert [14]. Häufig werden hier Remissionsgrenzwerte zwischen ≤ 8 und ≤ 12 angegeben [18, 20].

Ob Antidepressiva eine Wirksamkeit aufweisen, wurde in der Vergangenheit kontrovers beurteilt. Aktuelle Studien zeigen, dass sich bei vielen Patienten bereits innerhalb der ersten zwei Wochen antidepressive Effekte nachweisen lassen [21]. Es scheint individuelle Unterschiede zu geben, wobei im Mittel nach etwa zwei Wochen die antidepressive Wirkung einsetzt [13, 28]. Frühes Ansprechen auf eine antidepressive Behandlung gilt als hochprädiktiv für den späteren Behandlungserfolg [29, 30].

Die Angaben zu Wirksamkeitsunterschieden zwischen Antidepressiva sind uneinheitlich. In der aktuellen Ausgabe der S3-Praxisleitlinie Unipolare Depression wird konstatiert, dass es zwischen den einzelnen Antidepressiva-Wirkstoffgruppen keine nachweislichen Wirksamkeitsunterschiede gibt [6]. In anderen Untersuchungen konnten hingegen Wirksamkeitsvorteile einzelner Antidepressiva, wie etwa von Escitalopram, Venlafaxin oder Mirtazapin, belegt werden [4].

Studien, in denen die Wirksamkeit von Antidepressiva untersucht wird, werden häufig nach dem Prinzip einer randomisierten kontrollierten Studie (RCT = randomized controlled trial) durchgeführt. Studien dieser Art basieren auf einem quasi experimentellen Design mit hoher Patienten- und Ärzteselektion sowie der Standardisierung von Abläufen in der Therapie. In der Routineversorgung müssen allerdings unterschiedlichste Patienten mit oft ausgeprägter Komorbidität, hohem Nebenwirkungsrisiko und eingeschränkter Compliance be-

handelt werden [36]. Man schätzt, dass weniger als 30% der Patienten in der Routineversorgung die Aufnahmekriterien für eine RCT erfüllen [37]. Auch das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) empfiehlt daher die Durchführung randomisierter pragmatischer Studien (RPT = randomized pragmatic trial) [10, 25]. Hierbei handelt es sich um klinische Studien mit weiter gefassten Ein- und Ausschlusskriterien, bei denen Patienten randomisiert unterschiedliche Behandlungen zugewiesen werden [25].

Ziel der vorliegenden Studie war es, an einem randomisierten, stationär behandelten Patientenkollektiv der Routinepraxis die Wirksamkeit, das initiale Therapieansprechen, Unterschiede zwischen gängigen MADRS-Cut-off-Werten, die Korrelation der einzelnen Beurteilungsskalen sowie mögliche Unterschiede zwischen den eingesetzten Antidepressiva zu untersuchen.

Patienten und Methode

Nach Genehmigung durch die Ethikkommission der Ludwig-Maximilians-Universität München wurden im Zeitraum vom Oktober 2009 bis August 2012 am Zentralklinikum Wasserburg des kbo-Inn-Salzach-Klinikums aus der Gesamtheit der aufgenommenen Patienten mit der Hauptdiagnose einer Depression (nach ICD-10) 125 Patienten in die Studie zur Depressions-Spektrums-Störung aufgenommen (nachfolgend als DSS-Studie bezeichnet).

Auf Basis des Konzepts einer pragmatisch durchgeführten Studie wurde an einem auf einer spezialisierten Depressionsstation stationär behandelten Patientenkollektiv mit der Einweisungsdiagnose einer Depression unter Routinepraxisbedingungen der Therapieverlauf und das Ansprechen der Patienten (Response, Remission) auf die randomisiert eingesetzten Therapien evaluiert. Der Schwerpunkt der Untersuchung lag dabei auf der medikamentösen Therapie, nur diese wurde entsprechend variiert.

Einschlusskriterien für die Teilnahme an der Studie waren: freiwillige Teilnahme, schriftlich informierte Einverständniserklärung, Alter zwischen 18 und 65 Jahren und die Eingangsdiagnose einer Depression (F32.xx, F33.xx nach ICD-10).

Ausschlusskriterien waren: nicht einwilligungsfähige Patienten, Patienten während Schwangerschaft und Stillzeit, Minderjährige, akute Selbst- und Fremdgefährdung, Substanzmissbrauch bzw. Abhängigkeit, neurologische Erkrankungen und Intelligenzminderung.

Zur Überprüfung der Wirksamkeit der eingesetzten Therapien kamen fünf Fremdbeurteilungsskalen und zwei Selbstbeurteilungsskalen zum Einsatz. Hierbei handelte es sich um zwei Versionen der Fremdbeurteilungsskala Clinical Global Impressions Scale (CGI) zur globalen Erfassung des Krankheitsbildes der Patienten (CGI-S: Schweregrad [severity]; CGI-I Verbesserung [improvement]), des Weiteren die beiden Fremdbeurteilungsskalen Hamilton Depression Scale, ebenfalls in zwei

Versionen (HAMD-7, HAMD-17), und Montgomery Åsberg Depression Rating Scale (MADRS) zur Einschätzung des Schweregrads der Depression sowie die Selbstbeurteilungsskalen Befindlichkeits-Skala (Bf-S) zur Dokumentation des subjektiven Befindens und das Beck-Depressions-Inventar in der revidierten Fassung (BDI-II) zur Erfassung des Schweregrads der depressiven Symptomatik [5].

Die fünf Fremdbeurteilungsskalen sowie die Selbstbeurteilungsskala Bf-S wurden zu fünf Messzeitpunkten erhoben. Der zeitliche Abstand zwischen den ersten vier Messzeitpunkten (t0 = Aufnahme) betrug jeweils eine Woche. Der zeitliche Abstand zwischen den Messzeitpunkten t3 und t4 (t4 = Entlassung) war variabel und richtete sich nach dem Therapieerfolg. Die Selbstbeurteilungsskala BDI-II wurde nur zu drei Messzeitpunkten (t0 [Aufnahme], t2 und t4 [Entlassung]) erhoben. Die medikamentöse Behandlung der depressiven Symptomatik erfolgte mit verschiedenen Antidepressiva, die den Patienten zu Studienbeginn nach einer balancierten Blockrandomisierung zugeteilt wurden. Die Zuordnung wurde dabei anhand von Randomisierungslisten vorgenommen, welche mit dem Randomisierungsgenerator „Research-Randomizer“ (<https://www.randomizer.org>) erstellt wurden. In einzelnen Fällen, wie bei bekannter Non-Response auf die Substanzgruppe, musste von der gemäß Randomisierungsschema vorgesehenen Behandlung abgewichen werden. Auch gab es Patienten, bei denen die bisherige Medikation nicht sofort abgesetzt werden konnte. Die ursprünglich gleichmäßige Zuordnung der Antidepressiva wurde dadurch beeinflusst. Um mögliche Wirksamkeitsunterschiede zwischen den einzelnen Medikamenten aufzeigen zu können, wurden Antidepressiva aus unterschiedlichen Wirkstoffgruppen ausgewählt. Die Therapie erfolgte in den ersten drei Studienwochen im Sinne einer antidepressiven Monotherapie. Bei unzureichender Symptomreduktion oder starken Nebenwirkungen wurde die Therapie in diesem Zeitraum nur durch Dosisveränderungen oder vereinzelt auch durch Medikamentenwechsel angepasst. Als einzige kurzfristige Begleitmedikation kam in diesem Zeitraum bei schweren Schlafstörungen oder Erregungszuständen Lorazepam (bis 4 mg/Tag) zum Einsatz. Nach der dritten Studienwoche (zwischen den Messzeitpunkten t3 und t4 [Entlassung]) wurde bei unzureichender Symptomreduktion die Antidepressiva-Medikation einiger Patienten durch Augmentation mit einem Neuroleptikum (wie Pipamperon oder Quetiapin) bzw. durch Kombination mit einem trizyklischen Antidepressivum (wie Opipramol oder Trimipramin) erweitert.

Folgende Antidepressiva kamen in der Studie zum Einsatz:

- Aus der Gruppe der selektiven Serotonin-Wiederaufnahmehemmer (SSRI): Escitalopram, Citalopram, Paroxetin und Sertralin
- Aus der Gruppe der noradrenergen und spezifisch serotonergen Antidepressiva mit Alpha₂-Adrenozeptor-antagonistischer Wirkung (NaSSA): Mirtazapin

- Aus der Gruppe der selektiven Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI bzw. SSNRI): Venlafaxin und Duloxetin
- Aus der Gruppe der kombinierten selektiven Noradrenalin-Dopamin-Wiederaufnahmehemmer (NDRI): Bupropion
- Aus der Gruppe der selektiven Noradrenalin-Wiederaufnahmehemmer (NARI): Reboxetin
- Aus der Gruppe der Melatonin-Rezeptoragonisten und selektiven Serotonin-Rezeptorantagonisten: Agomelatin

Untersucht wurde, wie sich die Skalenwerte der eingesetzten Beurteilungsskalen im Verlauf der stationären Antidepressiva-Behandlung verändert haben. Mittels gepaarter t-Tests wurde überprüft, ob signifikante Abweichungen zwischen den Skalenwerten der einzelnen Erhebungszeitpunkte und dem Baseline-Wert nachgewiesen werden können. Zudem erfolgte die Kalkulation der Response- und Remissionsraten zu den einzelnen Messzeitpunkten.

Ein etwaiger Zusammenhang zwischen den einzelnen Fremdbeurteilungsskalen HAMD-17, HAMD-7 und MADRS sowie der Selbstbeurteilungsskala Bf-S wurde mittels der Korrelation nach Pearson überprüft.

Das Ausmaß der Übereinstimmung unterschiedlicher MADRS- und HAMD-17-Remissionsgrenzwerte wurde mittels der Kappa-Statistik errechnet.

Die eingesetzten Antidepressiva wurden im Verlauf der ersten drei Wochen miteinander verglichen. Mittels der Regressionsanalyse wurde die Abweichung jeder Wirkstoffgruppe vom Gesamtmittelwert errechnet. Zudem wurden auch hier für jede Antidepressiva-Wirkstoffgruppe die Response- und Remissionsraten zu einzelnen Messzeitpunkten kalkuliert.

Das Signifikanzniveau wurde aufgrund der vielen statistischen Tests, im Sinne der einfachen Bonferroni-Korrektur, von $\alpha = 0,05$ auf $\alpha = 0,0005$ erniedrigt [35].

Ergebnisse Patientencharakteristika

Im Zeitraum von 2009 bis 2012 wurden am Zentralklinikum Wasserburg des kbo-Inn-Salzach-Klinikums 5310 Patienten mit einer F3x.xx-Hauptdiagnose nach ICD-10 aufgenommen. Die Hauptdiagnosen F32.2 und F33.2 waren mit 27,3 % bzw. 32,2 % am häufigsten vertreten. Das mittlere Alter der Patienten betrug $51,5 \pm 18,9$ Jahre. Der Anteil weiblicher Patienten überwog mit 60,8 %.

Im gleichen Zeitraum erfolgte die Aufnahme von 125 Patienten mit der Einweisungsdiagnose einer Depression (F32.xx, F33.xx nach ICD-10) in die Datenbank der DSS-Studie. Bei einigen Patienten änderte sich im weiteren Studienverlauf die Hauptdiagnose. **Tabelle 1** zeigt die demographischen Daten aller DSS-Studienpatienten und der Subgruppe Medikation. Diese beinhaltet nur die Patienten, denen in der ersten Studienwoche eine Wirkstoffgruppe zugeordnet werden konnte. Die angegebene Hauptdiagnose ist die bei Entlassung der Patienten.

Tab. 1. Demographische Daten der DSS-Studien-Patienten und der Subgruppe Medikation

		DSS-Studie	Subgruppe Medikation
N		125	88
Geschlecht	Männlich	51,2%	52,3%
	Weiblich	48,8%	47,7%
Alter [Jahre]		47,1 ± 8,8	46,9 ± 8,7
Gewicht [kg]		81,4 ± 18,9	81,0 ± 17,8
BMI		27,5 ± 6,1	27,4 ± 6,0
Diagnose	F31.4	4,0%	4,5%
	F32.1	4,0%	4,5%
	F32.2	43,2%	43,2%
	F32.3	1,6%	1,1%
	F33.1	4,0%	4,5%
	F33.2	37,6%	36,4%
	F33.3	0,8%	1,1%
	F34.0	0,8%	0%
	F34.1	1,6%	1,1%
	F43.2	1,6%	2,3%
Nikotinkonsum	Raucher	44,8%	45,5%
	Nichtraucher	53,6%	52,2%
	Fehlende Daten	1,6%	2,3%
Frühere depressive Episoden [n]		1,77 ± 2,8	1,70 ± 2,2
Dauer der aktuellen Episode [Wochen]		19,6 ± 22,7	18,5 ± 19,2
Suizidversuche	Ja	16,0%	18,2%
	Nein	84,0%	81,8%
Ersterkrankungsalter [Jahre]		39,7 ± 13,0	39,1 ± 13,4
Medikation	Paroxetin	1,6%	2,3%
	Sertralin	2,4%	3,4%
	Citalopram	3,2%	4,5%
	Escitalopram	14,4%	20,5%
	Mirtazapin	15,2%	21,6%
	Venlafaxin	7,2%	10,2%
	Duloxetin	7,2%	10,2%
	Bupropion	3,2%	4,5%
	Reboxetin	3,2%	4,5%
	Agomelatin	12,8%	18,2%
	Valproinsäure	0,8%	0%
	Opipramol	0,8%	0%
	Keine Medikation	1,6%	0%
	Fehlende Daten	26,4%	0%

Die Tabelle verdeutlicht, dass das reduzierte Patientenkollektiv (Subgruppe Medikation) eine sehr hohe Übereinstimmung mit dem Studienpatientenkollektiv aufweist. Die Einteilung der verabreichten Antidepressiva erfolgte anhand der genannten sechs Wirkstoffgruppen. Die unterschiedlichen Fallzahlen der Medikationsgruppen sind dadurch bedingt, dass Patienten mit einer Vorbehandlung des randomisiert zugeordneten Medikaments ausgeschlossen wurden.

Escitalopram wurde dabei aufgrund der mehrfach beschriebenen Wirksamkeitsvorteile gegenüber anderen SSRI einer eigenen Gruppe zugeordnet. Die Wirkstoffgruppen NDRI und NARI wurden angesichts ihrer geringen Gruppenstärke zusammengefasst. Siehe hierzu **Tabelle 2**.

Therapieverläufe der Studienpatienten mit Response und Remissionsraten

Die Skalen CGI-S (Therapieverlauf), CGI-I (Zustandsänderung) und Bf-S (Befindlichkeit) bilden den globalen Therapieverlauf der Patienten ab. Die Skalen HAMD-17, HAMD-7, MADRS und BDI-II bilden den Depressionsverlauf der Patienten ab. Die **Abbildungen 1 bis 6** zeigen die Verteilung der Skalendaten über alle Erhebungszeitpunkte und für alle auswertbaren Studienpatienten anhand von Boxplot-Diagrammen. Bei diesen entspricht die Box dem Bereich, in dem die mittleren 50% der Daten liegen. Ein zusätzlicher Querbalken innerhalb der Box gibt den Median an, welcher die Anzahl der Daten in zwei gleiche Hälften teilt. Die T-Balken zeigen die weitere Verteilung der Daten. Zusätzliche Ausreißer sind mit einem Kreis und Extremwerte mit einem Stern gekennzeichnet.

Tabelle 3 zeigt den Anteil an respondierenden und remittierenden Patienten. Die Errechnung der Responderaten erfolgte dabei stets in Bezug auf den Skalen-Score bei Patientenaufnahme (Messzeitpunkt t0). Die Kalkulation der Remissionsraten erfolgte anhand der Remissionsgrenzwerte ≤ 7 der HAMD-17, ≤ 3 der HAMD-7, ≤ 12 der MADRS und < 13 des BDI-II.

Der Messzeitpunkt t0 entsprach dem Studienbeginn und der Aufnahme der Patienten, die nachfolgenden drei Datenerhebungen (Messzeitpunkte t1–t3) erfolgten jeweils im wöchentlichen Abstand. Der Entlassungszeitpunkt war variabel und richtete sich nach dem Therapieerfolg (Messzeitpunkt t4).

Da mit der Skala CGI-I nur Zustandsänderungen erfasst werden können, wurden bei dieser Skala die Skalendaten zum Messzeitpunkt t0 nicht miteingeschlossen.

Die CGI-S- und CGI-I-Skalendaten zeigten eine stetige Verbesserung des Zustands von Messzeitpunkt zu Messzeitpunkt bei der Mehrheit der Patienten. Diese war zu allen Messzeitpunkten signifikant (gepaarte t-Tests) und konnte bereits nach der ersten Behandlungswoche nachgewiesen werden (p < 0,0005). Gleiches galt auch für alle den Depressionsverlauf abbildenden Skalen (einschließlich des BDI-II). Bereits nach der ersten Studienwoche wiesen knapp 10% der Patienten Re-

Tab. 2. Anzahl der nach verabreichter Antidepressiva-Wirkstoffgruppe auswertbaren Patienten in der 1. Studienwoche der DSS-Studie

Wirkstoffgruppe	Antidepressiva	Patienten [n (%)]	Geschlecht [%]		Alter [Jahre]
			Männlich	Weiblich	
SSRI	Citalopram + Paroxetin + Sertralin	9 (10,2)	44,4	55,6	44,3 ± 6,8
SSRI	Escitalopram	18 (20,5)	50,0	50,0	43,9 ± 7,5
NaSSA	Mirtazapin	19 (21,6)	63,2	36,8	46,9 ± 11,3
SSNRI	Venlafaxin + Duloxetin	18 (20,5)	50,0	50,0	51,0 ± 7,2
NDRI + NARI	Bupropion + Reboxetin	8 (9,1)	25	75	42,8 ± 9,5
Melatonerges Antidepressivum	Agomelatin	16 (18,2)	62,5	37,5	49,4 ± 6,8
Gesamt		88			46,9 ± 8,7

Tab. 3. HAMD-17, HAMD-7, MADRS und BDI-II-Response- und Remissionsraten der Studienpatienten

Skala	Messzeitpunkt	N	Response*	Remission*
HAMD-17	t1	99	11,1%	10,1%
	t2	112	26,8%	9,8%
	t3	85	42,4%	16,5%
	t4	108	74,1%	62,0%
HAMD-7	t1	103	17,5%	12,6%
	t2	116	31,0%	12,1%
	t3	89	46,1%	11,2%
	t4	112	75,9%	58,9%
MADRS	t1	100	10,0%	10,0%
	t2	113	20,4%	11,5%
	t3	86	38,4%	22,1%
	t4	109	75,2%	70,6%
BDI-II	t2	107	32,7%	29,0%
	t4	103	67,0%	66,0%

*Response: mindestens 50% Besserung ggü. Ausgangswert (t0);
Remission: HAMD-17 ≤ 7, HAMD-7 ≤ 3, MADRS ≤ 12, BDI-II ≤ 13

response- oder gar Remissionswerte auf. Die Responderaten erhöhten sich im weiteren Verlauf der initialen Therapiephase in allen Skalen deutlich und kontinuierlich, die Remissionsraten erhöhten sich erst ab der dritten Studienwoche. Bei Entlassung

Tab. 4. Kappa-Werte der HAMD-17- und MADRS-Remissionsgrenzwerte

Messzeitpunkt	MADRS-Remissionsgrenzwert				
	MADRS ≤ 8	MADRS ≤ 9	MADRS ≤ 10	MADRS ≤ 11	MADRS ≤ 12
t0	0,319	0,559	0,559	0,559	0,559
t1	0,669	0,797	0,797	0,797	0,954
t2	0,564	0,597	0,567	0,567	0,623
t3	0,369	0,487	0,607	0,629	0,613
t4	0,767	0,822	0,822	0,822	0,799
Mittelwert	0,538	0,652	0,670	0,675	0,710

der Patienten wiesen je nach Skala etwa drei Viertel der Patienten Response- und etwa zwei Drittel der Patienten Remissionswerte auf. Hierbei ergaben sich mit der MADRS die höchsten Remissionsraten.

Im Unterschied dazu ließ sich anhand der Skalendatenverteilung der Bf-S keine eindeutige und kontinuierliche Verbesserung des Befindens der Patienten nachweisen. Die p-Werte der Teststatistik wiesen nur zu den Messzeitpunkten t2 und t4 auf eine signifikante Abnahme der Skalenwerte hin (gepaarte t-Tests; α = 0,0005).

Konkordanz der MADRS- und HAMD-17-Cut-off-Werte

Tabelle 4 zeigt die Kappa-Werte zwischen dem HAMD-17-Remissionsgrenzwert ≤ 7 und den einzelnen MADRS-Remissionsgrenzwerten ≤ 8, ≤ 9, ≤ 10, ≤ 11 und ≤ 12 zu den jeweiligen Messzeitpunkten.

Die Tabelle verdeutlicht, dass der mittlere Kappa-Wert über alle Messzeitpunkte hinweg für den MADRS-Remissionsgrenzwert ≤ 12 am höchsten war. Dieser Wert hatte anhand des hier vorliegenden Datensatzes die größte Übereinstimmung mit dem HAMD-17-Remissionsgrenzwert ≤ 7. Die Differenzen zu den übrigen Remissionsgrenzwerten wurden in absteigender Reihenfolge größer.

Korrelation der Fremd- und Selbstbeurteilungsskalen

Nachfolgend ist die Korrelation der eingesetzten Beurteilungsskalen anhand des Korrelationskoeffizienten nach Pearson dargestellt.

Dabei wird zwischen der Korrelation der einzelnen Fremdbeurteilungsskalen untereinander und der Korrelation der Fremdbeurteilungsskalen mit der Selbstbeurteilungsskala Bf-S unterschieden. Siehe hierzu **Abbildung 7**.

Die Ergebnisse der Korrelationsanalysen nach Pearson zeigten, dass zwischen den drei verwendeten Fremdbeurteilungsskalen HAMD-17, HAMD-7 und MADRS über alle Messzeitpunkte hinweg stets eine signifikante hohe Korrelation bestand (p < 0,0005). Die Werte der Korrelationskoeffizienten lagen bei Entlassung der Patienten bei jeder der drei Paarungen über 0,9.

Im Unterschied dazu korrelierte die Selbstbeurteilungsskala Bf-S mit keiner der genannten Fremdbeurteilungsskalen. In den ersten zwei Wochen waren

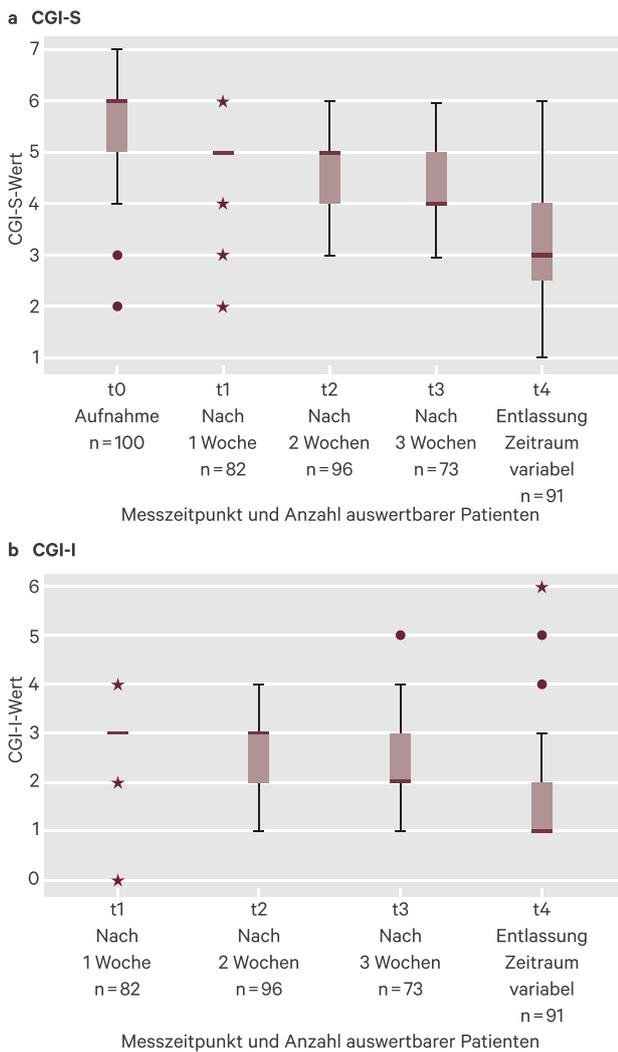


Abb. 1. Verlauf der CGI-S- und CGI-I-Krankheitswerte auswertbarer Studienpatienten (CGI-I = 0: nicht beurteilbar)

die Ergebnisse sogar leicht konträr. Im weiteren Verlauf der Behandlung zeigte sich eine gewisse Erhöhungstendenz der Korrelationen, die aber zu keinem Zeitpunkt signifikant war.

Antidepressiva-Wirkstoffgruppenvergleich

Die Auswertung des Antidepressiva-Wirkstoffgruppenvergleichs umfasste nur die ersten drei Wochen der Behandlung. Die Daten der Regressionsanalyse zwischen den Messzeitpunkten t1–t3 und dem Messzeitpunkt t0 (Aufnahme) ergaben, dass die Abweichungen einzelner Gruppen vom Gesamtskalenmittelwert zu keinem der Messzeitpunkte signifikant waren.

Die Errechnung der Response- und Remissionsraten erfolgte stets in Bezug auf den Skalen-Score bei Patientenaufnahme (Messzeitpunkt t0). Zur besseren Übersicht sind in den Tabellen 5 und 6 die durchschnittlichen Response- und Remissionsraten für jede Wirkstoffgruppe mit den entsprechenden Bereichen zu den einzelnen Messzeitpunkten dargestellt.

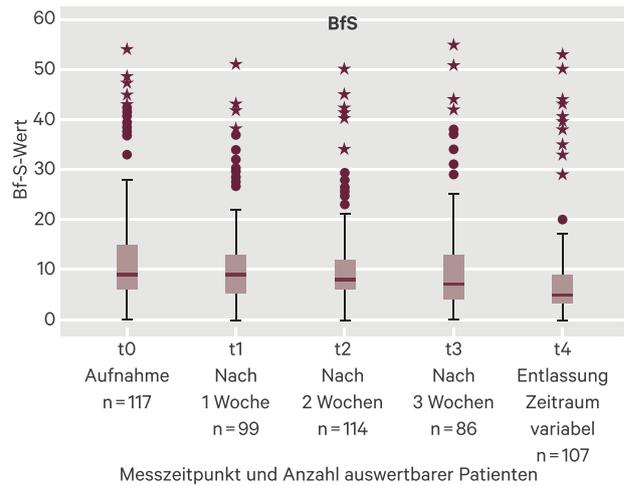


Abb. 2. Verlauf der Bf-S-Befindlichkeitswerte auswertbarer Studienpatienten

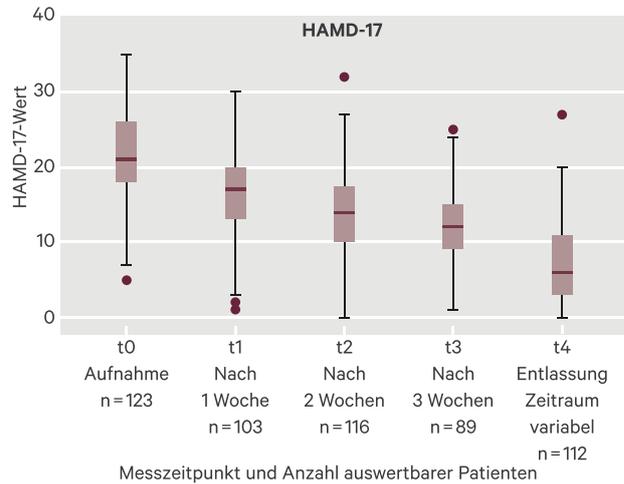


Abb. 3. Verlauf der HAMD-17-Depressionswerte auswertbarer Studienpatienten

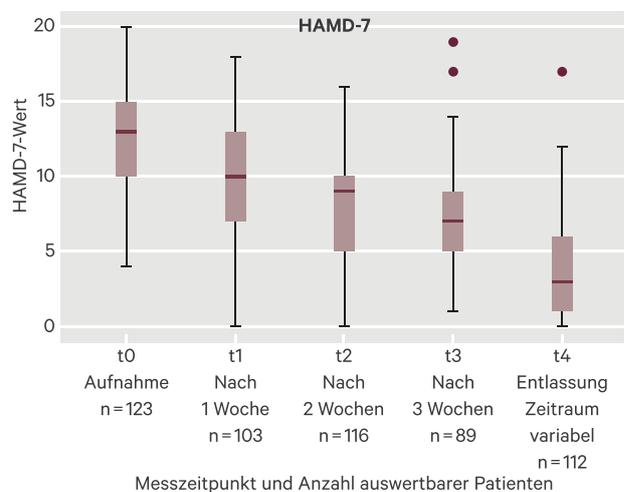


Abb. 4. Verlauf der HAMD-7-Depressionswerte auswertbarer Studienpatienten

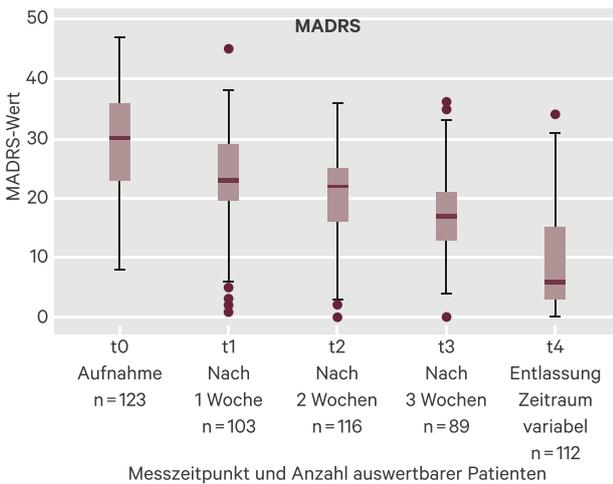


Abb. 5. Verlauf der MADRS-Depressionswerte auswertbarer Studienpatienten

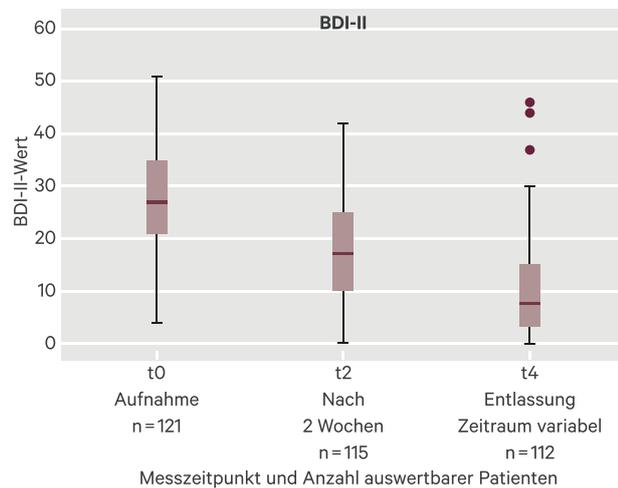


Abb. 6. Verlauf der BDI-II-Depressionswerte auswertbarer Studienpatienten

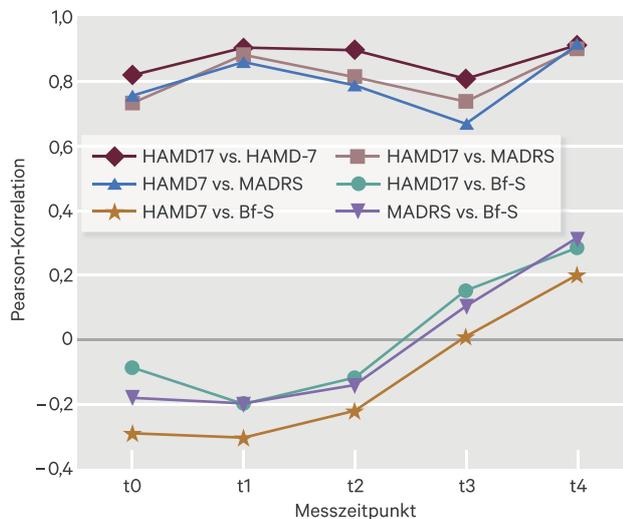


Abb. 7. Verlauf der kalkulierten Skalenkorrelation nach Pearson

Anhand der Response- und Remissionsraten schien die Wirkung in den Gruppen SSRI (Escitalopram), NaSSA und SSNRI schneller einzutreten, in den Gruppen SSRI (übrige), NDRI + NARI und melatonerges Antidepressivum hingegen erst verzögert.

Nach den ersten drei Behandlungswochen respondierten anteilig mehr Patienten in den Wirkstoffgruppen SSRI (Escitalopram) und NaSSA. Die höchsten Remissionsraten wiesen die Wirkstoffgruppen SSRI (übrige) und SSRI (Escitalopram) auf. In den Wirkstoffgruppen NDRI + NARI und SSNRI respondierten bzw. remittierten anteilig die wenigsten Patienten.

Diskussion

Aus der Gesamtheit der im Studienzeitraum am Zentralklinikum Wasserburg aufgenommenen 5310 Patienten wurden 125 in die DSS-Studie aufgenommen. Das Durchschnittsalter der Gesamtaufnahmen betrug 51,5 Jahre, mit 60,8% überwog der Anteil weiblicher Patienten. Frauen scheinen demnach stationäre Depressionsbehandlungen vermehrt in Anspruch zu nehmen.

Im DSS-Studienpatientenkollektiv lag der Anteil weiblicher Patienten nur noch bei 48,8%, das Durchschnittsalter betrug 47,1 Jahre. Dies deutet auf einen gewissen Selektionseffekt und eine damit einhergehende Verzerrung der Stichprobe hin. In einer vergleichbaren Studie von Seemüller et al. aus dem Jahre 2010 mit weitaus größerer Fallzahl entsprach die Geschlechterverteilung in etwa den hier beschriebenen Gesamtaufnahmen (62,6% weibliche Patienten). Durchschnitts- und Ersterkrankungsalter glichen allerdings dem Studienpatientenkollektiv (45,1 bzw. 37,9 Jahre, hier 47,1 bzw. 39,7 Jahre) [27]. Wegen unterschiedlicher Aufnahme- und Entlassungszeiträume konnte nur eine kontinuierliche Behandlung von drei Wochen evaluiert werden. Dieser Zeitraum ist allerdings klinisch relevant, da aus gesundheitsökonomischen Erwägungen den Kliniken eine Verweildauer in dieser Größenordnung vorgegeben wird, Patienten häufig auf kürzere stationäre Behandlung drängen und Antidepressiva-Effekte hauptsächlich in diesem Zeitraum auftreten. Allerdings ist in diesem Zeitraum nur von Response und nicht von Remission auszugehen.

Zum Entlassungszeitpunkt der Patienten aus der stationären Behandlung wiesen alle erhobenen Beurteilungsskalen eine signifikante Verringerung der Skalenwerte auf ($p < 0,0005$). Anhand der erhobenen Fremdbeurteilungsskalen CGI-S, CGI-I, HAMD-17, HAMD-7 und MADRS und der Selbstbeurteilungsskala BDI-II ließ sich diese Verringerung kontinuierlich ab der ersten Behandlungswoche (BDI-II ab der zweiten) nachweisen. Auch wenn der Therapieeffekt nicht vollständig der verabreichten Medikation zuzuschreiben war, ließ sich anhand dieser Skalen keine auffällige Wirklatenz der Antidepressiva-Therapie, wie sie teilweise beschrieben wurde, nachweisen [23]. Kritisch muss hierbei angemerkt werden, dass nur auswertbare Patienten in die Analyse mit einbezo-

Tab. 5. Durchschnittliche Responderaten der Fremdbeurteilungsskalen (HAMD-17, HAMD-7, MADRS und BDI-II [nur zu t2]) zu den Messzeitpunkten t1 bis t3

Wirkstoffgruppe	Durchschnittliche Responderate (Bereich)		
	Messzeitpunkt		
	t1	t2	t3
SSRI (übrige)	0 % (0 %)	19,4 % (11,1–22,2 %)	41,7 % (25–50 %)
SSRI (Escitalopram)	22,2 % (22,2 %)	40 % (29,4–60 %)	50 % (43,8–56,3 %)
NaSSA	16,7 % (15,8–17,6 %)	33,3 % (27,8–36,8 %)	53,7 % (46,7–64,3 %)
SSNRI	22,6 % (12,5–35,3 %)	21,6 % (13,3–37,5 %)	37,9 % (28,6–46,7 %)
NDRI + NARI	9,5 % (0–28,6 %)	25 % (14,3–57,1 %)	33,3 % (0–66,7 %)
Melatonerges AD	6,3 % (0–12,5 %)	31,7 % (20–40 %)	47,2 % (33,3–66,7 %)

Tab. 6. Durchschnittliche Remissionsraten der Fremdbeurteilungsskalen (HAMD-17, HAMD-7, MADRS und BDI-II [nur zu t2]) zu den Messzeitpunkten t1 bis t3

Wirkstoffgruppe	Durchschnittliche Remissionsrate (Bereich)		
	Messzeitpunkt		
	t1	t2	t3
SSRI (übrige)	0 % (0 %)	13,9 % (0–33,3 %)	20,8 % (12,5–25 %)
SSRI (Escitalopram)	22,2 % (22,2 %)	29,7 % (17,6–60 %)	25 % (18,8–31,3 %)
NaSSA	12,9 % (11,1–15,8 %)	19,8 % (10,5–40 %)	18 % (12,5–21,4 %)
SSNRI	16,4 % (12,5–23,5 %)	10,3 % (0–21,4 %)	16,6 % (14,3–20 %)
NDRI + NARI	0 % (0 %)	17,9 % (14,3–28,6 %)	5,6 % (0–16,7 %)
Melatonerges AD	6,3 % (6,3 %)	15 % (6,7–26,7 %)	19,4 % (0–33,3 %)

gen wurden. Eine zusätzliche LOCF(last observation carried forward)-Analyse zur Ergänzung der fehlenden Daten wurde nicht durchgeführt. Die Skalendaten der Selbstbeurteilungsskala Bf-S verringerten sich hier, im Gegensatz zu ähnlichen Studien, nicht kontinuierlich. Demnach deutete sich in der initialen Therapiephase eine auffällige Diskrepanz zwischen den erhobenen Selbst- und Fremdbeurteilungsskalen an. Erklären lässt sich diese damit, dass erste Symptomverbesserungen vom Beurteiler früher wahrgenommen werden, während es dem Patienten hingegen erst nach einer deutlichen Symptomverringerung gelingt, den eigenen Zustand differenzierter zu beurteilen und Verbesserungen wahrzunehmen [7, 17]. Auffällig ist zudem, dass die Medianwerte bereits bei Aufnahme der Patienten ungewöhnlich niedrig waren. Möglich ist, dass die Störung bei schweren Depressionen nicht ausreichend valide erfasst wurde – hier ist eine Ratingverzerrung durch die Klinikärzte anzunehmen. Zu geringe Kenntnis der deutschen Sprache sowie unterschiedliche Erhebungstageszeiten müssen ebenfalls in Betracht gezogen werden.

Weiterhin fiel auf, dass bei Entlassung der Patienten die Bewertung des CGI-I (starke Verbesserung) positiver gegenüber dem CGI-S (leicht krank) war. Zugrunde gelegt wird dieser Diskrepanz eine optimistischere Einschätzung der Skala CGI-I, da diese eher das allgemeine Wohlbefinden als die Stärke einzelner Symptome bewertet [12, 15].

Nach drei Behandlungswochen lagen die Responderaten der Patienten zwischen 38,4 % und 46,1 % und die Remissionsraten zwischen 11,2 % und 22,1 %.

Bei Entlassung der Patienten lagen die Responderaten der Patienten zwischen 67 % und 75,9 % und die Remissionsraten zwischen 58,9 % und 70,6 %. Vergleichbare Response- und Remissionsraten wurden mit 80,8 % und 57,9 % in der Studie MARS (Munich Antidepressant Response Signature) angegeben [9].

In der oben erwähnten Studie von Seemüller et al. lagen die HAMD-17-Response- und Remissionsraten mit 68,9 % (hier 74,1 %) bzw. 51,9 % (hier 62,0 %) ebenfalls etwa im Bereich der hier kalkulierten Ergebnisse. Auch zeigten sich in dieser Studie sehr schnell einsetzende Therapieeffekte [27].

MADRS-Remissionsgrenzwerte

Anhand des vorliegenden Datensatzes wies von den fünf untersuchten MADRS-Remissionsgrenzwerten ≤ 8 , ≤ 9 , ≤ 10 , ≤ 11 und ≤ 12 der MADRS-Remissionsgrenzwert ≤ 12 die höchste Übereinstimmung mit dem HAMD-17-Remissionsgrenzwert ≤ 7 auf (Kappa: 0,71). In einigen Studien kam dieser Remissionsgrenzwert ebenfalls zum Einsatz [1, 20]. In anderen wurden hingegen niedrigere Cut-off-Werte verwendet [2, 24, 39]. In der S3-Praxisleitlinie Unipolare Depression wird eine Depression sogar erst bei einem Skalenwert von ≤ 6 als remittiert definiert [6]. Auffällig ist allerdings, dass die hier kalkulierten MADRS-Remissionsraten zu den einzelnen Messzeitpunkten meist über denen der Skalen HAMD-17 und HAMD-7 liegen. Dies könnte ein Hinweis darauf sein, dass der Remissionsgrenzwert doch zu hoch gewählt wurde.

Korrelationen zwischen den Beurteilungsskalen

Die Korrelationen der drei Fremdbeurteilungsskalen HAMD-17, HAMD-7 und MADRS war sehr hoch und zu allen Messzeitpunkten signifikant ($p < 0,0005$). Der stärkste Zusammenhang bestand mit Werten zwischen 0,81 und 0,91 zwischen den Skalen HAMD-17 und HAMD-7. Die Korrelationskoeffizienten der Skalen HAMD-17 und MADRS waren mit Werten zwischen 0,74 und 0,90 etwas niedriger. Vergleichbare Studien kamen mit Werten zwischen 0,65 und 0,94 zu ähnlichen Ergebnissen [3, 16, 32]. Die in einer dieser Studien untersuchte Korrelation der Skalen HAMD-7 und MADRS wies mit Kor-

relationskoeffizienten zwischen 0,58 und 0,90 ebenfalls eine hohe Übereinstimmung mit den hier kalkulierten Werten zwischen 0,67 und 0,92 auf [16]. In den initialen drei Wochen der Therapie wich die Bewertung der Skala MADRS von der der Skalen HAMD-17 und HAMD-7 etwas ab. Erklären lässt sich dies mit Vorteilen der Skala MADRS gegenüber der Skala HAMD-17 beim Erfassen initialer Depressionssymptome [3]. Bei Entlassung der Patienten korrelierten mit Werten über 0,9 alle Fremdbeurteilungsskalen sehr stark miteinander.

Ein signifikanter Zusammenhang zwischen einer dieser Fremdbeurteilungsskalen und der Selbstbeurteilungsskala Bf-S ließ sich hingegen nicht nachweisen. Allerdings erhöhten sich die Korrelationskoeffizienten bei jeder Paarung im Verlauf der Therapie kontinuierlich. Die kalkulierten Werte liegen zwischen -0,20 und 0,29 (HAMD-17 vs. Bf-S), -0,30 und 0,20 (HAMD-7 vs. Bf-S) sowie -0,20 und 0,32 (MADRS vs. Bf-S). Zwischen der Einschätzung von Beurteilern (Fremdbeurteilung) und Patienten selbst (Selbstbeurteilung) konnte demnach kein Zusammenhang nachgewiesen werden. Erklären lässt sich dies damit, dass die Skalen im Sinne der multifaktoriellen Ätiopathogenese der Depression unterschiedliche Aspekte der Erkrankung bewerten. So dient die Skala Bf-S der Einschätzung des subjektiven Befindens, die Fremdbeurteilungsskalen hingegen der Einschätzung des Schweregrads einer Depression. Auch lassen sich manche Symptome besser durch einen Kliniker einschätzen (z. B. der psychomotorische Zustand), andere wiederum besser durch den Patienten selbst (z. B. sexuelle Symptome). Die Hauptursache wird aber darin gesehen, dass Kliniker besser zwischen einzelnen Symptomen unterscheiden können und diese stets im Kontext aller evaluierten Patienten sehen. Einzelne Symptomverbesserungen werden von ihnen daher sehr viel schneller registriert. Der Patient hingegen sieht nur seinen eigenen Zustand und kann meist schwer zwischen einzelnen Symptomen unterscheiden. Auch neigt ein Teil der Patienten dazu, sich selbst negativer einzuschätzen. Erst mit einer deutlichen Zustandsverbesserung im Verlauf der Therapie gelingt es ihnen, den eigenen Zustand differenzierter zu beurteilen. Die Übereinstimmung mit dem Kliniker nimmt dann nach und nach zu [5, 7, 17, 33]. Diese Annäherungstendenz zeigt sich auch im hier vorliegenden Datensatz. Wie allerdings oben beschrieben, muss hier auch in Betracht gezogen werden, dass die Depressionsschwere in der vorliegenden Studie nicht ausreichend valide erfasst wurde.

Wirksamkeitsunterschiede der Antidepressiva

Wirksamkeitsunterschiede zwischen den einzelnen Antidepressiva-Wirkstoffgruppen konnten in der initialen dreiwöchigen Therapiephase anhand der vorliegenden Daten nicht nachgewiesen werden. Die kalkulierten Werte der Regressionsanalyse zeigten zu keinem Zeitpunkt und bei keiner Wirkstoffgruppe signifikante Abweichungen einer Wirkstoffgruppe

vom Gesamtmittelwert. Die Ergebnisse decken sich demnach mit Studien, in denen postuliert wurde, dass zwischen den einzelnen Antidepressiva-Wirkstoffgruppen keine Wirksamkeitsunterschiede nachgewiesen werden können [8]. Anhand der unterschiedlichen initialen Response- und Remissionsraten deuteten sich allerdings tendenzielle Wirksamkeitsunterschiede zwischen den Antidepressiva-Wirkstoffgruppen an, die im hohen Maße mit den in der Literatur beschriebenen Unterschieden übereinstimmten. So deutete sich auch hier ein Wirksamkeitsvorteil für Escitalopram und Mirtazapin und ein Wirksamkeitsnachteil für Reboxetin an [4, 31]. Die Aussagekraft des Wirkstoffgruppenvergleichs war hier aber begrenzt. Problematisch waren neben den unterschiedlichen Gruppengrößen und der geringen Patientenzahl in den einzelnen Gruppen auch die verschiedenen Depressionsschweregrade der Patienten zu Studienbeginn. Zudem unterschied sich in den einzelnen Gruppen auch das Durchschnittsalter sowie die Geschlechter- und Hauptdiagnosenverteilung der Patienten. Auch konnten Wirksamkeitsunterschiede zwischen den hier zu einer Gruppe zusammengefassten Antidepressiva die Ergebnisse verfälschen. Hinsichtlich der differenzierten Antidepressiva-Bewertung können hier daher nur deskriptive Hinweise gegeben werden.

Fazit

Zusammenfassend belegten die Therapieverläufe sowie die Response- und Remissionsraten die Wirksamkeit einer stationären Antidepressiva-Therapie unter alltäglichen Bedingungen. Etwa zwei Drittel der Patienten profitierten von der Behandlung. Unterschiedliche Remissionsgrenzwerte können die Ergebnisse von Wirksamkeitsstudien stark verzerren und die Vergleichbarkeit von Ergebnissen einschränken. Es sollte daher dringend ein Konsens über einen geeigneten MADRS-Remissionsgrenzwert gefunden werden.

Die beschriebene hohe Korrelation der Fremdbeurteilungsskalen HAMD-17, HAMD-7 und MADRS zeigte sich auch anhand der vorliegenden Daten. Mit der Selbstbeurteilungsskala Bf-S korrelierte hier hingegen keine der drei Skalen. Neben der in Betracht zu ziehenden Erhebungsproblematik dieser Skala scheint zudem eine Skalenart die Heterogenität einer Depression nicht vollständig erfassen zu können. Die verschiedenen Aspekte der Depression können demnach sehr viel besser bewertet werden, wenn beide Skalenarten verwendet werden [7, 33]. In zukünftigen Studien sollte dies berücksichtigt werden und dementsprechend beide Skalenarten zum Einsatz kommen.

Der Fokus dieser Studie lag auf der Auswertung der Beurteilungsskalen und damit der Wirksamkeit aller eingesetzten Therapien insgesamt. Zur Bewertung einzelner Komponenten einer Behandlung war die vorliegende Studie nur bedingt geeignet. Hierbei handelte es sich um einen bekannten Nachteil randomisierter pragmatischer Studien [22]. Um zukünftig

sinnvolle Aussagen zur differenziellen Wirksamkeit von Antidepressiva in der Routinepraxis treffen zu können, sind demnach weitere randomisierte pragmatische Studien mit größeren Fallzahlen notwendig.

Interessenkonflikterklärung

CCH: Keine Interessenkonflikte.

GD: Keine Interessenkonflikte.

GL: Honorare für Vorträge und Beratungen oder unbeschränkte Subventionen für Schulungen von Bayer, Janssen-Cilag, Lundbeck, Pfizer und Servier.

Die Arbeit ist eine Zusammenfassung der abgeschlossenen Dissertation von CCH an der LMU München.

Literatur

- Boulenger JP, Huusom AK, Florea I, Baekdal T, 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.
- Carmody TJ, Rush AJ, Bernstein I, Warden D, et al. The Montgomery Asberg and the Hamilton ratings of depression: a comparison of measures. *Eur Neuropsychopharmacol* 2006;16:601–11.
- Carneiro AM, Fernandes F, Moreno RA. Hamilton depression rating scale and montgomery-asberg depression rating scale in depressed and bipolar I patients: psychometric properties in a Brazilian sample. *Health Qual Life Outcomes* 2015;13:42.
- Cipriani A, Furukawa TA, Salanti G, Geddes JR, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet* 2009;373:746–58.
- Collegium Internationale Psychiatriae Scalarum (CIPS) (Hrsg.). Göttingen: Beltz Test, 2015.
- DGPPN, BÄK, KBV, et al. S3-Praxisleitlinie/Nationale Versorgungsleitlinie unipolare Depression. 2. Auflage. Berlin: Springer, 2015.
- Dunlop BW, Li T, Kornstein SG, Friedman ES, et al. Correlation between patient and clinician assessments of depression severity in the PREVENT study. *Psychiatry Res* 2010;177:177–83.
- Gartlehner G, Hansen RA, Morgan LC, Thaler K, et al. Comparative benefits and harms of second-generation antidepressants for treating major depressive disorder: an updated meta-analysis. *Ann Intern Med* 2011;155:772–85.
- Hennings JM, Owashi T, Binder EB, Horstmann S, et al. Clinical characteristics and treatment outcome in a representative sample of depressed inpatients – findings from the Munich Antidepressant Response Signature (MARS) project. *J Psychiatr Res* 2009;43:215–29.
- IQWiG, GFR. Diskussionsforum zur Nutzenbewertung im Gesundheitswesen. https://www.iqwig.de/download/DLR_Nutzenbewert_07-11-22_Druckversion.pdf (Zugriff am 06.11.2015).
- Jacobi F, Hoyer M, Strehle J, Mack S, et al. Mental disorders in the general population: Study on the health of adults in Germany and the additional module mental health (DEGS1-MH). *Nervenarzt* 2014;85:77–87.
- Jiang Q, Ahmed S. An analysis of correlations among four outcome scales employed in clinical trials of patients with major depressive disorder. *Ann Gen Psychiatry* 2009;8:4.
- Katz MM, Tekell JL, Bowden CL, Brannan S, et al. Onset and early behavioral effects of pharmacologically different antidepressants and placebo in depression. *Neuropsychopharmacology* 2004;29:566–79.
- Kjaergaard M, Arfwedson Wang CE, Waterloo K, Jorde R. A study of the psychometric properties of the Beck Depression Inventory-II, the Montgomery and Asberg Depression Rating Scale, and the Hospital Anxiety and Depression Scale in a sample from a healthy population. *Scand J Psychol* 2014;55:83–9.
- Lee TS, Nkouibert Assam P, Gersing KR, Chan E, et al. The effectiveness of antidepressant monotherapy in a naturalistic outpatient setting. *Prim Care Companion CNS Disord* 2012;14: pii: PCC.12m01364.
- McIntyre RS, Konarski JZ, Mancini DA, Fulton KA, et al. Measuring the severity of depression and remission in primary care: validation of the HAMD-7 scale. *Can Med Assoc J* 2005;173:1327–34.
- Möller H-J. Outcome criteria in antidepressant drug trials: self-rating versus observer-rating scales. *Pharmacopsychiatry* 1991;24:71–5.
- Möller H-J. Outcomes in major depressive disorder: the evolving concept of remission and its implications for treatment. *World J Biol Psychiatry* 2008;9:102–14.
- Möller H-J, Laux G, Deister A (Hrsg.). *Duale Reihe Psychiatrie, Psychosomatik und Psychotherapie*. 6. Auflage. Stuttgart: Thieme, 2015.
- 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.
- Papakostas GI, Perlis RH, Scalia MJ, Petersen TJ, et al. A meta-analysis of early sustained response rates between antidepressants and placebo for the treatment of major depressive disorder. *J Clin Psychopharmacol* 2006;26:56–60.
- Patsopoulos NA. A pragmatic view on pragmatic trials. *Dialogues Clin Neurosci* 2011;13:217–24.
- Quitkin FM, McGrath PJ, Stewart JW, Taylor BP, et al. Can the effects of antidepressants be observed in the first two weeks of treatment? *Neuropsychopharmacology* 1996;15:390–4.
- Riedel M, Möller HJ, Obermeier M, Schennach-Wolff R, et al. Response and remission criteria in major depression – a validation of current practice. *J Psychiatr Res* 2010;44:1063–8.
- Ruppert T. Anwendungsbeobachtungen: Mehr Akzeptanz durch Transparenz. *Deutsche Zeitschrift für klinische Forschung: DZKF* 2010;2010:20–3.
- Rush AJ, Kraemer HC, Sackeim HA, Fava M, et al. Report by the ACNP Task Force on response and remission in major depressive disorder. *Neuropsychopharmacology* 2006;31:1841–53.
- Seemuller F, Riedel M, Obermeier M, Bauer M, et al. Outcomes of 1014 naturalistically treated inpatients with major depressive episode. *Eur Neuropsychopharmacol* 2010;20:346–55.
- Stassen HH, Angst J, Hell D, Scharfetter C, et al. Is there a common resilience mechanism underlying antidepressant drug response? Evidence from 2848 patients. *J Clin Psychiatry* 2007;68:1195–205.
- Szegedi A, Jansen WT, van Willigenburg AP, van der Meulen E, et al. Early improvement in the first 2 weeks as a predictor of treatment outcome in patients with major depressive disorder: a meta-analysis including 6562 patients. *J Clin Psychiatry* 2009;70:344–53.
- Tadic A, Helmreich I, Mergl R, Hautzinger M, et al. Early improvement is a predictor of treatment outcome in patients with mild major, minor or sub-syndromal depression. *J Affect Disord* 2010;120:86–93.
- Thase ME, Nierenberg AA, Vrijland P, van Oers HJ, et al. Remission with mirtazapine and selective serotonin reuptake inhibitors: a meta-analysis of individual patient data from 15 controlled trials of acute phase treatment of major depression. *Int Clin Psychopharmacol* 2010;25:189–98.
- Uher R, Perlis RH, Placentino A, Dernovsek MZ, et al. Self-report and clinician-rated measures of depression severity: can one replace the other? *Depress Anxiety* 2012;29:1043–9.
- Vares EA, Salum GA, Spanemberg L, Caldieraro MA, et al. Depression dimensions: integrating clinical signs and symptoms from the perspectives of clinicians and patients. *PLoS One* 2015;10:e0136037.
- Vos T, Flaxman AD, Naghavi M, Lozano R, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012;380:2163–96.
- Weiß C. *Basiswissen medizinische Statistik*. 5. Aufl. Heidelberg: Springer, 2010.
- Willich SN. Randomisierte kontrollierte Studien: Pragmatische Ansätze erforderlich. *Dtsch Arztebl Int* 2006;103:A-2524.
- Wisniewski SR, Rush AJ, Nierenberg AA, Gaynes BN, et al. Can phase III trial results of antidepressant medications be generalized to clinical practice? A STAR*D report. *Am J Psychiatry* 2009;166:599–607.
- Zimmerman M, Martinez JH, Young D, Chelminski I, et al. Severity classification on the Hamilton Depression Rating Scale. *J Affect Disord* 2013;150:384–8.
- Zimmerman M, Posternak MA, Chelminski I. Derivation of a definition of remission on the Montgomery-Asberg depression rating scale corresponding to the definition of remission on the Hamilton rating scale for depression. *J Psychiatr Res* 2004;38:577–82.