

Warum Antidepressiva (AD) nicht so heißen sollten – sie wirken vermutlich nicht besser als ein Placebo

Martin Plöderl

13.10.2018

Vorbemerkung: Warum diese kritischen Informationen?

Leider ist es so, dass psychiatrische Fachpublikationen den Wissensstand eher einseitig wiedergeben, und leider ist das auch in Leitlinien zur klinischen Behandlung der Fall, wie das vielfach nachgewiesen wurde (z.B. Cosgrove et al. 2017, Hengartner, 2017, Turner et al., 2008). Meiner Erfahrung nach ist das auch bei wissenschaftlichen Kongressen nicht anders.

Angesichts dieser fundamentalen Probleme, die letztendlich das Wohl der Patient*innen gefährden könnten, möchte ich interessierten Psychiater*innen und anderen Personen, die in der Versorgung von psychisch kranken Menschen arbeiten, kritische Texte zur Wirksamkeit von Psychopharmaka zur Verfügung stellen, die sie wohl kaum von Pharmareferent*innen oder hochrangigen Psychiater*innen auf Konferenzen und in Fortbildungen erfahren werden, damit sie sich selber ein Bild machen können. *Ich stelle diese Information als Privatperson zur Verfügung und vertrete damit auch meine private Meinung.*

Cosgrove, L., Krinsky, S., Wheeler, E. E., Peters, S. M., Brodt, M., & Shaughnessy, A. F. (2017). Conflict of interest policies and industry relationships of guideline development group members: a cross-sectional study of clinical practice guidelines for depression. *Accountability in research*, 24(2), 99-115. [Download](#)

Hengartner, M. P. (2017). Methodological flaws, conflicts of interest and scientific fallacies: Implications for the evaluation of antidepressants' efficacy and harm. *Frontiers in Psychiatry*, 8, 275. [Download](#)

Turner, E. H., Matthews, A. M., Linardatos, E., Tell, R. A., & Rosenthal, R. (2008). Selective publication of antidepressant trials and its influence on apparent efficacy. *New England Journal of Medicine*, 358(3), 252-260. [Download](#)

Wer sich mit depressiven Symptomen an Ärzt*innen wendet, wird mit hoher Wahrscheinlichkeit ein sogenanntes Antidepressivum empfohlen bzw. verordnet bekommen. Die Wirksamkeit dieser Medikamente ist jedoch schon seit längerem umstritten (z.B. Kirsch u. a., 2008).

Anfang 2018 erschien die vermutlich größte Meta-Analyse zur Wirksamkeit von verschiedenen Antidepressiva (AD) (Cipriani u. a., 2018), und in den Medien war zu hören und lesen, dass AD unumstritten wirken. Zum Beispiel schrieb die [BBC](#): „Forscher*innen sagen dass sie nun eine der größten Kontroversen in der Medizin geklärt haben, nachdem ihre riesige Studie zeigte dass Antidepressiva wirken“, und ähnliche Schlussfolgerungen waren im [ORF](#), deutschen [Ärzteblatt](#), oder in der [Apothekerzeitung](#) zu lesen.

Die Schlussfolgerung, dass AD wirken, ist angesichts der aktuellen Evidenzlage kaum haltbar. Vermutlich wirken sie nicht besser oder nur vernachlässigbar besser als ein Placebo, und was übrig bleibt, sind Nebenwirkungen, zu denen wahrscheinlich auch Suizide zählen. Im Folgenden möchte ich dies kurz erklären, greife dabei aber ohnehin nur auf Argumente zurück, die schon lange von kompetenten Forscher*innen vertreten werden (David Healy, Peter Gøtzsche, Irving Kirsch, Michael Hengartner). Ich kann nur versuchen, die Argumente kurz und anschaulich wiederzugeben und auch neuere Studien zu berücksichtigen. Zusammen mit Michael Hengartner sind einige

der folgenden Argument mittlerweile auch in Fachzeitschriften abgedruckt worden (Hengartner & Plöderl, 2018b, 2018a).

1. Was ist mit „Antidepressiva wirken“ eigentlich gemeint?

In ihrer bemerkenswerten Studie finden Cipriani et al. dass alle AD besser wirken als ein Placebo. Dazu zogen sie sogenannte randomisiert kontrollierte Studien heran, bei denen entweder ein Medikament mit einem anderen Medikament verglichen wurde, oder mehrere Medikamente mit Placebo. Bei diesen Studien werden Patient*innen zufällig in die Medikamenten- oder die Placebogruppe eingeteilt (siehe Abb. 2a), ohne dass sie wissen, in welcher Gruppe sie landen. Auch die behandelnden Ärzt*innen wissen das nicht (deswegen die Bezeichnung „Doppelblindstudie“). Solche Studien liefern die stärkste wissenschaftliche Evidenz, weil mit diesem strengen Studiendesign Unterschiede zwischen der Placebogruppe und der Medikamentengruppe nur auf die Wirkung des Medikamentes zurückzuführen sind (aber auch das ist umstritten, s.u.). In dieser Hinsicht gibt es an der Meta-Analyse von Cipriani et al. nichts auszusetzen. Die Evidenzlage der sogenannten Alternativmedizin ist bei weitem nicht so gut. Problematisch ist hingegen die Schlussfolgerung der Studie bzw. wie diese in vielen Medien interpretiert wurde.

Nur in einer gewissen Weise kann die Aussage „Antidepressiva wirken besser als Placebo“ überhaupt vertreten werden. In der Meta-Analyse wurde gefunden, dass Patient*innen unter AD nach einer Behandlungszeit von ca. 8 Wochen weniger depressive Symptome aufwiesen als Patient*innen unter Placebo. Um das zu messen, führten Kliniker*innen in den Studien regelmäßig mit den Patient*innen Interviews durch, bei denen die depressiven Symptome erfasst wurden, meist mit der „Hamilton Depression Rating Scale“ (HAMD). Durch die enorm hohe Anzahl an Patient*innen in der Meta-Analyse (n = 116.477) ist die Einschätzung der Wirksamkeit von AD sehr genau, und auch statistisch signifikant. Aber was heißt „Medikamente wirken besser als Placebo“ in diesem Fall? Oder technisch gesprochen, was bedeutet „statistisch signifikant“? Statistiker*innen wissen, dass die statistische Signifikanz kaum etwas über die Größe eines Unterschiedes aussagt. Bei einer extrem großen Stichprobe, wie das in der Studie von Cipriani et al. der Fall war, wird auch noch der kleinste Unterschied statistisch signifikant (für eine fachliche Kritik am Signifikanztest siehe J. Cohen, 1994). Was aber zählt, und da braucht man nur Hausverstand und keine Statistikausbildung, ist die Größe eines Effektes. Einfacher gesagt, man möchte doch wissen, um *wieviel* besser ein Medikament im Vergleich zu einem Placebo wirkt und ob das irgendwie *klinisch bedeutsam (klinisch signifikant)* ist.

Darüber ist in der Studie von Cipriani nur ein einziger Satz zu finden („and the summary effect sizes were mostly modest“, S. 7). Das könnte man mit „mäßig“ oder „bescheiden“ übersetzen. Dennoch ist das eine Übertreibung. Die fachlich korrekte Interpretation müsste lauten: *Die Wirksamkeit von AD im Vergleich zu Placebo ist klinisch bedeutungslos bzw. klinisch nicht signifikant.* Dies wird im folgenden erläutert.

2. Antidepressiva haben keine klinisch bedeutsame Wirksamkeit

Die Wirksamkeit von AD im Vergleich zu Placebo ist so klein, dass sie als vernachlässigbar, oder anders ausgedrückt, klinisch nicht relevant, zu erachten ist.

Um von „klinisch wirksam“ oder „klinisch relevant/signifikant“ zu sprechen, muss der Unterschied zwischen Medikament und Placebo eine bestimmte Mindestgröße erreichen. Über die Wahl der Mindestgröße wird zwar diskutiert, aber bei AD ist es so, dass sie in keinem Fall irgendeine der gängigen Mindestgrößen bezüglich klinischer Signifikanz erreichen. Die gefundenen Effekte sind einfach zu klein. Es gibt folgende üblichen Zugänge um festzustellen, ob eine Wirkung als „klinisch Signifikant“ zu beurteilen ist (Hengartner, 2018; siehe Jakobsen, 2017; Moncrieff & Kirsch, 2015)

2.1. Kriterium 1: *Effektstärke nach Cohen.*

Die Effektstärke d wurde von Jacob Cohen vorgeschlagen und berechnet sich als standardisierter Mittelwertsunterschied. Per Konvention ist eine Effektstärke von $d \leq 0,5$ als klein zu bezeichnen, $0,5 \leq d < 0,8$ als mittelgroß, und darüber hinaus spricht man von großen Effekten (Jacob Cohen, 1988). Bei Cipriani et al. wurde eine Effektstärke von $d = 0,3$ über alle Medikamente hinweg im Vergleich zu Placebo gefunden (siehe Abbildung 1). Diese kleine Effektstärke wird übrigens im Text nie genauer erläutert, und die entsprechende Grafik ist nur im Online-Appendix (S. 150) zu sehen. Die genannte Einteilung von Effektstärken ist von Cohen selbst kritisch gesehen worden, weil es sich um eine reine Konvention handelt, und das bedeutet inhaltlich noch nicht viel. Auch sehr kleine Effekte können in gewissen Anwendungsgebieten klinisch bedeutsam sein, z.B. wenn ein Medikament schwerwiegende Nebenwirkungen wie etwa Suizide verursacht.

2.2. Kriterium 2: *Effektstärke nach klinischer Bemerkbarkeit.*

Nach diesem Kriterium ist die Wirksamkeit eines Medikamentes dann klinisch relevant, wenn eine Veränderung durch das Medikament gegenüber Placebo auch beobachtbar/wahrnehmbar ist. Mittlerweile gibt es Studien darüber, welche Ausprägung an Symptomveränderungen von Ärzt*innen überhaupt wahrgenommen werden können. Wie schon erwähnt verwendeten die meisten Studien die Hamilton Depression Rating Scale mit 17 Fragen (HAMD-17), bei der man zwischen 0 und 52 Punkte erreichen kann. Wenn Ärzt*innen in den Studien den Eindruck hatten, dass sich Patient*innen „nicht verbessert“ hatten, dann entsprach das einer Punktedifferenz von 0-3 Punkten auf der HAMD-17. Anders ausgedrückt, eine Punktedifferenz ≤ 3 ist nicht als Veränderung wahrnehmbar (Moncrieff & Kirsch, 2015). Diese Wahrnehmbarkeitsgrenze entspricht einer Effektstärke von $d = 0,375$ und ist also weniger streng als die Klassifikation nach Cohen, nach der $d \leq 0,5$ kleine Effekte sind. *Dennoch liegt die Effektivität der AD unter diesem minimal geforderten Maß!* Diese Grenze ist wirklich lächerlich klein, denn was sind 3 Punkte Veränderung auf einer Skala von 0 bis 52 Punkten? Dennoch versuchten renommierte Psychiater, selbst diese minimale Grenze zu verteidigen (Hegerl & Mergl, 2010; Möller, 2008). Vielleicht würden sie das heute nicht mehr tun, seit dem bekannt wurde, dass erst ab einer Differenz von 7 Punkten (entspricht $d = 0,875$) überhaupt eine „minimale Verbesserung“ von Ärzt*innen wahrgenommen werden kann. Davon liegen die Effekte bei AD im Vergleich zu Placebo meilenweit entfernt, wie auch in Abbildung 1 ersichtlich ist. In der Meta-Analyse von Cipriani et al. betrug die Effektstärke nur $d = 0,3$. Der Unsicherheitsbereich (95% - Konfidenzintervall) lag dabei zwischen 0,26-0,34, d.h. auch die Obergrenze dieses Intervalls liegt unter der Grenze der klinischen Bedeutsamkeit. Anders

ausgedrückt, ÄrztInnen könnten nach 8 Wochen Behandlung ganz einfach nicht feststellen, welcher Patient aus der Placebo- oder Medikamentengruppe stammte. Wie Abbildung 1 zeigt, liegen auch fast alle einzelnen Medikamente klar unter der klinischen Signifikanzgrenze, und keines kommt auch nur annähernd in einen Bereich, wo es wahrnehmbar Placebo überlegen wäre. Interessant ist, dass das wirksamste Medikament Amitriptylin war, ein altes sogenanntes „Trizyklika“. Dieses Ergebnis basiert aber auf älteren und kleinen Studien mit hohem Risiko für Verzerrungseffekte, wie Cipriani et al. anmerkten, und es ist daher mit Vorbehalt zu interpretieren.

2.2.1. Kriterium 3: Anzahl der notwendigen Behandlungen (Number needed to treat, NNT): Die Effektstärke kann man auch zum einfacheren Verständnis in die Anzahl der notwendigen Behandlung übersetzen, die man machen muss, damit ein zusätzlicher Patient vom Medikament im Vergleich zu Placebo profitiert. Im Falle von AD ist NNT ca. 9, bei einer angenommenen Effektstärke von $d = 0.30$ und einer Placebo-Responderrate von 40% (T. A. Furukawa u. a., 2016). Das heißt man muss 9 depressive Patient*innen behandeln, damit einer dabei ist, der von der Behandlung profitiert, im Vergleich zu Placebobehandlungen. Das heißt aber auch, dass acht Patient*innen nicht vom Antidepressivum profitieren (bzw. gleich gut wie von Placebo). Diese acht Patient*innen sind jedoch dem Risiko für Nebenwirkungen ausgesetzt. Über Nebenwirkungen werden Patient*innen übrigens fast nie aufgeklärt (McLaren, 2018; Read, Cartwright, & Gibson, 2018).

Anm: der R-code lautet: $d = 0.3$; $CER = 0.4$; $NNT = 1 / (pnorm(d + qnorm(CER)) - CER)$; `print(NNT)`, siehe <http://rpsychologist.com/d3/cohend/> und (Toshi A. Furukawa & Leucht, 2011)

*Egal wie man es dreht und wendet, die Überlegenheit der AD gegenüber Placebo ist vernachlässigbar klein, keinesfalls von Kliniker*innen wahrnehmbar, und daher klinisch nicht relevant bzw. klinisch nicht signifikant.*

Es ist sehr bedenklich, dass die Autor*innen dieses so wichtige Faktum nicht erwähnten, obwohl sich unter ihnen sogar der sehr kritische Forscher John Ioannides befand, von dem annehmen kann, dass er sich dieser Limitationen bewusst war. Unter der Federführung von Michael Hengartner, einem Psychologen aus der Schweiz, schrieben wir die eben dargestellten Hauptargumente in einem Brief an den Herausgeber von „Lancet“. Der Brief wurde ohne Angabe von Gründen abgelehnt (siehe hier <http://drmartinploderl.zohosites.com/files/hengartner-ploederl-2018-response-cipriani-rejected-lancet.pdf>).

Diese kritischen Argumente sind nicht neu. Schon 2008 fanden Kirsch et al. in ihrer Meta-Analyse eine fast gleiche Effektstärke von $d = 0,32$ (Kirsch u. a., 2008). Letztes Jahr erschien eine Meta-Analyse von Jakobsen et al. (2017), die ebenfalls zu sehr ähnlichen Ergebnissen kam (siehe graue Linie in Abbildung 1). Im Unterschied zu Cipriani et al. gingen Jakobsen et al. den wissenschaftlich seriöseren Weg und definierten im Vorhinein, was sie für klinisch relevant halten und was nicht, denn eine statistische Signifikanz liefert diese Information ja nicht. Sie kamen zum Schluss, dass AD gegenüber Placebo klinisch nicht signifikant überlegen sind.

Wie kann es sein, dass nahezu gleiche Ergebnisse von zwei unterschiedlichen Forschungsgruppen

so unterschiedlich interpretiert werden? Vielleicht liegt es an Interessenskonflikten?

Die Autor*innen der Meta-Analyse von Jakobsen et al. scheinen unabhängig zu sein, während einige Autor*innen der Meta-Analyse von Cipriani et al. unter anderem von folgenden Pharmafirmen finanzielle Zuwendungen erhielten: Eli Lilly, Janssen, Merck Sharp & Dohme, Pfizer Lundbeck, SanofiAventis. Zudem kommt, dass Herausgeber hochrangiger Fachzeitschriften direkt Geld von Pharmafirmen bekommen. Henry Nasrallah, Editor in Chief von „Current Psychiatry“, erhielt von 2013 bis 2016 fast eine Million Dollar, siehe [hier](#).

Es gibt übrigens eine ähnliche Meta-Analyse zu AD bei Kindern und Jugendlichen, ebenfalls unter der Federführung von Cipriani (Cipriani u. a., 2016). Die Effektstärken waren in der Regel noch kleiner als die in der Meta-Analyse bei Erwachsenen und rangierten zwischen $d = 1.14$ (Placebo besser als Nortryptiline) und $d = -0.51$ (Flouxetine besser als Placebo). Allerdings waren die meisten Effektstärken kleiner als $d = 0.3$, nur Douloxetine ($d = -0.35$), Desimpramine ($d = -0.45$) und Flouxetine ($d = -0.51$) waren geringfügig besser, und bis auf Flouxetine nicht statistisch signifikant. Die Autoren schließen im Abstract auch skeptischer als in der Erwachsenen-Metaanalyse: „Unter Berücksichtigung des Risiko-Nutzen Profils von AD in der Akutbehandlung von Depression scheinen diese Medikamente keinen klaren Vorteil für Kinder und Jugendliche zu bringen“. Wenn man die Studien einzeln unter die Lupe nimmt, stellt sich heraus, dass alle 20 Studien, die zwischen 1990 und 2006 erschienen, negativ waren (im Sinne der Studienhypothese), und dass die meisten Studien auch mehr Suizidalität unter AD als unter Placebo fanden (Healy, Le Noury, & Jureidini, 2018). Auch die 15 Studien, die seit 2006 erschienen, fanden negative Ergebnisse und meist mehr Suizidalität unter AD.

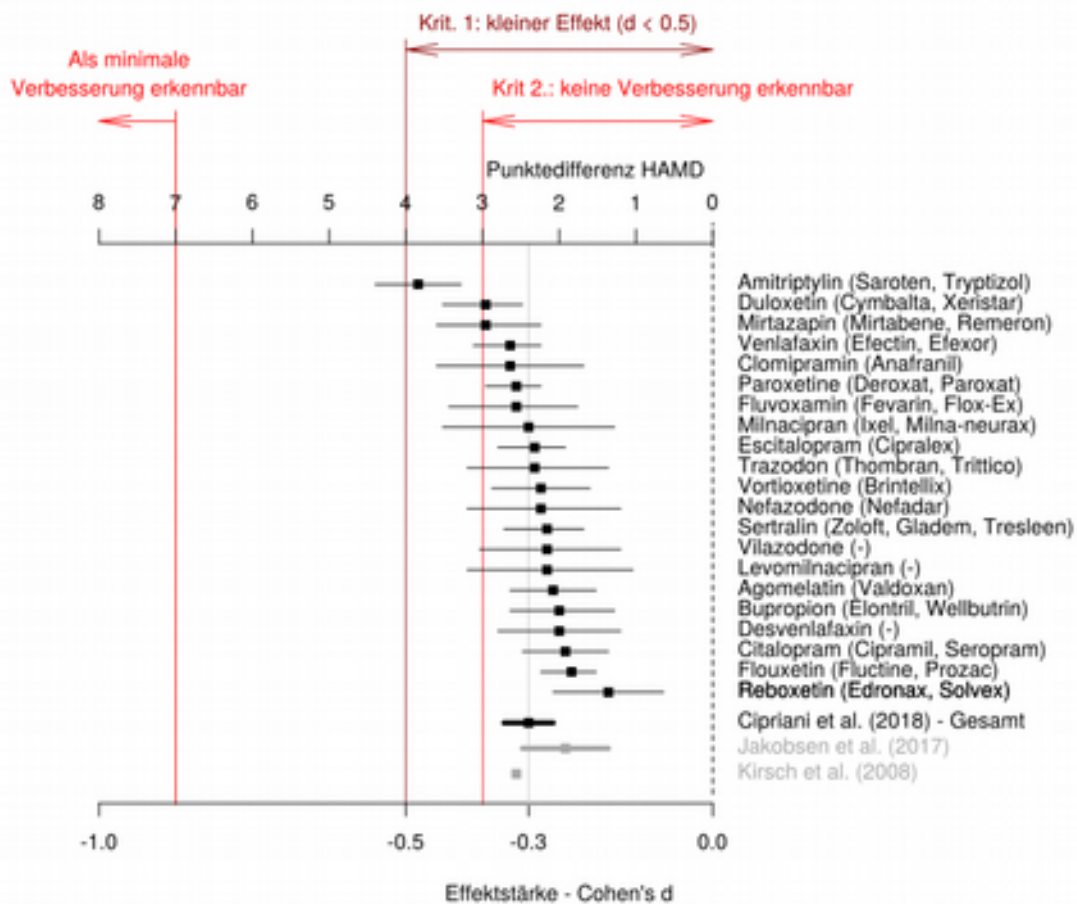


Abbildung 1. Ergebnisse der Meta-Analyse von Cipriani et al. (2018), angelehnt am Online-Appendix S. 150. Horizontale Linien entsprechen den 95% Konfidenzintervallen (bzw. den Bayesianischen Vertrauensintervallen in Cipriani et al.), die Quadrate der Punktschätzung der Effektivität von AD, jeweils im Vergleich zu Placebo. In Klammer sind exemplarisch Handelsnamen der Medikamente angeführt. Fette horizontale Linien sind die Ergebnisse über alle AD hinweg (gegenüber Placebo). Die grauen Linie ist das Gesamtergebnis der Meta-Analyse von Jakobsen et al. (2017), die Meta-Analyse von Kirsch et al. (2008) ist nur als Punktschätzung eingezeichnet, da kein Konfidenzintervall berichtet wurde.

3. Die minimale, klinisch unbedeutende Wirksamkeit von Antidepressiva ist ziemlich sicher eine Überschätzung der tatsächlichen Wirksamkeit.

Zwar liefern die klinischen Studien mit einem Randomisierungs- und Doppelblinddesign den höchsten Evidenzgrad, doch sind auch sie nicht frei von Fehlerquellen. Das wäre auch nicht problematisch, wenn die Fehlerquellen unsystematisch wären. *Die Fehlerquellen verzerren in der Regel aber die Ergebnisse der Studien so, dass AD besser abschneiden als Placebo.* Auch dazu gibt es mittlerweile eine ganze Forschungslinie, die sich mit diesen Fehlern auseinandersetzt (Ansari, Ansari, & Müller-Oerlinghausen, 2016; Gøtzsche, 2016; Hengartner, 2017). Im Folgenden eine Auflistung der wichtigsten Fehlerquellen.

3.1. *Fehlende Verblindung auf Ärzt*innenseite*: Im Vergleich zu Placebo haben AD bestimmte Nebenwirkungen, die unter Umständen von den Studienärzt*innen wahrgenommen werden können. Damit ist die Verblindung nicht mehr garantiert. Dies führt dazu, dass Ärzt*innen nun wissen, welche Patient*innen in der Medikamenten- oder Placebogruppe sind. Da es die Ärzt*innen sind, welche per Interview die depressive Symptomatik feststellen, könnten sie bewusst oder unbewusst dazu tendieren, die depressive Symptomatik bei Patient*innen mit Medikamenten besser zu beurteilen als bei jenen mit Placebo. Tatsächlich zeigen Studien, dass nicht-verblindete Studienleiter*innen die Wirksamkeit von Medikamenten um etwa 36% überschätzen (Hrobjartsson, z.n. Gøtzsche, 2016, S. 53) *Alleine diese Fehlerquelle könnte die gesamte gefundene Überlegenheit von AD gegenüber Placebo erklären*. Wie Gøtzsche in seinem Buch demonstriert, geht die Wirksamkeit von AD auf Null, wenn nur 5% durch fehlende Verblindung falsch eingeschätzt werden (Gøtzsche, 2016, S. 53). Die geringe Wirksamkeit von AD geht auch dann auf nahezu null, wenn anstelle von Interviews, die von Ärzt*innen durchgeführt werden, Fragebögen verwendet werden, die von Patient*innen selber ausgefüllt werden (siehe Gøtzsche, 2016, S. 58). Auch in einer Meta-Analyse an Kindern und Jugendlichen wurden kleine und signifikante Effekte bei Kliniker*innen-Einschätzungen gefunden (3,5 Punkte auf der CDRS-R Skala), aber nahezu Null-Unterschiede bei Selbsteinschätzungen (0,5 Punkte beim BDI) (Hetrick, McKenzie, Cox, Simmons, & Merry, 2012).

3.2. *Fehlende Verblindung auf Patient*innenseite*: AD verursachen wie gesagt spezifische Nebenwirkungen, die Placebos nicht verursachen. Daher wäre ein ideales Placebo ein solches, das die Nebenwirkungen des Medikamentes nachahmt, aber sonst keine Wirkung zeigt (ein sogenanntes aktives Placebo). Für die neueren AD gibt es leider keine Studien mit aktiven Placebos. Für die alten AD, die sog. Trizyklika, gab es ein paar Studien mit aktiven Placebos, welche ebenso wie die Trizyklika Mundtrockenheit bewirkten. In diesen Studien betrug der Unterschied zwischen AD und Placebos nur $d = 0.17$ (0.00-0.37) (unter Ausschluss einer Studie mit extrem großer Wirkung) (Moncrieff, Wessely, & Hardy, 2004). Eine Erklärung könnte sein, dass Patient*innen aufgrund der Nebenwirkung wissen, ob sie das Placebo oder das Medikament bekommen. Dadurch glauben Patient*innen, etwas Wirksames zu bekommen, was den Placeboeffekt verstärkt.

3.3. *Placebopatient*innen haben Absetzsymptome*: In Studien nehmen auch Patient*innen teil, die schon in einer laufenden AD-behandlung standen. Einige Tage vor der zufälligen Zuordnung der Patient*innen in die AD- oder die Placebogruppe, werden die Medikamente abgesetzt („lead-in Phase“, siehe Abb. 2a). Bekanntlich kommt es durch ein abruptes Absetzen von AD zu Absetzsymptomen, die denjenigen einer Depression sehr ähneln können. Diese Absetzsymptome treten durchaus häufig und zum Teil verzögert auf (Davies & Read, 2018; Fava u. a., 2018; Fava, Gatti, Belaise, Guidi, & Offidani, 2015). Das heißt, dass es in der Placebogruppe wahrscheinlich Personen mit Absetzsymptomen gibt, nicht aber in der AD-gruppe, weil dort die Absetzsymptome natürlich durch die Medikamente wieder verschwinden. Dies führt zu einem künstlich erzeugten günstigen Effekt für die Antidepressiva.

3.4. *Wirken Antidepressiva oder doch die Benzodiazepine?*

In vielen Studien sind auch Benzodiazepine erlaubt, so bei zumindest 71% der Studien im Fall von Schlafstörungen, laut einer Meta-Analyse (Sharma, Guski, Freund, & Götzsche, 2016) Es gibt meines Wissens noch keine Studien, die zeigen, inwieweit Benzodiazepine als Zusatzmedikation das Ergebnis beeinflusst, und in welche Richtung dies die Ergebnisse verzerrt. So könnten Nebenwirkungen der AD wie Unruhe zur mehr Verabreichungen von Benzodiazepin führen in der AD-Gruppe führen und so die Medikamentengruppe bevorteilen. Andererseits könnten durch die Absetzsymptome in der Placebogruppe Nebenwirkungen wie Akathisia mit Benzodiazepinen behandelt worden sein, was wiederum einen Vorteil in der Placebo-Gruppe darstellen könnte. Meine Vermutung ist, dass dies entweder noch nie systematisch ausgewertet wurde, oder es wurde ausgewertet aber nicht publiziert, weil sich somit herausgestellt hätte, dass Wirkung der AD überschätzt wird.

3.5. *Selektierte Patient*innen:*

Die meisten Wirksamkeitsstudien schließen Patient*innen, wie sie in der täglichen Praxis erscheinen, aus. Zum Beispiel wenn sie Suizidgedanken oder ein anderweitig erhöhtes Suizidrisiko haben, wie das in zumindest 63% der Studien in einer Meta-Analyse der Fall war (Sharma u. a., 2016). Die Generalisierbarkeit der klinischen Studien ist anzuzweifeln, denn Studien mit „normalen“ Patient*innen, wie etwa in der STAR*D Studie, fanden deutlich niedrigere Ansprechraten (34%) sowie sehr schlechte langfristige Erfolge (Pigott, 2011, 2015; Rush u. a., 2004). Die Veröffentlichungen zur STAR*D Studie sind ebenfalls wieder Negativbeispiele, die zeigen, dass Studienergebnisse verzerrt wiedergegeben werden, wie aus den Publikationen von Pigott hervorgeht. Übrigens werden auch in 86% der Studien PatientInnen ausgeschlossen, die sich in der im Schnitt 7-Tage langen Placebo-Einleitungsphase verbesserten. Wieviele PatientInnen das jeweils waren, wurde kaum berichtet (Sharma u. a., 2016) - auch das führt vermutlich zu einer Überschätzung der Wirksamkeit von AD.

3.6. *Ansprechquoten anstelle von tatsächlichen Unterschieden:*

Viele Meta-Analysen berichten nicht die Ausprägung der Depressivität am Behandlungsende sondern nur die Ansprechquoten, das heißt PatientInnen werden nach einer gewissen Behandlungsdauer in zwei Gruppen eingeteilt: jene die auf die Behandlung „angesprochen“ haben oder eben nicht, je nachdem ob sie unter- oder überhalb eines bestimmten Ausmaßes an Verbesserung liegen (meist 50% - Verbesserung beim HAMD-Wert). So können die Unterschiede in den Ansprechquoten groß werden, selbst wenn der absolute Unterschied in den kontinuierlichen Werten minimal ist (Kirsch & Moncrieff, 2007). Nachdem die Messergebnisse über die Behandlungsdauer schwanken, wurde manchmal selektiv jener Zeitpunkt zur Auswertung gewählt werden, bei dem der Unterschiede zwischen AD und Placebo gerade am größten sind (Khan, Leventhal, Khan, & Brown, 2002).

3.7. *Viele Studien werden nicht publiziert:*

Studien mit günstigen Effekten werden viel häufiger in Fachzeitschriften publiziert als jene, bei denen AD Placebo nicht überlegen sind (Eyding u. a., 2010; Turner, Matthews, Linardatos, Tell, & Rosenthal, 2008). Daher überschätzen die älteren Meta-Analysen, die sich nur auf publizierte Studien stützen, die Wirkeffekte von AD (und anderen Medikamenten). Auch entsprechen die

Ergebnisse zahlreicher publizierter Studien nicht dem, was unabhängige Forschende (bzw. Zulassungsbehörden) aus den Originaldaten berechnen würden. Zwar verwendeten Cipriani et al. auch nicht-publizierte Studien, dennoch ist davon auszugehen, dass einige von der Industrie gesponserte Studien mit negativen Ergebnissen nie oder viel zu spät publik werden (Anderson u. a., 2015). Die Möglichkeiten von Pharmafirmen, Studien zu manipulieren sind weitreichend und vieles bleibt intransparent, wie die Forscher*innengruppe um Peter Gøtzsche anhand von angeforderten Studienprotokollen feststellten (Marquardsen, Ogden, & Gøtzsche, 2018).

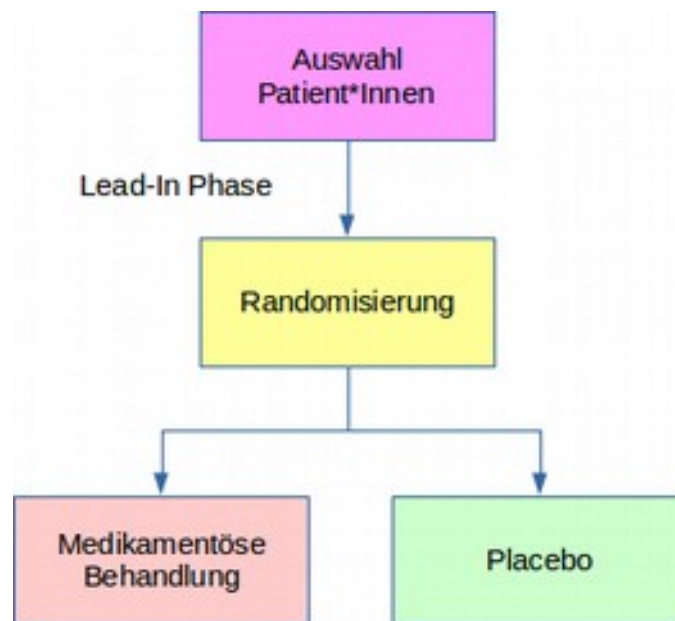


Abbildung 2a . Klassisches Design einer randomisierten kontrollierten Studie für Akutbehandlungen

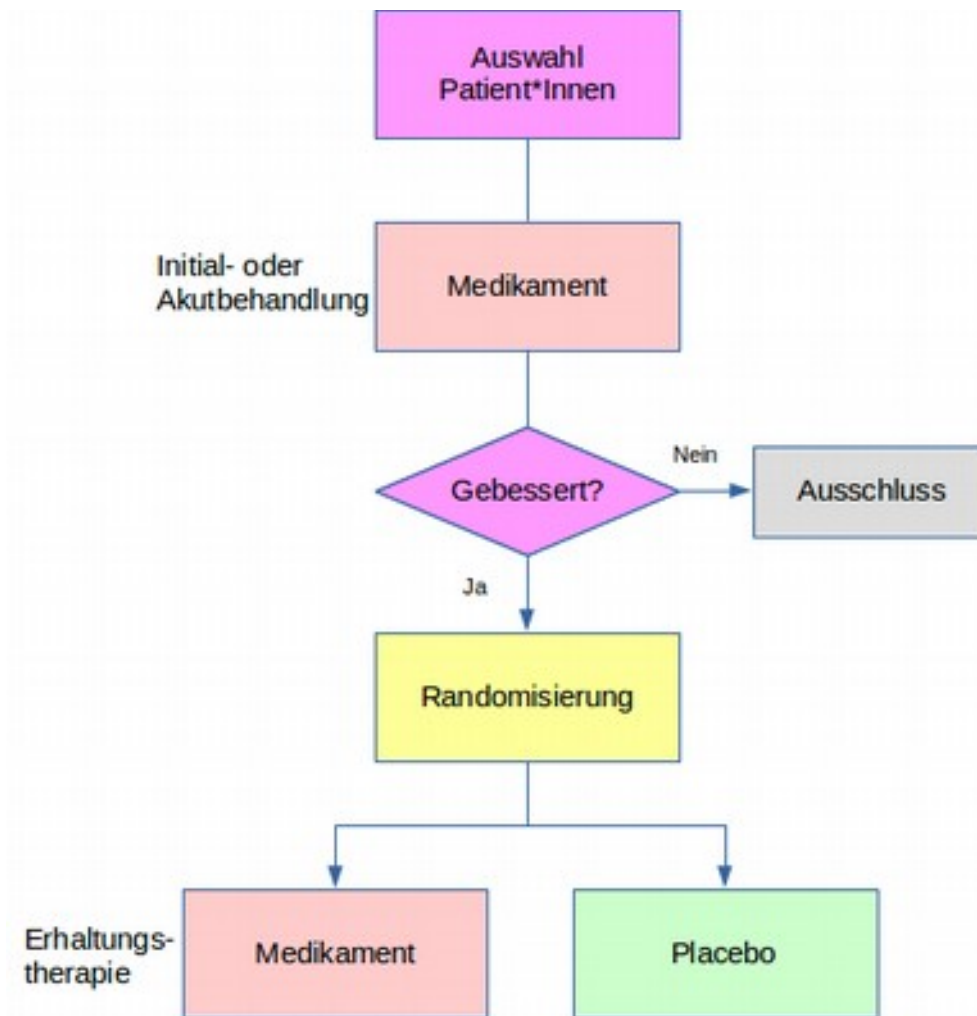


Abbildung 2b. Klassisches Design einer randomisierten kontrollierten Studie für Langzeitbehandlungen

3.8. Verzerrungseffekte, welche die Wirksamkeit von AD unterschätzen.

Es gibt einige wenige mögliche Verzerrungseffekte zu Ungunsten der Wirksamkeit von AD, die vor allem von Psychiatern mit offensichtlichen finanziellen Zuwendungen von der Pharmaindustrie vorgebracht werden. Ein Argument ist etwa, dass Patient*innen neben den AD auch Beruhigungsmittel und zusätzliche Therapien (z.B. Psychotherapie) erhalten, und so der reine Effekt von AD verschleiert wird. Ein weiteres Argument ist, dass manche Patient*innen in Studien teilnehmen, weil sie dafür Geld bekommen. Es könnte sein dass Patient*innen zunächst ihre Depressivität übertreiben um teilnehmen zu können, später ist dann kein Übertreiben mehr nötig (Hegerl & Mergl, 2010; Möller, 2008). Dieses Argument wurde meines Wissens noch nicht empirisch untersucht, erscheint aber wenig plausibel. Die Depressivität wird in der Regel im klinischen Interview festgestellt, wobei ÄrztInnen die Möglichkeit hätten, „Fake-Depressivität“ zu enttarnen oder zumindest übertriebene Depressivität zu relativieren. Hegerl und Mergl schlagen vor, dass man die Studienergebnisse nach Versicherungsstatus differenziert auswerten sollte (nicht-Versicherte wären demnach eher ScheinpatientInnen). Dies wäre sicher interessant. Aber wenn daran wirklich etwas dran wäre, dann kann man annehmen, dass dies schon längst ans Licht gekommen wäre, denn die Pharmafirmen haben großes Interesse an allen Argumenten, welche die

Wirksamkeit von AD günstiger darstellen.

Das Argument könnte und sollte man viel eher von der anderen Seite sehen: StudienärztInnen bekommen zum Teil viel Geld für PatientInnen, welche die Studie abschließen. So stehen die Ärzte in einem finanziellen Interessenkonflikt. Manchmal besteht auch ein Interessenkonflikt indirekt durch Vorgesetzte: „Als ich mal bei einer Studie als Arzt mitmachte, habe ich mir genau überlegt, was ich da ankreuze. Schließlich hatte der leitende Arzt einen wesentlichen Einfluss auf meine Karriere“. Dies ist eine Anekdote aus einem Gespräch mit einem bekannten Arzt. Daraus lässt sich natürlich nichts generalisieren lässt, aber bemerkenswert ist das schon.

Ein weiterer Verzerrungseffekt könnte durch StudienabbrecherInnen hervorgerufen werden. Diese Verzerrungseffekte gibt es (Wang u. a., 2018). Doch auch hier kann man sicher sein, dass die Analysen in den Studien so ausfallen, dass jene Methode gewählt wurde, die zugunsten von AD ausfällt (nämlich die LOCF Methode).

Möller diskutierte noch 2008, dass die Wirksamkeitseffekte in den Studien geringer ausfallen als in der Praxis, denn dort würde die Dosierung der Medikamente angepasst, wenn Patient*innen mit der Eingangsdosis nicht Ansprechen, oder es wird auf andere AD umgestellt. Mittlerweile weiß man, dass die Dosiserhöhung keinen Vorteil bringt (Dold, Bartova, Rupprecht, & Kasper, 2017; Rink u. a., 2018), und die STAR*D Studie und andere Studien zeigten, dass AD bei „echten“ PatientInnen *deutlich* weniger wirken (Pigott, 2011, 2015; Rush u. a., 2004).

4. Antidepressiva wirken auch bei starken Depressionen nicht

Die in den derzeitigen Leitlinien und von vielen mir bekannten Ärzt*innen vertretene Ansicht, dass AD bei schweren Depressionen eher helfen als bei mittleren und leichten Formen ist eine gängige, aber mittlerweile widerlegte Meinung. Vor 10 Jahren konnten Kirsch und Kolleg*innen in einer Meta-Analyse zeigen, dass der Unterschied zwischen AD und Placebo mit dem Schweregrad der Depression stieg. Dies lag aber nicht daran, dass AD bei schweren Depressionen besser wirkte, sondern dass Placebo bei schweren Depressionen weniger wirkte (Kirsch u. a., 2008). Auch in der neueren Meta-Analyse von Jakobsen et al. (2017) zeigte sich ein Unterschied der Wirksamkeit von AD je nach Schweregrad: 2,69 Punktedifferenz in Studien mit schwerer depressiven Patient*innen (Mittelwert HAMD > 23 Punkte) im Vergleich zu 1,29 Punkte Unterschied bei weniger depressiven Patient*innen (Mittelwert HAMD ≤ 23 Punkte). Die klinische Signifikanz von AD war aber in keinem Fall gegeben.

Das eigentliche Problem ist jedoch, dass die Vergleiche von Gruppenmittelwerten keine klare Aussage liefern können, ob die Wirksamkeit der AD vom Schweregrad der Depression abhängt. Dazu müssten die *individuellen* Verläufe der einzelnen Patient*innen beobachtet werden. Mittlerweile gibt es erste Meta-Analysen mit solchen individuellen Verlaufsdaten. Drei aktuelle Meta-Analysen mit mehreren tausenden Patient*innen fanden keinen nennenswerten Unterschied je nach Schweregrad (T. A. Furukawa u. a., 2018; Gibbons, Hur, Brown, Davis, & Mann, 2012; Rabinowitz u. a., 2016). Die Unterschiede zwischen AD und Placebo lagen in allen drei Metaanalysen unter 3 Punkte auf der HAMD Skala, auch bei schweren Depressionen, also in keinem Fall klinisch signifikant. Rabinowitz et al. machten mit den gleichen Daten auch den „klassischen“ Vergleich der Gruppenmittelwerte und siehe da, so wie in den alten Studien zeigte sich, dass die Differenz zwischen AD und Placebo mit dem Schweregrad der Depression zunahm. Dieser falsche Mythos ist somit als methodisches Artefakt enttarnt.

5. Antidepressiva haben Nebenwirkungen, vermutlich auch Suizidversuche und Suizide

5.1. Nebenwirkungen von Antidepressiva

In der Meta-Analyse von Jakobsen et al. (2017) hatten PatientInnen unter AD signifikant häufiger schwerwiegende Nebenwirkungen als unter Placebo (2.7% vs. 2.1%). Schwerwiegende Nebenwirkungen sind solche, die lebensbedrohlich oder tödlich sind, oder zu schwerwiegendem Funktionsverlust, Krankenhauseinweisungen oder längeren Krankenhausaufenthalten führen (S. 4). Der Unterschied, in Odds-Ratios ausgedrückt, beträgt 1,37 (1,08-1,75) und ist statistisch signifikant ($p = 0,009$). Per Konvention ist das ein kleiner Effekt. Nachdem es sich aber um schwerwiegende Nebenwirkungen handelt (auch Todesfälle), hat dieser Unterschied eine klinische Relevanz. Vor allem vor dem Hintergrund, dass AD gegenüber Placebo nach den oben angegebenen Argumenten keinen klinischen Nutzen haben. Auch die meisten anderen erfassten, weniger schwerwiegenden Nebenwirkungen traten signifikant häufiger unter AD auf. Folgende traten doppelt so häufig ($OR > 2$) und statistisch signifikant häufiger unter AD als unter Placebo auf: Geschmacksveränderung, Ejakulationsprobleme, reduzierte Libido, sexuelle Funktionsstörungen, Impotenz, Zittern, Schütteln, Anorexie, Übelkeit, Schläfrigkeit, Schwitzen, weniger Appetit, Gewichtsverlust, grippeähnliche Symptome, Mattigkeit. So ist es auch kaum verwunderlich, dass der Prozentsatz an Patient+innen, welche die Studie aufgrund von Nebenwirkungen abgebrochen hatte, bei allen Medikamenten höher war als bei Placebo, wie Cipriani et al. berichten ($1.20 < OR < 4.45$), und die Unterschiede waren hier bis auf ein Medikament (Agomelatin) auch statistisch signifikant.

Diese Daten beziehen sich nur auf die Akutphase der Behandlung, daher sagt das noch wenig über die langfristigen Nebenwirkungen auf. Darauf einzugehen sprengt den Rahmen, doch auf drei mögliche Langzeitschäden durch AD sollte schon hingewiesen werden. Erstens der unumkehrbare sexuelle Funktionsverlust (Hogan, Le Noury, Healy, & Mangin, 2014). Zweitens die Verschlechterung/Chronifizierung der depressiven Symptomatik. *Das mag absurd klingen, aber die Hinweise häufen sich, dass der Langzeitverlauf unter „Antidepressiva“ schlechter ist als unter Placebo* (Hengartner, Angst, & Rössler, 2018; Vittengl, 2017). Eine aktuelle 30-jährige Langzeitbeobachtung aus der Schweiz zeigte, dass Patient*innen, die AD verordnet bekamen, einen schlechteren Verlauf hatten (Hengartner, Angst, & Rössler, 2018). Weil es sich um eine Beobachtungsstudie handelte, könnte das Problem vorliegen, dass jene, die eine schlechtere Prognose hatten, häufiger AD verordnet bekamen, d.h. die Kausalrichtung könnte anders sein. Hengartner und Kolleg*innen hatten aber sorgfältig für kritische Variablen statistisch kontrolliert (soziodemografische Variablen, familiäre Depression, schwere Depressivität, affektive Störungen und Suizidalität), sodass diese Alternativerklärung nicht plausibel ist. Schon vor vielen Jahren hat Mauricio Fava anhand einer Literaturübersicht formuliert, dass AD erst recht den Verlauf der Depression verschlechtern könnten (Fava, 2003). Er fasste Studien zusammen, die folgende Effekte von AD fanden: Ungünstige Langzeiteffekte, Entstehung von Depression unter AD (z.B. bei Behandlung von Panikstörungen), Auslösung von manischen Episoden, Toleranzentwicklung (Rückfälle tauchen dann auch in der Erhaltungstherapie auf), Resistenz (ein AD wirkt beim zweiten Mal nicht mehr), Absetzsymptome und Abhängigkeit. Er spekuliert, dass AD im Gehirn gegensteuernde Mechanismen in Gang setzen. Nach Absetzen der Medikamente können diese neuen Mechanismen erst recht Symptome ähnlich wie bei der Depression auslösen. Eine Meta-Analyse von Absetzstudien bestätigte diese Hypothesen und kann so erklären, warum es nach einer

Behandlung mit AD zu deutlich mehr Rückfällen kommt (42-56% je nach Substanzgruppe) als in der Placebogruppe (25%) (Andrews, Kornstein, Halberstadt, Gardner, & Neale, 2011). Die derzeitigen Empfehlungen zur längerfristigen Einnahme von AD um Rückfälle zu vermeiden steht auch nach einer neuen Literaturübersicht nicht evidenzbasiert (Berwian, Walter, Seifritz, & Huys, 2017).

Drittens, die eben beschriebenen Nebenwirkungen könnten auch erklären, warum es unter AD im Vergleich zu Placebo häufiger zu Suiziden oder Suizidversuchen kommt. Dies wird im nächsten Punkt diskutiert.

5.2. Antidepressiva und Suizide bzw. Suizidversuche

5.2.1 Vorbemerkungen

Zwar schreiben manche Autor*innen über das erhöhte Suizidrisiko unter AD bzw. schließen ein solches nicht aus, relativieren dies aber dann durch Aussagen wie:

„anhand der vorhandenen Evidenz ist es so, dass jegliches erhöhtes Suizidrisiko welches durch SSRIs bei einer Untergruppe von depressiven Patient*innen ausgelöst werden könnte, durch den gesundheitlichen Nutzen einer häufigeren Diagnose und Behandlung in der Hausarztpraxis ausgeglichen wird“ (Hall, 2006, S. 948, meine Übers.).

„Solche Bedenken [erhöhtes Suizidrisiko durch AD, Anm.] müssen mit dem erhöhten Risiko einer unbehandelten Depression abgewogen werden, denn Suizide sind die dritthäufigste Todesursache bei Jugendlichen und mehr als 90% der Suizidfälle bei depressiven Jugendlichen waren zum Todeszeitpunkt nicht behandelt“ (Mann u. a., 2005, S. 2069, meine Übers.).

Bei dieser Argumentation wird angenommen, dass AD gegen Depression helfen was zu bezweifeln ist, wie ich in diesem Text und viele andere versucht haben zu zeigen. Leider finden sich solche unkritischen und angesichts der Evidenzlage sehr einseitigen Argumente häufig in der Literatur zur Suizidprävention, auch in sehr hochrangigen Fachzeitschriften wie etwa in Lancet und JAMA (Mann u. a., 2005; Zalsman u. a., 2016).

Daher ist es natürlich extrem kontrovers und provokant, wenn man behauptet, dass AD eine Zunahme von Suizidversuchen oder sogar Suiziden bewirkt. Ich habe selber auf der eMail-Liste der American Association for Suicidology die Expert*innen mit entsprechenden Daten konfrontiert, aber hauptsächlich Schweigen geerntet, oder es wurde als „Meinung“ abgetan. Wissenschaftliche Gegenargumente kamen so gut wie gar nicht, oder sie waren einfach zu widerlegen, und daraufhin gab es wieder Schweigen. Privat hatten mich jedoch ein paar bekannte Suizidforscher*innen kontaktiert und positiv reagiert. Diese Nicht-Reaktion oder „Evidenz-Resistenz“ begegnet mir leider häufig (siehe auch unten – Reaktion von Fachärzt*innen).

Aber wie sieht es nun wirklich mit der Evidenzlage aus? Dazu muss zunächst gesagt werden, dass die Evidenzlage eher dünn ist. Suizide sind statistisch seltene Ereignisse und es braucht sehr große Stichproben, um Unterschiede zwischen AD und Placebo nachzuweisen. Diese gäbe es mittlerweile, wie die Meta-Analyse von Cipriani et al. beweist. Das Problem sind aber

wiedereinmal Verzerrungseffekte, die im Falle von Suiziden schwerwiegend sind und daher vor dem Bericht der Evidenzlage erläutert werden müssen.

5.2.1 Methodische Probleme und Verschleierungsversuche

Suizide oder Suizidversuche werden in den klinischen Studien vermutlich nicht immer erfasst oder fehlklassifiziert, und auch regelrechte Verschleierungsversuche wurden bekannt, wie in den Büchern von Peter Gøtzsche und Peter und Sabine Ansari sehr gut recherchiert wurde (Ansari u. a., 2016; Gøtzsche, 2016). Vermutlich nicht-berichtete Suizid(versuche) wurden schon in einer Meta-Analyse aus dem Jahre 2005 diskutiert, bei der Suizidversuche (aber nicht Suizide) bei Erwachsenen häufiger unter AD als unter Placebo auftraten (OR = 1.57, 95%-CI 0.99-2.55) (Gunnell, Saperia, & Ashby, 2005). Die Autor*innen fanden, dass das Verhältnis von Suiziden zu Suizidversuchen in den Studien nur 1:10 betrug, während das in der Bevölkerung 1:30 beträgt. Vermutlich sind daher Suizidversuche in den Studien nicht berichtet worden. Dies bestätigte sich anderweitig, wie ein Zitat aus dem Buch der Ansaris (S. 27) verdeutlicht: *„Die Manipulationen um das Medikament Prozac sind besonders gründlich von den Psychiatern David Healy und Peter Breggin recherchiert worden. Sie erhielten als medizinische Gutachter Einblicke in konzerninterne Unterlagen und veröffentlichten mehrere Bücher darüber. Es stellte sich heraus, dass Lilly den Begriff »Suizid« aus seiner Datenbank der Nebenwirkungen gelöscht hatte und anstelle von »einem Suizidversuch unter Prozac« von einer Überdosierung sprach. Nicht allen Lilly-Mitarbeitern gefiel diese Vorgabe. Der deutsche Mitarbeiter Claude Bouchy schrieb in einem internen Memo: »Ich glaube nicht, dass ich dem Bundesgesundheitsamt, einem Richter, einem Reporter oder meiner Familie erklären könnte, warum wir das tun. Schon gar nicht, wenn es um so heikle Probleme wie Suizid und Suizidgedanken geht.“*

Peter Gøtzsche berichtet auch über fragwürdige Unstimmigkeiten, selbst in den Berichten der U.S. Regulierungsbehörde FDA „Food and Drug Administration“. Die FDA veröffentlichte im Jahre 2001 eine Meta-Analyse von Klinischen Studien, bei der sich 22 von 22.062 Patient*innen unter AD das Leben nahmen (interessanterweise ist das Dokument nicht mehr beim Elsevier Verlag erhältlich, obwohl es eine DOI-Nummer hat). Fünf Jahre später erschien eine größere Meta-Analyse dazu, bei der nur mehr noch 5 Suizide von 52960 Patient*innen unter AD vorkamen. Laut Marc Stone ergab sich die Diskrepanz, weil im früheren Bericht alle Ereignisse angeführt wurden (22 Suizide), in der 2009 publizierten Meta-Analyse aber nur mehr noch Suizide in der Doppelblindphase (siehe <https://www.bmj.com/content/352/bmj.i65/rr-4>). Gøtzsche erwiderte darauf, dass in dem ursprünglichen FDA-Bericht die Suizide entweder in der Behandlungsphase oder in den darauf folgenden 30 Tagen gezählt wurden, dass es aber sinnvoll ist, genau diese Ergebnisse auszuwerten. In der Placebogruppe waren nur 2 Suizide, was ein statistisch signifikantes Ergebnis ergibt. Die Diskrepanz zwischen dem früheren und späteren FDA Bericht bleibt laut Gøtzsche dennoch ungeklärt (siehe <https://www.bmj.com/content/352/bmj.i65/rr-7>).

In einer neueren Studie von Gøtzsche und KollegInnen wurden Studienprotokolle hinsichtlich negativer Effekte (inklusive Suizidalität) von AD untersucht (Sharma u. a., 2016). Dabei war nur ein Bruchteil der Studienprotokolle zugänglich. Sie konnten dennoch eindeutige schwerwiegende Fehlklassifikationen nachweisen, immer zugunsten der AD. Ein Patient, der sich in der laufenden AD-Behandlung strangulieren wollte und 5 Tage später im Krankenhaus starb, wurde

nicht mitgezählt sondern als Abbrecher klassifiziert. Umgekehrt wurde ein Patient der Placebogruppe, der sich 26 Tage nach dem Ende der Studie das Leben nahm, dennoch zur Placebogruppe mitgezählt.

Ein großes Problem ist, dass unbequeme Studienergebnisse nicht publik werden. Ein besonders dramatisches Beispiel ist „Study 329“, eine Studie an Jugendlichen, in der zwei AD (Paroxetin und Imipramin) mit Placebo verglichen wurden. Die Studie war als Langzeitstudie konzipiert, aber nur die Ergebnisse der Akutphase wurden publiziert, diese dafür in einer renommierten Zeitschrift (Keller u. a., 2001). Nur durch das Engagement von unabhängigen Forscher*innen wurden die längerfristigen Ergebnisse publik (Le Noury u. a., 2016), siehe auch den entsprechenden [Wikipedia Eintrag](#). Es zeigte sich, dass es längerfristig, d.h. in der Erhaltungstherapie und beim Ausschleichen, unter Paroxetine zu mehr Rückfällen und zu mehr suizidalen Ereignissen kam als unter Placebo. Die Studie wurde nie zurückgezogen. Und das, obwohl schon die publizierte Version über die Akutphase eine falsche Schlussfolgerung zog. Wie üblich wurden nur die Responderraten berichtet, bei der es unter AD signifikant häufiger zu einem Ansprechen kam als unter Placebo. Dabei wurde allerdings eine andere als die geplante Auswertung gemacht, sonst wäre selbst dieser Effekt nicht signifikant geworden. Ein Vergleich der HAMD-Punkte ergibt jedoch wie üblich einen sehr kleinen Effekt, der weder berichtet noch diskutiert wurde ($d = 0.21$, $p = 0.13$, meine Berechnung). In der AD-Gruppe kam es zu 5 suizidalen Ereignissen (Suizidgedanken oder suizidale Gesten), in der Placebogruppe zu keinem. Ein entsprechender Signifikanztest (Barnards-Test) ergibt sogar einen signifikanten Unterschied ($p = 0.03$, meine Berechnung). Dennoch steht in der Zusammenfassung der Studie, dass Paroxetine gut verträglich und effektiv ist. Die Publikation wurde bis heute nicht zurückgezogen.

5.2.2 Evidenzlage bezüglich Antidepressiva und Suizidrisiko

Schon 1993 zeigte sich in einer Meta-Analysen, dass es unter AD zu mehr Suiziden (OR 4.40, 95% CI 1.32–unendlich, $p = 0.01$) kam, und auch zu mehr Suizidversuchen (OR 2.39, CI 1.66–unendlich; $p \leq 0.0001$) (Healy & Whitaker, 2003). Bemerkenswert an dieser alten Meta-Analyse ist, dass jene Suizide oder Suizidversuche, die in der Wash-Out Phase statt fanden, herausrechneten. Viele Studien rechnen diese Ereignisse einfach der Placebo-Gruppe zu, was völlig fahrlässig ist. Neuere Meta-Analysen zur Akutbehandlung fanden gemischte Ergebnisse, auch in Abhängigkeit der Altersgruppen und Form der Suizidalität. So wurden bei Kindern und Jugendlichen in mehreren Meta-Analysen signifikant höhere Raten and Suizidgedanken und Suizidversuchen unter AD im Vergleich zu Placebo (Hetrick u. a., 2012; Sharma u. a., 2016; Stone u. a., 2009) gefunden. Zum Beispiel ein 1,6-fach erhöhtes Risiko für Suizidgedanken und Suizidversuche (95% CI 1.02 to 2.45) (Hetrick u. a., 2012), oder ein über 2-fach erhöhtes Risiko für Suizidversuche bis zu 25 Jährigen (Stone u. a., 2009) Bei Erwachsenen fallen die Ergebnisse gemischt aus, meist gibt es keine signifikanten Unterschiede. Mehr Suizidversuche unter AD als unter Placebo fanden sich in der ersten richtig großen Meta-Analyse an SSRI's versus Placebo (OR 2.28, 95% CI 1.14 – 4.55) (Fergusson u. a., 2005), aber nicht in einer neuen Meta-Analyse (Jakobsen, 2017). In der FDA-Meta-Analyse von Stone hatten Erwachsene Personen keine signifikant erhöhten Suizidversuchsraten unter AD, und bei älteren Personen war das Risiko sogar signifikant geringer als unter Placebo (Stone u. a., 2009) – doch gibt es bei dieser Meta-Analyse die oben beschriebenen

Ungereimtheiten.

Langzeiteffekte: Meines Wissens gibt es bisher nur zwei Meta-Analysen über randomisiert-kontrollierte Langzeitstudien zu Suiziden oder Suizidversuchen (Baldessarini, Lau, Sim, Sum, & Sim, 2016; Braun, Bschor, Franklin, & Baethge, 2016). Wie oben erwähnt muss man bedenken, dass aufgrund des Studiendesigns die Placebogruppe benachteiligt wird, weil die Patient*innen erst randomisiert werden, nachdem sie sich unter dem Medikament gebessert waren. Nicht gebesserte PatientInnen scheiden aus der Studien aus, d.h. es werden nur jene weiter untersucht, die das Medikament gut vertragen. Dazu kommt, dass durch Absetzsymptome in der Placebogruppe das Ergebnis verzerrt wird (siehe Abb. 2b). *Trotz der Bevorteilung von AD durch das Studiendesign wurden signifikant mehr Suizidversuche knapp nicht signifikant mehr Suizide unter AD als unter Placebo gefunden.* Weil in vielen Studien weder in der Placebo- noch in der AD-gruppe Suizide vorkamen, rechnete ich mit den Daten von Baldessarini und Kolleg*innen eine Bayesianische Random-Effect-Meta-Analyse zu Suizidversuchen und Suiziden (über die Meta-Analyse von Braun et al. wurde ich erst kürzlich informiert, die Ergebnisse ähneln sich aber). Der Bayesianische Ansatz hat kein Problem mit Null-Ereignissen. Die Ergebnisse der Analyse zeigen, dass Suizidversuche unter Placebo eindeutig weniger häufig vorkamen als unter AD: OR = 0,18 (95%-CI 0,00-0,90). Das ist ein großer Unterschied. Auch bei Suiziden war das so, aber das Ergebnis war knapp nicht signifikant, weil der Unsicherheitsbereich zu groß war (OR = 0,31, 95%-CI 0,01-1,10). Bei einem zusätzlichen Suizid in der AD-Gruppe wäre das Ergebnis eindeutig. Durch die oben berichteten Verzerrungseffekte und Fehlklassifikation ist es durchaus wahrscheinlich, dass ein Suizid in den AD-gruppen nicht mitgezählt wurde. Aber unabhängig davon sind die Ergebnisse auch so problematisch genug.

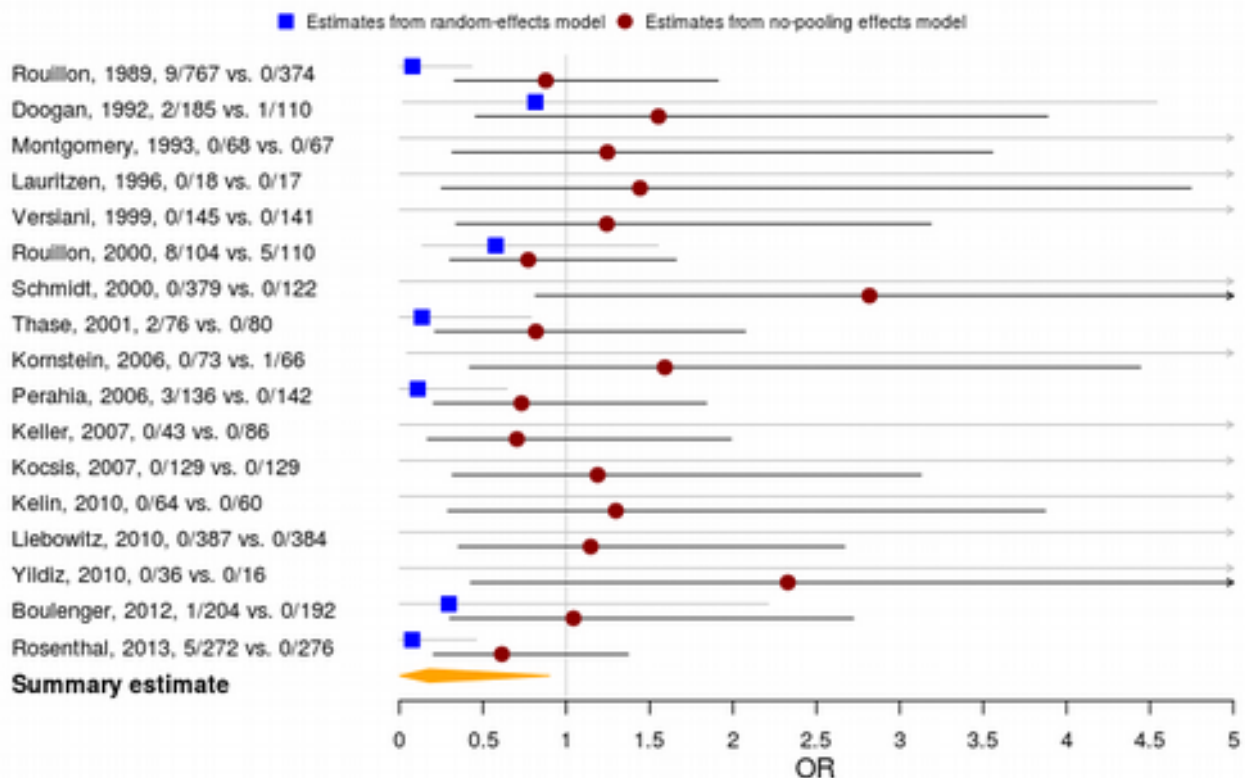


Abbildung 3a. Suizidversuche - Bayesianische Meta-Analyse der Daten von Baldessarini et al. (2016)

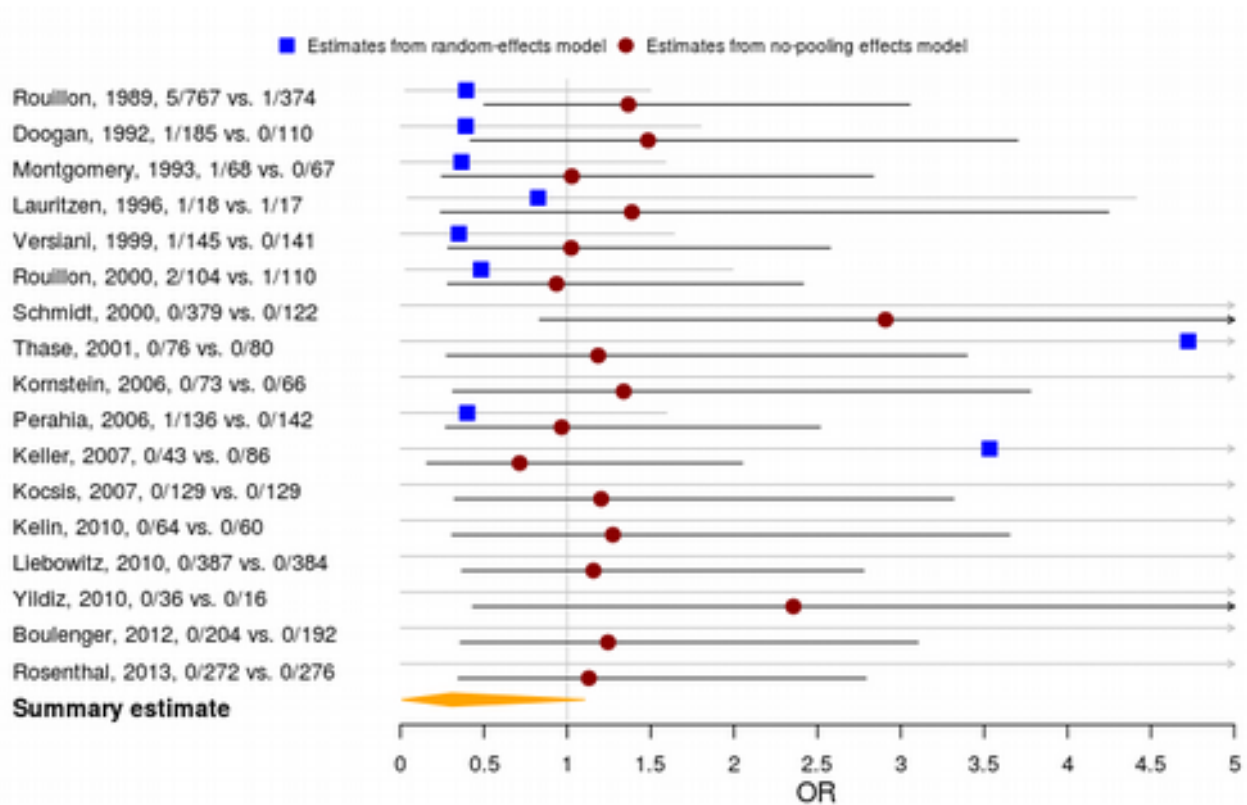


Abbildung 3b. Suizide - Bayesianische Meta-Analyse der Daten von Baldessarini et al. (2016)

AD wirken bei suizidalen PatientInnen noch weniger. Wie schon erwähnt, wird in der Suizidprävention argumentiert, dass das Risiko einer unbehandelten Depression größer ist als das eventuelle erhöhte Risiko bei manchen mit AD behandelten depressiven Menschen. Dabei wird davon ausgegangen, dass AD wirken, das tun sie aber nicht klinisch bedeutsam, vermutlich gar nicht (siehe oben). Wie AD bei suizidalen Patient*innen wirken ist kaum bekannt, weil diese in der Regel von Studien ausgeschlossen werden. Zum Beispiel in mindestens 63% der Studien in einer Meta-Analyse (Sharma u. a., 2016). Das ist absurd, denn genau hier wäre Hilfe dringlich, und meiner Erfahrung nach werden suizidgefährdeten Patient*innen viel eher Medikamente verordnet. AD scheinen bei suizidalen Patient*innen aber noch weniger zu wirken als bei nicht-suizidalen Patient*innen, wie zwei mir bekannte Studien zeigten.

In einer Studie wurden chronisch depressive Patient*innen in einer psychiatrischen Einrichtung 6 Wochen intravenös mit AD behandelt, nachdem die orale Behandlung vorher nicht wirkte (Pompili u. a., 2010). Das waren „echte“ Patient*innen, also nicht selektiert wie in sonstigen klinischen Studien. Es zeigte sich, dass bei Behandlungsbeginn suizidale Patient*innen (definiert als jene mit Suizidgedanken oder Suizidhandlungen, gemessen mit Item 4 der HAMD-17 Skala), fünfmal so häufig nicht auf die Behandlung ansprachen (Verbesserung < 20% auf der HAMD Skala, 75% vs. 43%) als nicht-suizidale Patient*innen, selbst nach Kontrolle für Substanzmissbrauch und

Länge der Krankheitsepisode (weil in diesen Punkten signifikante Unterschiede zwischen suizidalen und nicht-suizidalen Patient*innen gefunden wurden). Interessant war, dass zu Studienbeginn suizidale und nicht-suizidale Patient*innen gleich hohe Werte auf der HAMD Skala hatten, aber die Reduktion der Werte deutlich unterschiedlich ausfiel (13 Punkte bei suizidalen vs. 25 Punkte bei nicht-suizidalen Patient*innen). Eine weitere Beobachtungsstudie mit gut 4000 „echten“ Patient*innen fand ebenfalls, dass Suizidgedanken zu Behandlungsbeginn oder ein Suizidversuch in der Vorgeschichte mit einer signifikant geringeren Remissionsrate (HAMD-Wert ≤ 7) verbunden war, auch nach Kontrolle vieler Störvariablen (Lopez-Castroman, Jaussent, Gorwood, & Courtet, 2016; Nobile u. a., 2018). Insgesamt waren nach der 6-wöchigen Behandlungszeit mit vier verschiedenen AD 50% Responder, bei Personen mit Suizidversuchen oder bei Suizidgedanken nur ca. 40%. Die beiden Studien bestätigen übrigens auch, dass die hohen Responderraten in den randomisiert kontrollierten Studien mit sehr selektieren Patient*innen die Wirksamkeit der AD generell überschätzen, wie das ja auch in der STAR*D Studie gefunden wurde.

Resümee: Die Evidenzlage spricht ziemlich eindeutig dagegen, dass AD Suizide oder Suizidversuche verhindern. Meist wird entweder kein protektiver Effekt gefunden, oder sogar, dass Suizidversuche oder Suizide häufiger unter AD als unter Placebo vorkommen. Vor allem die Ergebnisse der Langzeitstudien sprechen für eine erhöhte Suizidgefährdung unter AD.

6. Reaktionen von Fachärzt*innen

Seit einiger Zeit habe ich angefangen, Ärzt*innen auf die Problematik der fehlenden Effektivität und die Risiken von AD hinzuweisen. Das ist schwierig. Viele meiner bekannten Ärzt*innen sind sehr motiviert und wollen Patient*innen wirklich helfen. Bei Depressionen AD zu verordnen war und ist eine jahrelange Selbstverständlichkeit, und plötzlich sollte das alles nichts gebracht oder sogar geschädigt haben? Zu den Reaktionen zählen Äußerungen und Reaktionen wie:

- Also ich habe den Eindruck dass AD bei manchen schon helfen
- Es gibt aber andere Meinungen
- Bei schweren Depressionen helfen AD sehr wohl
- Oft braucht es halt lange, bis die AD wirken
- Bei den wirklich depressiven Patient*innen helfen sie schon, es wird halt keine ordentlich Depressionsdiagnose mehr gemacht
- Das kann schon sein, aber ich muss die Leitlinien befolgen. Wenn ich nichts verschreibe, mache ich mich angreifbar
- Und was soll ich dann sonst tun? Sie sehen ja, draußen warten jede Menge Leute, soll ich mit denen alle reden?

Ich traf aber auch kritische Psychiater*innen, die von der Problematik Bescheid wussten und auch kritische Haltung vertreten – es besteht also Hoffnung.

Literatur

- Anderson, M. L., Chiswell, K., Peterson, E. D., Tasneem, A., Topping, J., & Califf, R. M. (2015). Compliance with Results Reporting at ClinicalTrials.gov. *New England Journal of Medicine*, 372(11), 1031–1039. <https://doi.org/10.1056/NEJMsa1409364>
- Andrews, P. W., Kornstein, S. G., Halberstadt, L. J., Gardner, C. O., & Neale, M. C. (2011). Blue Again: Perturbational Effects of Antidepressants Suggest Monoaminergic Homeostasis in Major Depression. *Frontiers in Psychology*, 2. <https://doi.org/10.3389/fpsyg.2011.00159>
- Ansari, P., Ansari, S., & Müller-Oerlinghausen, B. (2016). *Unglück auf Rezept: die Antidepressiva-Lüge und ihre Folgen*. Stuttgart: Klett-Cotta.
- Baldessarini, R. J., Lau, W. K., Sim, J., Sum, M. Y., & Sim, K. (2016). Suicidal Risks in Reports of Long-Term Controlled Trials of Antidepressants for Major Depressive Disorder II. *International Journal of Neuropsychopharmacology*, pyw092. <https://doi.org/10.1093/ijnp/pyw092>
- Berwian, I. M., Walter, H., Seifritz, E., & Huys, Q. J. M. (2017). Predicting relapse after antidepressant withdrawal – a systematic review. *Psychological Medicine*, 47(03), 426–437. <https://doi.org/10.1017/S0033291716002580>
- Braun, C., Bschor, T., Franklin, J., & Baethge, C. (2016). Suicides and Suicide Attempts during Long-Term Treatment with Antidepressants: A Meta-Analysis of 29 Placebo-Controlled Studies Including 6,934 Patients with Major Depressive Disorder. *Psychotherapy and Psychosomatics*, 85(3), 171–179. <https://doi.org/10.1159/000442293>
- Cipriani, A., Furukawa, T. A., Salanti, G., Chaimani, A., Atkinson, L. Z., Ogawa, Y., ... Geddes, J. R. (2018). Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *Lancet*, 391(10128), 1357–1366. [https://doi.org/10.1016/S0140-6736\(17\)32802-7](https://doi.org/10.1016/S0140-6736(17)32802-7)
- Cipriani, A., Zhou, X., Del Giovane, C., Hetrick, S. E., Qin, B., Whittington, C., ... Xie, P. (2016). Comparative efficacy and tolerability of antidepressants for major depressive disorder in children and adolescents: a network meta-analysis. *The Lancet*, 388(10047), 881–890. [https://doi.org/10.1016/S0140-6736\(16\)30385-3](https://doi.org/10.1016/S0140-6736(16)30385-3)
- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 997–1003.
- Cohen, Jacob. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale,

N.J: L. Erlbaum Associates.

- Davies, J., & Read, J. (2018). A systematic review into the incidence, severity and duration of antidepressant withdrawal effects: Are guidelines evidence-based? *Addictive Behaviors*.
<https://doi.org/10.1016/j.addbeh.2018.08.027>
- Dold, M., Bartova, L., Rupprecht, R., & Kasper, S. (2017). Dose Escalation of Antidepressants in Unipolar Depression: A Meta-Analysis of Double-Blind, Randomized Controlled Trials. *Psychotherapy and Psychosomatics*, 86(5), 283–291. <https://doi.org/10.1159/000477770>
- Eyding, D., Lelgemann, M., Grouven, U., Harter, M., Kromp, M., Kaiser, T., ... Wieseler, B. (2010). Reboxetine for acute treatment of major depression: systematic review and meta-analysis of published and unpublished placebo and selective serotonin reuptake inhibitor controlled trials. *BMJ*, 341(oct12 1), c4737–c4737. <https://doi.org/10.1136/bmj.c4737>
- Fava, G. A. (2003). Can Long-Term Treatment With Antidepressant Drugs Worsen the Course of Depression? *The Journal of Clinical Psychiatry*, 64(2), 123–133.
<https://doi.org/10.4088/JCP.v64n0204>
- Fava, G. A., Benasi, G., Lucente, M., Offidani, E., Cosci, F., & Guidi, J. (2018). Withdrawal Symptoms after Serotonin-Noradrenaline Reuptake Inhibitor Discontinuation: Systematic Review. *Psychotherapy and Psychosomatics*, 87(4), 195–203.
<https://doi.org/10.1159/000491524>
- Fava, G. A., Gatti, A., Belaise, C., Guidi, J., & Offidani, E. (2015). Withdrawal Symptoms after Selective Serotonin Reuptake Inhibitor Discontinuation: A Systematic Review. *Psychotherapy and Psychosomatics*, 84(2), 72–81. <https://doi.org/10.1159/000370338>
- Fergusson, D., Doucette, S., Glass, K. C., Shapiro, S., Healy, D., Hebert, P., & Hutton, B. (2005). Association between suicide attempts and selective serotonin reuptake inhibitors: systematic review of randomised controlled trials. *British Medical Journal*, 330, 396–403.
- Furukawa, T. A., Cipriani, A., Atkinson, L. Z., Leucht, S., Ogawa, Y., Takeshima, N., ... Salanti, G. (2016). Placebo response rates in antidepressant trials: a systematic review of published and unpublished double-blind randomised controlled studies. *The Lancet. Psychiatry*, 3(11), 1059–1066. [https://doi.org/10.1016/S2215-0366\(16\)30307-8](https://doi.org/10.1016/S2215-0366(16)30307-8)
- Furukawa, T. A., Maruo, K., Noma, H., Tanaka, S., Imai, H., Shinohara, K., ... Cipriani, A. (2018). Initial severity of major depression and efficacy of new generation antidepressants:

individual participant data meta-analysis. *Acta Psychiatrica Scandinavica*, 137(6), 450–458.
<https://doi.org/10.1111/acps.12886>

- Furukawa, Toshi A., & Leucht, S. (2011). How to obtain NNT from Cohen's d: comparison of two methods. *PloS One*, 6(4), e19070. <https://doi.org/10.1371/journal.pone.0019070>
- Gibbons, R. D., Hur, K., Brown, C. H., Davis, J. M., & Mann, J. J. (2012). Benefits From Antidepressants: Synthesis of 6-Week Patient-Level Outcomes From Double-blind Placebo-Controlled Randomized Trials of Fluoxetine and Venlafaxine. *Archives of General Psychiatry*, 69(6). <https://doi.org/10.1001/archgenpsychiatry.2011.2044>
- Gøtzsche, P. C. (2016). *Tödliche Psychopharmaka und organisiertes Leugnen: wie Ärzte und Pharmaindustrie die Gesundheit der Patienten vorsätzlich aufs Spiel setzen* (1. Auflage). München: riva.
- Gunnell, D., Saperia, J., & Ashby, D. (2005). Selective serotonin reuptake inhibitors (SSRIs) and suicide in adults: meta-analysis of drug company data from placebo controlled, randomised controlled trials submitted to the MHRA's safety review. *BMJ*, 330(7488), 385.
<https://doi.org/10.1136/bmj.330.7488.385>
- Hall, W. D. (2006). How have the SSRI antidepressants affected suicide risk? *The Lancet*, 367(9527), 1959–1962. [https://doi.org/10.1016/S0140-6736\(06\)68860-0](https://doi.org/10.1016/S0140-6736(06)68860-0)
- Healy, D., Le Noury, J., & Jureidini, J. (2018). Paediatric antidepressants: Benefits and risks. *International Journal of Risk & Safety in Medicine*, 1–7. <https://doi.org/10.3233/JRS-180746>
- Healy, D., & Whitaker, C. (2003). Antidepressants and suicide: risk-benefit conundrums. *Journal of Psychiatry & Neuroscience: JPN*, 28(5), 331–337.
- Hegerl, U., & Mergl, R. (2010). The clinical significance of antidepressant treatment effects cannot be derived from placebo-verum response differences. *Journal of Psychopharmacology*, 24(4), 445–448.
- Hengartner, M. P. (2017). Methodological Flaws, Conflicts of Interest, and Scientific Fallacies: Implications for the Evaluation of Antidepressants' Efficacy and Harm. *Frontiers in Psychiatry*, 8. <https://doi.org/10.3389/fpsy.2017.00275>
- Hengartner, M. P. (2018). What is the threshold for a clinical minimally important drug effect? *BMJ Evidence-Based Medicine*. <https://doi.org/10.1136/bmjebm-2018-111056>

- Hengartner, M. P., Angst, J., & Rössler, W. (2018). Antidepressant Use Prospectively Relates to a Poorer Long-Term Outcome of Depression: Results from a Prospective Community Cohort Study over 30 Years. *Psychotherapy and Psychosomatics*.
<https://doi.org/10.1159/000488802>
- Hengartner, M. P., & Plöderl, M. (2018a). False Beliefs in Academic Psychiatry: The Case of Antidepressant Drugs. *Ethical Human Psychology and Psychiatry*, 20(1), 6–16.
<https://doi.org/10.1891/1559-4343.20.1.6>
- Hengartner, M. P., & Plöderl, M. (2018b). Statistically significant antidepressant-placebo differences on subjective symptom-rating scales do not prove that antidepressants work: Effect size and method bias matter! *Frontiers in Psychiatry*, 9.
<https://doi.org/10.3389/fpsyt.2018.00517>
- Hetrick, S. E., McKenzie, J. E., Cox, G. R., Simmons, M. B., & Merry, S. N. (2012). Newer generation antidepressants for depressive disorders in children and adolescents. *The Cochrane Database of Systematic Reviews*, 11, CD004851.
<https://doi.org/10.1002/14651858.CD004851.pub3>
- Hogan, C., Le Noury, J., Healy, D., & Mangin, D. (2014). One hundred and twenty cases of enduring sexual dysfunction following treatment. *The International Journal of Risk & Safety in Medicine*, 26(2), 109–116. <https://doi.org/10.3233/JRS-140617>
- Jakobsen, J. C. (2017). Selective serotonin reuptake inhibitors versus placebo in patients with major depressive disorder. A systematic review with meta-analysis and Trial Sequential Analysis, 28.
- Keller, M. B., Ryan, N. D., Strober, M., Klein, R. G., Kutcher, S. P., Birmaher, B., ... McCafferty, J. P. (2001). Efficacy of Paroxetine in the Treatment of Adolescent Major Depression: A Randomized, Controlled Trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40(7), 762–772. <https://doi.org/10.1097/00004583-200107000-00010>
- Khan, A., Leventhal, R. M., Khan, S. R., & Brown, W. A. (2002). Severity of depression and response to antidepressants and placebo: an analysis of the Food and Drug Administration database. *Journal of Clinical Psychopharmacology*, 22(1), 40–45.
- Kirsch, I., Deacon, B. J., Huedo-Medina, T. B., Scoboria, A., Moore, T. J., & Johnson, B. T. (2008). Initial Severity and Antidepressant Benefits: A Meta-Analysis of Data Submitted to the Food

and Drug Administration. *PLoS Medicine*, 5(2), e45.

<https://doi.org/10.1371/journal.pmed.0050045>

Kirsch, I., & Moncrieff, J. (2007). Clinical trials and the response rate illusion. *Contemporary Clinical Trials*, 28(4), 348–351. <https://doi.org/10.1016/j.cct.2006.10.012>

Le Noury, J., Nardo, J. M., Healy, D., Jureidini, J., Raven, M., Tufanaru, C., & Abi-Jaoude, E. (2016). Study 329 continuation phase: Safety and efficacy of paroxetine and imipramine in extended treatment of adolescent major depression. *International Journal of Risk & Safety in Medicine*, 28(3), 143–161. <https://doi.org/10.3233/JRS-160728>

Lopez-Castroman, J., Jaussent, I., Gorwood, P., & Courtet, P. (2016). SUICIDAL DEPRESSED PATIENTS RESPOND LESS WELL TO ANTIDEPRESSANTS IN THE SHORT TERM: Research Article: Antidepressants Response and Suicidal Status. *Depression and Anxiety*, 33(6), 483–494. <https://doi.org/10.1002/da.22473>

Mann, J. J., Apter, A., Bertolote, J., Beautrais, A., Currier, D., Haas, A., ... Hendin, H. (2005). Suicide prevention strategies: a systematic review. *JAMA*, 294(16), 2064–2074. <https://doi.org/10.1001/jama.294.16.2064>

Marquardsen, M., Ogden, M., & Gøtzsche, P. C. (2018). Redactions in protocols for drug trials: what industry sponsors concealed. *Journal of the Royal Society of Medicine*, 111(4), 136–141. <https://doi.org/10.1177/0141076817750554>

McLaren, N. (2018, Mai 22). Questioning the Integrity of Psychiatry. Abgerufen 20. Juli 2018, von <https://www.madinamerica.com/2018/05/questioning-integrity-psychiatry/>

Möller, H. J. (2008). Isn't the efficacy of antidepressants clinically relevant? A critical comment on the results of the metaanalysis by Kirsch et al. 2008. *European archives of psychiatry and clinical neuroscience*, 258(8), 451–455.

Moncrieff, J., & Kirsch, I. (2015). Empirically derived criteria cast doubt on the clinical significance of antidepressant-placebo differences. *Contemporary Clinical Trials*, 43, 60–62. <https://doi.org/10.1016/j.cct.2015.05.005>

Moncrieff, J., Wessely, S., & Hardy, R. (2004). Active placebos versus antidepressants for depression. *Cochrane Database of Systematic Reviews*. <https://doi.org/10.1002/14651858.CD003012.pub2>

Nobile, B., Jaussent, I., Gorwood, P., Lopez Castroman, J., Olié, E., Guillaume, S., & Courtet, P.

(2018). Tianeptine is associated with lower risk of suicidal ideation worsening during the first weeks of treatment onset compared with other antidepressants: A naturalistic study. *Journal of Psychiatric Research*, 96, 167–170.

<https://doi.org/10.1016/j.jpsychires.2017.10.007>

Pigott, H. E. (2011). STAR*D: A Tale and Trail of Bias. *Ethical Human Psychology and Psychiatry*, 13(1), 6–28. <https://doi.org/10.1891/1559-4343.13.1.6>

Pigott, H. E. (2015). The STAR*D Trial: It is Time to Reexamine the Clinical Beliefs That Guide the Treatment of Major Depression. *The Canadian Journal of Psychiatry*, 60(1), 9–13. <https://doi.org/10.1177/070674371506000104>

Pompili, M., Baldessarini, R. J., Tondo, L., Innamorati, M., Tatarelli, R., Girardi, P., & De Pisa, E. (2010). Response to intravenous antidepressant treatment by suicidal vs. nonsuicidal depressed patients. *Journal of Affective Disorders*, 122(1–2), 154–158. <https://doi.org/10.1016/j.jad.2009.07.018>

Rabinowitz, J., Werbeloff, N., Mandel, F. S., Menard, F., Marangell, L., & Kapur, S. (2016). Initial depression severity and response to antidepressants v. placebo: patient-level data analysis from 34 randomised controlled trials. *British Journal of Psychiatry*, 209(05), 427–428. <https://doi.org/10.1192/bjp.bp.115.173906>

Read, J., Cartwright, C., & Gibson, K. (2018). How many of 1829 antidepressant users report withdrawal effects or addiction? *International Journal of Mental Health Nursing*. <https://doi.org/10.1111/inm.12488>

Rink, L., Braun, C., Bschor, T., Henssler, J., Franklin, J., & Baethge, C. (2018). Dose Increase Versus Unchanged Continuation of Antidepressants After Initial Antidepressant Treatment Failure in Patients With Major Depressive Disorder: A Systematic Review and Meta-Analysis of Randomized, Double-Blind Trials. *The Journal of Clinical Psychiatry*, 79(3). <https://doi.org/10.4088/JCP.17r11693>

Rush, A. J., Trivedi, M., Carmody, T. J., Biggs, M. M., Shores-Wilson, K., Ibrahim, H., & Crismon, M. L. (2004). One-year clinical outcomes of depressed public sector outpatients: a benchmark for subsequent studies. *Biological Psychiatry*, 56(1), 46–53. <https://doi.org/10.1016/j.biopsych.2004.04.005>

Sharma, T., Guski, L. S., Freund, N., & Gøtzsche, P. C. (2016). Suicidality and aggression during

antidepressant treatment: systematic review and meta-analyses based on clinical study reports. *BMJ*, i65. <https://doi.org/10.1136/bmj.i65>

Stone, M. B., Laughren, T., Jones, M. L., Levenson, M., Holland, P. C., Hughes, A., ... Rochester, G. (2009). Risk of suicidality in clinical trials of antidepressants in adults: analysis of proprietary data submitted to US Food and Drug Administration. *BMJ*, 339, 1–10.

Turner, E. H., Matthews, A. M., Linardatos, E., Tell, R. A., & Rosenthal, R. (2008). Selective publication of antidepressant trials and its influence on apparent efficacy. *The New England Journal of Medicine*, 358(3), 252–260. <https://doi.org/10.1056/NEJMsa065779>

Vittengl, J. R. (2017). Poorer Long-Term Outcomes among Persons with Major Depressive Disorder Treated with Medication. *Psychotherapy and Psychosomatics*, 86(5), 302–304. <https://doi.org/10.1159/000479162>

Wang, S.-M., Han, C., Lee, S.-J., Jun, T.-Y., Patkar, A. A., Masand, P. S., & Pae, C.-U. (2018). Efficacy of antidepressants: bias in randomized clinical trials and related issues. *Expert Review of Clinical Pharmacology*, 11(1), 15–25. <https://doi.org/10.1080/17512433.2017.1377070>

Zalsman, G., Hawton, K., Wasserman, D., van Heeringen, K., Arensman, E., Sarchiapone, M., ... Zohar, J. (2016). Suicide prevention strategies revisited: 10-year systematic review. *The Lancet Psychiatry*, 3(7), 646–659. [https://doi.org/10.1016/S2215-0366\(16\)30030-X](https://doi.org/10.1016/S2215-0366(16)30030-X)