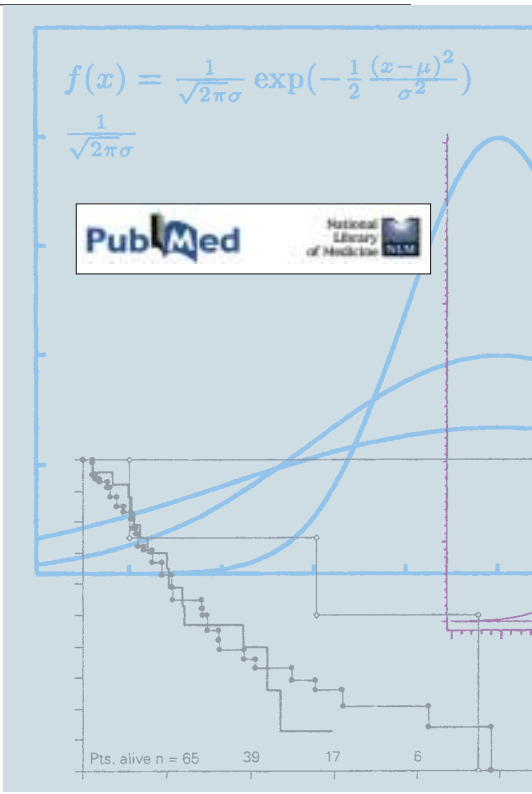
Wie evident ist Evidenz-basierte Medizin?

Eine Betrachtung (auch) über „atypische“ Neuroleptika



Mit Evidenz-basierter Medizin wird impliziert, dass die derzeit beste wissenschaftliche Erkenntnis systematisch in die medizinische Versorgung und die Beantwortung klinischer Fragestellungen einbezogen wird. Doch die Kriterien der Evidenz und die Qualität der zugrunde liegenden klinischen Studien sind nach wie vor Gegenstand zahlreicher Kontroversen: Prof. Dr. Matthias Dose beanstandet die empirische Evidenz und ihre Umsetzung im Bereich der Antipsychotika-Therapie, PD Dr. Stefan Leucht nimmt kritisch dazu Stellung.

Kritik an der Validität

Guyatt & Rennie, 2002, bezeichneten als „Evidenz“ die empirische Beobachtung der Beziehung von Ereignissen. Da die empirische Betrachtung subjektiven Einflüssen unterliegt, wird im wissenschaftlichen Bereich versucht, diesen „bias“ durch ein entsprechendes „Design“ wissenschaftlicher Studien auszuschalten oder zu minimieren. Dabei gilt die randomisierte, kontrollierte Studie der klinischen Beobachtung aus offenen Studien und Kasuistiken überlegen [21]. Bei Evidenz-basierter Medizin geht es um die systematische Bewertung publizierter Studienergebnisse, in der Regel in Form einer „Metaanalyse“ mittels Internet-Recherche (z. B. Medline, Current Contents, Cochrane Register) oder Literatur-Recherche aus Übersichtsarbeiten (z. B. Cochrane Reviews). Auch Pharmafirmen werden gebeten, bislang unpublizierte Studien vorzulegen [15]. Die so gesammelten Daten werden nach festgelegten Kriterien (Randomisierung, Plazebo-Kontrolle, definierte Ein- und Ausschlusskriterien der Studien, Behandlungsdauer, verwendete Rating-Instrumente, Therapieerfolgskriterien) erwählt und bewertet („Critical Appraisal“) [3]. Folgende Parameter sind dabei gängig:

▶ **Relative Risikoreduktion (RRR):** Wie verändert sich bezogen auf eine Kontrollgruppe das Risiko für das Eintreten eines Ereignisses (z. B. Rückfall)?

▶ **Absolute Risikoreduktion (ARR):** Um wieviel Prozent verringert sich das Risiko eines Ereignisses unter Medikation gegenüber Plazebo?

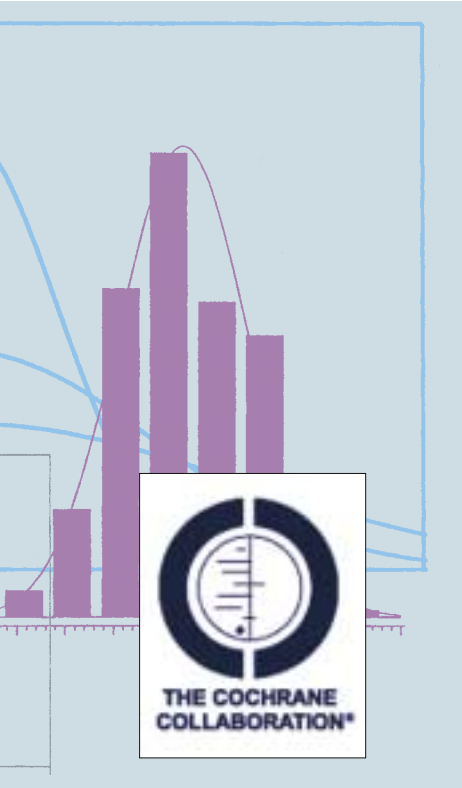
▶ **Number needed to treat (NNT):** Wie viele Patienten müssen behandelt werden, damit ein Ereignis eintritt oder nicht? (NNT von 10 bedeutet, dass zehn Patienten mit der Intervention behandelt werden müssen, damit das Ereignis bei einem Patienten eintritt oder nicht.)

▶ **Effektstärke:** Sie misst ausgehend von den „Outcome“-Maßen (z. B. BPRS, CGI, SANS, EPS-Skalen, AIMS) den Einfluss einer Behandlung auf abhängige Variablen (z. B. BPRS-Score). Effektstärken werden mit verschiedenen statistischen Methoden bestimmt, die wiederum zu sehr unterschiedlichen Interpretationen führen können

Evidenz-basierte Medizin in der Kritik

Auf Probleme, die mit der Metaanalyse mit Hinblick auf Evidenz-basierte (und folgerichtig in Behandlungsleitlinien umzusetzende) Medizin entstehen können, haben verschiedene Autoren [7, 12] schon vor Jahren hingewiesen: Zum einen beziehen die entsprechenden Literaturrecherchen ausschließlich englischsprachige Arbeiten in die Analysen ein, was beispielsweise erklärt, warum Perazin, ein erwiesenermaßen günstiges Neuroleptikum hinsichtlich extrapyramidalen Störungen EPS (< 10 % der behandelten Patienten) [8], kaum in einer aktuellen Metaanalyse auftaucht. Zum anderen werden Studien nach unterschiedlichen Kriterien (mit und ohne Plazebokontrolle, Symptombeseitigung versus Symptomreduktion als Responsekriterium) in die Analysen mit einbezogen. Daneben besteht das Problem, dass Studien mit neutralem beziehungsweise negativem Ergebnis in der Regel nicht publiziert werden, was zu einer völligen Verzerrung der Ergebnisse von Metaanalysen führen kann. Nicht umsonst hat der Vorsitzende der Arzneimittelkommission der Deutschen Ärzteschaft, Prof. Müller-Oerlinghausen, jüngst die Einrichtung einer zentralen Stelle zur Registrierung laufender Studien gefordert [10].

Auf „positivere“ Ergebnisse Pharmaindustrie-gesponsorter Studien gegenüber unabhängig durchgeführten Studien verweist eine im JAMA publizierte Untersuchung, die zeigt, dass bei Sponsoring in 51 % der Studien das geprüfte Medikament als „Medikament der Wahl“ empfohlen wurde, während in Studien



von unabhängigen Organisationen nur in 16% diese Empfehlung gegeben wurde. Im Rahmen einer Regressionsanalyse erwies sich „Sponsoring“ als signifikanter Prädiktor von Studienergebnissen, während Behandlungseffekte und unerwünschte Nebenwirkungen darauf weniger Einfluss nahmen [1].

Wenn gleichzeitig in Betracht gezogen wird, dass in den USA im Jahr 2000 der Anteil industrieller Mittel an der gesamten biomedizinischen Forschung 62% betrug [10], so stellt sich zu Recht die Frage, wer heute noch unabhängige Studien durchführt. Vor diesem Hintergrund verwundert es nicht, dass ein Teil der Ärzteschaft (z. B. Brune K., Deutsches Ärzteblatt, 19.3.2004; Niroomand F., Deutsches Ärzteblatt, 25.6.2004) kritisches Unbehagen gegen eine zu große Euphorie gegenüber der Evidenz-basierten Medizin artikuliert, während andere [20] für eine „grundsätzlich positive Einstellung“ plädieren, die erst zu notwendiger Weiterentwicklung und Verfeinerung führen könne. Sofern die „grundsätzlich positive Einstellung“ nicht einer kritischen Bewertung der Methoden und Aussagen der Evidenz-basierten Medizin und einer kritisch-abwägenden Haltung gegenüber auf deren Ergebnissen basierenden „Therapieleitlinien“ entgegensteht, wollen wir uns diesem Anliegen nicht verschließen.

Was heißt denn hier „atypisch“?

Als „atypisch“ gilt seit der Einführung von Clozapin eine antipsychotisch wirksame Substanz, die in therapeutisch wirksamen Dosierungen keine klinisch erkennbaren EPS hervorruft. Für die sowohl chemisch als auch hinsichtlich des Rezeptorprofils und der Wirkmechanismen heterogene Gruppe der ab 1995 auf den Markt gekommenen, neu entwickelten antipsychotischen Substanzen (Risperidon, Olanzapin, Quetiapin, Amisulprid, Ziprasidon und Aripiprazol) trifft dieses Kriterium (dosisabhängig) nur bedingt zu. Wenn sie dennoch als „atypisch“ angeboten werden, so entspricht dies mehr den Marketing-Strategien der Hersteller, als den tatsächlichen Gegebenheiten. Aus diesem Grund hat die „World Psychiatric Association“ (WPA) vorgeschlagen, die antipsychotischen Substanzen in „Generationen“ einzuteilen, wobei die oben genannte Gruppe der „atypischen“ Neuroleptika die zweite Generation, Aripiprazol die dritte und die „typischen“ Neuroleptika die erste Generation bilden [19].

Metaanalysen zu „atypischen“ Neuroleptika

In der Psychopharmakologie ist die Markteinführung der neuen „atypischen“ Neuroleptika (beginnend mit Risperidon, 1995) von einem wahren „Boom“ von Metaanalysen begleitet worden [4, 9, 15, 16, 23]. Eine erste vergleichende Metaanalyse „neuer“ Neuroleptika [23] befasste sich mit Studien zu Olanzapin, Quetiapin, Risperidon und Sertindol im Vergleich zu Haloperidol und Plazebo (publiziert ab 1992). Das wenig überraschende Ergebnis: Alle „neuen“ Substanzen erwiesen sich mit einer als „mäßig“ eingeschätzten Effektstärke wirksamer als Plazebo. Gegenüber Haloperidol zeigten sich Sertindol und Quetiapin bezüglich der schizophrenen Gesamtsymptomatik gleich wirksam, Risperidon und Olanzapin als gering überlegen. Bezüglich „Negativsymptomatik“ war auch Haloperidol wirksam (hier besser als Quetiapin), jedoch als nicht so effektiv wie Olanzapin und Risperidon. Der Anticholinergikaverbrauch war unter den „neuen“ Substanzen geringer im Vergleich zu Haloperidol, wobei sich Risperidon als weniger vorteilhaft erwies.

In der Diskussion der Ergebnisse decken die Autoren bemerkenswert offen die methodischen Schwachstellen ihrer Metaanalyse auf: Die „Überlegenheit“ von Risperidon und Olanzapin beruht auf minimalen (6–7%) Verbesserungen der PANSS-Scores. Bezüglich Negativsymptomatik weisen die Autoren darauf hin, dass die Metaanalyse nicht zwischen Wirkungen auf „primäre“ und „sekundäre“ (z. B. EPS-abhängige) Negativsymptome unterscheiden kann und dass (was auch den unterschiedlichen Anticholinergikaverbrauch erklärt) unter „konventionellen“ Neuroleptika signifikant häufiger EPS zu beobachten waren. Bezüglich der in den analysierten Studien verwendeten Vergleichsdosierungen von Haloperidol (bis 20 mg/Tag) verweisen sie auf die „relativ hohe“ Dosierung von Haloperidol in den meisten Studien, was das Auftreten von EPS „begünstigt“ habe. Lediglich eine Studie, bei der Haloperidol in einer Dosierung von 4 mg/Tag gegen Sertindol (dessen Zulassung gegenwärtig wegen QT-Zeit-Verlängerung ruht) geprüft worden war, ergab auch gegenüber dieser niedrigen Haloperidol-Dosis weniger EPS [22].

Als aufschlussreich für die Vergleichsdosis konventioneller Neuroleptika erwies sich eine Metaanalyse von 52 randomisierten Studien (n = 12.649 Patienten), in denen Amisulprid, Clozapin, Olanzapin, Quetiapin, Risperidon und Sertindol mit konventionellen Neuroleptika (in der Regel Haloperidol oder Chlorpromazin) oder untereinander verglichen wurden [9]: Sobald die zum Vergleich eingesetzte Dosis von konventionellen Neuroleptika die Äquivalenzdosis von 12 mg/Tag Haloperidol unterschritt, waren keine Unterschiede zwischen konventionellen und atypischen Neuroleptika mehr feststellbar, weder bei antipsychotischer Wirksamkeit noch bei der Verträglichkeit (mit Ausnahme geringerer EPS unter „Atypika“). Unterstützt wird die Hypothese, die in den gesponsorten Vergleichsstudien (die zur Zulassung führten) gefundene Überlegenheit „atypischer“ gegenüber herkömmlichen Neuroleptika könnte auf zu hohen Dosierungen der verwendeten Vergleichssubstanzen beruhen, durch eine Metaanalyse von Studien, in denen als herkömmliche Vergleichssubstanzen Neu-

roleptika mit niedrigerem EPS-Risiko (Chlorpromazin, Chlorprothixen, Perazin etc.) eingesetzt wurden [16]. Hier traten ausschließlich unter Clozapin signifikant weniger EPS als unter den Vergleichssubstanzen auf und es zeigte sich, dass sich bei Verwendung von weniger als 600 CPZ-Äquivalenten das EPS-Risiko herkömmlicher Substanzen nicht von dem der „Atypika" unterschied.

Vorsichtig schließen die Autoren (die im „Conflict of interest statement" zahlreiche Verbindungen zu Herstellern atypischer Neuroleptika offenlegen) aus ihrer Analyse, dass angemessen dosierte, niedrig-potente herkömmliche Neuroleptika den „atypischen" in der Schizophrenie-Therapie ebenbürtig sein könnten. Das Aufgreifen dieser Analyse im „Deutschen Ärzteblatt" („Zweifel an besserer Verträglichkeit", 12.5.2003) rief prompt Stellungnahmen zur Verteidigung der „guten und zum Teil besseren Wirksamkeit atypischer Neuroleptika" hervor [z. B. 20]. Widersprüchlich zu den Aussagen der Metaanalyse behaupten die Autoren, sie hätte „Hinweise" ergeben, dass ältere, niedrigpotente Substanzen in „individuell optimaler Dosierung" (davon ist keine Rede) vergleichbar wirksam seien und so „möglicherweise" nicht häufiger extrapyramidale Nebenwirkungen aufwiesen. Die Frage, ob niedrigpotente Neuroleptika im niedrigen Dosisbereich gleich wirksam seien wie Atypika, hätten die Autoren nicht beantworten wollen (was nicht den Aussagen der Arbeit von Leucht et al. entspricht). Um dem Leser gleich noch die gewünschte „Orientierung" zu geben, behaupten die Autoren, die „Leitlinie der DGPPN" (1998) würde atypische Neuroleptika in fünf Indikationen als „Therapie der ersten Wahl" empfehlen, obwohl die Leitlinie lautet: „Als Substanzen erster Wahl werden mittel- und hochpotente typische oder gleichrangig atypische Antipsychotika empfohlen (...). Auf Grund des günstigeren Nebenwirkungsprofils (...) ist der primäre Einsatz von atypischen Neuroleptika bei Ersterkrankten in Betracht zu ziehen" (DGPPN, 1998).

Zur Unterstützung ihrer Behauptungen führen Weinmann & Gaebel zwei weitere Metaanalysen an, die eine „zum Teil überlegene Wirkung atypischer Neuroleptika in der Akuttherapie" [4] und „eine geringfügige, aber signifikante Überlegenheit in der Verhinderung von Krankheitsrezidiven" [16] gezeigt hätten. In die Analyse von Davis et al. waren 124 kontrollierte Studien (1953–2002) zum Vergleich atypischer mit typischen (10 Studien) beziehungsweise zwischen atypischen Neuroleptika (18 Studien) einbezogen worden. Eine signifikante Überlegenheit gegenüber koventionellen Neuroleptika ergab sich lediglich für Clozapin, Amisulprid, Risperidon und Olanzapin, jedoch nicht für Aripiprazol, Quetiapin und Ziprasidon, woraus die Autoren den Schluss ziehen, atypische Neuroleptika seien keine homogene Gruppe. Ein gerne zitierter Nebenaspekt dieser Studie ist, dass die Autoren den von Geddes et al. (2000) gefundenen Effekt zu hoher Haloperidol-Dosen auf die scheinbare Überlegenheit atypischer Substanzen nicht bestätigen konnten. Dazu haben Geddes et al. (2004) deutlich gemacht, dass es die statistische Methodik ist, die den Einfluss der Dosis herkömmlicher Neuroleptika auf Wirksamkeit und Nebenwirkungen hervortreten lässt oder nicht. Während sie in ihrer Meta-Regressionsanalyse ein hierarchisches, multivariates Modell gewählt hätten, das die Dosierung den entsprechenden Patientengruppen zuordnete, hätten Davis et al. den Effekt der Dosierung als undifferenzierte Durchschnittsgröße einer univariaten Analyse unterzogen. Sehr deutlich äußern Geddes et al., dass diese Methode „optimal" gewählt sei – wenn man keinen Effekt der Dosis finden wolle. Auch die ins Feld geführte Metaanalyse zur rezidivprophylaktischen Wirkung atypischer im Vergleich zu typischen Neuroleptika [16] überzeugt nicht: Elf Vergleichsstudien (2032 Patienten; Haloperidol bis 30 mg/Tag) ergaben lediglich für Risperidon und Sertindol eine signifikant niedrigere Rückfallrate. In der Diskussion weisen die Autoren darauf hin, dass nicht zu bestimmen sei, ob die „bescheidene" Überlegenheit atypischer Neuroleptika auf bessere Wirksamkeit, bessere Compliance oder beides zurückzuführen sei. Sie diskutieren, dass Untersuchungen in den 70er- und 80er-Jahren gezeigt hätten, dass mit Dosierungen < 5 mg/Tag Haloperidol eine ausreichende Rückfallprophylaxe zu gewährleisten sei, dass aber (bis auf eine Ausnahme) keine der analysierten Studien diese Dosierungsstrategie angewendet hätte; vielmehr hätte die Mehrzahl der Patienten 10–20 mg/Tag Haloperidol erhalten. Über das schwache Ergebnis der Reduktion des absoluten Rückfallrisikos durch atypische Neuroleptika um 8% (wenn 1.000 Patienten ein Jahr mit Atypika behandelt werden, treten – gegenüber offensichtlich mit zu hohen Haloperidol-Dosen behandelten Patienten – 80 Rückfälle weniger auf) trösten die Autoren mit dem Verweis hinweg, das entspräche der Evidenz für den Einsatz von Acetylsalicylsäure zur Prävention von vaskulären Ereignissen. Das wiederum tröstet uns.

Neuere klinische Studien

Zusammenfassend zeigt sich, dass die auf Metaanalysen von Studien zum Vergleich atypischer mit herkömmlichen Neuroleptika basierende Evidenz derzeit noch wenig überzeugt, weil der Einfluss inhärenter Mängel der „randomisierten, Plazebo-kontrollierten" Studien (durch den Vergleich unangemessener Dosierungen [6]) die Ergebnisse hinsichtlich erwünschter als auch unerwünschter Wirkungen zu Ungunsten der herkömmlichen Substanzen (in der Regel Haloperidol) beeinflusst. Da sich zu hohe Dosierungen herkömmlicher Neuroleptika auch in Form „sekundärer Negativsymptome" (emotionaler und sozialer Rückzug, Anhedonie, Antriebslosigkeit bis hin zu kognitiven Störungen) abbilden können, steht auch die behauptete überlegene Wirkung der atypischen Neuroleptika auf „Negativsymptomatik" und kognitive Störungen zur Diskussion.

So ergab eine kürzlich in den USA durchgeführte Studie an ersterkrankten schizophrenen Patienten zwölf Wochen nach Behandlung eine nur geringe Überlegenheit von Olanzapin (9,63 mg/Tag) versus Haloperidol (4,63 mg/Tag) bezüglich neurokognitiver Funktionen [13]. Eine doppelblinde, randomisierte 1-Jahres-Studie zum Vergleich der Effektivität und Behandlungskosten von Olanzapin (5–20 mg/Tag) und Haloperidol (5–20 mg/Tag) und jeweils prophylaktischer Gabe eines Anticholinergikums [18] zeigte, dass Olanzapin bei Compliance, Symptomkontrolle, EPS und Lebensqualität nicht überlegen ist. Den Vorteilen, ge-

ringere Akathisie und kognitiver Symptome (die keine Auswirkungen auf die Lebensqualität hatten), wurden Gewichtszunahme und höhere Behandlungskosten (3.000–9.000 $ pro Jahr) unter Olanzapin gegenübergestellt.

Mit Hinblick auf das Argument, durch die geringeren Nebenwirkungen erhöhe sich die Compliance der Patienten und dadurch reduzierten sich Rückfälle, stationäre Wiederaufnahmen und Kosten haben Kilian & Angermeyer (2004) eine prospektive Beobachtungsstudie durchgeführt. Sie zeigt, dass Patienten unter herkömmlichen Neuroleptika, weniger stationäre Behandlungen benötigten und geringere Kosten verursachten.

Die Kosten der Schizophrenietherapie erhöhen sich auch durch die seit einigen Jahren zunehmende Verordnung mehrerer antipsychotischer Substanzen gleichzeitig (Polypharmazie). So berichtet eine amerikanische Studie mit 372 Patienten über eine Zunahme „antipsychotischer Polypharmazie" von 27,5 % auf 44,7 % zwischen 1998 und 2000, wobei die häufigste Kombination die eines herkömmlichen Neuroleptikum mit Clozapin oder Olanzapin ist [17].

Klinische Wirklichkeit

Welchen Einfluss Evidenz-basierte Medizin und auf ihr beruhende „Behandlungsleitlinien" auf die medikamentöse Therapie schizophrener Patienten in den vergangenen zehn Jahren genommen haben, wurde in einer Umfrage bei ärztlichen Leitern von acht psychiatrischen Pflichtversorgungskliniken untersucht [Berger & Fritz, 2004]. Der Fragebogen von 1990 zu verschiedenen Fallkonstellationen (z. B. Mittel der ersten Wahl bei akuter Schizophrenie mit/ohne Erregungszustand) wurde denselben Ärzten 2000 erneut vorgelegt. Es stellte sich heraus, dass auch 2000 bei Schizophrenie mit akuten Erregungszuständen an erster Stelle Haloperidol (mit Maximaldosis von 10 mg/Tag statt 20 mg/Tag) eingesetzt wurde und keine atypischen Neuroleptika. Auch bei akuten schizophrenen Psychosen ohne Erregungszustände wurde weiterhin Haloperidol als Mittel der ersten Wahl genannt, atypische Neuroleptika wurden bei Therapieversagen als Alternative erwogen. Die Autoren kontrastieren dies zu Recht mit Aussagen von Umfragen, dass 87 % aller Psychiater die Auffassung vertreten, atypische Neuroleptika seien zu bevorzugen.

Zu pragmatisch und kritisch?

Es stellt sich die Frage, woher es kommt, dass – entgegen den Bemühungen von Meinungsbildnern und „Leitlinien-Machern" – sich die behauptete Evidenz antipsychotischer Wirksamkeit und Überlegenheit der atypischen Neuroleptika in der Praxis so zögernd umsetzt. Fehlt es an der Vermittlung entsprechender Informationen, an Fort- und Weiterbildung? Oder sind die praktisch tätigen Kollegen nur so kritisch und pragmatisch, sich jenseits gesponsorter Studienergebnisse und glänzender „Folder" ein eigenes Urteil zu bilden und sich dem behaupteten Fortschritt (auch mit Hinblick auf noch unbekannte oder zu wenig bekannte Risiken wie Diabetes, Prolaktinerhöhung, vaskuläre Ereignisse) nur differenziert zurückhaltend nähern. Es sieht so aus, als sei letzteres der Fall. Nach Abklingen der „Einführungseuphorie" rücken neue Studien den Stellenwert der „atypischen" Neuroleptika realistisch zurecht: eine heterogene, antipsychotisch wirksame Substanzgruppe mit unterschiedlichem Rezeptor- und entsprechend unterschiedlichem Profil erwünschter und (hinsichtlich Häufigkeit und Schweregrad angesichts der kurzen Zulassungszeit noch unkalkulierbarer) unerwünschter Wirkungen. Wo sie im Instrumentarium der zur Verfügung stehenden Substanzen anzusiedeln sein werden, kann erst nach einem dosisadäquaten „Kopf-zu-Kopf"-Vergleich mit herkömmlichen Substanzen bestimmt werden. Herkömmliche Neuroleptika (auch eine heterogene Gruppe) einfach aus dem Repertoire der Behandlungsmöglichkeiten zu streichen, wie bereits in den APA-Richtlinien zur Behandlung schizophrener Psychosen (2004) geschehen, mag zwar den Herstellerfirmen „atypischer" Neuroleptika gefallen, ist aber angesichts der bestehenden Datenlage nicht „evident".

Literatur beim Verfasser

Prof. Dr. med. Matthias Dose
Bezirkskrankenhaus Taufkirchen,
Bräuhausstr. 5,
84416 Taufkirchen/Vils,
E-Mail: m.dose@bkh-taufkirchen.de

Plädoyer für die Integration in ärztliches Handeln

Karl Popper hätte der NeuroTransmitter vermutlich erfreut, weil auf seinen Seiten auch kontroverse Diskussion ausgetragen werden, die wissenschaftlichen Fortschritt fördern und Lesern oft mehr Erkenntnis bringen als perfekt austarierte Übersichtsartikel. Dass dies nicht häufiger geschieht, liegt sicher auch daran, dass nur wenige Autoren es wagen, ihre Meinung so dezidiert – und ohne Rücksicht auf „conflicts of interest" – auf den Punkt zu bringen wie Prof. Dr. Matthias Dose.

Der gelungenen Einleitung von Dose wäre hinzuzufügen, dass evidence-based-medicine (EBM) im Grunde nichts Neues ist. Schon immer gab es in der Medizin den Anspruch, Therapieoptionen auch wissenschaftlich zu überprüfen. Das Neue besteht nur darin, dass EBM fordert, systematisch vorzugehen und randomisierte Studien besonders zu gewichten. Die maßgeblich treibende Kraft war und ist dabei die Cochrane-Collaboration (www.cochrane.org), eine in den 80er-Jahren in England gegründete Organisation, mit dem Ziel, systematische Übersichtsarbeiten zu medizinischen Behandlungsformen zu erstellen, die auf dem derzeitigen Goldstandard, nämlich randomisiert-kontrollierten Studien, basieren. Die Schizophrenie-Gruppe, zu deren Herausgeberteam Stefan Leucht zählt, hat derzeit etwa 80 solcher systematischer Reviews erstellt [1]. Diese erscheinen viermal jährlich in elektronischer Form in der Cochrane Library publiziert (The Cochrane Collaboration 2005), sodass sie im Gegensatz zu Übersichtsartikeln in Zeitschriften regelmäßig auf den neuesten Stand gebracht werden können. Der Begriff „systematische Review" beschreibt dabei das im Gegensatz zu konventionellen Übersichtsarbeiten

systematische Vorgehen: Die Suchstrategie, Einschlusskriterien (Studienart, Patienten und Interventionen) und die Art der Synthese der Studienergebnisse werden vor Beginn der Review-Erstellung festgelegt. In einer konventionellen Review verläuft der Erstellungsprozess hingegen eher ungeordnet und bietet viele Fehlerquellen. Die Systematik ist also das Kernelement der EBM, weniger der Aspekt der „Metaanalyse". Metaanalyse meint die statistische Kombination der Ergebnisse von einzelnen Studien mit Hilfe von Effektstärken. Die metaanalytische Kombination von Ergebnissen kann sinnvoll sein, zum Beispiel um durch eine Erhöhung der Fallzahl die statistische Aussagekraft („Power") zu verbessern oder um anhand der Effektstärke eine objektivere Aussage über die Größe des Unterschieds zwischen zwei Interventionen machen zu können. Das Hauptziel – zumindest der Cochrane Schizophrenia Group – bleibt aber das systematische Sichten des Wissens bezüglich einer Fragestellung. Auch wenn es keine Studie zu einer bestimmten Frage gibt, ist dies ein wichtiges Ergebnis, weil die Review dann zeigt, dass dringender Forschungsbedarf besteht. Oft ist es aber noch nicht möglich, auf der Basis von systematischen Reviews allgemeingültige Therapievorschläge zu machen.

Ad 1: „Evidenz-basierte Medizin in der Kritik"

Es trifft nicht zu, dass in systematische Reviews nur englischsprachige Artikel eingehen. Vielmehr wurde der „Sprachbias" gerade von Vertretern der EBM intensiv erforscht. Egger und Kollegen zeigten 1997, dass deutschsprachige Autoren Studien ohne signifikante Ergebnisse häufiger in deutschsprachigen Zeitschriften publizieren, solche mit signifikanten Ergebnissen häufiger in englischsprachigen. Würden tatsächlich nur englischsprachige Artikel in eine systematische Review aufgenommen, so würde das Ergebnis zugunsten positiver Ergebnisse verzerrt. Diesen Befund versuchten wir kürzlich für die Neurowissenschaften zu replizieren, fanden aber nur einen nicht signifikanten gleichen Trend [12]. Bei dieser Arbeit stellte sich auch heraus, dass von deutschen Autoren leider ohnehin nur sehr wenige randomisiert-kontrollierte Studien durchgeführt werden. So liegt es auch nicht an einem Ausschluss deutschsprachiger Publikationen a priori , dass unser gutes altes Perazin selten in solchen Reviews erscheint. Für unsere von Herrn Dose kritisierte, Metaanalyse in The Lancet wurden sogar einige chinesische Studien mühsam übersetzt [21]. Vielmehr stellte sich in einer Cochrane-Review über Perazin heraus, dass seit der Marktzulassung vor etwa 40 Jahren nur fünf kleine randomisierte Studien mit insgesamt 288 Patienten veröffentlicht worden sind [19]. Noch erstaunlicher war die entsprechende Cochrane Review über Benperidol, für das sich sogar nur eine randomisierte Studie mit 40 Teilnehmern fand [18].

Der beschriebene Publikationsbias, also dass Studien mit neutralem oder negativem Ergebnis oft nicht publiziert werden und daher auch nicht in systematische Reviews eingehen können, ist trotzdem ein großes Problem der EBM. Ironisch wurde dies mit „evidence biased medicine" betitelt. Metaanalysen bieten aber – im Gegensatz zu konventionellen Reviews – statistische Methoden, wie ein solcher Publikationsbias aufgedeckt werden kann (z. B. die „funnel-plot" Methode oder „fail-safe" Berechnungen [6, 27]). Auch wird versucht, Studien sogar dann einzuschließen, wenn sie nur als Kongressposter oder Vortrag vorliegen („grey literature"). Und es sind gerade die führenden Vertreter der EBM, die auf die Einrichtung zentraler Stellen zur Registrierung sämtlicher laufender Studien drängen, um diesen Bias zu minimieren. In England ist ein solches Register bereits Realität. Auch die FDA verlangt seit Neuestem, dass Firmen bei der Einreichung alle Studien, auch die bis dato nicht publizierten, offen legen. Alle Vertreter der EBM hoffen mit Nachdruck, dass diese Forderungen bald überall in die Realität umgesetzt werden und jegliches Material zugänglich gemacht wird.

Es ist problematisch, dass heutzutage der überwiegende Teil der Medikamentenstudien von der Pharmaindustrie gesponsert wird. Diese Studien sind zwar von hervorragender Qualität (große Fallzahlen, multizentrisch, kontinuierliches professionelles Monitoring), die von industrieunabhängigen Studien nur selten erreicht wird. Aber es gibt zwangsläufig auch einen Interessenskonflikt. So fanden auch wir, ähnlich der von Dose zitierten Untersuchung in JAMA, einen Einfluss des pharmazeutischen Sponsors auf das Ergebnis von Direktvergleichen zwischen „atypischen" Antipsychotika [12]. Aber auch hier möchten wir betonen, dass das stärkste Drängen auf pharmaunabhängige Studien von Protagonisten der EBM ausgeht. Von Cochrane-Kollegen aus Brasilien und Indien wurden kürzlich mit minimaler finanzieller Förderung zwei große, industrieunabhängige Studien über Strategien zur Sedierung angespannter psychotischer Patienten in renommierten Zeitschriften publiziert [2, 13]. Der Vorwurf geht somit an die falsche Adresse. Auch wir würden systematische Reviews bevorzugen, die überwiegend auf öffentlich finanzierten Studien basieren. Im Bereich der Antipsychotika gibt es aber nur erschreckend wenige solche Studien.

Was fehlt, ist eine Diskussion über Alternativen zu EBM. Dies verwundert jedoch nicht, denn es kann unserer Meinung nach niemand ernsthaft behaupten, es gäbe eine alternative Möglichkeit. Die größten Gegenspieler der EBM wurden ironisch „Eminence-Based Medicine", „Experience-Based Medicine" und „Marketing-Based Medicine" genannt. Bei „Experience-Based-Medicine" verlässt man sich – überspitzt ausgedrückt – allein auf die eigene klinische Erfahrung. Der Einbezug der klinischen Erfahrung ist aber selbtsverständlich auch bei EBM wichtig. So definierten Berner et al. 2000: „Unter Evidence-based medicine versteht man die gewissenhafte und vernünftige Anwendung der besten zur Zeit vorhandenen wissenschaftlich-empirischen Belege in Kombination mit der individuellen klinischen Erfahrung (Expertise) bei medizinischen Entscheidungen.

EBM strebt keine „Kochbuch-Medizin" an; diese ist allein deshalb schon unmöglich, weil es für viele Behandlungsschritte keine gesicherte Evidenz gibt und viele Entscheidungen daher auf klinische Erfahrungen gründen müssen. Verlässt man sich aber nur auf die eigene Erfahrung oder die Behandlungstradition des Hauses, kann dies zu gravierenden Fehleinschätzungen führen. Als Beispiel sei die bis Ende der 80er-Jahre vielerorts praktizierte neuroleptische Hochdosis-

therapie genannt, die auf positiven individuellen Erfahrungen und der Hoffnung „viel hilft viel" beruhte, bis sich in Übersichtsarbeiten zeigte, dass Haloperidoldosierungen über 10 mg/Tag in aller Regel nicht vorteilhaft sind [3]. Auch ist es fraglich, ob wir uns nur an der „Eminence-Based Medicine" ausrichten sollten, das heißt Therapieentscheidungen vor allem an der Meinung renommierter Professoren festmachen sollten. Einerseits können auch „Eminenzen" irren und andererseits gibt es möglicherweise auch bei Ihnen Interessenskonflikte zum Beispiel durch Kooperationen mit der pharmazeutischen Industrie. Die Verschreibungsgewohnheiten ganz den Gesetzen des Marktes zu überlassen (ironisch „Marketing-Based-Medicine"), möchte wohl auch niemand. Es gibt also keine Alternativen zur EBM, wir können diese nur ständig verbessern und fortentwickeln.

Ad 2: Was heißt denn hier „atypisch"?

Wir stimmen völlig zu, dass der Begriff der „atypischen" Antipsychotika wissenschaftlich keinen Sinn ergibt. Schon Möller (2000) hat in einem Übersichtsartikel beschrieben, dass es sich bei der Abgrenzung atypisch/typisch nicht um eine Dichotomie, sondern vielmehr um ein Kontinuum handelt, nach dem auch einige alte Substanzen wie Perazin eher als Atypika eingeordnet werden müssen. Dies wurde von Fischer-Barnicol, Dose und Kollegen 2003 in einer großen naturalistischen Studie an den bayerischen Bezirkskrankenhäusern bestätigt.

Ad 3: Metaanalysen zu „atypischen" Neuroleptika

Dass es zu einem „Boom" an Metaanalysen über neue Antipsychotika gekommen ist, liegt an der Bedeutung der Fragestellung und der ständig steigenden Studienzahl in diesem Gebiet. Inzwischen wurden in mehr als 150 randomisierten Doppelblindstudien „atypische" Antipsychotika mit konventionellen Substanzen verglichen. Bei dieser Flut an Publikationen ist es unmöglich, mit konventionellen Review-Methoden, also der narrativen Aufzählung einzelner Studienergebnisse, den Überblick zu wahren und zu einer objektiven Gesamtbewertung zu kommen. Dieses Problem versuchen Metaanalysen durch die Berechnung von (mittleren) Effektstärken zu lösen, bei denen auch der Einfluss von Moderatorvariablen (z. B. Dosis) analysiert werden kann. Dieses Verfahren ist nicht der Weisheit letzter Schluss, es gibt eine Reihe methodischer Probleme. Es sei aber noch einmal betont, dass zumindest die Cochrane Schizophrenia Group den Aspekt „systematische Review" für wichtiger hält als „Metaanalyse".

Dass Haloperidol auch in niedriger Dosierung bis 4–6 mg/Tag mehr EPS verursacht als neue Antipsychotika [32] wurde in mindestens drei weiteren multizentrischen, randomisierten Studien bestätigt (Haloperidol 1–6 mg/Tag versus Risperidon 1–6 mg/Tag [8]; Olanzapin 9,1 mg/Tag versus Haloperidol 4,4 mg/Tag [22]; Risperidon versus Haloperidol jeweils 1–4 mg/Tag [28]). Auch in der von Weinmüller unter Supervision von Dose im Bezirkskrankenhaus Taufkirchen durchgeführten naturalistischen Studie kam es unter Haloperidol (ca. 6 mg/Tag) und Flupentixol (ca. 10 mg/Tag) zu mehr EPS als unter neuen Antipsychotika. Jedoch erhielten die Patienten unter Atypika mehr Benzodiazepine und Antidepressiva und es fanden sich keine Wirksamkeitsunterschiede. Dennoch müsste Haloperidol mit Hilfe der neuroleptischen Schwellenmethode [11] dosiert werden, um die EPS-Rate noch weiter zu senken. Es gibt Belege, dass die Dosierung nach der neuroleptischen Schwelle ausreichende Wirkung bei minimalen EPS erzielt (Haloperidol ca. 2–3 mg/Tag [23]). Die Methode müsste aber weiter erforscht werden; sie hat sich – obwohl in Deutschland seit etwa 50 Jahren bekannt – klinisch nie durchgesetzt.

Es wurde bereits beschrieben, dass auch einige alte Antipsychotika, wie Chlorpromazin, Thioridazin oder Perazin relativ wenig EPS verursachen und deshalb als „alte Atypika" bezeichnet werden könnten. Dies war der Anlass für unsere Metaanalyse, in die wir nur Vergleiche zwischen neuen und „niedrigpotenten" konventionellen Antipsychotika aufnahmen. Wenn die niedrigpotenten konventionellen Antipsychotika in mittleren Dosierungen unter 600 mg/Tag verwendet wurden, fand sich kein Unterschied in den EPS-Raten. Gleichzeitig sprachen aber etwa 11 % mehr Patienten auf die Behandlung mit neuen Antipsychotika an [20]. Tatsächlich haben wir diese Ergebnisse vorsichtig als Hypothesen-generierend und keinesfalls als Behandlungsempfehlungen diskutiert. Dies hat nichts mit einem potenziellen „Conflict of interest" zu tun, sondern mit der Qualität der zugrunde liegenden Daten. Die meisten Studien waren sehr alt, wiesen nur geringe Fallzahlen auf und hätten den heute üblichen methodischen Kriterien kaum genügt. Aus diesen Daten heute eine Behandlungsempfehlung abzuleiten, wäre nicht möglich gewesen. Insofern sind wir mit der Stellungnahme von Weinmann und Gaebel, 2004, einverstanden. Ferner sind die Ergebnisse nicht auf Deutschland übertragbar, weil Chlorpromazin hier seit Jahrzehnten nicht mehr verwendet wird und Perazin nach heutigen Standards ungenügend untersucht worden ist. Hier sind weitere Studien erforderlich, der diesbezügliche Schlusssatz unserer Publikation war deshalb nicht als reine Floskel gemeint.

Es ist zwar richtig, dass Haloperidol als Vergleichssubstanz einen Bias zugunsten der neuen Medikamente darstellt, weil es bekanntermaßen EPS verursacht. Andererseits war Haloperidol aber in vielen Ländern das Standardpräparat. Auch im Bezirkskrankenhaus Taufkirchen sind Haloperidol und Flupentixol die überwiegend verwendeten konventionellen Antipsychotika [31]. Um die EPS-Rate weiter zu reduzieren, könnte überlegt werden, Haloperidol-Patienten auf „billige Atypika" wie Perazin oder Sulpirid umzustellen. Aber für die Akutbehandlung mit diesen Präparaten liegen kaum Dosisfindungsstudien vor. Es ist ein entscheidender Vorteil der Atypika, dass für sie vorab Dosisfindungsstudien durchgeführt worden sind. Das Wissen, dass etwa die optimale Olanzapindosis bei 10–20 mg/Tag liegt und auch in höherer Dosierung nicht mit häufigen EPS gerechnet werden muss, erleichtert die Behandlung ungemein. Insofern müsste das Rad der Geschichte zurückgedreht werden, um viele Aspekte der alten Substanzen neu zu erforschen.

In diesem Kontext werden auch die beiden widersprüchlichen Metaanalysen von Geddes et al. (2000) und Davis et al. (2003) diskutiert. Geddes zeigte, dass Haloperidol (12 mg/Tag) genauso wirk-

sam ist wie „Atypika“. Auch die Drop-Out-Raten unterschieden sich nicht. Daher sollte, so Geddes, Haloperidol als Medikament der ersten Wahl verwendet werden. Diese Schlussfolgerung war bereits damals überzogen, da allgemeine Drop-Out-Raten als Nebenwirkungsmaß verwendet wurden. In den Studien scheiden aber die meisten Patienten nicht wegen Nebenwirkungen, sondern wegen nicht ausreichender Wirksamkeit aus.

Geddes fand mit Haloperidol deutlich mehr EPS, auch unter 12 mg/Tag. In die Metaanalyse von Davis gingen mehr als doppelt so viele Studien ein, der Dosiseffekt ließ sich aber nicht replizieren. Dass dies, wie Geddes behauptet, an der statistischen Methode liegt, lässt sich ohne konkrete Rohdaten nicht nachvollziehen.

Wir erläutern erneut die Ergebnisse unserer Metaanalyse (wegen heterogenen Studienmaterials noch nicht optimal) über Rückfallraten: Nach einem Jahr hatten 23 % der Patienten in der Typika- und 15 % in der Atypika-Gruppe einen Rückfall. Dies entspricht einer absoluten Reduktion des Rückfallrisikos um 8 %. Relativ gesehen ist das Risiko unter „Atypika“ aber um etwa 35 % niedriger (8 % geteilt durch das Basisrisiko unter Haloperidol von 23 % x 100). Je tiefer das Basisrisiko ist, umso weniger absoluter Risikounterschied zwischen den Substanzen ist möglich. In diesem Fall ist also maximal nur eine Reduktion der Rückfallraten um absolut 23 % erzielbar. Wenn sich die in der Metaanalyse gefundene Reduktion der Rückfallraten um absolut 8 % aber in weiteren Studien bestätigen sollte, wäre dies gesundheitsökonomisch sehr relevant. Denn die Medikamentenkosten machen nur einen Bruchteil der direkten Kosten der Schizophrenie aus. Den überwiegenden Teil stellen durch Rückfälle bedingte Rehospitalisierungen dar, sodass sich auch ein eher kleiner Unterschied insgesamt rechnen kann [17].

Ad 4: Neuere klinische Studien

Tatsächlich berichteten Keefe et al., 2004, von einer signifikanten, aber nur geringen Überlegenheit von Olanzapin gegenüber Haldol im Bezug auf kognitive Funktionen. Es kam unter Olanzapin aber zu deutlich weniger EPS, obwohl nur etwa 4,4 mg/Tag Haldol verwendet wurden (publiziert in Lieberman et al., 2003). Die Studie von Rosenheck et al. (2003) ist bemerkenswert: Dass sich hier kein Wirksamkeitsunterschied zwischen Haloperidol und Olanzapin fand, lässt sich eventuell damit erklären, dass die Patienten im Vergleich zu früheren Studien chronifizierter waren und sich Symptome in beiden Gruppen im Einjahresverlauf nur wenig besserten. Erstaunlich ist, dass sich trotz mittlerer Haloperidol-Dosis (am Ende 14 mg/Tag) keine unterschiedlichen EPS-Raten zeigten. Zwar erhielten alle Haloperidol-Patienten prophylaktisch ein Anticholinergikum, diese Vorsichtsmaßnahme reichte in früheren Studien aber nicht immer aus, um EPS unter Chlorpromazin oder Haloperidol zumindest im Vergleich zu Clozapin zu vermeiden [14, 26]. Ob der Anstieg der Polypharmazie tatsächlich mit dem Gebrauch von Atypika zusammenhängt und nicht vielmehr mit den sich ständig verkürzenden Hospitalisierungszeiten in den USA, muss vorerst Spekulation bleiben.

Ad 5: Klinische Wirklichkeit

Liegt es wirklich an einer überlegenen Wirksamkeit von Haloperidol, dass es in vielen deutschen Kliniken in der Akutphase weiterhin als Medikament der Wahl gegeben wird? Es gibt keine randomisierten Studien, die eine solche Überlegenheit beweisen. Natürlich ist die Übertragbarkeit von Studienergebnissen gerade für die „Hochakutphase“ eingeschränkt, weil Patienten ihre schriftliche Einwilligung geben müssen und hochgradig erregte Patienten hierzu gar nicht in der Lage sind. Auf der anderen Seite könnte es sich bei der nach klinischem Eindruck überlegenen Wirksamkeit von Haloperidol auch um einen Scheineffekt handeln. Geben wir einem Patienten initial Haloperidol, so stellen sich bald EPS ein und das beruhigt uns. Wir sehen, dass sich „etwas tut“ und setzen dies mit Wirksamkeit gleich, obwohl es eigentlich nur die Nebenwirkungen sind. Die im Vergleich zu den meisten Industrienationen in Deutschland verzögerte Umsetzung der Leitlinienempfehlungen zugunsten der „atypischen“ Antipsychotika liegt doch auch am besonderen Kostendruck unseres Systems, wie der drohenden Regresspflichtigkeit bei Budgetüberschreitungen beziehungsweise der Notwendigkeit, sich in diesem Fall vor einem Gremium rechtfertigen zu müssen.

Literatur bei den Verfassern

Fazit

Bei den „atypischen“ Antipsychotika handelt es sich um eine heterogene Substanzgruppe mit unterschiedlichen Rezeptoren und entsprechend unterschiedlichen Nebenwirkungsprofilen. Der Ruf nach dosisadäquaten Kopf-zu-Kopf-Vergleichen mit herkömmlichen Substanzen kommt aber eigentlich zu spät. Amisulprid ist bereits als Generikum verfügbar, Risperidon wird nächstes Jahr generisch. Für solche Vergleichsstudien müssten rechtzeitig öffentliche Mittel bereitgestellt werden und sie müssten bereits zu Zulassungsbeginn durchgeführt werden. Hierfür fehlt aber sowohl bei den öffentlichen Forschungsförderern als auch bei den Krankenkassen nach wie vor die Bereitschaft. Versuchsweise haben wir im Rahmen des Programms „Klinische Studien“ des Bundesgesundheitsministeriums Fördermittel für die Studie „Perazin versus „Atypika“ beantragt, diese wurden aber (natürlich) nicht bewilligt. Wir würden uns gerne mit anderen Bezirks- und Landeskrankenhäusern zusammenschließen, um (wie die genannten Kollegen in Brasilien und Indien) mit einfach durchführbaren randomisierten Studien Fragestellungen wie „Ist Perazin genauso atypisch und wirksam wie Risperidon?“ oder „Ist Haloperidol bei hochakuten Patienten wirksamer als Quetiapin?“ zu beantworten. Ob wir die dafür erforderlichen Mittel jedoch bekommen werden, ist fraglich. Solange solche Studien nicht in ausreichender Zahl und Qualität durchgeführt wurden, gibt es weiterhin gute Gründe „atypische“ Antipsychotika zu verwenden. Denn die Datenlage spricht insgesamt für sie und sie sind viel besser untersucht als die alten Neuroleptika, bei denen wir noch nicht einmal die Frage nach der optimalen Dosis evidenzbasiert beantworten können.

PD Dr. med. Stefan Leucht
Klinik und Poliklinik für Psychiatrie und Psychotherapie, Klinikum rechts der Isar, Ismaninger Str. 22, 81675 München,
E-Mail: stefan.leucht@lrz.tum.de

Dr. med. Werner Kissling, München