There is a deep mistrust of psychiatry fostered by reports suggesting that psychotropic drug efficacy is very small. Kirsch et al concluded that antidepressants should only be used in severely ill patients;1 the efficacy of cholinesterase inhibitors in Alzheimer’s disease and of lithium prophylaxis in bipolar disorder has been questioned;2,3 and we found a smaller antipsychotic drug–placebo difference in schizophrenia than we intuitively expected.4 These reviews inspired an article in *The New Yorker* summarising them,5 and fuelled a vocal antipsychiatry movement.6,7 Psychiatrists, patients, caregivers and the press are unsettled by these findings and some may think that psychiatric medication is not worth the bother. But is this small efficacy really true, and what about other medical interventions? As medicine is becoming highly specialised, few psychiatrists are familiar with the evidence of general medicine and psychiatric drugs. In this context we reviewed the efficacy of psychiatric pharmacotherapy in the perspective of standard medical drugs, making this paper the first attempt to provide a panoramic overview of major drugs. It is not possible to compare qualitatively different outcomes in qualitatively different diseases, but one can compare the percentages of patients helped with a drug or placebo, keeping in mind the differences in outcome for the mere purpose of perspective. We hasten to add a warning not to be overly concrete and to interpret this review as a qualitative perspective and not as a comparison. Therefore we discuss major factors that need to be taken into account in the interpretation of clinical trials and systematic reviews.

**Background**

The efficacy of psychopharmacological treatments has been called into question. Psychiatrists are unfamiliar with the effectiveness of common medical drugs.

**Aims**

To put the efficacy of psychiatric drugs into the perspective of that of major medical drugs.

**Method**

We searched Medline and the Cochrane Library for systematic reviews on the efficacy of drugs compared with placebo for common medical and psychiatric disorders, and systematically presented the effect sizes for primary efficacy outcomes.

**Results**

We included 94 meta-analyses (48 drugs in 20 medical diseases, 16 drugs in 8 psychiatric disorders). There were some general medical drugs with clearly higher effect sizes than the psychotropic agents, but the psychiatric drugs were not generally less efficacious than other drugs.

**Conclusions**

Any comparison of different outcomes in different diseases can only serve the purpose of a qualitative perspective. The increment of improvement by drug over placebo must be viewed in the context of the disease’s seriousness, suffering induced, natural course, duration, outcomes, adverse events and societal values.

**Declaration of interest**

In the past 3 years S.L. has received fees for consulting and/or lectures from the following companies: Bristol-Myers Squibb, Actelion, Sanofi-Aventis, Eli Lilly, Essex Pharma, AstraZeneca, MedAvante, Alkermes, Janssen/Johnson & Johnson, Lundbeck Institute and Pfizer, and grant support from Eli Lilly. W.K. has received fees for consulting and/or lectures from Janssen-Cilag, Sanofi-Aventis, Johnson & Johnson, Pfizer, Bristol-Myers Squibb, AstraZeneca, Lundbeck, Novartis and Eli Lilly. All authors work in psychiatry.

Background

The efficacy of psychopharmacological treatments has been called into question. Psychiatrists are unfamiliar with the effectiveness of common medical drugs.

Aims

To put the efficacy of psychiatric drugs into the perspective of that of major medical drugs.

Method

We searched Medline and the Cochrane Library for systematic reviews on the efficacy of drugs compared with placebo for common medical and psychiatric disorders, and systematically presented the effect sizes for primary efficacy outcomes.

Results

We included 94 meta-analyses (48 drugs in 20 medical diseases, 16 drugs in 8 psychiatric disorders). There were some general medical drugs with clearly higher effect sizes than the psychotropic agents, but the psychiatric drugs were not generally less efficacious than other drugs.

Conclusions

Any comparison of different outcomes in different diseases can only serve the purpose of a qualitative perspective. The increment of improvement by drug over placebo must be viewed in the context of the disease’s seriousness, suffering induced, natural course, duration, outcomes, adverse events and societal values.

Declaration of interest

In the past 3 years S.L. has received fees for consulting and/or lectures from the following companies: Bristol-Myers Squibb, Actelion, Sanofi-Aventis, Eli Lilly, Essex Pharma, AstraZeneca, MedAvante, Alkermes, Janssen/Johnson & Johnson, Lundbeck Institute and Pfizer, and grant support from Eli Lilly. W.K. has received fees for consulting and/or lectures from Janssen-Cilag, Sanofi-Aventis, Johnson & Johnson, Pfizer, Bristol-Myers Squibb, AstraZeneca, Lundbeck, Novartis and Eli Lilly. All authors work in psychiatry.

**Method**

Identification of diseases of interest and search strategy

We reviewed textbooks,5,9 identified common diseases by consensus (S.L., S.H. and J.M.D.) based on frequency, importance and available treatment, and consulted national and international guidelines to identify primary treatments. We hand-searched the Cochrane Library, and searched Medline combining MeSH terms for the medical and psychiatric disorders with the MeSH term for meta-analysis (no time or language limit, last search May 2009) and references of included reports for systematic reviews of randomised controlled trials that applied meta-analysis and compared monotherapy of these treatments with placebo.

We first excluded meta-analyses of studies of subgroups (e.g. elderly people) and chose reviews of classes of drugs rather than single drugs (e.g. any antipsychotic, rather than only haloperidol) if available, based on the assumption that the original reviewers had made an appropriate decision to pool the drugs. We then chose the most recent reviews, because even if methodologically better an older review would have certainly been out of date. This was a conservative decision, because old meta-analyses in psychiatry usually had higher effect sizes (see Discussion and online Table DS1). The rare exceptions were slightly older meta-analyses that reported the indices necessary for our analysis more completely. These usually were Cochrane reviews which were
preferred in case of doubt, because they use similar methodology and always fully report the data. To corroborate these decisions we always compared different reviews for consistency of results and contacted authors in the rare case that the results were discrepant. (These additional reviews are quoted in the footnotes of the tables in the online data supplement.) The quality of the included systematic reviews was evaluated with the AMSTAR score (range of possible values 0–11).10 Only primary efficacy outcomes in the areas of interest according to the treatment guidelines were extracted.

**Statistical analysis**

For continuous outcomes we extracted effect sizes and their 95% confidence intervals, presented both as differences in original units (mean difference) and as standardised mean differences (SMD). Mean differences were calculated according to the general formula (mean group A) – (mean group B), e.g. 75 kg in the drug group minus 70 kg in the placebo group gives a mean difference in body weight of 5 kg. Standardised mean differences (SMDs) provide a difference in standard deviation units (mean group A – mean group B)/standard deviation, e.g. (75 – 70)/10 = 0.50, using the values from the previous example.

For dichotomous outcomes we presented the percentage of participants improved in the drug and placebo groups, the absolute risk/response difference (ARD; % responder drug – % responder placebo); the relative risk reduction (RRR; 1 – (% risk drug / % risk placebo) or relative response (RR) ratio (% responder drug / % responder placebo); and the number needed to treat (NNT), with their 95% confidence intervals. We also presented the $P$ value, the number of studies and participants included and the average study duration (see online Table DS2 for a detailed description of these parameters).

Where our five standard parameters (mean difference, SMD, ARD, RRR, RR, NNT) were not reported in the studies, we transformed the existing data, or re-calculated meta-analyses by entering single study results using Review Manager version 5.0 or Comprehensive Meta-analysis version 2 for Windows.11,12 S.H. ran the searches, S.H. and S.L. selected the reports. S.H. extracted the data, S.L. independently verified them, disagreements were resolved by J.M.D. and W.K., and M.D. rated the AMSTAR score.

**Results**

The Medline searches yielded 6175 abstracts and we hand-searched 1830 titles of Cochrane reviews – see online Figs DS1–24 for Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) diagrams of the selection process.13 We included 94 meta-analyses of 48 drugs in 20 medical diseases (median AMSTAR score 9.0, 95% CI 8.2–9.2) and 33 meta-analyses of 16 drugs in 8 psychiatric disorders (median AMSTAR score 8.0, 95% CI 6.9–8.9). In the text we systematically present the raw numbers (for dichotomous outcomes the percentage responders in the placebo and drug groups; for continuous outcomes the average mean difference) and the average effect size (ARD and RRR/RR for dichotomous outcomes, SMD for continuous outcomes). Tables 1 and 2 present only some examples. Online Tables DS3 and DS4 present a comprehensive list including number of studies/participants, numbers needed to treat, $P$ values and confidence intervals for each outcome and each intervention. A positive sign means that a drug either increased a positive outcome (e.g. response) or reduced a negative outcome (e.g. relapse). All the effect sizes in online Tables DS3 and DS4 are presented in Fig. 1 to give the overall gestalt. For this purpose, effect sizes for dichotomous outcomes (ARD, RR/RRR) were converted to SMDs in Comprehensive Meta-analysis 2.12,14

This figure corresponds to online Fig. DS25 which indicates which dot relates to which study or outcome. Figures DS26 and DS28 present the same gestalt for relative and absolute risk/responder differences.

**Medical disorders**

In Tables 1 and DS3 the data are presented in an abbreviated ‘participants – intervention – comparator – outcome’ (PICO) format (the comparator is always placebo or no treatment).

Hypertension: antihypertensives for reduction of blood pressure, prevention of cardiovascular events and mortality

Several drug classes yielded similar results (Table DS3). Combining all agents, blood pressure was reduced by 9.4 mmHg systolic and 5.5 mmHg diastolic in the short term (SMDs 0.54 and 0.56 respectively).15 In the long term all drug classes significantly reduced cardiovascular events, e.g. angiotensin-converting enzyme (ACE) inhibitors reduced such events from 18% to 14% (ARD 4%, RRR 22%).16 A significant reduction of mortality has not been shown for all of them (Table DS3).

![Fig. 1 Summary of effect sizes.](image-url)

All effect sizes in online Tables DS3 and DS4 are presented except for duplicates (e.g. effect size on dichotomous response and continuous reduction of severity in schizophrenia). Online Fig. DS25 identifies which dot corresponds to which result, and Figs DS26–29 present the results of dichotomous outcomes as relative and absolute risk/responder differences. Data on older meta-analyses from Table DS1 are not included. Effect sizes of dichotomous outcomes were converted to standardised mean differences expressed as Hedges’ $g$. Effect sizes of general medicine medication are presented on the left-hand side (median 0.43, range 0.37–0.57) and psychiatric drugs on the right-hand side (median 0.41, range 0.37–0.53).
### Table 1: Examples of the efficacy of general medicine drugs v. placebo (full version is given in online Table DS3)

| Study/AMSTAR
c | Therapy | No. of studies | n | Duration, mean | Outcome (units) | Participants with outcome % |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Placebo</td>
<td>Drug</td>
<td>ARD(^{b}) %</td>
<td>RRR/RR %</td>
<td>MD</td>
<td>SMD(^{c})</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypertension</td>
<td>Effects on blood pressure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Law et al(^{a,75})</td>
<td>Any antihypertensive</td>
<td>94</td>
<td>17,641</td>
<td>8 wk</td>
<td>RR systolic (mmHg)</td>
<td>9.4***</td>
<td>0.56***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Law et al(^{a,75})</td>
<td>Any antihypertensive</td>
<td>94</td>
<td>17,641</td>
<td>8 wk</td>
<td>RR diastolic (mmHg)</td>
<td>5.5***</td>
<td>0.54***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long-term effects on cardiovascular events and mortality</td>
<td>BPLTTC(^{a,75})</td>
<td>ACE inhibitors</td>
<td>5</td>
<td>18,229</td>
<td>3.9 yr</td>
<td>Cardiovascular events</td>
<td>18.1</td>
<td>14.1</td>
<td>4 Ni</td>
<td>22***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute stroke</td>
<td>Wardlaw(^{a,79})</td>
<td>Thrombolysis</td>
<td>22</td>
<td>6,283</td>
<td>12–26 yr</td>
<td>Death/dependency</td>
<td>55.8</td>
<td>50.9</td>
<td>5**</td>
<td>11***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prevention of cardiovascular disease and stroke</td>
<td>BPLTTC(^{a,75})</td>
<td>ACE inhibitors</td>
<td>5</td>
<td>95,000</td>
<td>58 yr</td>
<td>Cardiovascular events</td>
<td>0.57/yr</td>
<td>0.51/yr</td>
<td>0.07/yr</td>
<td>12/yr***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prevention of cardiovascular disease and stroke</td>
<td>BPLTTC(^{a,75})</td>
<td>ACE inhibitors</td>
<td>16</td>
<td>17,000</td>
<td>58 yr</td>
<td>Cardiovascular events</td>
<td>8.2/yr</td>
<td>6.7/yr</td>
<td>1.5/yr</td>
<td>19/yr***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prevention of cardiovascular disease and stroke</td>
<td>Law et al(^{a,75})</td>
<td>ACE inhibitors</td>
<td>164</td>
<td>-38,000</td>
<td>2–6 wk</td>
<td>LDL cholesterol (mmol/l)</td>
<td>1.54**</td>
<td>Ni</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute stroke</td>
<td>Wardlaw(^{a,79})</td>
<td>Thrombolysis</td>
<td>22</td>
<td>6,283</td>
<td>12–26 yr</td>
<td>Death/dependency</td>
<td>55.8</td>
<td>50.9</td>
<td>5**</td>
<td>11***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rheumatoid arthritis</td>
<td>Suarez-Abreu et al(^{a,70})</td>
<td>Methotrexate</td>
<td>5</td>
<td>218</td>
<td>&gt;12 wk</td>
<td>Tender joints</td>
<td>NE</td>
<td>0.86***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Migraine</td>
<td>McCrory &amp; Gray(^{a,8})</td>
<td>Sumatriptan</td>
<td>8</td>
<td>2,221</td>
<td>2 h</td>
<td>Pain-free</td>
<td>8.5</td>
<td>29.5</td>
<td>20**</td>
<td>200***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rheumatoid arthritis</td>
<td>Suarez-Abreu et al(^{a,70})</td>
<td>Methotrexate</td>
<td>5</td>
<td>218</td>
<td>&gt;12 wk</td>
<td>Tender joints</td>
<td>NE</td>
<td>0.86***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asthma</td>
<td>Sin et al(^{a,77})</td>
<td>Corticosteroids</td>
<td>19</td>
<td>3,271</td>
<td>&gt;12 wk</td>
<td>FEV(_1), l</td>
<td>0.33***</td>
<td>0.56***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asthma</td>
<td>Sin et al(^{a,77})</td>
<td>Corticosteroids</td>
<td>11</td>
<td>8,999</td>
<td>&gt;12 wk</td>
<td>Exacerbations</td>
<td>Ni</td>
<td>Ni</td>
<td>Ni</td>
<td>54*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chronic obstructive pulmonary disease</td>
<td>Yang et al(^{a,70})</td>
<td>Inhaled corticosteroids</td>
<td>3</td>
<td>952</td>
<td>8–24 wk</td>
<td>FEV(_1), l</td>
<td>0.10**</td>
<td>0.36***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chronic obstructive pulmonary disease</td>
<td>Yang et al(^{a,70})</td>
<td>Inhaled corticosteroids</td>
<td>4</td>
<td>2,063</td>
<td>&gt;26 wk</td>
<td>Exacerbations</td>
<td>Ni</td>
<td>Ni</td>
<td>Ni</td>
<td>54*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diabetes</td>
<td>Saenz et al(^{a,71})</td>
<td>Metformin</td>
<td>12</td>
<td>1,537</td>
<td>21.5 yr</td>
<td>Fasting glucose (mmol/l)</td>
<td>21.7</td>
<td>14.6</td>
<td>7**</td>
<td>32**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chronic hepatitis C</td>
<td>Myers et al(^{a,70})</td>
<td>Interferon</td>
<td>8</td>
<td>4,109</td>
<td>&gt;26 wk</td>
<td>Virological response</td>
<td>1.0</td>
<td>38.3</td>
<td>35**</td>
<td>107**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reflux oesophagitis</td>
<td>Moayyedi et al(^{a,79})</td>
<td>Proton pump inhibitors</td>
<td>5</td>
<td>6,145</td>
<td>8 wk</td>
<td>Remission</td>
<td>28.3</td>
<td>83.2</td>
<td>58**</td>
<td>256***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ulcerative colitis</td>
<td>Sutherland &amp; MacDonald(^{a,79})</td>
<td>5-ASA</td>
<td>4</td>
<td>892</td>
<td>8 wk</td>
<td>Remission</td>
<td>10.0</td>
<td>19.9</td>
<td>8**</td>
<td>70**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td>Rice et al(^{a,70})</td>
<td>Interferon</td>
<td>3</td>
<td>9,19</td>
<td>20 yr</td>
<td>Exacerbation</td>
<td>69.5</td>
<td>55.2</td>
<td>14**</td>
<td>19***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast cancer</td>
<td>EBCCTC(^{a,74})</td>
<td>Polychemotherapy</td>
<td>60</td>
<td>28,764</td>
<td>15.0 yr</td>
<td>Mortality age &lt;50</td>
<td>42.4</td>
<td>32.4</td>
<td>10**</td>
<td>28***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-small cell lung cancer</td>
<td>Bria et al(^{a,78})</td>
<td>Adjuvant chemotherapy</td>
<td>21</td>
<td>7,408</td>
<td>4.5 yr</td>
<td>Mortality</td>
<td>Ni</td>
<td>Ni</td>
<td>Ni</td>
<td>9*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antibiotics for various diseases</td>
<td>Glaszou et al(^{a,70})</td>
<td>Ciprofloxacin</td>
<td>4</td>
<td>10,162</td>
<td>2–7 d</td>
<td>With pain</td>
<td>22.2</td>
<td>16</td>
<td>6.2 Ni</td>
<td>28***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antibiotics for various diseases</td>
<td>Glaszou et al(^{a,70})</td>
<td>Ciprofloxacin</td>
<td>4</td>
<td>10,162</td>
<td>3–17 d</td>
<td>Cure</td>
<td>25.7</td>
<td>61.8</td>
<td>36.1 Ni</td>
<td>19**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ACE, angiotensin-converting enzyme; ARD, absolute response or risk difference; ASA, acetylsalicylic acid; BPLTTC, Blood Pressure Lowering Treatment Trialists' Collaboration; d, day; EBCCTC, Early Breast Cancer Trialists' Collaborative Group; FEV\(_1\), forced expiratory volume in 1 s; h, hours; LDL, low-density lipoprotein; MD, mean difference; NE, not estimable; Ni, not indicated; NNT, number needed to treat; Ni, not significant; prev., prevention; RRR/RR, response ratio/relative risk reductions; SMD, standardised mean difference; wk, week; yr, year.

\(^{a}\) AMSTAR quality score (range of possible values 0–11).

\(^{b}\) Positive values always mean superiority of drug.

\(^{c}\) Italics indicate mean estimated values.

\(*P<0.05, \quad \star\star P<0.01, \quad \star\star\star P<0.001. Results on absolute and relative risk/responder differences do not always exactly match with the formulae presented in the manuscript due to weighting processes in meta-analyses.\)
Leucht et al.

Acute ischaemic stroke: thrombolysis, aspirin and heparin for prevention of death or dependency
Thrombolysis reduced death or dependency from 56% to 51% (ARD 5%, RRR 9%), but when administered after 4.5 h mortality is increased by haemorrhages. Aspirin reduced death or dependency from 46% to 45%, whereas heparin was ineffective.

Cardiovascular disease: aspirin for primary and secondary prevention of cardiovascular events and mortality
In secondary prevention, low-dose aspirin reduced serious cardiovascular events per year from 8.2% to 6.7% (ARD 1.5%, RRR 19%) and vascular mortality per year from 4.1% to 3.7% (ARD 0.29%, RRR 9%, P = 0.05). In primary prevention, aspirin reduced the number of cardiovascular events per year from 0.57% to 0.51%, but there was no effect on mortality because the reduction of occlusive events was balanced by an increase in major bleeds (mortality per year: placebo 0.19%, drug 0.19%).

Hypercholesterolaemia: statins for reduction of cholesterol levels and prevention of cardiovascular disease and mortality
In the short term, statins reduced low-density lipoprotein (LDL) cholesterol by 1.54 mmol/l or 31%. In the long term, cardiovascular events were reduced from 18% to 14% (primary and secondary prevention combined, ARD 4%, RRR 21%) and 5-year mortality from 9.7% to 8.5%.

Chronic heart failure: various drugs for reduction of mortality
Angiotensin-converting enzyme inhibitors, angiotensin receptor blockers, beta-blockers and diuretics respectively reduced mortality rates from 27% to 23% (ARD 4%, RRR 15%), from 18% to 11%, from 13% to 8% and from 12% to 3%. Digitalis significantly reduced hospital admission (from 33% to 25%) but not mortality.

Rheumatoid arthritis: antirheumatic drugs for the reduction of tender joints
Various immunosuppressants, corticosteroids and other agents reduced the number of tender joints with reasonably good SMDs between 0.33 and 1.33 (raw values for mean differences were not presented; Table DS3). In maintenance treatment (relapse: placebo 75%, drug 83%, ARD 58%, RR 256%), and maintenance of remission from 37% to 53% (ARD 18%, RR 50%).

Acute migraine: effects of sumatriptan and aspirin on the number of patients pain-free after 2 h
Sumatriptan increased the percentage of patients pain-free after 2 h from 9% to 30% (ARD 20%, RR 220%) and intravenous aspirin increased it from 15% to 27%.

Prophylaxis of migraine: effects of propanolol and anticonvulsants on responder rates and on the number of migraine attacks
Fifty-two per cent responded to propanolol prophylaxis and 31% to placebo (ARD 35%, RR 80%). Patients had approximately one migraine attack less (SMD 0.47). The results of anticonvulsants were similar.

Chronic asthma: effects of inhaled corticosteroids and beta-2-agonists on forced expiratory volume and on asthma exacerbations
The first-line drugs for chronic, severe asthma are inhaled corticosteroids and beta-2-agonists (short-acting as needed, long-acting in patients with refractory disease). Inhaled corticosteroids increased forced expiratory volume in 1 s (FEV1) by 330 ml (SMD 0.56). The addition of long-acting beta-2-agonists improved FEV1 by 190 ml (SMD 0.35), but the reduction of asthma exacerbations found by some meta-analyses is controversial, because another meta-analysis found more severe exacerbations.

Chronic obstructive pulmonary disease: effects of various agents on FEV1 and on disease exacerbations
Guidelines recommend anticholinergics, beta-2-agonists and inhaled corticosteroids. The anticholinergic tiotropium improved FEV1 by 200 ml (SMD 0.99). It reduced exacerbations from 31% to 23% (ARD 5%, RRR 17%). Inhaled corticosteroids improved FEV1 by 100 ml (SMD 0.36) and the number of exacerbations per patient and year by 0.26 (SMD 0.20). The data on long-acting beta-2-agonists are equivocal. They reduced exacerbations (e.g. Salpeter et al.), but one systematic review found them to increase respiratory deaths.

Type 2 diabetes: various antidiabetics for reduction of HbA1c and mortality
Metformin reduced HbA1c by 1% (SMD 0.97) and z-glucosidase inhibitors reduced it by 0.8% (SMD 0.64). In the long term, metformin reduced the death rate from 22% to 15% (ARD 7%, RRR 32%), but z-glucosidase inhibitors have not been shown to change the death rate.

Hepatitis C: effects of interferon and ribavirin on virological response/morbidity and mortality
Interferon increased the number of participants with no detectable virus at treatment end (virological response) from 1% to 38% (ARD 35%, RR 1070%). Ribavirin was only efficacious in combination with interferon.

Reflux oesophagitis: proton pump inhibitors for clinical remission and relapse prevention
Proton pump inhibitors are highly effective in acute treatment (response: placebo 28%, drug 83%, ARD 58%, RR 256%) and also in maintenance treatment (relapse: placebo 75%, drug 36%).

Ulcereative colitis: 5-aminosalicylic acid for clinical remission and maintenance of remission
Five-aminosalicylic acid (5-ASA) increased remission from 10% with placebo to 20% (ARD 8%, RR 70%), and maintenance of remission from 37% to 53% (ARD 18%, RR 50%).

Multiple sclerosis: corticosteroids for treatment of acute episodes and interferon for prevention of exacerbations
Acute treatment with corticosteroids increased the proportion of responders from 28% with placebo to 68% (ARD 41%, RR 140%). In the first 2 years, prevention with interferon beta reduced exacerbations from 70% to 55% (ARD 14%, RRR 19%).

Parkinson's disease: effects of levodopa on disease symptoms
There was no systematic review of the standard treatments levodopa or dopamine agonists with data compared with placebo. We parenthetically note that the National Institute for Health and Clinical Excellence (NICE) guideline based its recommendation on a pivotal 42-week trial in which levodopa produced 7 points more improvement in the Unified Parkinson's Disease Rating Scale.
Scale total score than placebo (SMD 0.93), but also a 7% stronger decline of striatal dopamine transporter density (SMD –0.44), suggesting a possible acceleration of nigrostriatal dopamine nerve terminal loss.

Breast and lung cancer: polychemotherapy for reduction of mortality
Breast cancer is the most frequent neoplasm in women and lung cancer is the leading cause of cancer death. Polychemotherapy reduced the 15-year breast cancer mortality in younger women (<50 years) from 42% to 32% (ARD 10%, RRR 24%) but in older women only from 50% to 47%.56 Tamoxifen added to polychemotherapy reduced the 15-year mortality in oestrogen receptor-positive patients from 35% to 26%.56 In the study by Bria et al, adjuvant chemotherapy led to a small reduction of 5-year lung-cancer mortality (ARD 3%, RRR 9%),57 confirming a landmark previous meta-analysis.58

Infectious diseases: antibiotics for rhinosinusitis, otitis media, uncomplicated cystitis and prophylaxis of wound infection after surgery
The effects of antibiotics depend on the infection. We did not find meta-analyses on severe infections such as pneumonia or on antivirals (monotherapy v. placebo) for HIV. A meta-analysis concluded against their general use in rhinosinusitis owing to small effect size (response: placebo 57%, drug 64%, ARD 7%, RRR 13%).59 The use of antibiotics in otitis media is debated, as within 2–7 days 78% of patients recovered spontaneously compared with 84% taking antibiotics (ARD 6%, RR 28%).60 In contrast, the efficacy in uncomplicated cystitis (response: placebo 26%, drug 62%) and for the prophylaxis of wound infections after major operations (infections: placebo 39%, antibiotics 10%) was clear.61,62

Psychiatric disorders
Full data are given in Table DS4; examples are summarised in Table 2.

Schizophrenia: antipsychotics for reduction of overall symptoms and relapse prevention
In acute treatment, second-generation antipsychotics increased the percentage responding from 24% with placebo to 41% (ARD 18%, RR 70%), and reduced the Brief Psychiatric Rating Scale/Positive and Negative Syndrome Scale total score by 9/10 points (SMD 0.51).76 Antipsychotic maintenance treatment reduced relapse rates from 57% to 22% within approximately 10 months (ARD 38%, RRR 65%).53

Bipolar disorder: mood stabilisers for acute mania, antidepressants for depression and mood stabilisers for relapse prevention
Various antimania agents increased the percentage responding from approximately 30% with placebo to approximately 50% within 3 weeks (response to lithium 52% v. placebo 34%, ARD 17%, RR 50%;76 response to valproate 47% v. placebo 21%, ARD 27%, RR 150%;76 response to carbamazepine 51% v. placebo 26%, ARD 25%, RR 100%;76 response to antipsychotics 50% v. placebo 31%, ARD 20%, RR 60%).66 In bipolar depression, antidepressants increased the response rate from 34% to 58% (ARD 34%, RR 130%).57 In maintenance treatment, lithium reduced relapse rates from 81% to 36% (ARD 53%, RRR 51%),68 or from 61% to 40% after excluding studies in which lithium was suddenly discontinued (ARD 24%, RRR 35%).69

Major depressive disorder: antidepressants for acute depression and relapse prevention
The absolute responder differences in recent meta-analyses of various selective serotonin reuptake inhibitors (SSRIs) (or tricyclic antidepressants used as an active comparator in SSRI v. placebo studies)70 v. placebo in major depressive disorder were 10–15% (Table DS4). For example, paroxetine increased the percentage responding from 42% to 53% (ARD 10%, RR 20%) and reduced the Hamilton Rating Scale for Depression score by 3 points (SMD 0.32).71 These studies were currently primarily conducted in out-patients with less severe disorder (e.g. 90% of the sample were out-patients in the meta-analysis by Barbui et al).71 Maintenance treatment with any antidepressant reduced the relapse rate from 41% with placebo to 18% (ARD 23%, RRR 58%).72 Consistent with another meta-analysis restricted to new antidepressants (placebo 48% v. drug 26%, ARD 22%, RRR 44%).73

Obsessive-compulsive disorder: effects of SSRIs on responder rates and overall symptoms
Selective serotonin reuptake inhibitors increased the proportion of patients responding in the acute phase from 23% to 43% (ARD 20%, RRR 84%).74 These drugs reduced the mean Yale–Brown Obsessive Compulsive Scale score by 3.2 points (SMD 0.44).74

Panic disorder: tricyclic antidepressants, SSRIs and benzodiazepines for anxiety symptoms
The mean SMDs (raw differences in rating scale scores or responder rates were not indicated) of tricyclic antidepressants, SSRIs and benzodiazepines in acute treatment were 0.40–0.41.75

Alzheimer’s disease: cholinesterase inhibitors for prevention of cognitive decline
Within 6 months, cholinesterase inhibitors increased the percentage of participants unchanged or improved from 17% to 24% (ARD 7%, RRR 43%).76 The cognitive subscore of the Alzheimer’s Disease Assessment Scale was better by 2 points (SMD 0.41).76

Attention-deficit hyperactivity disorder: effects of various drugs on symptoms
Methylphenidate (SMD 0.78), amphetamines (SMD 1.00) and atomoxetine (SMD 0.64) showed robust effect sizes in overall reduction of attention-deficit hyperactivity disorder symptoms (raw differences in rating scale scores or responder rates were not indicated).77–79

Discussion
Any comparison of treatments for different diseases can only be qualitative in nature and therefore Fig. 1 is no more than a way to place psychiatric drugs in the perspective of general medicine medication. Some general medical drugs have very high effect sizes, but those obtained by psychiatric drugs are in the same range as most general medical pharmacotherapeutics. This said, the increment of improvement by a drug must be viewed in the context of the seriousness of the disease, the suffering induced, the outcome in question, societal values and the natural course including the duration of the disease. In the following paragraphs we discuss a number of these issues which readers should take into account in interpreting the results.
# Examples of efficacy of psychiatric drugs v. placebo (full version is given in online Table DS4)

<table>
<thead>
<tr>
<th>Study/AMSTAR(^a)</th>
<th>Therapy</th>
<th>No. of studies</th>
<th>(n)</th>
<th>Weeks, mean</th>
<th>Outcome (units)</th>
<th>Placebo</th>
<th>Drug</th>
<th>ARD(^b) %</th>
<th>RRR/RR %</th>
<th>MD</th>
<th>SMD(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Schizophrenia</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Acute treatment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leucht et al(^a)/10</td>
<td>SGAs</td>
<td>28</td>
<td>4498</td>
<td>9</td>
<td>Responders</td>
<td>23.7</td>
<td>40.6</td>
<td>18***</td>
<td>70***</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>35</td>
<td>5668</td>
<td>10</td>
<td>PANSS/BPRS</td>
<td>19.5</td>
<td>29.3</td>
<td>12%***</td>
<td>60%**</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>1440</td>
<td>6</td>
<td>Responders</td>
<td>19.5</td>
<td>29.3</td>
<td>12%***</td>
<td>60%**</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>11</td>
<td>1540</td>
<td>6</td>
<td>PANSS/BPRS</td>
<td>19.5</td>
<td>29.3</td>
<td>12%***</td>
<td>60%**</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td><strong>Maintenance treatment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leucht et al(^a)/10</td>
<td>Antipsychotics</td>
<td>62</td>
<td>6392</td>
<td>42</td>
<td>Relapse</td>
<td>57.0</td>
<td>22.0</td>
<td>38%***</td>
<td>65***</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td><strong>Bipolar affective disorder</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Acute manic episode</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Storosum et al(^a)/6</td>
<td>Lithium</td>
<td>6</td>
<td>811</td>
<td>3</td>
<td>Responders</td>
<td>34.0</td>
<td>52.0</td>
<td>17***</td>
<td>50**</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>7</td>
<td>1165</td>
<td>3</td>
<td>YMRS/MRS</td>
<td>21.1</td>
<td>47.1</td>
<td>27%***</td>
<td>150%*</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td>Smith et al(^a)/19</td>
<td>Valproate</td>
<td>4</td>
<td>782</td>
<td>3</td>
<td>YMRS/MRS</td>
<td>25.5</td>
<td>51.1</td>
<td>29%***</td>
<td>100%**</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td>Smith et al(^a)/8</td>
<td>Carbamazepine</td>
<td>2</td>
<td>443</td>
<td>3</td>
<td>Responders</td>
<td>30.8</td>
<td>49.9</td>
<td>20***</td>
<td>60***</td>
<td>0.73</td>
<td></td>
</tr>
<tr>
<td>Scher et al(^a)/10</td>
<td>SGAs and haloperidol</td>
<td>12</td>
<td>2939</td>
<td>3</td>
<td>Responders</td>
<td>30.8</td>
<td>49.9</td>
<td>20***</td>
<td>60***</td>
<td>0.73</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>2939</td>
<td>3</td>
<td>YMRS/MRS/MVS</td>
<td>30.8</td>
<td>49.9</td>
<td>20***</td>
<td>60***</td>
<td>0.73</td>
<td></td>
</tr>
<tr>
<td><strong>Depressive episode</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gisman et al(^a)/9</td>
<td>Antidepressants</td>
<td>4</td>
<td>662</td>
<td>7</td>
<td>Responders</td>
<td>34.1</td>
<td>57.7</td>
<td>34***</td>
<td>130**</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>Davis et al(^a)/9</td>
<td>Lithium</td>
<td>9</td>
<td>421</td>
<td>Ni</td>
<td>Any relapse</td>
<td>81.4</td>
<td>36.2</td>
<td>53**</td>
<td>51***</td>
<td>1.12</td>
<td></td>
</tr>
<tr>
<td>Geddes et al(^a)/8</td>
<td>Lithium</td>
<td>5</td>
<td>770</td>
<td>73</td>
<td>Any relapse</td>
<td>610</td>
<td>40.0</td>
<td>24**</td>
<td>35**</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>Macnab et al(^a)/10</td>
<td>Valproate</td>
<td>1</td>
<td>281</td>
<td>52</td>
<td>Any relapse</td>
<td>38.3</td>
<td>24.1</td>
<td>14%*</td>
<td>37%*</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td><strong>Major depressive disorder</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Acute episode</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barbui et al(^a)/8</td>
<td>Paroxetine</td>
<td>22</td>
<td>5112</td>
<td>7.5</td>
<td>Responders</td>
<td>42.4</td>
<td>53.2</td>
<td>10***</td>
<td>20***</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>Storosum et al(^a)/4</td>
<td>TCAs (new AD studies)</td>
<td>32</td>
<td>4314</td>
<td>6</td>
<td>Responders</td>
<td>31.0</td>
<td>46.0</td>
<td>15**</td>
<td>50***</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>Bech et al(^a)/5</td>
<td>Fluoxetine</td>
<td>16</td>
<td>2761</td>
<td>6</td>
<td>Responders</td>
<td>24.2</td>
<td>37.8</td>
<td>13.6%***</td>
<td>65%***</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>Kirsch et al(^a)/7</td>
<td>New ADs</td>
<td>35</td>
<td>5133</td>
<td>6</td>
<td>Responders</td>
<td>41.0</td>
<td>18.0</td>
<td>23 Ni</td>
<td>58***</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td><strong>Maintenance treatment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geddes et al(^a)/2</td>
<td>Antidepressants</td>
<td>30</td>
<td>5032</td>
<td>63</td>
<td>Relapse</td>
<td>41.0</td>
<td>18.0</td>
<td>23 Ni</td>
<td>58***</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>Hansen et al(^a)/10</td>
<td>New ADs</td>
<td>11</td>
<td>3326</td>
<td>68</td>
<td>Recurrence</td>
<td>48.0</td>
<td>26.0</td>
<td>22 Ni</td>
<td>44***</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>Davis et al(^a)</td>
<td>Lithium</td>
<td>9</td>
<td>227</td>
<td>Ni</td>
<td>Relapse</td>
<td>75.0</td>
<td>36.0</td>
<td>39 Ni</td>
<td>53***</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td><strong>Obsessive–compulsive disorder</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soomro et al(^a)/11</td>
<td>SSRIS</td>
<td>17</td>
<td>3097</td>
<td>10</td>
<td>Y–BOCS</td>
<td>22.6</td>
<td>43.3</td>
<td>20***</td>
<td>84***</td>
<td>3.21</td>
<td>0.64***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13</td>
<td>2709</td>
<td>10</td>
<td>Anxiety</td>
<td>22.6</td>
<td>43.3</td>
<td>20***</td>
<td>84***</td>
<td>3.21</td>
<td>0.64***</td>
</tr>
<tr>
<td><strong>Panic disorder</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mitte(^a)/4</td>
<td>TCAs, SSRIS, BZDs</td>
<td>17–25</td>
<td>Ni</td>
<td>8</td>
<td>Anxiety</td>
<td>22.6</td>
<td>43.3</td>
<td>20***</td>
<td>84***</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td><strong>Dementia</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birks(^a)/8</td>
<td>Cholinesterase inhibitors</td>
<td>10</td>
<td>4236</td>
<td>26</td>
<td>ADAS-Cog</td>
<td>23.8</td>
<td>24.3</td>
<td>0.37**</td>
<td>0.14**</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>3118</td>
<td>28</td>
<td>MMSE</td>
<td>1.33**</td>
<td>0.39**</td>
<td>0.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>3402</td>
<td>26</td>
<td>Not worse</td>
<td>16.8</td>
<td>24.4</td>
<td>7***</td>
<td>43***</td>
<td>0.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Attention-deficit hyperactivity disorder</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schachter et al(^a)/9</td>
<td>Methylphenidate</td>
<td>22</td>
<td>963</td>
<td>3.3</td>
<td>Hyperactivity(^f)</td>
<td>22.6</td>
<td>43.3</td>
<td>20***</td>
<td>84***</td>
<td>0.53</td>
<td></td>
</tr>
</tbody>
</table>

---

\(^a\) AMSTAR quality score (range of possible values 0–11).
\(^b\) Italic indicates mean estimated values.
\(^c\) YMRS only.
\(^d\) Teacher rated.
\(^e\) Updated and supplemented by our own searches.

AD, antidepressant; ADAS-Cog, Alzheimer’s Disease Assessment Scale Cognitive Subscale; ARD, absolute response or risk difference; BPRS, Brief Psychiatric Rating Scale; BZD, benzodiazepine; HRSD, Hamilton Rating Scale for Depression; MD, mean difference; MMSE, Mini-Mental State Examination; MRS, Mania Rating Scale; MS, Mania Scale; Ni, not indicated; NNT, number needed to treat; PANSS, Positive and Negative Syndrome Scale; RRR/RR, response ratio/relative risk reduction; SMD, standardised mean difference; SGA, second-generation antipsychotic; SIR, selective serotonin reuptake inhibitor; TCA, tricyclic antidepressant; Y–BOCS, Yale–Brown Obsessive Compulsive Scale; YMRS, Young Mania Rating Scale.

---

* \(P < 0.05\), ** \(P < 0.01\), *** \(P < 0.001\). Results on absolute and relative risk/responder differences do not always exactly match with the formulae presented in the manuscript due to weighting processes in meta-analyses.
Outcomes
Psychiatry is often criticised for using rating scales which are subjective and considered ‘soft’ outcomes, whereas many medical treatments prevent ‘hard’ outcomes such as death or major events (stroke, heart attack, etc.). High blood pressure or cholesterol levels per se do not lead to suffering, therefore they should not be the primary outcome, rather their long-term consequences. Sometimes an intermediate outcome is improved but mortality increases; for example, in a large multicentre effectiveness trial for asthma \( n = 26,000 \), long-acting beta-2-agonists increased respiratory-related deaths.\(^{80}\) In diabetes, aggressive glycaemic control reduced glucose levels compared with standard care, but increased mortality rates \( n = 10,251 \).\(^{81}\)

Other drugs reduce the symptoms and suffering originating directly from the disease such as oesophagitis or migraine, but their pathophysiological disease processes do not progress to death. Psychiatric drugs fall in this category. Therefore, reduction of disease severity (e.g. degree of delusions and hallucinations in schizophrenia) and prevention of future episodes are primarily outcomes, and it is not entirely appropriate to criticise psychiatry for using ‘soft’ outcomes. This said, there is considerable room for improvement in psychiatric outcome measures.\(^{82}\) and death or suicide should be always reported. The example of lithium shows that some psychiatric drugs may reduce suicide rates.\(^{83,84}\)

Placebo effects
Readers may be surprised that many effect sizes in both areas were not larger. The median of all effect sizes was 0.40, similar to that found in another analysis of Cochrane reviews \( (0.32) \).\(^{85}\) In this context there is a general misconception that with placebo all patients will have a poor outcome, but many patients will recover spontaneously owing to the natural course of the disorder (for example, a manic episode will remit by itself) and placebo effects.

Effect sizes for dichotomous and continuous outcomes
For dichotomous outcomes both relative and absolute risk reductions should be considered. There is substantial evidence showing that clinicians tend to overestimate treatment effects presented as relative risk reductions.\(^{86}\) For example, statins showed that clinicians tend to overestimate treatment effects \( (1 – (0.14/0.18)) \times 100 \) is more impressive than the absolute risk difference of 4% \( (14\% – 18\% = [–4\%]) \). On the other hand, if the risk in the placebo group is low, the maximally possible absolute risk reduction must be lower than the base rate (here 18%), making the relative risk reduction more important.

In continuous outcomes the standardised mean difference (Cohen’s \( d \), Hedges’ \( g \) etc.) is necessary when different instruments are used to measure the same concept (e.g. two depression scales) or if the original unit is difficult to interpret intuitively (e.g. the score of an unknown rating scale). As the SMD is relative to the pooled standard deviation, large variability will reduce it. In psychiatry this often occurs with rating scales in somewhat ill-defined, ‘variable’ diseases such as depression, whereas in general medicine the measure may be a highly accurate laboratory test (e.g. serum cholesterol concentration) in a well-defined disease entity. Cohen’s rule that an SMD of 0.2 is a small effect size, 0.5 medium and 0.8 a large effect size is often used, but Cohen hastened to say that the interpretation depends on the context.\(^{87}\) a small SMD for a fatal disease is more important than a large SMD for a transitory rash. In the future, quality-adjusted life years (QALYs) could be a uniform measure for comparisons across treatments, but these are not yet available for all drugs and we did not find this outcome in the meta-analyses. In addition, there is much debate about the validity of QALYs (see, for example, studies by Schlander\(^{88}\) and Griebsch et al \(^{89}\)).

Sample size
Meta-analyses in somatic medicine sometimes include impressively large patient numbers, e.g. 95,000 participants in studies of the primary prevention of cardiovascular events with aspirin.\(^{81}\) Aspirin reduced the risk of a cardiovascular event from 0.57% per year to 0.51% per year. Angiotensin-converting enzyme inhibitors for hypertension reduced 5-year mortality from 10.4% to 9.2% in 18,229 participants.\(^{90}\) In such situations, large sample sizes are needed for two reasons: first, the aspirin \( v \). placebo difference was 0.07% event and the ACE inhibitors \( v \). placebo difference was 1.2% events, requiring large sample sizes for statistical significance; second, the base rate (equivalent to the risk in the placebo group) was very low (e.g. 0.57% per year without aspirin), limiting the drug effect to a maximum 0.57% per year. Nevertheless, for mortality even a small difference can be clinically meaningful. In psychiatry the difference in percentages of those responding to drug or placebo is usually higher and it has been shown that here meta-analyses with at least 1000 participants are robust.\(^{90}\)

Drug effects could accumulate over time
The mean duration of the studies included in a meta-analysis should always be considered. For example, treated or not, few patients with hypertension will die in the course of a year. Thus, to obtain a large difference in mortality, studies of many years’ duration would be necessary, but such studies are almost impossible to conduct for many reasons. Therefore, shorter studies are performed which show only small differences. Although only very long-term studies could prove this, it is likely that the reduction of mortality accumulates over time. In this context, many psychiatric drugs not only improve the acute episode but also prevent further episodes. Patients with severe recurrent depression might have 20 episodes in their lifetime, which could be reduced by medication to 10.\(^{72}\)

Has drug efficacy decreased over the decades?
To be systematic we generally chose the most up-to-date systematic reviews, but there is an impression that earlier meta-analyses in psychiatry yielded higher effect sizes (see online Table DS1 for some examples). In the first 103 double-blind studies in depression, summarised in 1993, approximately two-thirds responded to tricyclic antidepressants or monoamine esterase inhibitors compared with a third responding to placebo.\(^{91}\) The large National Institute of Mental Health schizophrenia trial, published in 1964, reported that 69% responded to antipsychotics and 24% to placebo (NNT 2, effect size 1.31).\(^{92}\) In the first large obsessive–compulsive disorder trial, published in 1991, half the sample responded to clomipramine and only 5% to placebo.\(^{93}\) Recent meta-analyses found much smaller effect sizes for both the new SSRIs and clomipramine.\(^{94,95}\) The reasons for decreasing effect sizes are not entirely understood. The early trials were often small and single-centre, and methodology less well developed (blinding, scales, external auditing, statistical methods). There may also have been more publication bias, as efforts to control it have expanded only in the past two decades. Modern trials are often large, multicentre, but have other problems such as the impossibility of recruiting severely ill patients with truly acute disorders because of ethical concerns, the availability of effective medication leaving few drug-naïve patients, and the phenomenon of symptomatic volunteers answering an
advertisement for free medication and thereby increasing placebo response. It is possible that there are similar temporal trends in general medicine and the phenomenon needs thorough examination.

Limitations

We made a considerable effort to be systematic, but for the reasons stated below we could not meet all criteria of a systematic review. We did not examine a single drug but put different medications in perspective, for which an established methodology does not exist.

First, we could not present a complete collection, but we chose common diseases by consensus based on frequency, importance and available treatment. It would be difficult to operationalise the selection. For example, there are diseases that are frequent but not severe (an extreme example is the common cold). Others are extremely severe but rare (e.g. certain cancers). The selection was made a priori, and once chosen all diseases and drugs were presented. We feel that the selection is representative and that the major diseases of the industrialised world are included; nevertheless, the selection process may have introduced bias.

Second, in the selection among reviews, we emphasised up-to-dateness and full presentation, but we compared the results of different meta-analyses on the same topic which were usually consistent. Third, a review of reviews is observational by nature: our unit of analysis was published meta-analyses, which does not exist for all drugs/indications, and the included reports differed in the exact methods, publication dates, inclusion criteria, etc. Fourth, many meta-analyses did not present the data in a consistent manner, resulting in a major challenge for us. We made substantial efforts to present the results in a consistent way by back-calculating indices, but stringent following of the PRISMA statement would facilitate future attempts.

Fifth, we did not address side-effects. These are a serious problem of many psychotropic drugs, although improvements have been made. For example, SSRIs have much less serious toxicity than tricyclic antidepressants. General medicine drugs also have important side-effects, for example death induced by bleeding from thrombolyis or aspirin or cancer chemotherapies. It would have been simply impossible to describe side-effects as well and to balance them with efficacy, because there are many subjective judgement calls. Finally, publication bias is a major problem for meta-analyses. For example, Turner et al (see Table DS4) showed that the inclusion of unpublished antidepressant trials reduced the effect size. Publication bias exists in general medicine as well (see, for example, Rising et al), and we are not aware of evidence comparing its degree in different fields.

There are many reasons why doctors, patients and caregivers are and should be critical about psychotropic drug treatment, such as unclear disease aetiology, lack of diagnostic tests, commercial conflict of interest, unclear mechanism of drug action and side-effects. Moreover, some people think that psychiatric disorders are purely psychological conditions that should be treated exclusively with psychotherapy. However, the efficacy of psychotropic drugs is supported by randomised controlled trials. In this context we have put psychiatric drugs in the perspective of general medicine medication.

Stefan Leucht, MD, Sandra Hierl, MD, Werner Kissling, MD, Markus Dold, MD, Department of Psychiatry and Psychotherapy, Munich Technical University, Klinikum rechts der Isar, Munich, Germany. John M. Davis, Department of Psychiatry, University of Illinois, Chicago, Illinois, USA

Correspondence: Professor Stefan Leucht, Department of Psychiatry and Psychotherapy, Technische Universität München, Klinikum rechts der Isar, Ismaningerstr. 22, 81675 München, Germany. Email: Stefan.Leucht@lrz.tum.de

First received 16 May 2011, final revision 22 Aug 2011, accepted 15 Sep 2011

Acknowledgements

We thank Drs Malcom Law, Toshi Furukawa, Corrado Ibarbi and Shelley Sapeter for replying to our requests about their studies.

References

5. Menand L. Head case: can psychiatry be a science? The New Yorker 2010; 1 March: 68–74.


