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; the efficacy of cholinesterase inhibitors in Alzheimer’s disease and of lithium prophylaxis in bipolar disorder has been questioned; and we found a smaller antipsychotic drug–placebo difference in schizophrenia than we intuitively expected. These reviews inspired an article in The New Yorker summarising them, and fuelled a vocal antipsychiatry movement. 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.

Method

Identification of diseases of interest and search strategy

We reviewed textbooks, 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then chose the most recent reviews, because even if methodologically better an older review would have certainly been out of date.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). 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/responder 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.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. 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. 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). 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%). 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 0.37, mean 0.45, 95% CI 0.37–0.53) and psychiatric drugs on the right-hand side (median 0.41, mean 0.49, 95% CI 0.41–0.57).
# Table 1: Examples of the efficacy of general medicine drugs v. placebo (full version is given in online Table DS3)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Therapy</th>
<th>No. of studies</th>
<th>Duration, mean</th>
<th>Outcome (units)</th>
<th>Participants with outcome %</th>
<th>ARD&lt;sup&gt;a&lt;/sup&gt;</th>
<th>RRR/RR %</th>
<th>MD</th>
<th>SMD&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hypertension</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effects on blood pressure</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&lt;sup&gt;a&lt;/sup&gt;/15</td>
<td>Any antihypertensive</td>
<td>94</td>
<td>17 641</td>
<td>RR systolic (mmHg)</td>
<td>94 17 641 8 wk</td>
<td>9.4***</td>
<td>0.56***</td>
<td>5.5***</td>
<td>0.54***</td>
</tr>
<tr>
<td>Law et al&lt;sup&gt;a&lt;/sup&gt;/15</td>
<td>Any antihypertensive</td>
<td>94</td>
<td>17 641</td>
<td>RR diastolic (mmHg)</td>
<td>94 17 641 8 wk</td>
<td>5.5***</td>
<td>0.46***</td>
<td>3.5***</td>
<td>0.43***</td>
</tr>
<tr>
<td><strong>Acute stroke</strong></td>
<td>Wardenlaw&lt;sup&gt;a&lt;/sup&gt;/9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Prevention of cardiovascular disease and stroke</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baigent et al&lt;sup&gt;a&lt;/sup&gt;/5</td>
<td>ACE inhibitors</td>
<td>5</td>
<td>18 229</td>
<td>Cardiovascular events</td>
<td>18.1 14.1 4 Ni</td>
<td>22***</td>
<td>0.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Chronic heart failure</strong></td>
<td>Flather et al&lt;sup&gt;a&lt;/sup&gt;/6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rheumatoid arthritis</strong></td>
<td>Suarez-Azmanor et al&lt;sup&gt;a&lt;/sup&gt;/10</td>
<td>5</td>
<td>218</td>
<td>Mortality</td>
<td>26.8 23.0 4**</td>
<td>15***</td>
<td>0.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Migraine</td>
<td>McCory &amp; Gray&lt;sup&gt;a&lt;/sup&gt;/8</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&lt;sup&gt;a&lt;/sup&gt;/17</td>
<td></td>
<td></td>
<td>Fever</td>
<td>19 3271 &gt; 12 wk FIPEV, l</td>
<td>0.33***</td>
<td>0.56***</td>
<td>NE</td>
<td></td>
</tr>
<tr>
<td>Chronic obstructive pulmonary disease</td>
<td>Yang et al&lt;sup&gt;a&lt;/sup&gt;/9</td>
<td></td>
<td></td>
<td>Exacerbations</td>
<td>4 2063 &gt; 26 wk</td>
<td>0.26***</td>
<td>0.20***</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>Diabetes</td>
<td>Sarenz et al&lt;sup&gt;a&lt;/sup&gt;/11</td>
<td></td>
<td></td>
<td>Fasting glucose (mmol/l)</td>
<td>12 1587 21.5 wk</td>
<td>1.84***</td>
<td>0.87***</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>Chronic hepatitis C</td>
<td>Myers et al&lt;sup&gt;a&lt;/sup&gt;/10</td>
<td></td>
<td></td>
<td>Mortality</td>
<td>1 753 10.7 yr</td>
<td>2.27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reflux oesophagitis</td>
<td>Moayedi et al&lt;sup&gt;a&lt;/sup&gt;/9</td>
<td></td>
<td></td>
<td>Virological response</td>
<td>8 409 &gt; 26 wk</td>
<td>2.27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ulcerative colitis</td>
<td>Sutherland &amp; MacDonald&lt;sup&gt;a&lt;/sup&gt;/9</td>
<td>4</td>
<td>892</td>
<td>Remission</td>
<td>4 892 8 wk</td>
<td>1.39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td>Rice et al&lt;sup&gt;a&lt;/sup&gt;/10</td>
<td></td>
<td></td>
<td>Exacerbation</td>
<td>3 919 2.0 yr</td>
<td>0.44</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast cancer</td>
<td>EBCTCG&lt;sup&gt;a&lt;/sup&gt;/4</td>
<td></td>
<td></td>
<td>Mortality age &lt; 50</td>
<td>60 28764 15.0 yr</td>
<td>0.24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-small cell lung cancer</td>
<td>Bria et al&lt;sup&gt;a&lt;/sup&gt;/8</td>
<td></td>
<td></td>
<td>Mortality</td>
<td>21 7408 4.5 yr</td>
<td>0.22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antibiotics for various diseases</td>
<td>Glazouz et al&lt;sup&gt;a&lt;/sup&gt;/10</td>
<td></td>
<td></td>
<td>Mortality</td>
<td>4 1062 2-7 d</td>
<td>0.22</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; EBCTCG, Early Breast Cancer Trialists’ Collaborative Group; FEV<sub>1</sub>, 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; NS, not significant; prev., prevention; RR, relative risk; RRR, response ratio; SMD, standardised mean difference; wk, week; yr, year.

*<sup>a</sup> AMSTAR quality score (range of possible values 0–11).
*<sup>b</sup> Positive values always mean superiority of drug.
*<sup>c</sup> Italics indicate mean estimated values.

*<sup>P</sup> < 0.05, **<sup>P</sup> < 0.01, ***<sup>P</sup> < 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.
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).

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 α-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 α-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%).

Ulcerative 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 than dopamine agonists.
Igno
Table 2: Examples of efficacy of psychiatric drugs vs. placebo (full version is given in online Table DS4)

<table>
<thead>
<tr>
<th>Study/AMSTAR</th>
<th>Therapy</th>
<th>No. of studies</th>
<th>n</th>
<th>Weeks, mean</th>
<th>Outcome (units)</th>
<th>Participants with outcome</th>
<th>Placebo</th>
<th>Drug</th>
<th>ARD (%)</th>
<th>RRR/RR (%)</th>
<th>MD</th>
<th>SMD (%)</th>
</tr>
</thead>
</table>

**Schizophrenia**

**Acute treatment**

Leucht et al 2010 SGAs 28 4498 9 Responders 23.7 40.6 18*** 70*** 0.43
Leucht et al 2010 Haloperidol 10 1440 6 Responders 19.5 29.3 12*** 60%*** 0.30

**Maintenance treatment**

Leucht et al 2010 Antipsychotics 62 6392 42 Relapse 57.0 22.0 38*** 65*** 0.92

**Bipolar affective disorder**

**Acute manic episode**

Storosum et al 2010 Lithium 6 811 3 Responders 34.0 52.0 17*** 50** 0.41
Smith et al 2010 Valproate 4 782 3 Responders 21.1 47.1 27*** 150* 0.66
Smith et al 2010 Carbamazepine 2 331 3 YMRS/BPRS NI 0.51***

Scherk et al 2010 SGAs and haloperidol 12 29.39 3 Responders 30.8 49.9 20*** 60*** 0.44

**Depressive episode**

Giyman et al 2010 Antidepressants 4 662 7 Responders 34.1 57.7 34*** 130** 0.53

**Major depressive disorder**

**Acute episode**

Barbui et al 2010 Paroxetine 22 5112 7.5 Responders 42.4 53.2 10*** 20*** 0.24
Storosum et al 2010 TCAs (new AD studies) 32 4314 6 Responders 31.0 46.0 15*** 50** 0.35
Bech et al 2010 Fluoxetine 16 2761 6 Responders 24.2 37.8 13.6*** 65%*** 0.35
Kirsch et al 2010 New ADs 35 5133 6 HPSD 8.1 12.0 43*** 1.60*** 0.32***

**Maintenance treatment**

Geddes et al 2010 Antidepressants 35 5032 63 Relapse 41.0 18.0 23 NI 58*** 0.64
Hansen et al 2010 New ADs 11 3326 68 Recurrence 48.0 26.0 22 NI 44*** 0.53
Davis et al 2010 Lithium 9 227 NI Relapse 75.0 36.0 39 NI 53*** 0.92

**Obsessive–compulsive disorder**

Soomro et al 2010 SSRIs 17 3097 10 Y-BOCS 3.21*** 0.44***

**Panic disorder**

Mitte 2010 TCAs, SSRIs, BZDs 17–25 NI 8 Anxiety NI 0.41*

**Dementia**

Birks 2010 Cholinesterase inhibitors 10 4236 26 ADAS-Cog 2.38*** 0.41***
9 3118 28 MMSE 1.33*** 0.39***
8 3402 26 Not worse 16.8 24.4 7*** 43*** 0.26

**Attention-deficit hyperactivity disorder**

Schachter et al 2010 Methylphenidate 22 963 3.3 Hyperactivity 0.78***

*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.

AD, antidepressant; ADAS-Cog, Alzheimer’s Disease Assessment Scale Cognitive Subscale; ARD, absolute response or risk difference; BPRS, Brief Psychiatric Rating Scale; BZD, benzodiazepine; HPSD, 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; PANS, Positive and Negative Syndrome Scale; RRR/RR, response ratio/relative risk reduction; SMD, standardised mean difference; SGA, second-generation antipsychotic; SSR, selective serotonin reuptake inhibitor; TCA, tricyclic antidepressant; Y-BOCS, Yale-Brown Obsessive Compulsive Scale; YMRS, Young Mania Rating Scale.

a. AMSTAR quality score range (score 3–27; higher score indicates better methodology).
b. Updated and supplemented by our own searches.
c. Teacher rated.

Note: Estimated values.

\[ \text{NNT} = \frac{1}{\text{MD}} \]

\[ \text{MD} = \frac{\text{Drug} - \text{Placebo}}{\text{Placebo}} \]

\[ \text{SMD} = \frac{\text{Drug} - \text{Placebo}}{\text{Placebo}} \]
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 \( n = 26000 \), 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 \( n = 10251 \).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 primary 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, and death or suicide should be always reported. The example of lithium shows that some psychiatric drugs may reduce suicide rates.82,83

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 showing that clinicians tend to overestimate treatment effects.86 For dichotomous outcomes both relative and absolute risk reductions should be considered. There is substantial evidence that some psychiatric drugs may reduce suicide rates.82,83

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.82 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.80

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 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 studies 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, 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 thrombolysis 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.

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

British Journal of Psychiatry (2012)
200, 97–106. doi: 10.1192/bjp.bp.111.096594

Table DS1: Examples of systematic reviews on the efficacy of psychiatric drugs versus placebo that were mainly based on older studies

Table DS2: Calculation of effect sizes for continuous and dichotomous outcomes

Table DS3: Systematic reviews on the efficacy of general medicine drugs versus placebo (full version of Table 1 in the print version, with more drugs and meta-analyses, confidence intervals and numbers needed to treat (NNTs))

Table DS4: Systematic reviews on the efficacy of psychiatric drugs versus placebo (full version of Table 2 in the print version, with more drugs and meta-analyses, confidence intervals and NNTs)

Figs DS1–24: PRISMA diagrams on search process

Fig. DS25: Systematic presentation of the effect sizes in Fig. 1 labelled by ‘Disease - Drug – Outcome’

Figs DS26, DS27: Summary of the percentage relative risk reductions/response ratios presented in Tables DS3 and DS4

Figs DS28, DS29: Summary of the percentage absolute risk/response differences presented in Tables DS3 and DS4

Explanation of statistical indices presented in the text and Tables DS3 and DS4

References to supplemental material
Table DS1 Examples of systematic reviews on the efficacy of psychiatric drugs versus placebo that were mainly based on older studies

<table>
<thead>
<tr>
<th>Study (Ref)</th>
<th>Therapy</th>
<th>Outcome</th>
<th>Mean dur. wks (range)</th>
<th>N</th>
<th>n</th>
<th>% PBO</th>
<th>% Drug</th>
<th>ARR/ARD (95%CI)</th>
<th>NNT/NNH (95%CI)</th>
<th>RRR/RR (95%CI)</th>
<th>SMD (95%CI)</th>
<th>WMD (95%CI)</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>
<td></td>
</tr>
<tr>
<td>128</td>
<td>Haloperidol1</td>
<td>Response</td>
<td>(0-24)2</td>
<td>8</td>
<td>409</td>
<td>15.4</td>
<td>50.7</td>
<td>36% (25-46)***</td>
<td>3 (2-4)</td>
<td>203% (97-414)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>129</td>
<td>Chlorpromazine1</td>
<td>Response</td>
<td>(0-26)2</td>
<td>24</td>
<td>555</td>
<td>24.7</td>
<td>39.1</td>
<td>26% (17-34)***</td>
<td>78% (48-115)***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>130</td>
<td>Antipsychotics</td>
<td>Relapse</td>
<td>(4-104)</td>
<td>35</td>
<td>3720</td>
<td>55.0</td>
<td>21.0</td>
<td>34% (n.i.)</td>
<td>62% (n.i.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>131</td>
<td>Antipsychotics</td>
<td>Relapse</td>
<td>(2-104)</td>
<td>66</td>
<td>4365</td>
<td>53.0</td>
<td>16.0</td>
<td>37% (n.i.)</td>
<td>70% (n.i.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Depression</strong></td>
<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>132</td>
<td>Tricyclic ADs</td>
<td>Response</td>
<td>n.i.</td>
<td>79</td>
<td>5159</td>
<td>36.0</td>
<td>63.0</td>
<td>27% (n.i.)***</td>
<td>4 (n.i.)</td>
<td>75% (n.i.)***</td>
<td>0.67 (n.i.)</td>
<td>n.i.</td>
</tr>
<tr>
<td>133</td>
<td>MAO-inhibitors</td>
<td>Response</td>
<td>n.i.</td>
<td>33</td>
<td>1944</td>
<td>34.0</td>
<td>65.0</td>
<td>32% (n.i.)***</td>
<td>3 (n.i.)</td>
<td>78% (n.i.)***</td>
<td>0.75 (n.i.)</td>
<td>n.i.</td>
</tr>
<tr>
<td>134</td>
<td>Phenezine</td>
<td>Response</td>
<td>5 (3-6)</td>
<td>9</td>
<td>1108</td>
<td>24.7</td>
<td>39.1</td>
<td>26.5% (n.i.)</td>
<td>62% (n.i.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>135</td>
<td>Tricyclic ADs</td>
<td>Var. scales</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td></td>
</tr>
<tr>
<td>136</td>
<td>Tricyclic ADs</td>
<td>Var. scales</td>
<td>8.1</td>
<td>54</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</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>
<td></td>
</tr>
<tr>
<td>137</td>
<td>Clomipramine</td>
<td>OC sym.</td>
<td>9 (5-10)</td>
<td>9</td>
<td>668</td>
<td>36.0</td>
<td>63.0</td>
<td>27% (n.i.)***</td>
<td>4 (n.i.)</td>
<td>75% (n.i.)***</td>
<td>1.31 (1.15 to 1.47)**</td>
<td>n.i.</td>
</tr>
<tr>
<td>138</td>
<td>Any SRI</td>
<td>OC sym.</td>
<td>(6-13)</td>
<td>12</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td></td>
</tr>
<tr>
<td>139</td>
<td>Any SRI</td>
<td>OC sym.</td>
<td>9 (4-13)</td>
<td>15</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td></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>
<td></td>
</tr>
<tr>
<td>140</td>
<td>Antidepressants</td>
<td>Var. scales</td>
<td>16</td>
<td>580</td>
<td>51.0</td>
<td>72.0</td>
<td>21% (n.i.) ***</td>
<td>5 (n.i.)</td>
<td>40% (n.i.)</td>
<td>0.66 (0.17-0.82)*</td>
<td>n.i.</td>
<td></td>
</tr>
<tr>
<td>141</td>
<td>Benzodiazepines</td>
<td>Var. scales</td>
<td>7 (5-8)</td>
<td>6</td>
<td>696</td>
<td>51.0</td>
<td>72.0</td>
<td>21% (n.i.) ***</td>
<td>5 (n.i.)</td>
<td>40% (n.i.)</td>
<td>0.37 (0.24-0.89)*</td>
<td>n.i.</td>
</tr>
<tr>
<td>142</td>
<td>Antidepressants</td>
<td>Var. scales</td>
<td>5</td>
<td>n.i.</td>
<td>30.0</td>
<td>80.0</td>
<td>50% (n.i.)</td>
<td>2 (n.i.)</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td></td>
</tr>
<tr>
<td>143</td>
<td>Benzodiazepines</td>
<td>Var. scales</td>
<td>4</td>
<td>n.i.</td>
<td>30.0</td>
<td>80.0</td>
<td>50% (n.i.)</td>
<td>2 (n.i.)</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td></td>
</tr>
<tr>
<td>144</td>
<td>Tricyclic ADs</td>
<td>Response</td>
<td>7</td>
<td>1072</td>
<td>51.0</td>
<td>72.0</td>
<td>21% (n.i.) ***</td>
<td>5 (n.i.)</td>
<td>40% (n.i.)</td>
<td>0.66 (0.17-0.82)*</td>
<td>n.i.</td>
<td></td>
</tr>
<tr>
<td>145</td>
<td>SSRIs</td>
<td>Response</td>
<td>7</td>
<td>1486</td>
<td>45.0</td>
<td>72.0</td>
<td>26% (n.i.) ***</td>
<td>4 (n.i.)</td>
<td>n.i.</td>
<td>n.i.</td>
<td>n.i.</td>
<td></td>
</tr>
</tbody>
</table>

Ref = reference, N = number of studies, n = number of participants, % PBO = percentage of patients with the outcome in the placebo group, % Drug = percentage of patients with the outcome in the drug group, ARD = absolute response or risk difference, CI = 95% confidence interval, NNT/NNH = number needed to treat or number needed to harm, RRR/RR: ‘Negative outcomes’ (mortality, relapse, exacerbation, hospitalization, dropout etc) are presented as relative risk reductions (RRR), while positive outcomes (response to treatment, improvement, remission) are presented as percentage response ratios (RR). Positive values mean superiorit of drug, SMD = standardized mean difference, WMD = weighted mean difference, SE = standard error, *** = p<0.001, ** = p<0.01, * = p<0.05, ns = not significant, n.i = not indicated, d = days, var. = various, AD = antidepressant, MAO-inhibitors = mono-amino-oxidase inhibitors, OC symptoms = symptoms of obsessive compulsive disorder, SRI = serotonin-reuptake-inhibitor, SSRIs = selective-serotonin-reuptake-inhibitors, SleepOL = sleep onset latency, min = minutes, TST = total sleep time, Benzod. =benzodiazepine; 1) mainly based on studies from the 1960s to 80s, 2) we combined short-term and medium-term results, 3) various forms of depression, not exclusively major depressive disorder.
<table>
<thead>
<tr>
<th>Index/measure</th>
<th>Formula</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><em><em>Effect sizes for continuous data (e.g. weight, rating scale scores</em>)</em>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference of means (DM)</td>
<td>Mean group A – Mean group B</td>
<td>75kg bodyweight at endpoint in drug group and 70kg in placebo, DM = 5kg</td>
</tr>
<tr>
<td>Standardised difference of means (SDM)</td>
<td>(Mean group A – Mean group B)/pooled standard deviation (SD)</td>
<td>75kg drug, 70kg placebo, pooled SD 10, SMD = 5/10 = 0.50</td>
</tr>
<tr>
<td><strong>Effect sizes for dichotomous data (“yes/no”, e.g. death, relapse)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risk or response rate</td>
<td>Number of participants in a group with an event divided by total number of participants in this group</td>
<td>1 out of 100 participants died, mortality risk = 1/100 = 1%</td>
</tr>
<tr>
<td>Absolute risk or response difference (ARD)</td>
<td>Risk or percentage responders in group A – Risk or percentage responders in group B</td>
<td>1% deaths in drug - 3% deaths in placebo, ARD = -2%</td>
</tr>
<tr>
<td>Relative risk reduction (RRR)</td>
<td>1- (Risk or percentage responders in group A – Risk or percentage responders in group B)</td>
<td>1-(1%/3%) = 67%</td>
</tr>
<tr>
<td>Percentage response ratio (RR)</td>
<td>Percentage responders group A / percentage responders group B</td>
<td>50% drug responders / 31% to placebo, RR = 1.61 times or 61% more responders.</td>
</tr>
<tr>
<td>Number needed to treat (NNT)</td>
<td>1/absolute risk or response difference</td>
<td>1 / 2% = 1/0.02, NNT = 50</td>
</tr>
</tbody>
</table>

* mean values of psychiatric ratings scales are not really continuous data, but are treated as such in meta-analyses
## Table DS3 Systematic reviews on the efficacy of general medicine drugs versus placebo (full version of Table 1 in the print version with more drugs and meta-analyses, confidence intervals and NNTs)

<table>
<thead>
<tr>
<th>Study (Ref)</th>
<th>AM-STAR score</th>
<th>Therapy</th>
<th>Outcome</th>
<th>Mean weeks</th>
<th>N</th>
<th>n</th>
<th>% PBO</th>
<th>% Drug</th>
<th>ARD (CI)</th>
<th>NNT/H (CI)</th>
<th>RR (CI)</th>
<th>SDM (CI)</th>
<th>WMD (CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<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 – 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>
</tr>
<tr>
<td>1/5†1</td>
<td>Any antihyp.</td>
<td>RR systolic (mmHg)</td>
<td>8</td>
<td>94</td>
<td>17641</td>
<td>25 (mi)</td>
<td>22% (17-27)***</td>
<td>0.56 (0.52-0.58)***</td>
<td>9.4 (8.9-9.9)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RR diastolic (mmHg)</td>
<td>8</td>
<td>94</td>
<td>17641</td>
<td>83 (mi)</td>
<td>12% (4-19)**</td>
<td>0.07 (-0.02-0.13)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/5†1</td>
<td>ACE-inhibitors</td>
<td>RR systolic (mmHg)</td>
<td>8 (2-14)</td>
<td>39</td>
<td>6601</td>
<td>50 (mi)</td>
<td>18% (5-29)**</td>
<td>0.13 (0.04-0.22)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RR diastolic (mmHg)</td>
<td>8 (2-14)</td>
<td>39</td>
<td>6601</td>
<td>125 (mi)</td>
<td>11% (-5-25)</td>
<td>0.07 (-0.03-0.17)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/5†1</td>
<td>ARBs</td>
<td>RR systolic (mmHg)</td>
<td>8 (2-14)</td>
<td>28</td>
<td>11715</td>
<td>100 (mi)</td>
<td>12% (2-22)**</td>
<td>0.08 (-0.02-0.14)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RR diastolic (mmHg)</td>
<td>8 (2-14)</td>
<td>28</td>
<td>11715</td>
<td>100 (mi)</td>
<td>12% (2-22)**</td>
<td>0.08 (-0.02-0.14)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/5†1</td>
<td>Beta-blockers</td>
<td>RR systolic (mmHg)</td>
<td>8 (2-14)</td>
<td>19</td>
<td>3018</td>
<td>65 (mi)</td>
<td>24% (17-31)**</td>
<td>0.06 (0.05-0.06)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RR diastolic (mmHg)</td>
<td>8 (2-14)</td>
<td>19</td>
<td>3018</td>
<td>65 (mi)</td>
<td>24% (17-31)**</td>
<td>0.06 (0.05-0.06)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypertension – long term effects on cardiovascular events and mortality</td>
<td></td>
<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>1/5†1</td>
<td>ACE- inhibitors</td>
<td>CV events</td>
<td>3.9y</td>
<td>5</td>
<td>18229</td>
<td>18.1</td>
<td>14.1</td>
<td>4% (mi)</td>
<td>25 (mi)</td>
<td>22% (17-27)***</td>
<td>0.16 (0.12-0.21)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mortality</td>
<td>3.9y</td>
<td>5</td>
<td>18229</td>
<td>10.4</td>
<td>9.2</td>
<td>1% (mi)</td>
<td>83 (mi)</td>
<td>12% (4-19)**</td>
<td>0.07 (-0.02-0.13)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/5†1</td>
<td>Ca-antagonists</td>
<td>CV events</td>
<td>2.8y</td>
<td>3</td>
<td>6656</td>
<td>10.3</td>
<td>8.3</td>
<td>2% (mi)</td>
<td>50 (mi)</td>
<td>18% (5-29)**</td>
<td>0.13 (0.04-0.22)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mortality</td>
<td>2.8y</td>
<td>4</td>
<td>7482</td>
<td>7.1</td>
<td>6.3</td>
<td>1% (mi)</td>
<td>125 (mi)</td>
<td>11% (-5-25)</td>
<td>0.07 (-0.03-0.17)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/10†1</td>
<td>Beta-Blockers</td>
<td>CV events</td>
<td>3.9y</td>
<td>4</td>
<td>23613</td>
<td>6.5</td>
<td>5.7</td>
<td>1% (0-2)</td>
<td>100 (mi)</td>
<td>12% (2-22)**</td>
<td>0.08 (-0.02-0.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mortality</td>
<td>3.9y</td>
<td>4</td>
<td>23613</td>
<td>5.2</td>
<td>5.0</td>
<td>0% (0-1)</td>
<td>100 (mi)</td>
<td>12% (2-22)**</td>
<td>0.08 (-0.02-0.14)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/8</td>
<td>Diuretics</td>
<td>CV events</td>
<td>3.9y</td>
<td>42</td>
<td>192478</td>
<td>mi</td>
<td>mi</td>
<td>mi</td>
<td>mi</td>
<td>24% (17-31)**</td>
<td>ne</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mortality</td>
<td>3.9y</td>
<td>42</td>
<td>192478</td>
<td>mi</td>
<td>mi</td>
<td>mi</td>
<td>mi</td>
<td>10% (4-16)**</td>
<td>ne</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute stroke</td>
<td></td>
<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>1/9</td>
<td>Thrombolysis</td>
<td>Death/dependency</td>
<td>12-26</td>
<td>22</td>
<td>6283</td>
<td>55.8</td>
<td>50.9</td>
<td>5% (1-9)**</td>
<td>20 (11-100)</td>
<td>9% (3-14)**</td>
<td>0.11 (-0.05-0.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/11</td>
<td>ASA</td>
<td>Death/dependency</td>
<td>4-26</td>
<td>4</td>
<td>41291</td>
<td>46.0</td>
<td>45.0</td>
<td>1% (0-2)*</td>
<td>100 (mi)</td>
<td>2% (0-1.4)*</td>
<td>0.02 (-0.01-0.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/11</td>
<td>Anticoagulants</td>
<td>Death/dependency</td>
<td>&gt;4</td>
<td>8</td>
<td>22152</td>
<td>59.9</td>
<td>59.4</td>
<td>3% (-1-7)</td>
<td>33 (H50-T14)</td>
<td>5% (-4-14)</td>
<td>0.01 (-0.02-0.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prevention of cardiovascular diseases and stroke</td>
<td></td>
<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>1/5◦</td>
<td>ASA (prim.prev.)</td>
<td>Serious vasc. ev.</td>
<td>5.8y</td>
<td>6</td>
<td>95000</td>
<td>0.57y</td>
<td>0.51y</td>
<td>0.07% (0.5)</td>
<td>1429y (mi)</td>
<td>12% (6-18)***</td>
<td>0.06 (-0.03-0.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study (Ref) AM-STAR score</td>
<td>Therapy</td>
<td>Outcome</td>
<td>Mean weeks</td>
<td>N</td>
<td>n</td>
<td>% PBO</td>
<td>% Drug</td>
<td>ARD (CI)</td>
<td>NNT/H (CI)</td>
<td>1RRR/RR (CI)</td>
<td>SDM (CI)</td>
<td>WMD (CI)</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>---------</td>
<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>7/5</td>
<td>ASA (sec.prev.)</td>
<td>Vascular mort.</td>
<td>5.8y</td>
<td>6</td>
<td>95000</td>
<td>0.19y</td>
<td>0.19y</td>
<td>0.0% (y (ni))</td>
<td>ne</td>
<td>3% (y (-9.13))</td>
<td>0.00 (-0.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7/5</td>
<td>Statins</td>
<td>Vascular mort.</td>
<td>ni</td>
<td>16</td>
<td>17000</td>
<td>4.07/</td>
<td>3.07/</td>
<td>0.29% (y (ni))</td>
<td>358</td>
<td>9% (y (0.18))</td>
<td>0.06 (-0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9/6</td>
<td>Statins</td>
<td>LDL-cholesterol (mmol/l)</td>
<td>2-6</td>
<td>164</td>
<td>~38000</td>
<td>72</td>
<td>30</td>
<td>14</td>
<td>5</td>
<td>4% (3-4)**</td>
<td>27 (25-33)</td>
<td>21% (19-23)**</td>
<td>0.15 (0.13-0.17)</td>
</tr>
</tbody>
</table>

### Chronic heart failure

<table>
<thead>
<tr>
<th>Study (Ref) AM-STAR score</th>
<th>Therapy</th>
<th>Outcome</th>
<th>Mean weeks</th>
<th>N</th>
<th>n</th>
<th>% PBO</th>
<th>% Drug</th>
<th>ARD (CI)</th>
<th>NNT/H (CI)</th>
<th>1RRR/RR (CI)</th>
<th>SDM (CI)</th>
<th>WMD (CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7/6</td>
<td>ACE-inhib. short-term</td>
<td>Mortality</td>
<td>&gt; 8</td>
<td>32</td>
<td>7105</td>
<td>21.9</td>
<td>15.8</td>
<td>6% (4-8)**</td>
<td>16 (13-23)</td>
<td>19% (10-28)**</td>
<td>0.22 (0.16-0.29)</td>
<td></td>
</tr>
<tr>
<td>7/6</td>
<td>ACE-inhib. long-term</td>
<td>Mortality</td>
<td>&gt; 8</td>
<td>30</td>
<td>6988</td>
<td>32.6</td>
<td>22.4</td>
<td>10% (8-12)**</td>
<td>10 (8-13)</td>
<td>27% (19-34)**</td>
<td>0.28 (0.23-0.34)</td>
<td></td>
</tr>
<tr>
<td>7/9</td>
<td>ARBs</td>
<td>Hospitalization</td>
<td>2.9/</td>
<td>5</td>
<td>12763</td>
<td>26.8</td>
<td>23.0</td>
<td>4% (2-5)**</td>
<td>26 (20-50)</td>
<td>15% (12-20)**</td>
<td>0.11 (0.07-0.16)</td>
<td></td>
</tr>
<tr>
<td>7/9</td>
<td>ARBs</td>
<td>Hospitalization</td>
<td>2.9/</td>
<td>5</td>
<td>12763</td>
<td>18.9</td>
<td>13.7</td>
<td>5% (ni)**</td>
<td>19 (ni)</td>
<td>29% (22-34)**</td>
<td>0.21 (0.16-0.26)</td>
<td></td>
</tr>
<tr>
<td>7/1</td>
<td>Beta-Blockers (partly add on)</td>
<td>Mortality</td>
<td>36</td>
<td>22</td>
<td>10135</td>
<td>13.2</td>
<td>8.4</td>
<td>5% (3-6)**</td>
<td>21 (17-33)</td>
<td>32% (18-44)**</td>
<td>0.28 (0.21-0.35)</td>
<td></td>
</tr>
<tr>
<td>7/10</td>
<td>Diuretics</td>
<td>Hospitalization</td>
<td>36</td>
<td>22</td>
<td>10135</td>
<td>15.5</td>
<td>10.2</td>
<td>5% (3-6)**</td>
<td>19 (14-25)</td>
<td>32% (18-43)**</td>
<td>0.26 (0.20-0.33)</td>
<td></td>
</tr>
<tr>
<td>7/9</td>
<td>Digitalis</td>
<td>Hospitalization</td>
<td>36</td>
<td>22</td>
<td>10135</td>
<td>13.2</td>
<td>8.4</td>
<td>5% (3-6)**</td>
<td>21 (17-33)</td>
<td>32% (18-44)**</td>
<td>0.28 (0.21-0.35)</td>
<td></td>
</tr>
</tbody>
</table>

### Rheumatoid arthritis

<table>
<thead>
<tr>
<th>Study (Ref) AM-STAR score</th>
<th>Therapy</th>
<th>Outcome</th>
<th>Mean weeks</th>
<th>N</th>
<th>n</th>
<th>% PBO</th>
<th>% Drug</th>
<th>ARD (CI)</th>
<th>NNT/H (CI)</th>
<th>1RRR/RR (CI)</th>
<th>SDM (CI)</th>
<th>WMD (CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7/10</td>
<td>Methotrexate</td>
<td>No. tend. joints</td>
<td>&gt; 12</td>
<td>5</td>
<td>218</td>
<td>58% (0-4.1)**</td>
<td>33 (H4-T7)</td>
<td>0.49% (0.03-8.84)**</td>
<td>0.96 (0.35-1.57)</td>
<td>0.52 (0.3-1.0)**</td>
<td>ne</td>
<td>ne</td>
</tr>
<tr>
<td>7/10</td>
<td>Steroids short-term</td>
<td>No. tend. joints</td>
<td>&lt; 4</td>
<td>2</td>
<td>182</td>
<td>58% (0-4.1)**</td>
<td>33 (H4-T7)</td>
<td>0.49% (0.03-8.84)**</td>
<td>0.96 (0.35-1.57)</td>
<td>0.52 (0.3-1.0)**</td>
<td>ne</td>
<td>ne</td>
</tr>
<tr>
<td>7/9</td>
<td>Steroids mod.-term</td>
<td>No. tend. joints</td>
<td>&gt; 12</td>
<td>5</td>
<td>304</td>
<td>58% (0-4.1)**</td>
<td>33 (H4-T7)</td>
<td>0.49% (0.03-8.84)**</td>
<td>0.96 (0.35-1.57)</td>
<td>0.52 (0.3-1.0)**</td>
<td>ne</td>
<td>ne</td>
</tr>
<tr>
<td>7/9</td>
<td>Azathioprine</td>
<td>No. tend. joints</td>
<td>&gt; 26</td>
<td>3</td>
<td>81</td>
<td>58% (0-4.1)**</td>
<td>33 (H4-T7)</td>
<td>0.49% (0.03-8.84)**</td>
<td>0.96 (0.35-1.57)</td>
<td>0.52 (0.3-1.0)**</td>
<td>ne</td>
<td>ne</td>
</tr>
<tr>
<td>7/10</td>
<td>Cyclosporine</td>
<td>Change no. tend. joints</td>
<td>&lt; 52</td>
<td>1</td>
<td>144</td>
<td>58% (0-4.1)**</td>
<td>33 (H4-T7)</td>
<td>0.49% (0.03-8.84)**</td>
<td>0.96 (0.35-1.57)</td>
<td>0.52 (0.3-1.0)**</td>
<td>ne</td>
<td>ne</td>
</tr>
<tr>
<td>7/9</td>
<td>Cyclophosphamide</td>
<td>No. tend. joints</td>
<td>&gt; 26</td>
<td>2</td>
<td>70</td>
<td>58% (0-4.1)**</td>
<td>33 (H4-T7)</td>
<td>0.49% (0.03-8.84)**</td>
<td>0.96 (0.35-1.57)</td>
<td>0.52 (0.3-1.0)**</td>
<td>ne</td>
<td>ne</td>
</tr>
<tr>
<td>Study (Ref)</td>
<td>Therapy</td>
<td>Outcome</td>
<td>Mean weeks</td>
<td>N</td>
<td>n</td>
<td>% PBO</td>
<td>% Drug</td>
<td>ARD (CI)</td>
<td>NNT/H (CI)</td>
<td>RRR/RR (CI)</td>
<td>SDM (CI)</td>
<td>WMD (CI)</td>
</tr>
<tr>
<td>------------</td>
<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>27/10/10</td>
<td>Sulfasalazine</td>
<td>No. tend. joints</td>
<td>&gt; 26</td>
<td>6</td>
<td>256</td>
<td>32.0</td>
<td>10.3</td>
<td>22% (10-43)**</td>
<td>5 (3-6)</td>
<td>67% (39-2)**</td>
<td>0.49 (0.24-0.75)**</td>
<td>ne</td>
</tr>
<tr>
<td>27/8/8</td>
<td>Penicillamine</td>
<td>No. tend. joints</td>
<td>&gt; 12</td>
<td>5</td>
<td>316</td>
<td>5.7</td>
<td>2.8</td>
<td>3% (-1.7)***</td>
<td>33 (100-114)</td>
<td>56% (-33.85)***</td>
<td>0.41 (-0.33-1.04)</td>
<td>ne</td>
</tr>
<tr>
<td>27/8/8</td>
<td>Auranofin</td>
<td>No. tend. joints</td>
<td>&gt; 26</td>
<td>7</td>
<td>750</td>
<td>19.6</td>
<td>6.5</td>
<td>14% (3-5)***</td>
<td>7 (4-33)</td>
<td>60% (31.7)***</td>
<td>0.69 (0.47-0.92)</td>
<td>ne</td>
</tr>
<tr>
<td>27/10/10</td>
<td>Leflunomide</td>
<td>No. tend. joints</td>
<td>&gt; 26</td>
<td>8</td>
<td>1049</td>
<td>26.3</td>
<td>52.2</td>
<td>26% (15-37)**</td>
<td>4 (4-7)</td>
<td>50% (36-69)**</td>
<td>0.62 (0.35-0.88)</td>
<td>ne</td>
</tr>
<tr>
<td>27/9/9</td>
<td>Antimalarials</td>
<td>No. tend. joints</td>
<td>&gt; 26</td>
<td>4</td>
<td>571</td>
<td>7% (0.1-14)**</td>
<td>14 (7-1000)</td>
<td>40% (6-62)**</td>
<td>0.34 (0.05-0.62)</td>
<td>ne</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27/9/9</td>
<td>Celecoxib (200mg)</td>
<td>No. tend. joints</td>
<td>12</td>
<td>1</td>
<td>466</td>
<td>28.6</td>
<td>43.8</td>
<td>15% (7-24)**</td>
<td>7 (4-14)</td>
<td>53% (19-97)**</td>
<td>0.37 (0.15-0.58)</td>
<td>ne</td>
</tr>
<tr>
<td>27/10/10</td>
<td>Adalimumab (40mg)</td>
<td>No. tend. joints</td>
<td>28</td>
<td>1</td>
<td>213</td>
<td>20.5</td>
<td>52.9</td>
<td>32% (19-46)**</td>
<td>3 (2-5)</td>
<td>160% (60-410)**</td>
<td>0.81 (0.44-1.17)</td>
<td>ne</td>
</tr>
</tbody>
</table>

**Migraine - acute treatment**

<table>
<thead>
<tr>
<th>Study (Ref)</th>
<th>Therapy</th>
<th>Outcome</th>
<th>Mean weeks</th>
<th>N</th>
<th>n</th>
<th>% PBO</th>
<th>% Drug</th>
<th>ARD (CI)</th>
<th>NNT/H (CI)</th>
<th>RRR/RR (CI)</th>
<th>SDM (CI)</th>
<th>WMD (CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>27/8/8</td>
<td>Sumatriptan</td>
<td>Pain-free</td>
<td>2 hrs</td>
<td>8</td>
<td>2221</td>
<td>8.5</td>
<td>29.5</td>
<td>20% (15-24)**</td>
<td>5 (4-7)</td>
<td>220% (150-310)**</td>
<td>0.41 (0.25-0.56)</td>
<td>ne</td>
</tr>
<tr>
<td>27/3/3</td>
<td>Aspirin</td>
<td>Pain-free</td>
<td>2 hrs</td>
<td>3</td>
<td>1246</td>
<td>15.1</td>
<td>27.1</td>
<td>12% (nil)**</td>
<td>8 (ni)</td>
<td>80% (ni)**</td>
<td>0.83 (0.69-0.97)</td>
<td>ne</td>
</tr>
</tbody>
</table>

**Migraine – prophylaxis**

<table>
<thead>
<tr>
<th>Study (Ref)</th>
<th>Therapy</th>
<th>Outcome</th>
<th>Mean weeks</th>
<th>N</th>
<th>n</th>
<th>% PBO</th>
<th>% Drug</th>
<th>ARD (CI)</th>
<th>NNT/H (CI)</th>
<th>RRR/RR (CI)</th>
<th>SDM (CI)</th>
<th>WMD (CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>27/9/9</td>
<td>Propanolol</td>
<td>Mig. freq</td>
<td>13</td>
<td>4</td>
<td>205</td>
<td>30.9</td>
<td>52.3</td>
<td>35% (5-69)**</td>
<td>3 (1-20)</td>
<td>80% (3-210)**</td>
<td>0.49 (0.18-0.81)</td>
<td>0.47 (0.12-0.83)**</td>
</tr>
<tr>
<td>27/7/7</td>
<td>Anticonvulsants</td>
<td>Mig. freq</td>
<td>12.3</td>
<td>14</td>
<td>1773</td>
<td>20.6</td>
<td>47.0</td>
<td>26% (ni)**</td>
<td>4 (ni)</td>
<td>130% (90-180)**</td>
<td>0.68 (0.56-0.79)</td>
<td>0.55 (0.26-0.85)**</td>
</tr>
</tbody>
</table>

**Asthma** (data on short-acting-beta-2 agonists as needed are not presented, because we only found a systematic review compared to continuous treatment)**

**Inhaled corticosteroids**

<table>
<thead>
<tr>
<th>Study (Ref)</th>
<th>Therapy</th>
<th>Outcome</th>
<th>Mean weeks</th>
<th>N</th>
<th>n</th>
<th>% PBO</th>
<th>% Drug</th>
<th>ARD (CI)</th>
<th>NNT/H (CI)</th>
<th>RRR/RR (CI)</th>
<th>SDM (CI)</th>
<th>WMD (CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>27/7/7</td>
<td>Corticosteroids</td>
<td>FEV1(l)</td>
<td>&gt; 12</td>
<td>19</td>
<td>3271</td>
<td>54% (38-66)**</td>
<td>0.56 (0.45-0.66)**</td>
<td>0.33 (0.26-0.40)**</td>
<td>ne</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exacerbation</td>
<td>&gt; 12</td>
<td>11</td>
<td>8999</td>
<td>ni</td>
<td>ni</td>
<td>ni</td>
<td>ni</td>
<td>54% (38-66)**</td>
<td>ne</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study (Ref) AM-STAR score</td>
<td>Therapy</td>
<td>Outcome</td>
<td>Mean weeks</td>
<td>N</td>
<td>n</td>
<td>% PBO</td>
<td>% Drug</td>
<td>ARD (CI)</td>
<td>NNT/H (CI)</td>
<td>^RRR/RR (CI)</td>
<td>SDM (CI)</td>
<td>WMD (CI)</td>
</tr>
<tr>
<td>--------------------------</td>
<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>77/10</td>
<td>Beclomethasone</td>
<td>FEV1(l)</td>
<td>14</td>
<td>6</td>
<td>612</td>
<td></td>
<td></td>
<td>11% (3-18)**</td>
<td>9 (6-33)</td>
<td>71% (54-81)**</td>
<td>0.42 (0.22-0.49)**</td>
<td>0.36 (0.26-0.58)**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DO exacerbation</td>
<td>14</td>
<td>10</td>
<td>1185</td>
<td>15.4</td>
<td>3.1</td>
<td>11% (3-18)**</td>
<td>9 (6-33)</td>
<td>71% (54-81)**</td>
<td>0.42 (0.22-0.49)**</td>
<td>0.36 (0.26-0.58)**</td>
</tr>
<tr>
<td>3/10</td>
<td>Fluticasone</td>
<td>FEV1 (l)</td>
<td>14</td>
<td>21</td>
<td>4790</td>
<td></td>
<td></td>
<td>11% (1-21)**</td>
<td>9 (5-100)</td>
<td>80% (58-90)**</td>
<td>1.01 (0.56-1.46)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DO exacerbation</td>
<td>14</td>
<td>4</td>
<td>702</td>
<td>11.4</td>
<td>2.0</td>
<td>11% (1-21)**</td>
<td>9 (5-100)</td>
<td>80% (58-90)**</td>
<td>1.01 (0.56-1.46)</td>
<td></td>
</tr>
<tr>
<td><strong>Long-acting beta-2-agonists</strong></td>
<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>77/10</td>
<td>LAB2(add on)</td>
<td>FEV1 (l or %)</td>
<td>19</td>
<td>17</td>
<td>3926</td>
<td></td>
<td></td>
<td>5% (3.7)**</td>
<td>20 (14-33)</td>
<td>17% (7-25)**</td>
<td>0.15 (0.07-0.23)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Exacerbation</td>
<td>19</td>
<td>17</td>
<td>4027</td>
<td>27.4</td>
<td>22.4</td>
<td>-0.7% (-1.3-0.1)**</td>
<td>-143 (-106-1000)</td>
<td>-114% (-15-297)**</td>
<td>0.58 (0.26-0.90)</td>
<td></td>
</tr>
<tr>
<td>77/9</td>
<td>LAB2(partly add on)</td>
<td>Hospitalization</td>
<td>26</td>
<td>12</td>
<td>5091</td>
<td>0.6</td>
<td>1.7</td>
<td>-0.7% (-1.3-0.1)**</td>
<td>-143 (-106-1000)</td>
<td>-114% (-15-297)**</td>
<td>0.58 (0.26-0.90)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LAB2</td>
<td>FEV 1 (%)</td>
<td>&gt; 12</td>
<td>13</td>
<td>3888</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Exacerbation</td>
<td>&gt; 12</td>
<td>9</td>
<td>2854</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LAB2(add on)</td>
<td>Exacerbation</td>
<td>12</td>
<td>24</td>
<td>7549</td>
<td>8.3</td>
<td>4.9</td>
<td>2.5% (1.4-3.6)**</td>
<td>40 (28-71)**</td>
<td>35% (20-45)**</td>
<td>0.31 (0.21-0.41)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hospitalization</td>
<td>&gt; 12</td>
<td>7</td>
<td>4375</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Third line treatments (we did not find a systematic review on theophylline)</strong></td>
<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>77/7</td>
<td>Leukot. Antg.</td>
<td>FEV1 (%)</td>
<td>&gt; 12</td>
<td>7</td>
<td>4375</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Exacerbation</td>
<td>&gt; 12</td>
<td>7</td>
<td>4375</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Chronic obstructive pulmonary disease (COPD)</strong></td>
<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>77/10</td>
<td>Tiotropium</td>
<td>FEV1 (l)</td>
<td>&gt; 4</td>
<td>4</td>
<td>1735</td>
<td></td>
<td></td>
<td>5% (3.7)**</td>
<td>20 (14-33)</td>
<td>17% (10-24)**</td>
<td>0.21 (0.15-0.28)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Exacerbations</td>
<td>&gt; 4</td>
<td>8</td>
<td>5644</td>
<td>30.8</td>
<td>23.2</td>
<td>5% (3.7)**</td>
<td>20 (14-33)</td>
<td>17% (10-24)**</td>
<td>0.21 (0.15-0.28)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mortality</td>
<td>&gt; 4</td>
<td>2</td>
<td>1723</td>
<td>10.6</td>
<td>9.8</td>
<td>1% (0-2.1)*</td>
<td>100 (50-∞)</td>
<td>47% (-39-20)*</td>
<td>0.05 (-0.12-0.22)</td>
<td></td>
</tr>
<tr>
<td>77/10</td>
<td>Anticholinergics</td>
<td>Hospitalization</td>
<td>1.7y</td>
<td>3</td>
<td>3552</td>
<td>8.4</td>
<td>5.7</td>
<td>2.7% (1.5-4)*</td>
<td>37 (26-60)</td>
<td>33% (14-47)**</td>
<td>0.23 (0.09-0.37)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mortality</td>
<td>1.7y</td>
<td>5</td>
<td>7881</td>
<td>0.3</td>
<td>0.05</td>
<td>0.25% (0.4-1)*</td>
<td>400 (250-700)</td>
<td>73% (19-91)*</td>
<td>0.99 (0.16-1.82)</td>
<td></td>
</tr>
<tr>
<td>77/10</td>
<td>Inh. corticosteroids</td>
<td>FEV1 (l)</td>
<td>8-24</td>
<td>3</td>
<td>952</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Exacerbations/p/y</td>
<td>&gt; 24</td>
<td>4</td>
<td>2063</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mortality</td>
<td>&gt; 24</td>
<td>9</td>
<td>8390</td>
<td>7.6</td>
<td>7.7</td>
<td>0.0% (-0.01-0)*</td>
<td>ne</td>
<td>-1% (-15-14)*</td>
<td>-0.01 (-0.10-0.08)</td>
<td></td>
</tr>
<tr>
<td>77/10</td>
<td>Short-act β-2-ag.</td>
<td>FEV1 (l)</td>
<td>3</td>
<td>6</td>
<td>196</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Exacerbation</td>
<td>3</td>
<td>5</td>
<td>198</td>
<td>46.5</td>
<td>22.2</td>
<td>26% (12-40)**</td>
<td>4 (3-8)</td>
<td>47% (14-67)**</td>
<td>0.61 (0.27-0.95)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Resp. Mortality</td>
<td>1.7y</td>
<td>4</td>
<td>2404</td>
<td>0.7</td>
<td>1.6</td>
<td>-0.9% (0.4-1)*</td>
<td>-111 (106-500)</td>
<td>-147% (-12-445)*</td>
<td>0.46 (0.01-0.91)</td>
<td></td>
</tr>
<tr>
<td>77/6</td>
<td>Long-act β-2-ag.</td>
<td>Exacerbation</td>
<td>18</td>
<td>9</td>
<td>4198</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mortality</td>
<td>18</td>
<td>9</td>
<td>4198</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study (Ref)</td>
<td>Therapy</td>
<td>Outcome</td>
<td>Mean weeks</td>
<td>N</td>
<td>n</td>
<td>% PBO</td>
<td>% Drug</td>
<td>ARD (CI)</td>
<td>NNT/H (CI)</td>
<td>RRR/RR (CI)</td>
<td>SDM (CI)</td>
<td>WMD (CI)</td>
</tr>
<tr>
<td>------------</td>
<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>AM-STAR score</td>
<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>7/10</td>
<td>Theophylline</td>
<td>FEV1 (l)</td>
<td>&gt;1</td>
<td>13</td>
<td>244</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Exacerbations</td>
<td>n.i.</td>
<td>2</td>
<td>45</td>
<td>ni</td>
<td>ni</td>
<td>13% (1-26)*</td>
<td>8 (4-100)</td>
<td>67% (-14 - 90)m</td>
<td>ne</td>
<td></td>
</tr>
<tr>
<td>Diabetes</td>
<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>8/11</td>
<td>Metformin</td>
<td>HbA1c (%)</td>
<td>21.5</td>
<td>12</td>
<td>1587</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fasting glucose (mmol/l)</td>
<td>21.5</td>
<td>12</td>
<td>1587</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mort.(vs conv. treat)</td>
<td>10.7</td>
<td>1</td>
<td>753</td>
<td>21.7</td>
<td>14.6</td>
<td>7% (2-13)**</td>
<td>14 (8-50)</td>
<td>32% (7.51)**</td>
<td>0.27</td>
<td>(0.06-0.47)**</td>
</tr>
<tr>
<td>8/11</td>
<td>α-gluc.-inhib.</td>
<td>HbA1c (%)</td>
<td>30</td>
<td>28</td>
<td>2831</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fasting glucose (mmol/l)</td>
<td>30</td>
<td>28</td>
<td>2831</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mortality</td>
<td>1.7</td>
<td>2</td>
<td>385</td>
<td>2.2</td>
<td>2.5</td>
<td>0% (-2.3)ns</td>
<td>ne</td>
<td>-10% (-301-70)ns</td>
<td>-0.07</td>
<td>(-0.80-0.66)ns</td>
</tr>
<tr>
<td>8/8</td>
<td>GLP-1 anal.</td>
<td>HbA1c (%)</td>
<td>&gt;12</td>
<td>6</td>
<td>1285</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DPP4 inhib.</td>
<td>HbA1c (%)</td>
<td>&gt;12</td>
<td>16</td>
<td>4109</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8/10</td>
<td>Meglitinide add-on</td>
<td>HbA1c (%)</td>
<td>12</td>
<td>1</td>
<td>54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chronic hepatitis 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>
</tr>
<tr>
<td>7/10</td>
<td>Interferon</td>
<td>Virol. resp.</td>
<td>&gt;26</td>
<td>8</td>
<td>409</td>
<td>1.0</td>
<td>38.3</td>
<td>35% (23.47)**</td>
<td>3 (2-4)</td>
<td>1070% (370-2850)***</td>
<td>2.27</td>
<td>(1.49-3.04)**</td>
</tr>
<tr>
<td>7/10</td>
<td>Ribavirin</td>
<td>Virol. resp.</td>
<td>45</td>
<td>10</td>
<td>511</td>
<td>1.3</td>
<td>1.4</td>
<td>0% (-2.2)ns</td>
<td>ne</td>
<td>0% (-2.0-340)ns</td>
<td>0.04</td>
<td>(-0.79-0.87)ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Morb. and Mort.</td>
<td>45</td>
<td>11</td>
<td>521</td>
<td>0.4</td>
<td>0.7</td>
<td>-0% (-2.3)ns</td>
<td>ne</td>
<td>-34% (-90-10)ns</td>
<td>-0.31</td>
<td>(-1.64-1.02)ns</td>
</tr>
<tr>
<td>8/10</td>
<td>Interf. + Ribav.</td>
<td>Virol. resp.</td>
<td>30</td>
<td>52</td>
<td>8354</td>
<td>13</td>
<td>37</td>
<td>20% (16.24)**</td>
<td>5 (4-6)</td>
<td>160% (120-210)***</td>
<td>0.75</td>
<td>(0.69-0.82)**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Morb. and Mort.</td>
<td>29</td>
<td>79</td>
<td>9991</td>
<td>0.44</td>
<td>0.20</td>
<td>0% (-0.6)ns</td>
<td>ne</td>
<td>51% (4.75)ns</td>
<td>0.44</td>
<td>(0.02-0.85)ns</td>
</tr>
<tr>
<td>Reflux oesophagitis</td>
<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>8/9</td>
<td>PPI</td>
<td>Clin. Remission</td>
<td>8</td>
<td>5</td>
<td>645</td>
<td>28.3</td>
<td>83.2</td>
<td>58% (47.68)**</td>
<td>2 (2-2)</td>
<td>256% (111-500)***</td>
<td>1.39</td>
<td>(1.18-1.60)**</td>
</tr>
<tr>
<td>8/9</td>
<td>PPI (maint. dose)</td>
<td>Relapse</td>
<td>26-51</td>
<td>5</td>
<td>1465</td>
<td>75.4</td>
<td>36.1</td>
<td>39% (35.44)**</td>
<td>3 (2-3)</td>
<td>54% (43-62)**</td>
<td>0.93</td>
<td>(0.81-1.06)**</td>
</tr>
<tr>
<td>8/9</td>
<td>PPI (healing dose)</td>
<td>Relapse</td>
<td>26-51</td>
<td>10</td>
<td>1385</td>
<td>78.8</td>
<td>21.7</td>
<td>57% (53.62)**</td>
<td>2 (2-3)</td>
<td>74% (64-81)**</td>
<td>1.43</td>
<td>(1.29-1.57)**</td>
</tr>
<tr>
<td>Ulcerative colitis</td>
<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>8/9</td>
<td>5-ASA</td>
<td>Clin. Remission</td>
<td>8</td>
<td>4</td>
<td>892</td>
<td>10.0</td>
<td>19.9</td>
<td>8% (4.13)**</td>
<td>13 (8-25)</td>
<td>70% (10-160)***</td>
<td>0.44</td>
<td>(0.23-0.66)***</td>
</tr>
<tr>
<td>8/9</td>
<td>5-ASA</td>
<td>Maint. Remission</td>
<td>26</td>
<td>5</td>
<td>881</td>
<td>36.7</td>
<td>52.9</td>
<td>18% (12-24)**</td>
<td>6 (4-8)</td>
<td>50% (38-70)***</td>
<td>0.36</td>
<td>(0.22-0.51)**</td>
</tr>
<tr>
<td>Multiple Sclerosis</td>
<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>8/10</td>
<td>Corticosteroids</td>
<td>Improvement</td>
<td>&lt;5</td>
<td>3</td>
<td>93</td>
<td>27.9</td>
<td>68.0</td>
<td>41% (23.59)**</td>
<td>2 (2-4)</td>
<td>140% (40-290)***</td>
<td>0.93</td>
<td>(0.45-1.42)**</td>
</tr>
<tr>
<td>8/10</td>
<td>Interferon</td>
<td>Exacerbation</td>
<td>2.0</td>
<td>3</td>
<td>919</td>
<td>69.5</td>
<td>55.2</td>
<td>14% (8.20)**</td>
<td>7 (5-13)</td>
<td>19% (11-26)***</td>
<td>0.34</td>
<td>(0.19-0.49)**</td>
</tr>
<tr>
<td>Parkinson’s disease</td>
<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>

8
<table>
<thead>
<tr>
<th>Study/Ref (AMSTAR score)</th>
<th>Therapy</th>
<th>Outcome</th>
<th>Mean weeks</th>
<th>N</th>
<th>n</th>
<th>% PBO</th>
<th>% Drug</th>
<th>ARD CI</th>
<th>NNT/H CI</th>
<th>RRR RR CI</th>
<th>SDM CI</th>
<th>WMD CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Levodopa</td>
<td>UPDRS</td>
<td>42</td>
<td>1</td>
<td>311</td>
<td></td>
<td></td>
<td>0.93</td>
<td>(0.65-1.20)***</td>
<td>7.01</td>
<td>(5.00-9.01)***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Striatal CIT uptake change (%)</td>
<td>40</td>
<td>1</td>
<td>116</td>
<td></td>
<td></td>
<td>-0.44</td>
<td>(-0.01 to 0.88)</td>
<td>-4.22</td>
<td>(-0.10 to 8.34)</td>
<td></td>
</tr>
<tr>
<td>Breast cancer</td>
<td>PCT (pg 1692 fig. 2)</td>
<td>Mort. (age &lt;50)</td>
<td>15.0y</td>
<td>60</td>
<td>28764</td>
<td>42.4</td>
<td>32.4</td>
<td>10% (ni)**</td>
<td>10 (ni)</td>
<td>24% (ni)**</td>
<td>0.24</td>
<td>(0.21-0.26)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mort. (age 50-69)</td>
<td>15.0y</td>
<td>50.4</td>
<td>47.4</td>
<td>3% (ni)**</td>
<td>33 (ni)</td>
<td></td>
<td>6% (ni)**</td>
<td>0.07</td>
<td>(0.04-0.09)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tamox. (pg 1704 fig.8)</td>
<td>Mort.</td>
<td>15.0y</td>
<td>12</td>
<td>10386</td>
<td>34.8</td>
<td>25.6</td>
<td>9% (ni)**</td>
<td>11 (ni)</td>
<td>26% (ni)**</td>
<td>0.24</td>
<td>(0.20-0.29)</td>
</tr>
<tr>
<td>Non-small cell lung cancer</td>
<td>adjuvant CT</td>
<td>Mortality</td>
<td>4.5y</td>
<td>21</td>
<td>7408</td>
<td></td>
<td></td>
<td></td>
<td>3% (3-4y)</td>
<td>30 (28-32)</td>
<td>9% (3-15y)</td>
<td>be</td>
</tr>
<tr>
<td>Antibiotics for various diseases</td>
<td>Rhinosinusitis</td>
<td>Cure</td>
<td>11.8d</td>
<td>10</td>
<td>2785</td>
<td>56.6</td>
<td>63.9</td>
<td>7% (ni)</td>
<td>15 (T=7; H=1090y)</td>
<td>13% (6-20)***</td>
<td>0.17</td>
<td>(0.08-0.25)</td>
</tr>
<tr>
<td></td>
<td>Otitis media</td>
<td>With pain</td>
<td>2-7d</td>
<td>10</td>
<td>2791</td>
<td>22.2</td>
<td>16</td>
<td>6.2% (ni)</td>
<td>17 (ni)</td>
<td>28% (26-30)**</td>
<td>0.22</td>
<td>(0.12-0.33)</td>
</tr>
<tr>
<td></td>
<td>Cystitis</td>
<td>Cure</td>
<td>3-17d</td>
<td>4</td>
<td>1062</td>
<td>25.7</td>
<td>61.8</td>
<td>36.1% (ni)</td>
<td>3 (ni)</td>
<td>13% (74-195)**</td>
<td>0.85</td>
<td>(0.71-0.99)</td>
</tr>
<tr>
<td></td>
<td>Colorectal surgery</td>
<td>Wound infection</td>
<td>ni</td>
<td>10</td>
<td>813</td>
<td>38.6</td>
<td>10.2</td>
<td>28.4% (23-34)**</td>
<td>4 (3-4)</td>
<td>70% (59-78)**</td>
<td>0.94</td>
<td>(0.73-1.15)</td>
</tr>
</tbody>
</table>

Ref = reference, AMSTAR = AMSTAR quality score (range of possible values 0-11), N = number of studies, n = number of participants, % PBO = percentage of patients with the outcome in the placebo group, % Drug = percentage of patients with the outcome in the drug group, ARD = absolute response or risk difference, CI = 95% confidence interval, NNT/H = number needed to treat or number needed to harm, RRR/RR = ‘Negative outcomes’ (mortality, relapse, exacerbation, hospitalization, dropout etc) are presented as percentage relative risk reductions (RRR), while positive outcomes (response to treatment, improvement, remission) are presented as percentage response ratios (RR). Positive values mean superiority of drug, SDM = standardized difference of means, WMD = weighted mean difference, *** = p<0.001, ** = p<0.01, * = p<0.05 or statistically significant but p-value not indicated, ne= not estimable, Antihyp. = antihypertensive drug, ACE = angiotensin-converting enzyme, ARBs = Angiotensin receptor blockers, CV events = cardiovascular events, vasc. ev. = vascular events, mort. = mortality, maj. = major, ASA = acetylsalicylic acid, prim.prev. = primary prevention, sec.prev. = secondary prevention, mod.-term = moderate term, no. tend. joints = number of tender joints, DO = dropout, ACR20 response = 20% improvement in American College of Rheumatology criteria, ifx=infliximab, (+ Mtx) = added to methotrexat, Mig. freq. = migraine frequency, FEV1 (l) = forced expiratory volume in one second and in liters, Leukot. Antg. = leukotriene-antagonists, Anti-IgE = Anti-IgE antibodies, Short-act β-2-ag. = short acting β-2 agonists, long-act β-2-ag. = long-acting β-2 agonists, Mort.(vs conv. treat) = mortality versus conventional treatment, Exacerb/p/y = Exacerbation per patient and year, HbA1c = glycated haemoglobin, α-gluc.-inhib. = alpha glucosidase inhibitors, GLP-1 analogues = glucagon-like peptide analogues, DPP4 inhibitors = dipeptidyl peptidase 4 inhibitors, Virol. resp. = virological response, Morb. and Mort. = morbidity and mortality, PPI = proton pump inhibitors, maint. = maintenance, 5-ASA = 5 aminosalicylic acid, clin. = clinical, maint. remission = maintenance of remission, UPDRS = Unified Parkinson's Disease Rating Scale, PCT = adjuvant polychemotherapy, Tamox. = tamoxifen

The following footmarks correspond to systematic reviews which were of similar quality and were similarly up-to-date as the included ones and yielded comparable results: 1) Wald et al. 2009; 2) BLTTC 2008. This update was based on even more participants, but presented results only based on the subgroups of younger and older participants separately; 3) Lindholm et al. 2005, 4) Berger et al. 2006 (presented results on men and women separately); 5) O’Regan et al. 2008, Ward et al. 2007, Cheung et al. 2004, 6) Saha et al. 2007; 7) Lechat et al. 1998;
Shibata et al. 200177, Bouzamondo et al. 200178, Heidenreich et al. 199779, Krum et al. 2005 (in ACE-inhibitor naïve patients30); 8) Weinblatt et al. 199931; 9) Ferrari et al. 2001 and 200282,83; 10) Ducharme et al. 200484; 11) Holgate et al. 200185, Bousquet et al. 200486; 12) Drummond et al. 200887 and Gartlehner et al. 200688; 13) Rodrigo et al. 200789; 14) Miller et al. 200090; 15) Filippini et al. 200391; 16) a more recent article the same group92 focussed on oestrogen-receptor-poor breast cancer and found polychemotherapy to be effective, as well, while tamoxifen was not; 17) Vouloumanolou et al. 200993 and Rovers et al. 200694.
Table DS4 Systematic reviews on the efficacy of psychiatric drugs versus placebo (full version of Table 2 in the print version with more drugs and meta-analyses, confidence intervals and NNTs)

<table>
<thead>
<tr>
<th>Study (Ref)</th>
<th>Therapy</th>
<th>Outcome</th>
<th>Mean weeks</th>
<th>N</th>
<th>n</th>
<th>% PBO</th>
<th>% Drug</th>
<th>ARD (CI)</th>
<th>NNT/H (CI)</th>
<th>RRR/RR (CI)</th>
<th>SDM (CI)</th>
<th>WMD (CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<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><strong>Schizophrenia – 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>
<td></td>
</tr>
<tr>
<td>95/10</td>
<td>SGAs</td>
<td>Response</td>
<td>9</td>
<td>28</td>
<td>4498</td>
<td>23.7</td>
<td>40.6</td>
<td>18% (14-22)***</td>
<td>6 (5-7)</td>
<td>70% (50-90)***</td>
<td>0.43 (0.36-0.51)</td>
<td></td>
</tr>
<tr>
<td></td>
<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>95/10</td>
<td>Haloperidol1)</td>
<td>Response</td>
<td>6</td>
<td>10</td>
<td>1440</td>
<td>19.5</td>
<td>29.3</td>
<td>12% (7-17)***</td>
<td>9 (6-15)</td>
<td>60% (30-90)***</td>
<td>0.30 (0.16-0.43)</td>
<td></td>
</tr>
<tr>
<td></td>
<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><strong>Schizophrenia – 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>
<td></td>
</tr>
<tr>
<td>96/10</td>
<td>Antipsychotics</td>
<td>Relapse</td>
<td>42</td>
<td>62</td>
<td>6392</td>
<td>57.0</td>
<td>22.0</td>
<td>38% (33-43)***</td>
<td>3 (2-3)</td>
<td>65% (59-69)***</td>
<td>0.92 (0.86-0.97)</td>
<td></td>
</tr>
<tr>
<td><strong>Bipolar – 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>
<td></td>
</tr>
<tr>
<td>97/6</td>
<td>Lithium</td>
<td>Response</td>
<td>3</td>
<td>6</td>
<td>811</td>
<td>34.0</td>
<td>52.0</td>
<td>17% (8-27)***</td>
<td>6 (4-13)</td>
<td>50% (20-100)***</td>
<td>0.41 (0.25-0.57)</td>
<td></td>
</tr>
<tr>
<td></td>
<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>98/92)5)</td>
<td>Valproate</td>
<td>Response</td>
<td>3</td>
<td>2</td>
<td>182</td>
<td>21.1</td>
<td>47.1</td>
<td>27% (14-40)***</td>
<td>4 (3-7)</td>
<td>150% (10-490)***</td>
<td>0.66 (0.30-1.02)</td>
<td></td>
</tr>
<tr>
<td></td>
<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>98/92)5)</td>
<td>Carbamazepine</td>
<td>Response</td>
<td>3</td>
<td>2</td>
<td>443</td>
<td>25.5</td>
<td>51.1</td>
<td>25% (12-38)***</td>
<td>4 (3-8)</td>
<td>100% (60-160)***</td>
<td>0.61 (0.39-0.83)</td>
<td></td>
</tr>
<tr>
<td></td>
<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>99/10</td>
<td>SGAs and haloperidol</td>
<td>Response</td>
<td>3</td>
<td>12</td>
<td>2939</td>
<td>30.8</td>
<td>49.9</td>
<td>20% (15-24)***</td>
<td>5 (4-7)</td>
<td>60% (50-80)***</td>
<td>0.44 (0.36-0.55)</td>
<td></td>
</tr>
<tr>
<td></td>
<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><strong>Bipolar disorder – 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>
<td></td>
</tr>
<tr>
<td>100/9</td>
<td>ADs</td>
<td>Response</td>
<td>7</td>
<td>4</td>
<td>662</td>
<td>34.1</td>
<td>57.7</td>
<td>34% (15-53)***</td>
<td>4 (2-7)</td>
<td>130% (30-280)**</td>
<td>0.53 (0.36-0.71)</td>
<td></td>
</tr>
<tr>
<td><strong>Bipolar disorder – maintenance therapy</strong></td>
<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>101/7)</td>
<td>Lithium</td>
<td>AR</td>
<td>n.i.</td>
<td>9</td>
<td>421</td>
<td>81.4</td>
<td>36.2</td>
<td>53% (n.i.)***</td>
<td>2 (n.i.)</td>
<td>51% (n.i.)***</td>
<td>1.12 (0.88-1.37)</td>
<td></td>
</tr>
<tr>
<td>102/8)5)</td>
<td>Lithium</td>
<td>AR</td>
<td>73</td>
<td>5</td>
<td>770</td>
<td>61.0</td>
<td>40.0</td>
<td>24% (8-39)***</td>
<td>5 (3-13)</td>
<td>35% (16-50)***</td>
<td>0.47 (0.31-0.63)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MR</td>
<td>73</td>
<td>4</td>
<td>565</td>
<td>23.6</td>
<td>13.8</td>
<td>10% (1-18)*</td>
<td>10 (5-100)</td>
<td>38% (5-60)*</td>
<td>0.36 (0.12-0.60)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>103/105)</td>
<td>Valproate</td>
<td>AR</td>
<td>52</td>
<td>1</td>
<td>281</td>
<td>38.3</td>
<td>24.1</td>
<td>14% (3-26)*</td>
<td>7 (4-33)</td>
<td>37% (10-56)*</td>
<td>0.37 (0.09-0.65)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MR</td>
<td>52</td>
<td>1</td>
<td>281</td>
<td>22.3</td>
<td>17.6</td>
<td>5% (5-15)*</td>
<td>20 (10-27)</td>
<td>21% (29-51)*</td>
<td>0.16 (0.06-0.49)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major depressive disorder – acute episode</td>
<td>DR</td>
<td>52</td>
<td>1</td>
<td>281</td>
<td>16.0</td>
<td>6.4</td>
<td>10% (1-18)*</td>
<td>10 (6-100)</td>
<td>60% (18-80)*</td>
<td>0.56 (0.12-1.01)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------------------</td>
<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>10/8/4</td>
<td>Paroxetine</td>
<td>Response</td>
<td>7.5</td>
<td>22</td>
<td>5112</td>
<td>42.4</td>
<td>53.2</td>
<td>10% (7-13)**</td>
<td>10 (8-14)</td>
<td>20% (20-30)**</td>
<td>0.24 (0.18-0.30)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>HAM-D</td>
<td>7.5</td>
<td>34</td>
<td>5764</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.31 (0.22-0.40)**</td>
<td>2.62 (2.00-3.25)**</td>
</tr>
<tr>
<td>10/4</td>
<td>TCAs (new AD studies)</td>
<td>Response</td>
<td>6</td>
<td>32</td>
<td>4314</td>
<td>31.0</td>
<td>46.0</td>
<td>15% (11-17)**</td>
<td>7 (5-8)</td>
<td>50% (n.i.)***</td>
<td>0.35 (0.28-0.42)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>HAM-D</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.33 (0.27-0.39)**</td>
<td>2.65 (2.17-3.13)**</td>
</tr>
<tr>
<td>10/7</td>
<td>Fluoxetine</td>
<td>Response</td>
<td>6</td>
<td>16</td>
<td>2761</td>
<td>24.2</td>
<td>37.8</td>
<td>13.6% (n.i.)***</td>
<td>7 (n.i.)</td>
<td>65% (44-85)**</td>
<td>0.35 (0.26-0.45)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>HAM-D</td>
<td>6</td>
<td>7</td>
<td>n.i.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.30 (0.21-0.39).</td>
<td>n.i.</td>
</tr>
<tr>
<td>10/7/7</td>
<td>New ADs</td>
<td>HAM-D</td>
<td>6</td>
<td>35</td>
<td>5133</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.32 (0.25-0.40)**</td>
<td>1.80 (n.i.)***</td>
</tr>
<tr>
<td>10/11</td>
<td>TCAs (low-dose)</td>
<td>Response</td>
<td>4</td>
<td>22</td>
<td>1119</td>
<td>29.6</td>
<td>46.5</td>
<td>27% (17-37)**</td>
<td>4 (3-6)</td>
<td>64% (35-98)**</td>
<td>0.40 (0.26-0.54)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Severity</td>
<td>4</td>
<td>16</td>
<td>861</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.40 (0.21-0.59).</td>
<td>n.e.</td>
</tr>
<tr>
<td>Major depressive disorder – maintenance treatment</td>
<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>10/9</td>
<td>ADs</td>
<td>Relapse</td>
<td>63</td>
<td>35</td>
<td>5032</td>
<td>41.0</td>
<td>18.0</td>
<td>23% (n.i.)u</td>
<td>4 (n.i.)</td>
<td>58% (49-68)**</td>
<td>0.64 (0.56-0.71)</td>
<td></td>
</tr>
<tr>
<td>10/10</td>
<td>New ADs</td>
<td>Recurrence</td>
<td>68</td>
<td>11</td>
<td>3326</td>
<td>48.0</td>
<td>26.0</td>
<td>22% (n.i.)u</td>
<td>5 (4-6)</td>
<td>44% (34-52)**</td>
<td>0.53 (0.45-0.61)</td>
<td></td>
</tr>
<tr>
<td>10/7</td>
<td>Lithium</td>
<td>Relapse (UpD)</td>
<td>n.i.</td>
<td>9</td>
<td>227</td>
<td>75.0</td>
<td>36.0</td>
<td>39% (n.i.)u</td>
<td>3 (n.i.)</td>
<td>53% (63-79)**</td>
<td>0.92 (0.61-1.23)</td>
<td></td>
</tr>
<tr>
<td>Obsessive compulsive disorder</td>
<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>11/11</td>
<td>SSRIs</td>
<td>YBOCS</td>
<td>10</td>
<td>17</td>
<td>3097</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.44 (0.36-0.52)**</td>
<td>3.21 (2.57-3.84)**</td>
</tr>
<tr>
<td></td>
<td>Response</td>
<td>10</td>
<td>13</td>
<td>2709</td>
<td>22.6</td>
<td>43.3</td>
<td>20% (17-24)**</td>
<td>5 (4-6)</td>
<td>84% (56-117)**</td>
<td>0.53 (0.44-0.62)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11/3/3</td>
<td>Clomipramine</td>
<td>YBOCS</td>
<td>8 (13)</td>
<td>7</td>
<td>808</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.48 (0.34-0.62)**</td>
<td>8.19 (5.85-10.53)**</td>
</tr>
<tr>
<td></td>
<td>Various SSRIs</td>
<td>YBOCS</td>
<td>8 (13)</td>
<td>18</td>
<td>1794</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.31 (0.21-0.41)**</td>
<td>1.85 (1.27-2.43)**</td>
</tr>
<tr>
<td>Panic disorder</td>
<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>11/4/4</td>
<td>TCAs</td>
<td>Anxiety</td>
<td>8</td>
<td>23</td>
<td>n.i.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.41 (n.i.)*</td>
<td>n.i.</td>
</tr>
<tr>
<td>11/4/10</td>
<td>SSRIs</td>
<td>Anxiety</td>
<td>8</td>
<td>17</td>
<td>n.i.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.41 (n.i.)*</td>
<td>n.i.</td>
</tr>
<tr>
<td>11/7</td>
<td>Benzodiazepines</td>
<td>Anxiety</td>
<td>8</td>
<td>25</td>
<td>n.i.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.40 (n.i.)*</td>
<td>n.i.</td>
</tr>
<tr>
<td>Dementia</td>
<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>11/8/11</td>
<td>ChE inhibitors</td>
<td>ADAS-cog</td>
<td>26</td>
<td>10</td>
<td>4236</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.41 (0.30-0.51)**</td>
<td>2.38 (1.79-2.97)**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MMSE</td>
<td>28</td>
<td>9</td>
<td>3118</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.39 (0.21-0.57)**</td>
<td>1.33 (0.73-1.92)**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UoI</td>
<td>26</td>
<td>8</td>
<td>3402</td>
<td>16.8</td>
<td>24.4</td>
<td>7% (3-11)**</td>
<td>14 (9-33)</td>
<td>43% (18-73)**</td>
<td>0.26 (0.17-0.35)</td>
<td></td>
</tr>
<tr>
<td>Attention-deficit/hyperactivity disorder</td>
<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>11/7/9</td>
<td>Methylphenidate</td>
<td>Hyperactivity</td>
<td>3.3 (0.5-28)</td>
<td>22</td>
<td>963</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.78 (0.64-0.91)**</td>
<td>n.i.</td>
</tr>
<tr>
<td>11/7/5</td>
<td>Amphetamine</td>
<td>ADHD symptoms</td>
<td>6</td>
<td>6</td>
<td>384</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00 (0.91-1.10)***</td>
<td>n.i.</td>
</tr>
</tbody>
</table>
Ref = reference, AMSTAR = AMSTAR quality score (range of possible values 0-11), N = number of studies, n = number of participants, % PBO = percentage of patients with the outcome in the placebo group, % Drug = percentage of patients with the outcome in the drug group, ARD = absolute response or risk difference, CI = 95% confidence interval, NNT/H = number needed to treat or number needed to harm, RRR/RR: ‘Negative outcomes’ (mortality, relapse, exacerbation, hospitalization, dropout etc) are presented as relative risk reductions (RRR), while positive outcomes (response to treatment, improvement, remission) are presented as percentage response ratios (RR). Positive values mean superiority of drug, SDM = standardized difference of means, WMD = weighted mean difference, *** = p<0.001, ** = p<0.01, * = p<0.05 or statistically significant but p-value not indicated, Ns= not significant, ni = not indicated, SGA = second generation antipsychotics, PANSS/BPRS = total score of either the Positive and Negative Syndrome Scale or the Brief Psychiatric Rating Scale, YMRS/MRS/MS = Young Mania Rating Scale or Mania Rating Scale or Mania Scale, ADs = Antidepressants, AR = Any relapse, DR = Depressive relapse, MR = Manic relapse, H = harm, T = treat, HAM-D = Hamilton Rating Scale for Depression, YBOCS = Yale-Brown Obsessive Compulsive Scale, MMSE = Mini Mental State Examination, SSRIs = selective serotonin reuptake inhibitors, TCAs = Tricyclic Antidepressants, new AD studies = here TCAs were active comparators in studies comparing new antidepressants with placebo, SleepOL = sleep onset latency, d = days, min = minutes, UoI = Unchanged or improved, UpD = Unipolar depression, ChE = Cholinesterase, Benzod. = Benzodiazepines, ADAS-cog = change from baseline of the Alzheimer’s Disease Assessment Scale - cognitive subscale; 1) in studies on SGAs, 2) Only studies based on the YMRS 3) updated and supplemented by our own searches, 4) consistent with Burgess et al. 2001118 and Beynon et al. 2008119, 5) consistent with Beynon et al. 2008119, 6) consistent with Katzman et al. 2007120, 7) consistent with Turner et al. 2008121, 8) all studies published after 1989, 9) consistent with Gould et al. 1995122 and Australian and New Zealand College 2003123, 10) consistent with Otto et al. 2001124, 11) consistent with Raina et al. 2008125 and Hansen 2008126, 12) after correspondence with the primary author a clear outlier with extremely positive results was excluded, 13) teacher rated, parent rated result was 0.54 (0.40-0.67), 14) consistent with the most recent meta-analysis which was poorly reported 127
Potentially relevant SRs identified and screened for retrieval:
- Medline (N = 503)
- Cochrane Hypertension Group (N=17)

Reports excluded by title or abstract
- Medline (N=477)
- Cochrane (N=16)

Reports retrieved for more detailed evaluation (N=27)

Reports excluded, with reasons (N =23)
- subgroup only (N=5)
- wrong intervention (N=3)
- wrong diagnosis (N=1)
- no meta-analysis (N=1)
- out of date (N=6)
- Cochrane preferred (N=4)
- pre-post meta-analysis (N=3)

Reports included (N =4)
Potentially relevant reports identified and screened for retrieval:
- Medline (N = 452)
- Cochrane Stroke Group (N=106)

Reports excluded by title or abstract
- Medline (N= 422)
- Cochrane (N= 103)

Reports retrieved for more detailed evaluation (N=33)

Reports excluded, with reasons (N =30):
- wrong diagnosis (N=13)
- no meta-analysis (N=2)
- wrong intervention (N=10)
- out of date (N=4)
- Cochrane preferred (N=1)

Included (N =3 )
Potentially relevant reports identified and screened for retrieval:
- Medline (N = 912)
- Cochrane metabolic and endocrine disorders Group, Cochrane stroke Group, Cochrane heart Group (N=277)

Reports excluded by title or abstract
- Medline (N= 876)
- Cochrane (N= 277)

Reports retrieved for more detailed evaluation (N=36)

Reports excluded, with reasons (N =34)
- out of date (N=11)
- wrong intervention (N=3)
- subgroup only (N=6)
- wrong diagnosis (N=8)
- no placebo-group (N=1)
- data more poorly reported than in other report (N=3)
- selective outcome (N=1)

Reports included (N =3)
Fig. DS4: PRISMA diagram - heart failure

(MEDLINE search term: "Heart failure"[Mesh] AND "Meta-Analysis"[Publication Type])

- Potentially relevant Reports identified and screened for retrieval:
  - Medline (N = 204)
  - Cochrane Heart Group (N = 71)

- Reports excluded by title or abstract
  - Medline (N = 180)
  - Cochrane (N = 71)

- Reports retrieved for more detailed evaluation (N = 24)

- Reports excluded, with reasons (N = 18)
  - subgroup only (N = 2)
  - wrong intervention (N = 8)
  - out of date (N = 5)
  - Wrong outcome (N = 3)

- Reports included (N = 6)
Fig. DS5: PRISMA diagram - rheumatoid arthritis

(MEDLINE search term: "Arthritis, rheumatoid"[Mesh] AND "Meta-Analysis"[Publication Type])
Fig. DS6: PRISMA diagram - migraine

(MEDLINE search term: „Migraine Disorders"[Mesh] AND "Meta-Analysis"[Publication Type])

Potentially relevant Reports identified and screened for retrieval:
- Medline (N=85)
- Cochrane Pain Group (N=90)

Reports excluded by title or abstract
- Medline (N=71)
- Cochrane (N=88)

Reports retrieved for more detailed evaluation (N=16)

Reports excluded, with reasons (N=12)
- out of date (N=1)
- Cochrane preferred (N=2)
- wrong question (placebo-response, N=2)
- no meta-analysis (N=1)
- single anti-migraine drug only (N=6)

Reports included (N=4)
Potentially relevant reports identified and screened for retrieval:
- Medline (N = 293)
- Cochrane Airways Group (N=210)

Reports excluded by title or abstract
- Medline (N= 277)
- Cochrane (N= 206)

Reports retrieved for more detailed evaluation (N=21)

Reports excluded, with reasons (N =15)
- subgroup only (N=4)
- wrong intervention (N=6)
- wrong diagnosis (N=2)
- no meta-analysis (N=2)
- out of date (N=1)

Reports included (N =6)
Potentially relevant reports identified and screened for retrieval:
- Medline \( (N = 127) \)
- Cochrane Airways Group \( (N=210) \)

Reports excluded by title or abstract
- Medline \( (N=109) \)
- Cochrane \( (N= 206) \)

Reports retrieved for more detailed evaluation \( (N=22) \)

Reports excluded, with reasons \( (N =16): \)
- subgroup only \( (N=2) \)
- wrong intervention \( (N=5) \)
- wrong diagnosis \( (N=1) \)
- Cochrane preferred \( (N=6) \)
- double publication \( (N=1) \)
- out of date \( (N=1) \)

Included reports \( (N =6) \)
Fig. DS9: PRISMA diagram - diabetes mellitus

(MEDLINE search term: „Diabetes mellitus"[Mesh] AND "Meta-Analysis"[Publication Type])

Potentially relevant reports identified and screened for retrieval:
- Medline (N = 266)
- Cochrane Metabolic and Endocrine Disorders Group (N=60)

Reports excluded by title or abstract
- Medline (N = 257)
- Cochrane (N = 57)

Reports retrieved for more detailed evaluation (N = 12)

Reports excluded, with reasons (N = 8)
- wrong question (N = 1)
- wrong intervention (N = 5)
- out of date (N = 1)
- no placebo-group (N = 1)

Reports included (N = 4)
Potentially relevant reports identified and screened for retrieval:
- Medline (N =125)
- Cochrane Hepato-Biliary Group (N=86)

Reports excluded by title or abstract
- Medline (N= 114)
- Cochrane (N=83)

Reports retrieved for more detailed evaluation (N=14)

Reports excluded, with reasons (N =11)
- subgroup only (N=2)
- wrong question (N=2)
- out of date (N=5)
- same as Cochrane (N=1)
- wrong intervention (N=1)

Reports included (N =3)

Fig. DS10: PRISMA diagram - hepatitis C

(MEDLINE search term: „Hepatitis C"[Mesh] AND "Meta-Analysis "[Publication Type])
Potentially relevant reports identified and screened for retrieval:
- Medline (N =12)
- Cochrane (N=258)

Reports excluded by title or abstract
- Medline (N=9)
- Cochrane (N=258)

Report retrieved for more detailed evaluation (N=3)

Reports excluded, with reasons (N=1)
- subgroup only (N=1)

Reports included (N =2)
Potentially relevant reports identified and screened for retrieval:
- Medline (N = 58)
- Cochrane Inflammatory Bowel Disease Group (N=45)

Reports excluded by title or abstract
- Medline (N= 51)
- Cochrane (N=43)

Reports retrieved for more detailed evaluation (N=9)

Reports excluded, with reasons (N =7)
- out of date (N=3)
- wrong intervention (N=3)
- no meta-analysis (N=1)

Reports included (N =2)
Fig. DS13: PRISMA diagram - multiple sclerosis

(MEDLINE search term: "Multiple Sclerosis"[Mesh] AND "Meta-Analysis"[Publication Type])

Potentially relevant reports identified and screened for retrieval:
- Medline (N =77)
- Cochrane Multiple Sclerosis Group (N=21)

Reports excluded by title or abstract
- Medline (N= 74)
- Cochrane (N=19)

Reports retrieved for more detailed evaluation (N=5)

Reports excluded, with reasons (N =3)
- same as Cochrane (N=1)
- no meta-analysis (N=1)
- Cochrane preferred (N=1)

Reports included (N =2)
Potentially relevant reports identified and screened for retrieval:
- Medline (N =96 )
- Cochrane Movement Disorders Group (N=39)

Reports excluded by title or abstract
- Medline (N= 95)
- Cochrane (N=39)

Reports retrieved for more detailed evaluation (N=1)

Reports excluded, with reasons (N =1)
No useable data on dopa-agonists versus placebo

No reports included. Results of a pivotal RCT are mentioned in the text
Potentially relevant reports identified and screened for retrieval:
- Medline (N=527)
- Cochrane Breast Cancer Group (N=36)

Reports excluded by title or abstract
- Medline (N= 494)
- Cochrane (N= 35)

Reports retrieved for more detailed evaluation (N=34)

Reports excluded, with reasons (N =33)
- wrong intervention (N=14)
- no meta-analysis (N=6)
- out of date (N=11)
- not peer-reviewed (N=1)
- subgroup only (N=1)

Reports included (N =1)
Fig. DS16: PRISMA diagram - lung cancer

(MEDLINE search term: "Lung Neoplasms"[Mesh] AND "Meta-Analysis " [Publication Type])

Potentially relevant reports identified and screened for retrieval:
- Medline (N = 94)
- Cochrane Lung Cancer Group (N=24)

Reports excluded by title or abstract
- Medline (N = 84)
- Cochrane (N=23)

Reports retrieved for more detailed evaluation (N=11)

Reports excluded, with reasons (N =10)
- subgroup only (N=1)
- no meta-analysis (N=1)
- wrong intervention (N=2)
- out of date (N=4)
- poorly reported (N=1)
- no placebo (N=1)

Reports included (N =1)
Potentially relevant Reports identified and screened for retrieval:
- Medline (N = 505)
- Cochrane acute respiratory infectious group,
- Cochrane colorectal surgery group,
- Cochrane ear, nose and throat group,
- Cochrane infectious disease group (N=299)

Reports excluded by title or abstract:
- Medline (N=483)
- Cochrane (N=297)

Reports retrieved for more detailed evaluation (N=24)

Excluded Reports, with reasons (N=20)
- not vs placebo (N=6)
- out of date (N=3)
- Cochrane preferred (N=2)
- wrong diagnosis (N=8)
- no meta-analysis (N=1)

Reports included (N=4)

(MEDLINE search term: „Antibacterial agents”[Mesh] AND "Meta-Analysis "[Publication Type])
Potentially relevant reports identified and screened for retrieval:
- Medline (N = 449)
- Cochrane Schizophrenia Group (N = 127)

Reports excluded by title or abstract
- Medline (N = 444)
- Cochrane (N = 125)

Reports retrieved for more detailed evaluation (N = 7)

Reports excluded, with reasons (N = 5)
- Out of date (N = 3)
- No true meta-analysis (N = 1)
- Only on single antipsychotics (haloperidol or chlorpromazine), mainly old studies,
  - added to table on older meta-analyses (N = 2)

Reports included (N = 2)
Potentially relevant reports identified and screened for retrieval:
- Medline (N =129)
- Cochrane Depression, anxiety and neurosis Group (N=93)

Reports excluded by title or abstract
- Medline (N=118)
- Cochrane (N= 92)

Reports retrieved for more detailed evaluation (N=12)

Reports excluded, with reasons (N =5)
- wrong question (N=1)
- Wrong outcome (N=2)
- Less completely reported than other review (N=1)
- More specific report preferred (N=1)

Reports included (N =7)
Fig. DS20: PRISMA diagram – obsessive–compulsive disorder

(MEDLINE search term: „Obsessive-compulsive disorder”[Mesh] AND "Meta-Analysis"
[Publication Type])
Potentially relevant Reports identified and screened for retrieval:
- Medline (N = 535)
- Cochrane Depression, anxiety and neurosis Group (N = 93)

Reports excluded by title or abstract
- Medline (N = 494)
- Cochrane (N = 93)

Reports retrieved for more detailed evaluation (N = 40)

Reports excluded, with reasons (N = 32)
- subgroup only (N = 3)
- wrong intervention (N = 2)
- out of date (N = 9)
- no meta-analysis (N = 6)
- wrong question (N = 3)
- no placebo-group (N = 5)
- Specific question (N = 1)
- not peer reviewd (N = 1)
- wrong outcome (N = 1)
- other better reported (N = 1)

Reports included (N = 8)
Fig. DS22: PRISMA diagram - panic disorder

(MEDLINE search term: „Panic disorder”[Mesh] AND "Meta-Analysis "[Publication Type])

Potentially relevant reports identified and screened for retrieval:
- Medline (N =27)
- Cochrane Depression, anxiety and neurosis Group (N=93)

Reports excluded by title or abstract
- Medline (N=9)
- Cochrane (N=93)

Reports retrieved for more detailed evaluation (N=18)

Reports excluded, with reasons (N =17)
- out of date (N=4)
- no meta-analysis (N=7)
- pre-post meta-analysis (N=5)
- effect sizes not presented, only p-value (N=1)

Reports included (N =1)
Fig. DS23: PRISMA diagram - dementia

(MEDLINE search term: "Dementia"[Mesh] AND "Meta-Analysis "[Publication Type])

Potentially relevant reports identified and screened for retrieval:
- Medline (N =306)
- Cochrane Dementia Group (N=80)

Reports excluded by title or abstract
- Medline (N=301)
- Cochrane (N=79)

Report retrieved for more detailed evaluation (N=6)

Reports excluded, with reasons (N =5)
- wrong intervention (N=1)
- out of date (N=1)
- no meta-analysis (N=1)
- cochrane preferred (N=1)
- wrong question (N=1)

Reports included (N =1)
Potentially relevant reports identified and screened for retrieval:
- Medline (N = 131)
- Cochrane (N = 75)

Reports excluded by title or abstract
- Medline (N = 113)
- Cochrane (N = 75)

Report retrieved for more detailed evaluation (N = 18)

Report excluded, with reasons (N = 15)
- no meta-analysis (N = 2)
- subgroup only (N = 6)
- wrong intervention (N = 1)
- wrong diagnosis (N = 2)
- no placebo control (N = 2)
- incompletely reported (N = 1)
- out of date (N = 1)

Reports included (N = 3)
Fig. DS25 Systematic presentation of the effect sizes in Figure 1 labelled by ‘Disease – Drug – Outcome’. This figure presents the same results as Fig. 1 in the print version, but indicates exactly which result corresponds to which result in the text and Tables 1 and 2. To enable verification Tables 1 and 2 the bars are consistently labelled by ‘Disease – Drug – Outcome’. 
Fig. DS26  Summary of the percentage relative risk reductions/response ratios presented in Tables DS3 and DS4: dot plot. All relative risk reductions in Tables DS3 and DS4 are presented. Data on older meta-analyses from Table DS1 are not included. Effect sizes of general medicine medication are presented on the left-hand side as black dots (median 29, mean 56, 95% CI 29–84), psychiatric drugs on the right-hand side as red dots (median 61, mean 58, 95% CI 48–73).
Fig. DS27  Summary of the percentage relative risk reductions/response ratios presented in Tables DS3 and DS4: bar chart. This figure presents the same results as Fig. DS26, but indicates exactly which corresponds to which result in Tables DS3 and DS4.
Fig. DS28  **Summary of the percentage absolute risk/response differences presented in Tables DS3 and DS4: dot plot.** All absolute risk/response differences in Tables DS3 and DS4 are presented. Data on older meta-analyses from Table DS1 are not included. Effect sizes of general medicine medication are presented on the left-hand side as black dots (median 5, mean 10.1, 95% CI 7.2–12.9), psychiatric drugs on the right-hand side as red dots (median 20.0, mean 20.8, 95% CI 16.0–25.5). One outlier – interferon for hepatitis B (relative risk reduction 1070%) – could not be presented for graphical reasons.
Fig. DS29  Summary of the percentage absolute risk/response differences presented in Tables DS3 and DS4: bar chart. This figure presents the same results as Fig. DS28, but it indicates exactly which corresponds to which result in Tables DS3 and DS4.
Explanation of statistical indices

The following text explains the parameters and indices that are presented in Tables DS1 and DS4 summarising the results. To understand their meaning is important for the interpretation of meta-analytic results. Table DS3 presents the formula.

**Statistical significance**

Statistical significance means that a result is unlikely to have occurred by chance. For example, p=0.03 means that there is only a 3% probability that the null-hypothesis (no difference between groups) has been wrongly rejected. If a result is not statistically significant it may be due to chance alone, but it does not tell about the magnitude of the difference or clinical importance. This magnitude of the difference is addressed by effect size.

**Effect size**

1) continuous data

The simplest effect size is the **Difference of the means (DM)** which used the raw units. For example, 75kg mean bodyweight at study end in drug and 70kg mean bodyweight in placebo, DM = 5kg.

The **Standardised difference of means (SDM)** is DM divided by the pooled standard deviation of both groups. This measure thus expresses the Difference in means in standard deviation units. The general formula is (mean group A – mean group B)/pooled standard deviation. This formula is sometimes slightly modified to account for specific situations (Cohen’s D, Hedges’s g etc). SDM is useful in two situations: when in the single studies of a meta-analysis different instruments are used to measure the same concept (e.g. two schizophrenia scales). For example, a 10 point difference in the Positive and Negative Syndrome Scale (PANSS) is not equivalent to a 10 point difference in the Brief Psychiatric Rating Scale (BPRS), because the PANSS has 30 items and its total score goes from 30 to 210, while the BPRS has 18 items and goes from 18 to 126. The other situation is when an outcome unit is not intuitive for the reader. For example, general practitioners will not know whether a 5 points difference in PANSS total score is a large or small difference. In this situation SDM’s might be easier to interpret. According to Cohen (8) a SDM of 0.2 is a small, 0.5 a medium and 0.8 a large effect size, but Cohen described this as a rule of thumb only and the interpretation depends on the context.

2) dichotomous data

Dichotomous (binary) outcomes can be classified as “yes or no”, such as death, relapse or remission. We presented the **percentage patients with an outcome in the drug and the placebo groups**. The knowledge of these percentages is crucial for the interpretation of the effect sizes presented below, and the examples will illustrate this point.

The **absolute risk or response difference (ARD)** subtracts the percentage in the drug group from that of the placebo group, e.g. 3% mortality in placebo and 1% mortality in drug, thus 3% - 1% = 2% ARD. This is the most straightforward effect size for dichotomous outcomes, but its use in meta-analyses can be problematic when the baseline risk in the different studies varies. The relative risk (reduction) partly accounts for such differences in baseline risk.

The **relative risk reduction (RRR)** divides the absolute risk reduction by the
percentage in the placebo group, thus 2%/3% = 67% (in decimals: 0.02/0.03 = 0.67). Positive outcomes such as response were presented as a percent response ratio (RR) in a similar fashion. For example, in mania, 50% responded to antipsychotics and 31% to placebo, thus 50%/31% = 1.61 times or 61% (RR) more responders.

The number-needed-to-treat (NNT) indicates how many patients must be treated with an intervention to avoid one bad outcome (e.g. death). It was calculated as the inverse of the absolute risk difference, in the example above 1/2% = 1/0.02 = 50. Thus one out of 50 treated patients will not die. There are ways to incorporate assumptions about the baseline risk in the calculation of NNT, but we always used the baseline risk in the trials for the calculation of NNT.

Importantly, ARD, RRR/RR, NNT are based on the same numbers, but a 67% relative risk reduction looks much more impressive than a 2% absolute risk difference or a NNT of 50. As the relative risk reduction is often larger than the absolute risk reduction, authors often prefer to present the former. On the other hand, the maximal absolute risk reduction can not exceed 3% (3% placebo – 0% drug = 3% ARD). Therefore, all these indices must be interpreted in the context of the percentage patients with an outcome in the drug and the placebo group.
References


(68) Wald DS, Law M, Morris JK, Bestwick JP, Wald NJ. Combination therapy


(96)  Leucht S, Tardy M, Komossa K, Heres S, Kissling W, Davis JM. Maintenance treatment with antipsychotic drugs for schizophrenia. Submitted for publication.


