Augmentation therapies for treatment-resistant depression: systematic review and meta-analysis†

Rebecca Strawbridge, Ben Carter, Lindsey Marwood, Borwin Bandelow, Dimosthenis Tsapekos, Viktoriya L. Nikolova, Rachael Taylor, Tim Mantingh, Valeria de Angel, Fiona Patrick, Anthony J. Cleare and Allan H. Young

Background
Depression is considered to have the highest disability burden of all conditions. Although treatment-resistant depression (TRD) is a key contributor to that burden, there is little understanding of the best treatment approaches for it and specifically the effectiveness of available augmentation approaches.

Aims
We conducted a systematic review and meta-analysis to search and quantify the evidence of psychological and pharmacological augmentation interventions for TRD.

Method
Participants with TRD (defined as insufficient response to at least two antidepressants) were randomised to at least one augmentation treatment in the trial. Pre-post analysis assessed treatment effectiveness, providing an effect size (ES) independent of comparator interventions.

Results
Of 28 trials, 3 investigated psychological treatments and 25 examined pharmacological interventions. Pre-post analyses demonstrated N-methyl-D-aspartate-targeting drugs to have the highest ES (ES = 1.48, 95% CI 1.25–1.71). Other than aripiprazole (four studies, ES = 1.32, 95% CI 1.23–1.44) and lithium (three studies, ES = 1.00, 95% CI 0.81–1.20), treatments were each investigated in less than three studies. Overall, pharmacological (ES = 1.19, 95% CI 1.08–1.30) and psychological (ES = 1.43, 95% CI 0.50–2.36) therapies yielded higher ESs than pill placebo (ES = 0.78, 95% CI 0.66–0.91) and psychological control (ES = 0.94, 95% CI 0.36–1.52).

Conclusions
Despite being used widely in clinical practice, the evidence for augmentation treatments in TRD is sparse. Although pre-post meta-analyses are limited by the absence of direct comparison, this work finds promising evidence across treatment modalities.

Declaration of interest
In the past 3 years, A.H.Y. received honoraria for speaking from AstraZeneca, Lundbeck, Eli Lilly and Sunovion; honoraria for consulting from Allergan, Livanova and Lundbeck, Sunovion and Janssen; and research grant support from Janssen. In the past 3 years, A.J.C. received honoraria for speaking from AstraZeneca and Lundbeck; honoraria for consulting with Allergan, Janssen, Livanova, Lundbeck and Sandoz; support for conference attendance from Janssen; and research grant support from Lundbeck. B.B. has recently been (soon to be) on the speakers/advisory board for Hekel, Lilly, Lundbeck, Mundipharma, Pfizer, and Servier. No other conflicts of interest.

Keywords
Treatment-resistant depression; augmentation; psychological therapy; psychopharmacology; systematic review.

Copyright and usage
© The Royal College of Psychiatrists 2018.

The burden of treatment-resistant depression (TRD) is challenging to quantify. It has eluded a universal definition but is prevalent and encompasses considerably greater severity, chronicity, recurrence, admission to hospital and comorbidity with both psychiatric and non-psychiatric disorders than non-resistant major depressive disorder (MDD). Despite this, TRD has been a neglected area of research with numerous reviews calling for more comprehensive evidence. Indeed, many of these reviews have considered people as treatment resistant if they have failed one previous treatment trial (in contrast with the most popular guidelines1), in part because this represents the inclusion criteria frequently used in clinical trials. One such example examined the pharmacological augmentation treatments that the majority of TRD patients are treated with in practice. Only when using the less-stringent criteria of TRD was there sufficient evidence for a network meta-analysis in 2015, and the authors reported significant efficacy of quetiapine, aripiprazole, lithium and thyroid hormone compared with placebo. However, this evidence may not apply to people with more severe TRD. Pre-post analyses have the benefit of not requiring a placebo arm and the ability to compare effectiveness estimates between heterogeneous treatment approaches. Additionally, pre-post effect sizes (ESs) provide good clinical face validity as an estimate of the magnitude of effects seen with treatment in practice, incorporating both those specific to the individual modality as well as non-specific effects and the passage of time.

Objectives
This review aimed to qualify and quantify the evidence of augmentation treatments for TRD by using the most common clinical definition (i.e. two or more failed treatments in current episode) and to compare ESs across psychological and pharmacological interventions. To our knowledge, this is the first meta-analysis comparing pre-post treatment effects for all augmentation therapies across the two most popular treatment classes for depression in clinical practice. Specifically, our objectives were to:

(a) determine the efficacy of adjunctive interventions for TRD through comparisons between treatment category (i.e. pharmacological or psychological), class (e.g. antipsychotics, mood stabilisers) and individual treatments;
Methods

Criteria for considering studies for the review

The protocol for this systematic review was published via PROSPERO (registration code CRD42018088009), where full details of the search are available and reported consistently following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) reporting guidelines.

Types of included studies

Only randomised controlled trials (RCTs) of at least 10 participants and at least one suitable augmentation treatment were included.

Types of participants

Participants must have been adults with TRD, defined as unremitted depression despite at least two courses of treatment of adequate dose and duration undertaken in the current episode (current best-practice guidelines). Both within-class (in addition to between-class) switching of antidepressants and psychological treatments were permitted as they are considered valid contributors to a TRD definition. Studies including patients with psychotic or bipolar depression were excluded because of clear treatment distinctions.

Types of interventions

Participants must have been taking at least one continuation treatment prior to randomisation to a new (augmentation) intervention. The same eligibility criteria were employed for both continuation and augmentation treatments: permitted pharmacological treatments were any included in the Maudsley Treatment Inventory and augmentation treatments: permitted pharmacological treatments and pharmacological comparator; another psychological agent, another psychological intervention, waiting list, active control or treatment as usual (TAU).

Types of outcome measures

Primary outcome: The primary outcome was clinical improvement (ES) between pre- and post-treatment time points for each eligible treatment/comparator arm. One efficacy measurement was assessed, prioritising validated, clinician-rated measures of depression severity (and if not available, a patient-rated depression scale or assessment of global improvement if no depression symptom scale was reported).

Secondary outcomes: A measure of adherence (e.g. trial dropout due to any cause or treatment adherence data) and a measure of tolerability (e.g. adverse event or side-effects data) were recorded where available.

Search methods for identification of studies

MEDLINE and Institute for Scientific Information (ISI) Web of Science were searched in addition to citation lists from notable papers, available reviews and included articles. The following medical subject headings or text word terms were used for the electronic database search (all fields): (depress* OR MDD OR major depress*) AND (resistan* OR refractor* OR non-respon* OR non-respon* OR un-respon* OR unresp* OR TRD OR fail* OR inadequate OR difficult OR intractable) AND (augment* OR adjunct* OR add-on OR comb* OR co-administ*) AND (randomi* OR RCT) AND (treatment OR intervention OR trial). No language restriction was made.

Data collection and analyses

Article review and data extraction

All search results were evaluated against inclusion criteria independently by pairs of review authors (R.S., L.M., R.T., T.M., V.d.A., D.T., V.L.N. and F.P.), with disparities addressed by consensus with additional review authors (A.H.Y., A.I.C., B.C.). Following inclusion, data extraction was conducted by authors as above.

Quality assessment

The methodological quality assessment was examined using the Scottish Intercollegiate Guidelines Network and the Cochrane Risk of Bias (RoB) tools. Studies were assessed by two reviewers (rated as RoB high, low or unclear) for nine domains: sequence generation, allocation concealment, blinding of outcome assessors, use of intention-to-treat analysis, comparability of randomised groups at baseline, inter-site differences in findings, the potential for selective outcome reporting and presence of for-profit bias (allegiance). Using individual criterion ratings, each study was given an overall RoB rating of low, moderate or high (see Supplementary Table 1 available at https://doi.org/10.1192/bjp.2018.233).

Measures of treatment effect

Continuous data describing treatment effectiveness were extracted (e.g. pre- and post-severity scores or longitudinal change in severity scores) and presented as a standardised mean difference (Hedges’ g ES). Using a random-effects model, meta-analyses computed a pooled ES with 95% confidence intervals (CIs), P-values and the I² statistic. Statistical heterogeneity was considered important if I² exceeded 60% and explored using subgroups. The following comparisons were planned to assess the primary outcome:

(a) pooled effects of augmentation intervention/comparator categories (i.e. psychological treatment, psychological comparator, pharmacological treatment and pharmacological comparator);
(b) pooled effects of augmenters by class (e.g. selective serotonin reuptake inhibitor, serotonin–noradrenaline reuptake inhibitor, antipsychotic, mood stabiliser);
(c) pooled effects of individual treatment interventions within above categories.

Additional comparisons

The following secondary outcomes were quantitatively or qualitatively explored: acceptability, tolerability and an exploration of pairwise active-control comparisons to provide an indicated effect of treatment versus comparator trial arm, validating findings against the current gold standard.

Subgroups used to explore heterogeneity

Planned subgroups used to explore statistical heterogeneity included study quality (RoB) and trial duration, as well as participant treatment-resistance definition, continuation treatments, comorbidities, depression severity, duration of episode and treatment setting.

Changes made since protocol registration

The permitted range of treatment duration was amended from a range of 6–26 weeks to include any duration where expectations exceed 60% and explored using subgroups. The following comparisons were planned to assess the primary outcome:
of clinical efficacy were reported. This was to account for the variable windows of clinical efficacy between different treatment mechanisms (e.g. ketamine, which has well-documented rapid antidepressant effects). Excluding ketamine, the Maudsley Treatment Inventory recommends durations of 6 weeks for full clinical effect; therefore we selected to subgroup included trials of less than 6 weeks as ‘short-term’ (this excludes rapid-onset treatments such as ketamine) and those more than 26 weeks as ‘long-term’ treatment durations.

Results

Systematic search results

After duplicates were removed, 2246 manuscripts from the MEDLINE and ISI Web of Science databases (all years to 6 February 2018) and hand searches were screened. Of 297 full texts reviewed, 39 articles describing 28 studies were eligible for inclusion. A PRISMA flow chart presents a breakdown of the search process (Fig. 1).

Characteristics of included studies

Within the 28 included RCTs, 5461 TRD participants were randomised. All analysed interventions were of parallel-group studies, with ten trials (36%) conducted in North America, seven (25%) in Europe, six (21%) in Asia, four (14%) across multiple continents and one (4%) in South America. The mean study size was 199 (s.d. = 270, range 20–1293). The duration of interventions ranged from 5 days (ketamine) to 18 months (long-term psychoanalytic psychotherapy), with a median duration of 6 weeks (interquartile range = 2).

Characteristics of participants

Participants had a median age of 45 years (interquartile range = 4) and 66% were female. All analysed individuals had unremitted depression despite at least two adequate treatment trials in the current episode. A total of 15 studies defined TRD fully retrospectively (using a minimum duration of previous treatments of 4 or 6 weeks), whereas 12 required at least one unsuccessful treatment retrospectively and one prospectively. One study undertook two treatment trials to determine treatment resistance fully prospectively. Most studies did not consider psychological treatments to contribute to TRD definition; only Fonagy et al required one pharmacological and one psychological treatment failure as a minimum TRD criterion for study entry. Table 1 contains further details.

Quality assessment

Supplementary Table 1 contains the RoB ratings across criteria and studies. A total of 12 studies were rated as having a low RoB, 2,11,14–23 12 had a moderate RoB and 4 had a high RoB. The most

---

Fig. 1 Study flow diagram showing the Preferred Reporting Items for Systematic Reviews and Meta-Analyses. *A further 11 articles were included in the systematic review as they contained additional information but were pooled analyses or subgroup analyses. ISI, Institute for Scientific Information; TRD, treatment-resistant depression.
Table 1  Details of all studies included in this review

<table>
<thead>
<tr>
<th>Reference</th>
<th>Intervention</th>
<th>Continent of trial</th>
<th>TRD definition used (number of unsuccessful treatments)</th>
<th>Setting</th>
<th>Comorbidity</th>
<th>Continuation treatment(s)</th>
<th>Continuation therapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mood stabilisers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nierenberg et al[29]</td>
<td>Lithium/placebo</td>
<td>North America</td>
<td>1+ adequate ADM (retrospective), plus 6 weeks prospective ADM</td>
<td>OP</td>
<td>Some</td>
<td>NOR</td>
<td>Mono</td>
</tr>
<tr>
<td>Girlanda et al[35]</td>
<td>Lithium/TAU</td>
<td>Europe</td>
<td>2+ adequate ADM</td>
<td>OP</td>
<td>Frequent psychiatric</td>
<td>TAU</td>
<td>Mono/Poly</td>
</tr>
<tr>
<td>Schindler et al[34]</td>
<td>Lithium/famotidine</td>
<td>Europe</td>
<td>2+ ADM &gt;6 weeks</td>
<td>IP</td>
<td>Some</td>
<td>NR</td>
<td>Mono</td>
</tr>
<tr>
<td>Barbee et al[32]</td>
<td>Lamotrigine/placebo</td>
<td>North America</td>
<td>1+ ADM &gt;6 weeks (retrospective), plus 8 weeks prospective ADM</td>
<td>NR</td>
<td>None prominent</td>
<td>PAR</td>
<td>Mono</td>
</tr>
<tr>
<td>Santos et al[33]</td>
<td>Lamotrigine/placebo</td>
<td>South America</td>
<td>2+ adequate ADM</td>
<td>OP</td>
<td>Some</td>
<td>Any ADM</td>
<td>Mono</td>
</tr>
<tr>
<td>Berman et al[34]</td>
<td>Aripiprazole/placebo</td>
<td>North America</td>
<td>1+ ADM &gt;6 weeks (retrospective), plus 8 weeks prospective ADM</td>
<td>OP</td>
<td>None prominent</td>
<td>ESC, FLU, PAR, SER, VEN</td>
<td>Mono</td>
</tr>
<tr>
<td>Marcus et al[32]</td>
<td>Aripiprazole/placebo</td>
<td>North America</td>
<td>1+ ADM &gt;6 weeks (retrospective), plus 8 weeks prospective ADM</td>
<td>OP</td>
<td>None prominent</td>
<td>ESC, FLU, PAR, SER, VEN</td>
<td>Mono</td>
</tr>
<tr>
<td>Ozaki et al[34]</td>
<td>Aripiprazole/placebo</td>
<td>Asia</td>
<td>1+ ADM &gt;6 weeks (retrospective), plus 8 weeks prospective ADM</td>
<td>OP and IP</td>
<td>None prominent</td>
<td>PAR, FLU, SER, MIL, DUL</td>
<td>Mono</td>
</tr>
<tr>
<td>Thase et al[33]</td>
<td>Brexipiprazole/placebo</td>
<td>North America/Europe</td>
<td>1+ ADM (retrospective), plus 1 prospective ADM</td>
<td>OP</td>
<td>None prominent</td>
<td>ESC, FLU, PAR, SER, DUL, VEN</td>
<td>Mono</td>
</tr>
<tr>
<td>Thase et al[33]</td>
<td>Brexipiprazole/placebo</td>
<td>North America/Europe</td>
<td>1+ ADM (retrospective), plus 1 prospective ADM</td>
<td>OP</td>
<td>None prominent</td>
<td>ESC, FLU, PAR, SER, DUL, VEN</td>
<td>Mono</td>
</tr>
<tr>
<td>Shelton et al[35]</td>
<td>Olanzapine/placebo</td>
<td>North America</td>
<td>2+ ADM &gt;4 weeks (retrospective), plus 6 weeks prospective ADM</td>
<td>OP</td>
<td>None prominent</td>
<td>FLU</td>
<td>Mono</td>
</tr>
<tr>
<td>Dunner et al[36]</td>
<td>Ziprasidone</td>
<td>North America</td>
<td>1+ ADM &gt;4 weeks (retrospective), plus 6 weeks prospective ADM</td>
<td>OP</td>
<td>None prominent</td>
<td>SER</td>
<td>Mono</td>
</tr>
<tr>
<td>Psychological therapies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eisendrath et al[37]</td>
<td>MBCT/HEP</td>
<td>North America</td>
<td>2+ ADM &gt;4 weeks</td>
<td>OP</td>
<td>Some</td>
<td>TAU (ADM)</td>
<td>Poly</td>
</tr>
<tr>
<td>Fonagy et al[35]</td>
<td>LTPP/TAU</td>
<td>Europe</td>
<td>2+ treatments (≥1 ADM, ≥1 psychological)</td>
<td>OP</td>
<td>Some</td>
<td>TAU</td>
<td>Mono/Poly</td>
</tr>
<tr>
<td>Hauksson et al[33]</td>
<td>CBT/TAU</td>
<td>Europe</td>
<td>2+ adequate ADM</td>
<td>IP</td>
<td>Frequent psychiatric</td>
<td>Multimodal rehabilitation</td>
<td>Mono/Poly</td>
</tr>
<tr>
<td>NMDA targets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heresco-Levy et al[34]</td>
<td>D-cycloserine/placebo</td>
<td>Asia</td>
<td>2+ adequate ADM</td>
<td>OP</td>
<td>None prominent</td>
<td>TAU (ADM)</td>
<td>Mono/Poly</td>
</tr>
<tr>
<td>Husain et al[37]</td>
<td>Minocycline/placebo</td>
<td>Asia</td>
<td>2+ adequate ADM</td>
<td>OP</td>
<td>None prominent</td>
<td>TAU</td>
<td>Mono/Poly</td>
</tr>
<tr>
<td>Su et al[37]</td>
<td>Ketamine/placebo</td>
<td>Asia</td>
<td>2+ adequate ADM</td>
<td>OP</td>
<td>None prominent</td>
<td>TAU</td>
<td>Mono/Poly</td>
</tr>
<tr>
<td>Other pharmacological</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moller et al[35]</td>
<td>Dexamfetamine/placebo</td>
<td>All except Oceania</td>
<td>1+ ADM &gt;6 weeks (retrospective), plus 8 weeks ADM prospective</td>
<td>OP</td>
<td>None prominent</td>
<td>ESC, CIT, FLU, PAR, SER, DUL, VEN</td>
<td>Mono</td>
</tr>
<tr>
<td>Multiple treatment classes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bauer et al[38]</td>
<td>Quetiapine/lithium</td>
<td>Europe</td>
<td>2+ ADM (retrospective)</td>
<td>OP and IP</td>
<td>None prominent</td>
<td>Any ADM</td>
<td>Mono</td>
</tr>
<tr>
<td>Fang et al[39]</td>
<td>Risperidone/sodium valproate/buspirone/trazodone/thyroid hormone</td>
<td>Asia</td>
<td>2+ adequate ADM (stage 2, Thase and Rush criteria)</td>
<td>OP and IP</td>
<td>Some</td>
<td>PAR</td>
<td>Mono</td>
</tr>
</tbody>
</table>

(Continued)
common individual criteria rated as high RoB were being funded and/or conducted by an industrial sponsor (12 trials) and not applying or reporting an intention-to-treat analysis (seven trials). Blinding was not always maintained but was often maximised where possible, i.e. in the ketamine trial (reportedly double-blind\(^{11}\)), psychological trials (two out of three trials report blinding of outcome assessors\(^{12,18}\) and open-label studies (all but one\(^{36}\) reporting blinded outcome raters).

**Effectiveness of augmentation treatment**

There was clinical diversity in the design (see Table 1); intervention and outcomes were reported (see Supplementary Table 2) across studies.

**Primary outcomes**

Pre-post meta-analyses indicated improvements in depression with all interventions examined \((P < 0.001)\). From 23 studies including 3246 participants, pharmacological treatments yielded an overall ES of 1.15 \((95\% CI 1.01–1.29, I^2 = 82.7)\). Psychological therapies as a category comprised 3 studies totalling 276 participants, showing similar effects \((ES = 1.43, 95\% CI 0.50–2.36, I^2 = 95.3)\). For the majority of initial analyses conducted, severe heterogeneity limited the interpretability of comparisons (see Supplementary Table 3). The three studies with a high RoB contributed substantially to this heterogeneity, demonstrating either low\(^{36}\) or high\(^{13,38}\) outlier ESs and the subgroup of active treatments trialled for a short-term duration \((lithium,^{24} metyrapone^{14})\) showed an ES of 0.61 \((95\% CI 0.37–0.85, I^2 = 0)\); their removal from meta-analyses notably reduced heterogeneity. In contrast, long-term treatment trials of lithium\(^{15}\) and psychoanalytic psychotherapy\(^{12}\) were homogeneous \((ES = 0.67, 95\% CI 0.44–0.90, I^2 = 4.6)\) and did not affect heterogeneity of main analyses so were not excluded. Effects of all placebo trials \((pill ES = 0.78, psychological ES = 0.94)\) exhibited findings similar to the sub-therapeutic-duration pharmacological studies \((ES = 0.61)\) and were consistently lower than active treatments; see Fig. 2 and Table 2. All active treatment effects are displayed in Supplementary Fig. 1 and control arms in Supplementary Fig. 2.

Pharmacological treatment classes

Pharmacological interventions without high RoB and trialled for a therapeutic duration had an ES of 1.19 \((95\% CI 1.08–1.30; I^2 = 64.6)\).

\(N\)-methyl-\(d\)-aspartate (NMDA)-targeting drugs showed the most consistent and large ES of the pharmacological classes \((ES = 1.48, 95\% CI 1.25–1.71, I^2 = 0)\), despite the individual agents having different mechanisms of action.

Mood stabilisers demonstrated an overall ES of 1.12 \((95\% CI 0.92–1.31, I^2 = 23.6)\), exhibiting low heterogeneity only. Lithium was the most frequently investigated mood stabiliser and had a slightly smaller ES than the overall class without heterogeneity (three studies; ES = 1.00, 95\% CI 0.81–1.20, \(I^2 = 0\)).

Antipsychotics also had an ES of 1.12 \((95\% CI 0.98–1.26, I^2 = 75.0)\) and exhibited heterogeneity, likely due to differences between treatments within this class. Aripiprazole was the most frequently assessed antipsychotic and provided a consistent effect across four studies \((ES = 1.33, 95\% CI 1.23–1.44, I^2 = 0)\).

Medications not falling into the above mechanisms were grouped together \((trazodone, buspirone, thyroid hormone and dexamfetamine)\) and showed an ES of 1.36 \((95\% CI 1.09–1.63, I^2 = 46.4)\), comparable in terms of heterogeneity and ES to the other pharmacological treatments.
Psychological treatment classes

The overall ES of psychological therapies (three studies; ES = 1.43, 95% CI 0.50–2.36) contained substantial heterogeneity ($I^2 = 95.3$), likely due to different therapeutic modalities that we were not able to subgroup further due to lack of studies. Within this analysis, cognitive–behavioural therapy (CBT) had the highest ES of all

<table>
<thead>
<tr>
<th>Treatment class</th>
<th>k</th>
<th>ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMDA-targeting agents</td>
<td>3</td>
<td>1.48</td>
</tr>
<tr>
<td>Pharmacological (other*)</td>
<td>4</td>
<td>1.36</td>
</tr>
<tr>
<td>Mood stabilisers</td>
<td>8</td>
<td>1.12</td>
</tr>
<tr>
<td>Antipsychotics</td>
<td>10</td>
<td>1.12</td>
</tr>
<tr>
<td>Psychological therapies</td>
<td>3</td>
<td>1.43</td>
</tr>
<tr>
<td>Pill placebo</td>
<td>16</td>
<td>0.78</td>
</tr>
<tr>
<td>Psychological placebo</td>
<td>3</td>
<td>0.94</td>
</tr>
<tr>
<td>Short-term treatments</td>
<td>2</td>
<td>0.61</td>
</tr>
</tbody>
</table>

**Table 2** Results of meta-analyses assessing primary outcome

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Intervention</th>
<th>k</th>
<th>n</th>
<th>ES</th>
<th>s.e.</th>
<th>95% CI</th>
<th>Heterogeneity ($I^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pharmacological</td>
<td>All</td>
<td>23</td>
<td>2965</td>
<td>1.19</td>
<td>0.05</td>
<td>1.08–1.30</td>
<td>64.6</td>
</tr>
<tr>
<td>Mood stabilisers</td>
<td>All</td>
<td>6</td>
<td>260</td>
<td>1.12</td>
<td>0.10</td>
<td>0.92–1.31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lithium</td>
<td>3</td>
<td>157</td>
<td>1.00</td>
<td>0.10</td>
<td>0.81–1.20</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>Lamotrigine</td>
<td>2</td>
<td>64</td>
<td>1.13</td>
<td>0.17</td>
<td>0.79–1.47</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>Sodium valproate</td>
<td>1</td>
<td>39</td>
<td>1.64</td>
<td>0.25</td>
<td>1.16–2.12</td>
<td>n/a</td>
</tr>
<tr>
<td>Antipsychotics</td>
<td>All</td>
<td>10</td>
<td>1503</td>
<td>1.12</td>
<td>0.07</td>
<td>0.98–1.26</td>
<td>75.0</td>
</tr>
<tr>
<td></td>
<td>Aripiprazole</td>
<td>4</td>
<td>655</td>
<td>1.33</td>
<td>0.05</td>
<td>1.23–1.44</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>Brexpiprazole</td>
<td>2</td>
<td>638</td>
<td>0.96</td>
<td>0.06</td>
<td>0.85–1.06</td>
<td>17.3</td>
</tr>
<tr>
<td></td>
<td>Ziprasidone</td>
<td>1</td>
<td>41</td>
<td>0.65</td>
<td>0.17</td>
<td>0.31–0.99</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>Risperidone</td>
<td>1</td>
<td>45</td>
<td>1.15</td>
<td>0.19</td>
<td>0.77–1.53</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>Quetiapine</td>
<td>1</td>
<td>114</td>
<td>1.05</td>
<td>0.12</td>
<td>0.82–1.28</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>Olanzapine</td>
<td>1</td>
<td>10</td>
<td>0.98</td>
<td>0.38</td>
<td>0.66–1.36</td>
<td>n/a</td>
</tr>
<tr>
<td>Other mechanisms</td>
<td>NMDA targets</td>
<td>3</td>
<td>81</td>
<td>1.48</td>
<td>0.12</td>
<td>1.25–1.71</td>
<td>0.0</td>
</tr>
<tr>
<td>SARIT</td>
<td>All</td>
<td>4</td>
<td>1121</td>
<td>1.36</td>
<td>0.14</td>
<td>1.09–1.63</td>
<td>46.4</td>
</tr>
<tr>
<td>Anxiolytic</td>
<td>All</td>
<td>4</td>
<td>17</td>
<td>1.47</td>
<td>0.13</td>
<td>1.22–1.72</td>
<td>n/a</td>
</tr>
<tr>
<td>Thyroid</td>
<td>All</td>
<td>4</td>
<td>17</td>
<td>1.47</td>
<td>0.46</td>
<td>0.57–2.36</td>
<td>n/a</td>
</tr>
<tr>
<td>Nicotinic</td>
<td>All</td>
<td>4</td>
<td>17</td>
<td>1.54</td>
<td>0.29</td>
<td>0.97–2.11</td>
<td>n/a</td>
</tr>
<tr>
<td>Pill placebo</td>
<td>SARIT</td>
<td>16</td>
<td>1552</td>
<td>0.78</td>
<td>0.07</td>
<td>0.66–0.91</td>
<td>68.8</td>
</tr>
<tr>
<td>Psychological therapies</td>
<td>All</td>
<td>3</td>
<td>276</td>
<td>1.43</td>
<td>0.23</td>
<td>0.50–2.36</td>
<td>95.3</td>
</tr>
<tr>
<td>Psychological comparators</td>
<td>All</td>
<td>3</td>
<td>162</td>
<td>0.94</td>
<td>0.30</td>
<td>0.36–1.52</td>
<td>89.1</td>
</tr>
<tr>
<td>Short-term treatments (excluded from main analyses)</td>
<td>All</td>
<td>2</td>
<td>79</td>
<td>0.61</td>
<td>0.12</td>
<td>0.37–0.85</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Results of meta-analyses assessing treatment effectiveness at a category, class and individual intervention level for studies with data available for meta-analyses without a high risk of bias accounting for therapeutic duration of interventions trialled (for results including high risk of bias and short-term durations, see Supplementary Table 3). Bold text indicates pooled effects of each treatment category k, number of studies; n, number of participants; ES, effect size (Hedges’ $g$); n/a, not applicable; NMDA, $N$-methyl-$D$-aspartate; SARI, serotonin antagonist and reuptake inhibitor; MBCT, mindfulness-based cognitive therapy; LTPP, long-term psychoanalytic psychotherapy; CBT, cognitive–behavioural therapy; TAU, treatment as usual.

a. Pill placebo, without the two short-term studies whose active treatments had been excluded from their respective analyses, was as effective without these two studies (ES = 0.82, 95% CI 0.69–0.95) and as heterogeneous ($I^2 = 68.3$).
individual treatments (one study; $^{25}$ ES = 1.74), whereas psychoanalytic psychotherapy had the smallest (one study; $^{24}$ ES = 0.59).

Publication bias was not apparent (detail available from the author on request).

Secondary outcomes
Active versus control (pairwise) meta-analyses
Only three treatment classes were examinable due to data availability. Despite heterogeneity of therapies and studies, they proposed that psychological treatments were more beneficial than usual care or an active control (three studies; ES = 0.45, 95% CI 0.09−0.81, $I^2 = 63.8$). Antipsychotics showed effectiveness when compared to placebo (seven studies; ES = 0.38, 95% CI 0.18−0.58, $I^2 = 59.4$). The number of mood-stabiliser studies was lower in the pairwise comparison than in pre-post analyses due to a paucity of placebo-controlled trials, and they were not significantly more effective than placebo (four studies; ES = 0.13, 95% CI −0.14 to 0.39, $P = 0.34$, $I^2 = 0$).

Tolerability and acceptability
Tolerability and acceptability were defined differently between studies and were not sufficiently homogeneous to consider quantitatively in meta-analyses.

Eight studies reported the total number of adverse events occurring in each arm; they were higher in active versus placebo arms for most interventions but equal between active and placebo arms in the $\alpha$-cycliserine$^{26}$ and minocycline$^{1}$ trials. This rate might be heavily influenced by a large number of adverse events occurring in a minority of patients, and of seven studies reporting the percentage of participants experiencing at least one adverse event, most were similar between treatment arms.

The highest dropout rate was in the ziprasidone intervention (41% in the lower-dose arm).$^{26}$ There was a discrepancy of more than 10% in participant dropout between arms in this study, as well as in the studies by Heresco-Levy et al$^{15}$ ($\alpha$-cycliserine 23% versus placebo 11%) and Husain et al$^{17}$ (minocycline 24% versus 10% placebo). No dropouts were reported in the CBT trial arms of TAU and individual CBT (two participants withdrew from group CBT)$^{25}$ or from the 1-week lithium placebo-controlled study (either trial arm).$^{24}$

Discussion
Meta-analytic estimates of treatment effects for resistant and non-resistant depression
In contrast to TRD, progress is evolving regarding the comparative effectiveness of common treatments for MDD, as exemplified by a recent, extensive network meta-analysis: Cipriani et al$^{20}$ identified over 500 double-blind randomised trials of antidepressant mono-therapy for MDD (in contrast to the 28 that we found for TRD augmentation) and found all to be significantly more effective than placebo. Another meta-analysis of pharmacological augmentation treatments for depression non-responsive to one or more antidepressant reported comparable ESs.$^{3}$ We anticipate smaller ESs to pharmacological interventions, according to a meta-analysis of direct comparisons.$^{42}$ The most recent review investigating psychological treatments for TRD identified only two randomised studies, both underpowered and defining TRD loosely; one had found comparable benefits of CBT and antidepressants, whereas the other reported clinical benefits of CBT but not antidepressants.$^{43}$ The importance of building the psychological evidence base is clear and we predict that over the next decade growing efforts in this field will reduce the current uncertainty of their effectiveness for this population.$^{44}$

Many psychological trials were excluded from the current review as they focused on chronicity or recurrence of depression rather than on the number of failed treatments. This limitation reflects the lack of integration between psychological and pharmacological fields and the difficulty in operationalising a measure of treatment response, particularly for past psychological therapies (including treatment adequacy, adherence, dose, duration, intensity and other factors likely to influence outcome). The CoBalT RCT has been seminal in the field, finding CBT adjunct to usual care as clinically effective (odds ratio of 3.26),$^{41}$ but was not eligible for
inclusion in the current review due to only requiring non-response to 6 weeks of one ongoing antidepressant. It is important to note that for most people with TRD, a combination of pharmacological and psychological approaches may be the most effective treatment both in terms of acute response and relapse prevention, although only pharmacological continuation treatments were focused on in the original studies included in this review.

Limitations and strengths

This work highlights the weakness of the evidence base for augmentation treatments for TRD. Inconsistency of TRD definition excluded a large number of studies, and mediating and moderating factors (such as TRD or baseline severity, continuation treatments and case mix of included patients) limited the ability to control confounders. However, 23 studies comprising a total of 5034 participants exhibited consistency of findings. Limited comparable data were available on the tolerability and acceptability alongside effectiveness and we were not able to consider the influence of patient/investigator blinding, intention-to-treat analyses or allegiance effects in meta-analyses. These factors may have influenced ESs, although they have not notably affected similar results in other reviews. Due to the limited number of psychological studies included, uncertainty remains on the benefit of CBT, psychoanalytic therapy and mindfulness-based cognitive therapy in this population.

Meta-analytic comparisons between treatment types have been deemed unsuitable (unless compared directly in original studies), but pre-post meta-analysis provides indications of effectiveness that can be compared between modalities. The pre-post analysis approach may show larger ESs due to spontaneous or natural remission or patient expectations of effectiveness, but the likelihood of this is attenuated in TRD populations who have experienced non-effective treatments and have a lower natural recovery rate than MDD as a whole. These also therefore reflect effects as seen in real-world clinical practice. Pre-post analysis has the advantage of permitting comparisons between different treatment types and controls, which is not suitable for traditional meta-analysis (e.g. drug placebo pills have a larger effect than a waiting-list control, although no waiting-list controls were examined in the present studies). In spite of these advantages it must be highlighted that indirectly comparing ESs between treatments in this way does not account for between-study variability (including but not limited to sociodemographic and clinical differences between recruited participants, the adequacy and delivery of treatment, and other procedural and analytic distinctions).

Clinical implications

There has been continued controversy surrounding the comparison of psychological and medication-based treatment for depression. We have not found strong evidence that either method is more effective in TRD specifically, although we highlight an urgent need for more intensive investigation of psychological therapy programmes. This study also illustrates that a short duration of treatment affects outcomes more than differences between treatment modalities. However, our results indicate that both psychological and pharmacological treatments are more effective than either pill or psychological control, even for already resistant people. Far from being ‘lost causes’, our findings demonstrate that more therapeutically is needed to achieve an optimal response for this subpopulation. Specifically, clinicians should not rule out CBT if it is being delivered with sufficient intensity and by skilled therapists. Our findings also confirms previous work indicating that aripiprazole and – to a lesser extent – lithium are effective treatments, supporting their current recommendation as first-line therapies.

Although the measured ESs with these two pharmacotherapies are similar to other options, the fact that they have been more thoroughly investigated in a larger number of studies underlines their status as first-choice options. Although unconfirmed, even if some medication-based treatments are shown to have greater efficacy overall in TRD, treatment decisions should necessarily remain a clinical judgement. Clinicians need to balance difficulties with tolerability of medications in addition to the durability of effects and, vitally, patient preference when deciding on the most appropriate treatments to use.

Rebecca Strawbridge, BSc, MSc, PhD, Post-Doctoral Research Associate, Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, UK; Ben Carter, BSc, MSc, PhD, Senior Lecturer in Biostatistics, Department of Biostatistics, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, UK; Lindsey Marwood, BSc, MSc, PhD, Post-Doctoral Trial Manager, Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, UK; Borwin Bandelow, MD, PhD, Professor of Psychiatry and Neurology, Department of Psychiatry and Psychotherapy, University Medical Centre, Göttingen, Germany; Dimosthenis Tsapokos, BSc, MSc, PhD Student, Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, UK; Rachael Taylor, BA, MSc, PhD Student, Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, UK; Tim Mantingh, BSc, MSc, Research Assistant, Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, UK; Allan H. Young, MBChB, MPhil, PhD, Professor of Psychopharmacology, Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King’s College London and South London and Maudsley National Health Service Foundation Trust, UK; Valeria de Angel, BSc, MSc, Research Assistant, Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King’s College London and South London and Maudsley National Health Service Foundation Trust, UK; Fiona Patrick, BSc, MSc, PhD Student, Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, UK; Anthony J. Cleare, BSc, MBBS, PhD, FRCpsych, Professor of Psychopharmacology, Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King’s College London and South London and Maudsley National Health Service Foundation Trust, UK; Allan H. Young, MBChB, MPhil, PhD, FRCpsych, FRCPsych, FRISB, Professor of Mood Disorders, Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King’s College London and South London and Maudsley National Health Service Foundation Trust, UK.

Correspondence: Rebecca Strawbridge, Centre for Affective Disorders, PO14 Institute of Psychiatry, Psychology and Neuroscience, King’s College London, 103 Denmark Hill, London SE5 8AZ, UK. Email: beccs.strawbridge@kcl.ac.uk

First received 31 May 2018, final revision 26 Sep 2018, accepted 4 Oct 2018

Supplementary material

Supplementary material is available online at https://doi.org/10.1192/bjp.2018.233.

Funding

This study represents independent research part funded by the National Institute for Health Research (NIHR) Biomedical Research Centre (BRC) at South London and Maudsley National Health Service Foundation Trust (SLaM) and King’s College London. The NIHR BRC had no involvement in study design, data collection, analysis or the decision to submit for publication. The views expressed are those of the authors and not necessarily those of the National Health Service, the NIHR or the Department of Health.

References


https://doi.org/10.1192/bjp.2018.233 Published online by Cambridge University Press
Strawbridge et al


2 Conway CR, George MS, Sackeim HA. Toward an evidence-based, operational definition of treatment-resistant depression: when enough is enough. JAMA Psychiatry 2017; 74: 9–10.

3 National Institute for Health and Clinical Excellence (NICE). Depression in Adults: Treatment and Management. NICE (UK), 2017.


patients with treatment resistant depression: results of the CoBalT randomised controlled trial.


...stumbling in the dark midnight of ignorance...

Timothie Bright (1550/1–1615), Cambridge graduate and Doctor of Phisicke, published A Treatise of Melancholie in 1586 ‘from little S. Bartelemewes by Smithfield’. Robert Burton, a melancholic Oxonian scholar and clergyman, repeatedly signals his indebtedness to Bright in his unsurpassable The Anatomy of Melancholy, published in 1621. Bright was a physician at St Bartholomew’s Hospital from 1585 until his dismissal – for neglect of his duties – and departure on 29 September 1591. Subsequently, Bright took holy orders. He died in Shrewsbury and was probably buried there in St Mary’s Church.

A Treatise of Melancholie is an early holistic English essay in psychiatry that recognises, e.g. ‘the instrument of reason, the braine’, phrenses, madness, lunacies, and melancholie’, ‘the strange effects it worketh in our minds and bodies’, and integrates medicine with the religion of the era. The book was popular, with two editions in 1586 and another in 1613, and in Shakespeare’s day it was the most important work on the subject. Bright addresses his (fictitious?) ‘melancholickes’ friend:

‘Your request is not onely that I should minister unto you, what my slender skill either in divinite or physicke may afford, but that I would at large declare unto you the nature of melancholie, what causeth it, what effects it worketh, how cured, and farther to lay open whatsoever may serve for the knowledge thereof, with such companions of feare, sadness, desparation, tears, weeping, sobbing, sighing, as follow that mournefull traine, yea oftimes, unbridled laughter rising not from any comforte of the heart, or gladnes of spirit, but from a disposition in such some altered, as by errour of conceite, that gesture in a counterfet manner bestowed upon that disagreeing passion, whose nature is rather to extinguish it selfe with teares, then asswaged by the sweete breath of chearfulnes, otherwise to receive refreshing.’

Characterie. An Arte of Shorte, Swifte, and Secrete Writing by Character. An Arte of Shorte, Swifte, and Secrete Writing by Character

‘...How weary, stale, flat, and unprofitable/Seem to me all the uses of this world’: ‘O God, a beast that wants discourse of reason/Would have mourned longer married with my uncle.’

Hamlet refers explicitly to ‘my melancholy’ (Act II, Scene 2) and Claudius observes: ‘There’s something in his soul, O’er which his melancholy sits on brood’ (Act III, Scene 1). Loening’s Die Hamlet-Tragödie Shakespeare’s points to Bright as the source of Shakespeare’s psychology both in Hamlet and other plays. Loening also notes that A Treatise was published by Bright’s publisher, Vautrollier, at his shop in Blackfriars, which is in Shakespeare’s immediate neighbourhood. When Vautrollier died in 1587, his print shop and stock passed to his son-in-law, Richard Field, fellow townsmen and publisher of Shakespeare’s Venus and Adonis (1593) and The Rape of Lucrece (1594). O’Sullivan and Dover Wilson support the idea that Bright’s Treatise influenced Shakespeare, and particularly Hamlet; they cite two score parallels not only of thought but also of phrase, indicating that Shakespeare had read A Treatise. Dover Wilson quotes a passage summarising most of Hamlet’s moods (whether or not feigned):

‘The perturbations of melancholly are for the most parte, sadder and fearful, and such as rise of them: as distrust, doubt, diffidence, or despair, sometimes furious, and sometimes merry in appereance, though a kinde of Sardonian, and false laughter.’

O’Sullivan also summarises Bright’s references to the connection between melancholy and procrastination, which is so germane to the character of Hamlet.