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

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

1 Department of Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, London, UK.
2 Department of Biostatistics, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, London, UK.
3 Department of Psychiatry and Psychotherapy, University Medical Centre, Göttingen, Germany
4 South London & Maudsley NHS Foundation Trust, London, UK.

* Corresponding author:
Email address: becci.strawbridge@kcl.ac.uk
Telephone number: +44 207 848 5305
Full postal address: Centre for Affective Disorders, PO74 Institute of Psychiatry, Psychology & Neuroscience, King’s College London, 103 Denmark Hill, London, SE5 8AZ

Keywords: treatment-resistant depression; augmentation; psychological therapy; psychopharmacology; systematic review; meta-analysis.
Abstract

**Background:** Depression is now 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 those who do not respond adequately to antidepressant treatments and specifically the effectiveness of available augmentation approaches.

**Aims:** We conducted a systematic review and meta-analysis (Prospero registration CRD42018088009) aiming to search and quantify the evidence of psychological and pharmacological augmentation interventions for TRD.

**Methods:** Trials where patients with TRD were randomised to at least one augmentation treatment were included, where treatment resistance was defined as insufficient response to at least two antidepressant treatments in the current episode. Pre-post analysis assessed treatment effectiveness, providing an effect size (ES) independent of comparator interventions.

**Results:** Of 28 included trials, only 3 investigated psychological treatments, while 25 examined pharmacological interventions. Assessing treatment classes, pre-post analyses demonstrated N-methyl-D-aspartate targeting drugs to have the highest effect size (ES=1.48, 95%CI 1.25–1.71). Other than aripiprazole (4 studies, ES=1.33, 95%CI 1.23-1.44) and lithium (3 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 1.80-1.30) and psychological (ES=1.43, 95%CI 0.50-2.36) therapies yielded higher effect sizes 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**
Prof Young has received honoraria for speaking from Astra Zeneca, Lundbeck, Eli Lilly, Sunovion; honoraria for consulting from Allergan, Livanova and Lundbeck, Sunovion, Janssen; and research grant support from Janssen, in the last 3 years. Prof Cleare has received honoraria for speaking from Astra Zeneca and Lundbeck, honoraria for consulting with Allergan, Janssen, Livanova, Lundbeck and Sandoz and support for conference attendance from Janssen and research grant support from Lundbeck, in the last 3 years. Dr Bandelow has (recently/soon) to be on the speakers/advisory board for: Hexal, Lilly, Lundbeck, Mundipharma, Pfizer, and Servier. No other conflicts of interest are declared.

**Relevance Statement**
This manuscript is submitted for the “treatment-resistant mood disorders” Themed Issue. Despite having substantial burden, treatment-resistant depression (TRD) has received little research attention and clinicians do not have consistent guidelines for treating the illness. We systematically reviewed the evidence of augmentation treatment effectiveness, since most TRD patients are treated with augmenters. In contrast to previous meta-analyses, we focus on the most common clinical TRD definition and provide pre-post effect sizes for psychological and pharmacological therapies. Although the evidence is scant (28 randomised trials), findings demonstrate effectiveness of aripiprazole and lithium, and particular early-stage promise for ketamine, minocycline and intensive CBT.
Introduction

The burden of treatment resistant depression (TRD) is challenging to quantify, TRD having eluded a universal definition\(^1\) but being prevalent and encompassing considerably greater severity, chronicity, recurrence, hospitalisation and comorbidity with both psychiatric and non-psychiatric disorders than non-resistant major depressive disorder (MDD)\(^2\). 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 patients as treatment resistant if they have failed one previous treatment trial (in contrast with the most popular guidelines\(^1\)), in part because this represents the inclusion criteria frequently used in clinical trials. One such example examined pharmacological augmentation treatments which the majority of TRD patients are treated with in practice\(^3\). Only when using the less stringent criteria of TRD was there sufficient evidence for a network meta-analysis in 2015\(^3\), and the authors reported significant efficacy of quetiapine, aripiprazole, lithium and thyroid hormone compared to placebo. However, this evidence may not apply to patients 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\(^4\). Additionally, pre-post effect sizes 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\(^4\).

Objectives

This review aimed to qualify and quantify the evidence of augmentation treatments for TRD, using the most common clinical definition (i.e. ≥2 failed treatments in current episode), and to compare effect sizes (ES) 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:

1. 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.
2. Provide an indication of the acceptability and tolerability of these treatments.

Methods

Criteria for considering studies for the review

The protocol for this systematic review was published via PROSPERO\(^5\), where full details of the search are available and reported consistently following the PRISMA reporting guidelines.

Types of included studies

Only randomised controlled trials (RCT) of at least 10 participants, to at least one suitable augmentation treatment were included.

Types of participants

Participants must have been adults with TRD, defined as unremitting depression despite at least two courses of treatment of adequate dose and duration undertaken in the current episode (current best practice guidelines\(^6\)). It has been considered that both within-class (in addition to between-class) switching of antidepressants, and psychological treatments are valid contributors to a TRD definition\(^1\); as such, these were permitted. Due to clear treatment distinctions, studies including patients with psychotic or bipolar depression were excluded.

Types of interventions

Patients 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\(^1\) and psychological treatments from the NICE depression guidelines\(^2\) or those with multiple meta-analyses supporting
use in depression. Eligible comparator treatments included pill placebo, another pharmacological agent, another psychological intervention, waiting list, active control, or treatment-as-usual (TAU).

Types of outcome measures
Primary outcome: Clinical improvement (ES) between pre- and post-treatment time points, for each eligible treatment/comparator arm. One efficacy measurement was selected, 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/compliance (e.g. any-cause trial dropout 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 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 nonrespon* OR un-respon* OR unrespon* OR TRD OR fail* OR inadequate OR difficult OR intractable) AND (augment* OR adjunct* OR add-on OR combin* 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 (RS, LM, RT, TM, VA, DT, VN, FP) with disparities addressed by consensus with additional review authors (AHY, AJC, BC). Following inclusion, data extraction was conducted by authors as above.

Quality assessment
The methodological quality assessment was examined using the Scottish Intercollegiate Guidelines Network (SIGN) 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 (ITT) 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 RoB (see Supplementary Table 1).

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 (CI), p-values and the $I^2$ statistic. Statistical heterogeneity was considered important if $I^2$ exceeded 60%, and explored using subgroups.

The following comparisons were planned to assess the primary outcome:
1. Pooled effects of augmentation intervention/comparator categories (i.e. psychological treatment, psychological comparator, pharmacological treatment and pharmacological comparator).
2. Pooled effects of augmenters by class (e.g. SSRI, SNRI, antipsychotic, mood stabiliser).
3. Pooled effects of individual treatment interventions within above categories.

Additional comparisons
We planned to explore secondary outcomes quantitatively or qualitatively, comprising: 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 currently considered 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 6-26 weeks to include any duration where expectations 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 MTI recommends durations of 6 weeks for full clinical effect\(^1\); therefore we selected to subgroup included trials of less than 6 weeks as “short-term” (this excludes rapid-onset treatments such as ketamine)\(^1\) and those more than 26 weeks as “long-term” treatment durations.

--- Figure 1 about here ---

Results

Systematic search results
After duplicates were removed, 2246 manuscripts from the MEDLINE and ISI Web of Science databases (all years to 6\(^{th}\) 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 (Figure 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 (SD=270, range 20–1293). The duration of interventions ranged from five days (ketamine\(^{11}\)) to 18 months (long-term psychoanalytic psychotherapy\(^{12}\)), with a median duration of six weeks (IQR=2).

Characteristics of participants
Participants studied had a median age of 45 years (IQR=4), and 66% were female. All patients analysed had unremitted depression despite at least two adequate treatment trials in the current episode. Fifteen studies defined TRD fully retrospectively (using a minimum duration of previous treatments of 4 or 6 weeks) while twelve required at least one unsuccessful treatment retrospectively and one prospectively. One study undertook two treatment trials to determine treatment-resistance fully prospectively\(^{13}\). 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 criteria for study entry\(^{12}\). Table 1 contains further details.

Quality assessment
Supplementary Table 1 contains the RoB ratings across criteria and studies. Twelve studies were rated as having a low RoB\(^{11,12,14–23}\), twelve had a moderate RoB\(^{24–35}\), and four had a high RoB\(^{13,36–38}\). The most common individual criteria rated as a high RoB were being funded and/or conducted by an industrial sponsor (twelve 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\(^{38}\) reporting blinded outcome raters).

--- Table 1 about here ---
Effectiveness of augmentation treatment

There was clinical diversity in the design (see Table 1), intervention and outcomes 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 patients, pharmacological treatments yielded an overall ES of 1.15 (95%CI 1.01-1.29, I²=82.7). Psychological therapies as a category comprised 3 studies totalling 276 patients, showing similar effects (ES=1.43, 95%CI 0.50-2.36, I²=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 effect sizes and the subgroup of active treatments trialled for a short-term duration (lithium\(^{15}\), metyrapone\(^{14}\)) showed an ES of 0.61 (95%CI 0.37-0.85, I²=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, I²=4.6) and did not affect heterogeneity of main analyses so were not excluded from analyses. 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 consistently lower than active treatments; see Figure 2 and Table 2. All active treatment effects are displayed in Supplementary Figure 1 and control arms in Supplementary Figure 2.

Pharmacological treatment classes

Pharmacological interventions without high RoB trialled for a therapeutic duration had an effect size of 1.19 (95%CI 1.08-1.30; I²=64.6).

N-methyl-D-aspartate (NMDA) targeting drugs showed the most consistent and large effect size of the pharmacological classes (ES=1.48, 95%CI 1.25-1.71, I²=0), despite the individual agents included having different mechanisms of action.

Mood stabilisers demonstrated an overall effect size of 1.12 (95%CI 0.92-1.31, I²=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 (3 studies, ES=1.00, 95%CI 0.81-1.20, I²=0).

Antipsychotics also had an effect size of 1.12 (95%CI 0.98-1.26, I²=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, I²=0).

Medications not falling into the above mechanisms were grouped together (trazodone, buspirone, thyroid hormone and dexmecamylamine), showing an ES of 1.36 (95%CI 1.09-1.63, I²=46.4), comparable in terms of heterogeneity and effect size to the other pharmacological treatments.

Psychological treatment classes

The overall effect size of psychological therapies (3 studies; ES=1.43, 95%CI 0.50-2.36) contained substantial heterogeneity (I²=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 effect size of all individual treatments (one study\(^{25}\); ES=1.74) while psychoanalytic psychotherapy had the smallest (one study\(^{12}\), ES=0.59).

Publication bias was not apparent (detail available on request).
Secondary outcomes

Active versus control (pairwise) meta-analyses

In order to validate the post-adjudicated method against the currently considered gold standard, we conducted pairwise active/control comparisons. Due to data availability, only three treatment classes were examinable. Despite heterogeneity of therapies and studies, these proposed that psychological treatments were more beneficial than usual care or an active control (3 studies, ES=0.45, 95%CI 0.09-0.81, I²=63.8). Antipsychotics showed effectiveness when compared to placebo (7 studies, ES=0.38, 95%CI 0.18-0.58, I²=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 were not significantly more effective than placebo (4 studies, ES=0.13, 95%CI -0.14 to 0.39, p=0.34, I²=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 (AEs) occurring in each arm, higher in active versus placebo arms for most interventions but equally between active and placebo arms in the d-cycloserine and minocycline trials. This rate might be heavily influenced by a large number of AEs occurring in a minority of patients, and of 7 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). There was a >10% discrepancy in participant dropout between arms in this study, as well as Heresco-Levy et al. (d-cycloserine 23% versus placebo 11%) and Husain et al. (minocycline 24% versus 10% placebo). No dropouts were reported in the CBT trial arms TAU and individual CBT (two patients withdrew from group CBT) or from the 1-week lithium placebo-controlled study (either trial arm).

Discussion

We included 28 studies, most containing low to moderate RoB, reporting effect estimates for the most prevalent TRD augmentation treatment strategies, using the definition of TRD most often used in clinical practice.

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, exemplified by a recent, extensive network meta-analysis: Cipriani et al. (2018) identified over 500 double-blind randomised trials of antidepressant monotherapy for MDD, in contrast with 28 we found for TRD augmentation, finding all to be significantly more effective than placebo. Another meta-analysis of pharmacological augmentation treatments for depression non-responsive to ≥1 antidepressant reported comparable effect sizes. We anticipate smaller effect sizes within TRD populations. The greatest pre-post effect of augmentation that we report is for medications targeting the NMDA receptor, comprising of ketamine (antagonist), d-cycloserine (partial agonist) and minocycline (antagonist). This finding supports increasing attention towards drugs acting on this pathway, as illustrated by a network meta-analysis of pharmacological and somatic treatments for non-responsive depression reporting ketamine to have the strongest short-term efficacy of treatments studied. It is notable however, that this finding was based on three studies only; population or design differences between studies may have yielded stronger effects in these trials than if directly compared with other interventions. Ketamine produced the highest effect size of the NMDA medications and is particularly challenging to maintain interviewer blinding, although Su et al. reported the trial as double blind. Based on a larger number of studies, our findings also indicate that for patients with a history of two unsuccessful treatments in the current episode, aripiprazole is effective, but it is important to note that all trials investigating aripiprazole had a potential allegiance effect. The evidence is less certain (often assessed in open label designs) but promising for lithium. The World Federation of Societies of Biological Psychiatry Task Force recommends lithium as the first-line augmentation option for TRD, and quetiapine or aripiprazole as alternatives; however, we identified only one randomised
quetiapine trial in the current review (found to be non-inferior to lithium). As such, it is clear that much more work in this field is required.

**Effects of interventions versus placebo in randomised studies for TRD**

Even a pill placebo response is variable under some methodological conditions, suggesting that there is some small scope for improvement for patients with TRD without augmenting with a new active treatment. The effect size and confidence intervals for placebo were heterogeneous across studies (as displayed in Supplementary Figure 2), demonstrating that indeed there are limitations to inferring the relative effects of interventions across diverse investigations. Placebo and active treatment outcomes will have been influenced by a multitude of factors which differed across trials (including but not limited to the maintenance of blinding, analyses undertaken, inclusion criteria relating to comorbidities, severity, etc.). Notwithstanding, it does appear that as a whole both psychological and pharmacological treatments are more effective than either pill or psychological controls, even for already resistant patients. Specifically, the treatment classes whose pre-post confidence interval did not overlap with the pill placebo estimates were mood stabilisers, antipsychotics, NMDA drugs and medications with ‘other’ mechanisms. This was not the case for psychological treatments which contained a wide confidence interval, or for short-term treatment durations.

**Effectiveness of psychological versus pharmacological intervention**

For MDD, psychological therapies demonstrate overall comparable effect sizes to pharmacological interventions, according to a meta-analysis of direct comparisons. 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, while the other reported clinical benefits of CBT but not antidepressants. 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 current uncertainty of their effectiveness for this patient population.

Many psychological trials were excluded from the current review, as they focused on chronicity or recurrence of depression rather than 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), 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 patients 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, 5034 participants from 23 studies 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, ITT analyses or allegiance effects in meta-analyses. These factors may have influenced effect sizes, although have not notably affected similar results in other reviews. Due to the limited number of psychological studies included, uncertainty remains over 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 effect sizes 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 traditional meta-analysis is not suitable for (e.g. drug placebo pills have a larger effect than a waiting list control\(^a\), although no waiting list controls were examined in the present studies). In spite of these advantages it must be highlighted that indirectly comparing effect sizes between treatments in this way does not account for between-study variability (including but not limited to sociodemographic and clinical differences between patients recruited, 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 one or the other is more effective in TRD specifically, although we highlight an urgent need for more intensive investigation of psychological therapy programs. 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 patients. Far from being 'lost causes', our findings demonstrate that more therapeutic work is needed to achieve an optimal response for this subpopulation of patients. Specifically, clinicians should not rule out CBT if it is being delivered with sufficient intensity and skilled therapists\(^b\). Our findings also support previous work indicating that aripiprazole and, to a lesser extent, lithium are effective treatments, supporting their current recommendation as first-line therapies\(^c\). Although the measured effect sizes 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 possess greater efficacy overall in TRD, treatment decisions should necessarily remain a clinical judgement, in which 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. We continue to advocate pre-post meta-analyses and network meta-analyses following future primary RCTs to provide further assistance to clinicians for predicting the optimal treatment modality/ies for patients with TRD.

**Summary**

Despite advances in the treatment of affective disorders, both clinical response to and tolerance of current pharmacological agents is often poor\(^d\). This is particularly so for patients with TRD, for whom there is a wide range of treatment options that may be suitable, but very little consensus on which are the most effective and tolerable\(^e\). Our analyses provide both absolute (pre-post) and relative (pairwise) effect estimates for augmentation treatment strategies investigated for treatment-resistant depression, using the definition of TRD most often used in clinical practice. Based on our results, ketamine and other NMDA-targeting drugs, as well as buspirone and trazodone, hold particular promise for the future of evidence-based TRD treatments.

**References**


Acknowledgements
This study represents independent research part funded by the National Institute for Health Research (NIHR) Biomedical Research Centre (BRC) at South London and Maudsley NHS Foundation Trust (SLaM) and King’s College London (KCL). 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 author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Full details of authors
Rebecca Strawbridge BSc MSc PhD. Post-doctoral Research Associate. Department of Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, PO74 103 Denmark Hill, London, SE5 8AZ, UK.
Ben Carter BSc MSc PhD. Senior Lecturer in Biostatistics. Department of Biostatistics, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, De Crespigny Park, London, SE5 8AF, UK.

Lindsey Marwood BSc MSc PhD. Post-doctoral Trial Manager. Department of Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, PO74 103 Denmark Hill, London, SE5 8AZ, UK.

Borwin Bandelow MD, PhD. Professor of Psychiatry and Neurology. Department of Psychiatry and Psychotherapy, University Medical Centre, Göttingen, Germany.

Dimosthenis Tsapekos BSc MSc. PhD student. Department of Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, PO74 103 Denmark Hill, London, SE5 8AZ, UK.

Viktoriya Nikolova BSc MRes. Research Assistant. Department of Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, PO72 De Crespigny Park, London, SE5 8AF, UK.

Rachael Taylor BA MSc. PhD student. Department of Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, PO74 103 Denmark Hill, London, SE5 8AZ, UK.

Tim Mantingh BSc MSc. Research Assistant. Department of Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, PO74 103 Denmark Hill, London, SE5 8AZ, UK.

Valeria de Angel BSc MSc. Research Assistant. Department of Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, PO74 103 Denmark Hill, London, SE5 8AZ, UK.

Fiona Patrick BSc MSc. PhD student. Department of Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, PO74 103 Denmark Hill, London, SE5 8AZ, UK.

Anthony J Cleare BSc MBBS PhD FRCPsych. Professor of Psychopharmacology. Department of Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, PO74 103 Denmark Hill, London, SE5 8AZ, UK.

Allan H Young MBChB MPhil PhD FRCPsych FRCP FRSB. Professor of Mood Disorders. Department of Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, PO72 De Crespigny Park, London, SE5 8AF, UK.

Author contributions

All authors approved the publication of the manuscript. The additional contributions by individual authors are summarised below.

RS contributed to study design and methodology, data collection and analysis, interpretation of findings, drafting and revising the manuscript.

BC contributed to data analysis methodology, study inclusion, data extraction and analysis, interpretation and critical revision of the article.

LM contributed to study design, study inclusion and data collection, and critical revision of the manuscript.

BB contributed to the study conception and design, presentation of results and critical revision of the manuscript.

DT contributed to data collection and interpretation, drafting and revisions of the manuscript.

VN, RT, TM, VA and VP contributed to data collection and interpretation and manuscript revisions.

AJC contributed to study design, study inclusion and critical revisions of the manuscript.

AHY was responsible for study conception and supervised design, study inclusion and data interpretation as well as article drafting and revisions.