Ageing Research Reviews 107 (2025) 102710

Contents lists available at ScienceDirect

Ageing Research Reviews

journal homepage: www.elsevier.com/locate/arr

Review article

Systematic review of clinical effectiveness of interventions for treatment resistant late-life depression

Beatriz Pozuelo Moyano [a,*], Denise Gomez Bautista [b], Karla Jocelyn Porras Ibarra [c], Christoph Mueller [d,e], Armin von Gunten [a], Pierre Vandel [a], Setareh Ranjbar [a], Robert Howard [c], Allan H. Young [d], Robert Stewart [d], Suzanne Reeves [c], Vasiliki Orgeta [c], on behalf of the European Task Force for treatment resistant depression in older people

[a] Service of Old Age Psychiatry, Department of Psychiatry, Lausanne University Hospital (CHUV) and University of Lausanne, Prilly, Switzerland
[b] Division of Psychiatry, Faculty of Brain Sciences, University College London, London, UK
[c] Cognition and Behavior Unit, National Institute of Neurology and Neurosurgery "Manuel Velasco Suárez", Mexico City 14269, Mexico
[d] King's College London, Institute of Psychiatry, Psychology and Neuroscience, London, UK
[e] South London and Maudsley NHS Foundation Trust, London, UK

ARTICLE INFO

Keywords:
Late life depression
Resistant depression
Treatment
Older people
Randomized controlled trials

ABSTRACT

Background: Treatment-resistant late-life depression (TRLLD) affects nearly half of older adults with major depression. This systematic review evaluates published evidence of effectiveness of both pharmacological and non-pharmacological treatments for TRLLD.

Methods: A search of MEDLINE, EMBASE, CINAHL, PsycINFO, the Cochrane Library, and online trial registries up to March 2024 was conducted to identify randomized controlled trials (RCTs) evaluating pharmacological and non-pharmacological interventions for TRLLD.

Results: Seven studies assessed the effectiveness of pharmacological interventions (antidepressants, antipsychotics, mood stabilizers, or ketamine) and another seven examined non-pharmacological approaches (psychotherapy, electroconvulsive therapy, repetitive transcranial magnetic stimulation (rTMS), and computerized cognitive remediation). Aripiprazole (2 studies), venlafaxine (1 study), ketamine (1 study), and lithium (1 study) were associated with a reduction in depressive symptoms post-treatment compared to the comparator treatment group. rTMS (2 studies), sequential bilateral theta burst stimulation (1 study) and cognitive remediation (1 study) also showed significant improvements in depressive symptoms post-treatment compared to a comparator treatment group. Quality of evidence varied from very low to medium among the included studies. Most studies reported data on small sample sizes.

Conclusions and Implications: We identified a small number of RCTs evaluating treatments for TRLLD. Aripiprazole augmentation appears to be an effective treatment based on two studies, with an acceptable side effect profile. Other treatments may be effective, but the evidence is based on very low-quality evidence. Future large-scale RCTs are urgently needed to draw firm conclusions.

1. Introduction

Patients whose depressive disorder does not respond adequately to appropriate treatment are considered to have a more difficult-to-treat form of depression, commonly known as treatment-resistant depression (TRD) (Thase and Rush, 1997). In a recent meta-analysis, only 50.7 % of patients aged 65 years and older with late-life depression (LLD) achieved a reduction of at least 50 % on depressive symptoms after receiving an antidepressant at a mean dosage equivalent to 30 mg/day of fluoxetine (Gutsmiedl et al., 2020; Lenze et al., 2015). TRD in late life (TRLLD) is often accompanied by other medical comorbidities, such as anxiety, and sleep disturbances which impact older peoples' quality of life (Gareri et al., 2002; Soysal et al., 2017; Steffens, 2024; Subramanian et al., 2023). People with TRD are also more likely to experience higher rates of suicide and suicide attempts (Kern et al., 2023).

* Correspondence to: Department of Psychiatry, Lausanne University Hospital (CHUV) and University of Lausanne, Route de Cery, 60, Prilly 1008, Switzerland.
E-mail address: Beatriz.pozuelo-moyano@chuv.ch (B. Pozuelo Moyano).

https://doi.org/10.1016/j.arr.2025.102710
Received 20 December 2024; Received in revised form 17 February 2025; Accepted 24 February 2025
Available online 28 February 2025

Treatment of TRLLD presents a significant challenge. Patients who do not respond to two antidepressant trials, most commonly selective serotonin reuptake inhibitors (SSRIs) or serotonin–norepinephrine reuptake inhibitors (SNRIs), are also unlikely to respond to subsequent drug treatments (Buchalter et al., 2019). A US cohort study based on Medicare data showed that, for patients aged ≥ 65 years, all-cause total healthcare costs for TRD were approximately 1.5 times higher than for non-TRD major depressive disorder (MDD) (Benson et al., 2020). Due to limited access to geriatric psychiatric care, TRLLD is frequently managed in primary care settings, which can often be suboptimal (Hamm et al., 2022). The most common definition of TRD requires a lack of response to two adequate treatments (Gaynes et al., 2020; Steffens, 2024). However, other authors including Cooper et al., (2011) have used less stringent criteria, defining TRD by a lack of response to one or more adequate courses of treatment (Cooper et al., 2011).

There are currently no comprehensive clinical guidelines specifically dedicated to the treatment of TRLLD (Steffens, 2024). The 2021 Canadian Guidelines on the Prevention, Assessment, and Treatment of Depression Among Older Adults recommend that clinicians prioritize assessing medication adherence as a first step in managing partial response or treatment resistance (Canadian Coalition for Seniors' Mental Health, 2021). Additionally, ECT is suggested as a viable option for continuation or maintenance therapy in older patients who exhibit partial response, treatment resistance, or intolerance to pharmacotherapy during the acute phase of treatment (Canadian Coalition for Seniors' Mental Health, 2021). The only systematic review of effectiveness of treatments for refractory LLD was published in 2011 by Cooper et al. (2011). This review concluded that 52 % of older people (95 % CI=42–62; N = 381) with refractory MDD (defined as not responding to at least one course of treatment for depression), respond to pharmacological treatments. The only treatment found to be clinically effective for TRLLD was lithium augmentation to another antidepressant (when patients had not previously responded to antidepressant monotherapy) with a response rate of 42 % (95 % CI=21–65; N = 57).

Since then, only narrative reviews have been published synthesizing the evidence, with studies using heterogeneous definitions, trial designs, and treatments (Blaszczyk et al., 2023; Patrick et al., 2024; Roose and Brown, 2022; Subramanian et al., 2023). As a result, there are currently no updated systematic reviews with detailed evidence from existing RCTs reporting on effectiveness of treatments, their designs, definitions of TRLLD used and quality of trials conducted to date. Our objective therefore was to undertake a systematic review of worldwide evidence reporting on the clinical effectiveness of treatments for TRD and to comment on the quality of the evidence.

2. Methods

We searched all related LLD, resistant, and treatment terms in major databases (MEDLINE, EMBASE, CINHAL, PsycINFO and the Cochrane library). We also searched for ongoing trials, through national and international trial registers, and specialized databases of psychological treatments of depression up to March 2024 (see Supplementary Online Material for details of search terms). We registered our review in PROSPERO (protocol registration number: CRD42024510508).

Inclusion criteria were as follows: 1) participants were older people (> 60 years) with a formal diagnosis of major depression according to widely accepted clinical criteria, such as DSM-5 (Diagnostic and Statistical Manual of Mental Disorders) (American Psychiatric Association, 2013), ICD-10 (International Classification of Diseases) (World Health Organization, 2004), or a comparable diagnostic system, 2) resistant major depression, defined as depression not responding to at least one course of antidepressant treatment, 3) reporting effectiveness or efficacy data based on an RCT, on 4) patients receiving any type of pharmacological or non-pharmacological treatment.

Three reviewers (BPM, DGB and KJP) worked independently to identify studies that met inclusion criteria and extracted data. Disagreements were discussed with the last author (VO). We contacted authors of primary trials if there were any missing data and employed the Cochrane Risk of Bias Tool 2 for Systematic Reviews to assess the quality of the evidence.

2.1. Definition of treatment-resistant depression

For a study to be eligible for inclusion in this systematic review, all the participants had to meet criteria for TRLLD at the point at which they were randomly allocated to treatment. Given the lack of a consensus definition of TRD (McIntyre et al., 2023), we adopted a broad definition by including all studies that used the terms "resistant" or "refractory" LLD. We included clinical trials assessing the effectiveness or efficacy of any treatment reporting on clinical improvement of depressive symptoms after the completion of at least one adequate antidepressant treatment course. We categorized our results into two groups: patients who had failed to respond to two antidepressant treatments (meeting the most used criteria for TRLLD) and those who had received only one course of antidepressant treatment (see Table 1).

2.2. Types of interventions

The experimental interventions were based on the 'next step' approach of the management of LLD that had not responded to prior treatment.

As the main objective of this systematic review is to measure the degree of beneficial effect under "real world" clinical settings, we addressed the practical question of an intervention as it would occur in routine clinical practice; which treatment leads to the greatest reduction in depressive symptoms in a population with TRLLD, and used the term "effectiveness" (Godwin et al., 2003; Tunis et al., 2003).

We divided treatments into pharmacological and non-pharmacological interventions:

Within the pharmacological interventions group, we included:

- switching to placebo
- switching to a different pharmacological treatment
- switching to a different non-pharmacological intervention
- augmenting with placebo
- augmenting treatment with another pharmacological treatment
- augmenting treatment with a non-pharmacological intervention

Pharmacological treatments included antidepressants, antipsychotics, mood stabilizers or ketamine.

Studies examining non-pharmacological interventions included:

- adding psychotherapy
- adding cognitive remediation
- adding rTMS
- adding ECT

2.3. Types of outcome measures

2.3.1. Primary outcome

Change in depressive symptoms was measured by validated rating scales, either clinician-rated (e.g. HAM-D, Montgomery-Asberg Depression Rating Scale (MADRS), or self-report (e.g. PHQ-9).

2.3.2. Secondary outcomes

Treatment response was usually defined as at least a 50 % reduction on depressive symptoms on the HAM-D or MADRS (Riedel et al., 2010), but we accepted a study's original definition of response if reported. Depression remission was typically based on scoring at a specific cut off on a depression measure (i.e. 7 or less on the HAM-D), but again we accepted a study's alternative definition if this was reported.

Table 1
Characteristics of the included studies.

Author, year & country	Design	Arms and objectives	N, sex, Mean age	Inclusion criteria/ resistant depression definition	Change from baseline in depression rating	Outcomes response and/or remission	Absolute value of effect sizes	GRADE Score, certainty
Studies including patients with TRLLD defined as ≥ 2 unsuccessful antidepressant treatments								
Pharmacological trials								
Lenze 2023 USA	2 step RCT, 10 weeks	Step 1: A aripiprazole vs A bupropion vs S to bupropion Step 2: A Lithium vs S to nortriptyline	STEP 1: n = 619 (F=413; M=206) STEP 2: n = 248 (F=173; M= 75) STEP 1: Mean age = 69.2 (7.1) STEP 2: Mean age= 68.5 (5.85)	a) ≥ 60 years b) MDD (DSM−5) c) PHQ−9 initially ≥ 6 and then changed to ≥ 10 d) no dementia (Short Blessed ≥ 10 and/or clinical evidence of dementia) ≥ 2 AD	Change from baseline in MADRS scores (LS mean): Step 1: A aripiprazole −7.60 vs A bupropion −7.23 vs S to bupropion −4.14 Step 2: A lithium −4.63 vs S to nortriptyline −5.33	Remission (≤ 10 on the MADRS or ≤ 5 on the PHQ−9) Step 1: A aripiprazole 28.9 % vs A bupropion 28.2 % vs S to bupropion 19.3 % Step 2: A lithium 18.9 % vs S to nortriptyline 21.5 % RR compared to S to bupropion (95 % CI): Step 1: A aripiprazole 1.50 vs A bupropion 1.49 vs S to bupropion 1.00 RR compared to S to nortriptilyne (95 % CI): Step 2: A lithium 0.84 vs S to nortriptyline 1	Change in MADRS from baseline between 2 groups: Step 1: A aripiprazole vs A to bupropion: d= 0.031 A bupropion vs. S to bupropion: d = 0.257 A aripiprazole vs. S to bupropion: d= 0.289 Step 2 A lithium vs S nortriptyline: d = 0.057	Change from baseline MADRS: Step 1: ⨁⨁⨁⨁ Step 2: ⨁⨁⨁⨁
Mazeh 2007 Israel	Single-blind RCT, 8 weeks	S to paroxetine vs S to venlafaxine	N = 30 (F= 17; M= 13) Mean age = 75.9 (3.87)	a) ≥ 65 years b) MDD (DSM-IV) c) HAM-D (21-item) ≥ 18 d) no dementia (7-min screen for dementia) ≥ 2 AD	Changes from baseline in HAM-D scores: S to paroxetine –12.5 vs S to venlafaxine –19.1 HAM-D: F(1207) = 14.4; P < 0.0003 Changes from baseline in GDS scores: S to paroxetine –3.2 vs S to venlafaxine –6.0 GDS: F(1207) = 1.29; P < 0.2	Remission (≤7 on the HAM-D) S to paroxetine 33 % vs S to venlafaxine 60 % Response (< 50 % on the HAM-D) S to paroxetine 20 % vs S to venlafaxine 20 %	Change in HAM-D from baseline between 2 groups d = 0.526 Change in GDS from baseline between 2 groups d = 0.158	Change from baseline HAM-D: ⨁◯◯◯
Ochs-Ross 2020 USA	RCT, BD, PC multicenter study, 4 weeks	S to esketamine (flexibly-dosed) + AD vs S to placebo + AD	N = 137 (F= 85; M= 72) Mean age= 70 (4.52)	a) ≥ 65 years b) recurrent moderate to severe MDD (DSM−5) c) IDS-C30 ≥ 31 d) without impaired cognition (MMSE <25 (or <22 if less than the equivalent of a high school education) ≥ 2 AD	Changes from baseline in MADRS scores: S to esketamine + AD −10.0 vs S to placebo + AD −6.3 Difference of LS Means MADRS −3.6 (p = 0.059)	Remission (≤12 on the MADRS) S to esketamine + AD 17.5 % vs S to placebo + AD 6.7 % Response (≥50 % change in MADRS) S to esketamine + AD 27.0 % vs S to placebo + AD 13.3 %	Change in MADRS from baseline between 2 groups: d = 0.336	Change from baseline MADRS: ⨁⨁⨁◯
Luzny, (2013) Czech Republic	RCT	Citalopram iv vs UL ECT	N = 20 (F= 15; M= 5) Mean age= 67.8	a) ≥ 65 years b) MDD (ICD−10)) ≥ 2 AD		Effect (≤ 50 % on the HAM-D) citalopram iv 83.3 % vs UL ECT 87.5 % (p = 0.112)	Not enough data provided to calculate effect sizes	Effect (≤ 50 % on the HAM-D): ⨁◯◯◯
Non pharmacological trials								
Zhao 2019 China	RCT, 4 weeks	A rTMS vs no TMS	N = 88 (F=41; M=47) Mean age: 65.6 (4.05)	a) ≥ 60 years b) MDD (DSM-IV) c) HAM-D (24-item) > 17 d) excluded if cerebral organic	Changes from baseline in HAM-D scores: HAM-D scores W1: A rTMS (20.5–22.21) −1.71 vs A placebo (20.5–23.48) −2.98		Not enough data provided to calculate effect sizes	Change from baseline HAM-D: ⨁◯◯◯

(*continued on next page*)

Table 1 (*continued*)

Author, year & country	Design	Arms and objectives	N, sex, Mean age	Inclusion criteria/ resistant depression definition	Change from baseline in depression rating	Outcomes response and/or remission	Absolute value of effect sizes	GRADE Score, certainty
				disorders (dementia not specifically mentioned) ≥ 2 AD	HAM-D scores W2: A rTMS (17.60–22.21) −4.61 vs A placebo (20.45–23.48) −3.03 ($p < 0.05$) HAM-D scores W3: A rTMS (14.5–22.21) −7.71 vs A placebo (20.45–23.48) −3.03 ($p < 0.05$) HAM-D scores W4: A rTMS (12.6–22.21) −9.61 vs A placebo (21–23.48) −2.48 ($p < 0.05$)			
Studies including patients with TRLLD defined as ≥ 1 unsuccessful antidepressant treatments								
Pharmacological trials								
Lenze 2015 USA	RCT PC, 12 weeks	A aripiprazole vs A placebo	N = 181 (F=103; M=78) Mean age: 66 ((25th/ 75th Centile: 62.8/70.5)	a) ≥ 60 years b) MDD (DSM-IV) c) MADRS ≥ 15 d) no dementia (medical records, cognitive screening, formal dementia criteria) No remission after ≥ 12 weeks of treatment with venlafaxine at ≥ 4 weeks at the highest tolerated dose (150–300 mg/d)	Changes from baseline in HAM-D (17-item) scores: A aripiprazole −5.78 vs A placebo −4.02 The aripiprazole group showed a significantly greater improvement (F [1170]=5.1, p = 0.03) Changes from baseline in MADRS scores: A aripiprazole −7.5 vs A placebo −5 The aripiprazole group showed a significantly greater improvement.	Remission (≤ 10 on the MADRS) A aripiprazole 44.0 % vs A placebo 28.9 % Odds ratio (95 % CI) remission Aripiprazole vs. Placebo: 2.0 (1.1–3.7), p: 0.03	Change in HAM-D (17-item) from baseline between 2 groups d = 0.345 Change in MADRS from baseline between 2 groups d = 0.098	Change from baseline HAM-D: ⨁⨁⨁◯
Kok 2007 The Netherlands	RCT, 6 weeks	A lithium vs S to phenelzine	N = 29 (F=22; M=7) Mean age: 73.12	a) ≥ 60 years b) MDD (DSM-IV) c) MADRS ≥ 20 d) no dementia (MMSE <15 or DSM-IV criteria for dementia) ≥ 1 AD	Changes from baseline in MADRS scores: A lithium −13.1 ($p < 0.001$) vs S to phenelzine 0.7 Changes from baseline in HAM-D (17-items) scores: A lithium −10 ($p < 0.001$) vs S to phenelzine −0.1	Remission (≤ 10 on the MADRS) A lithium 33.3 % vs S to phenelzine 0 % (p.0.042) Response (≤ 50 % on the MADRS) A lithium 46.7 % vs S to phenelzine 7.1 % (p 0.035) Remission (≤ 7 HAM-D (17-items)) A lithium 33.3 % vs S to phenelzine 0 % (p 0.042) Response (≤ 50 % on the HAM-D (17-items)) A lithium 46.7 % vs S to phenelzine 7.1 % (p 0.035)	Change in MADRS from baseline between 2 groups: d = 1.475 Change in HAM-D (17-items) from baseline between 2 groups: d = 1.580	Change from baseline MADRS: ⨁⨁◯◯
George 2017 Australia	DB controlled multiple-crossover study, 7 days	A subcutaneous ketamine HCl in ascending doses vs A single dose of control treatment, (midazolam) within the first 3 treatment sessions.	N = 16 (F=6; M=10) Mean age = 65.6 (5.7)	a) ≥ 60 years b) MDD or BP (DSM-IV) c) depressive episode of duration ≥ 4 w d) MADRS≥ 20 e) patients with dementia non excluded Insufficient therapeutic response	Changes from baseline in MADRS scores: ketamine HCl: −0.1 mg/kg −9.71 ($p = 0.06$) −0.2 mg/kg −6.1 ($p < 0.01$) −0.3 mg/kg −3.75 ($p < 0.001$) −0.4 mg/kg −8.7 ($p < 0.001$)	Remission (≤ 10 on the MADRS) at least at one point during the trial (across all dose levels and time points) A Ketamine 68.8 % vs A midazolam 14 % Response (≤	Main effect of each 0.1 mg/kg increase in ketamine dose: f = 0.34	Change from baseline MADRS: ⨁◯◯◯

(*continued on next page*)

Table 1 (*continued*)

Author, year & country	Design	Arms and objectives	N, sex, Mean age	Inclusion criteria/ resistant depression definition	Change from baseline in depression rating	Outcomes response and/or remission	Absolute value of effect sizes	GRADE Score, certainty
				to ≥ 1 adequate trials of an AD medication during the current episode	compared with midazolam −1.1	50 % on the MADRS) at least at one point during the trial (across all dose levels and time points) A Ketamine 68.8 % vs A midazolam 29 % No participant sustained response or remission to midazolam treatment at Day 7		
Non pharmacological trials								
Kaster 2018 Canada	RCT DB PC, 4 weeks	A rTMS vs A sham rTMS	N = 52 (F=20; M=32) Mean age: 65.2 (5.45)	a) 60–85 years b) MDD (DSM-IV) c) HAM-D (24-item) ≥ 22 d) no dementia (MMSE <26 and clinical evidence) ≥ 1 AD	Changes from baseline in HAM-D (24-item) scores: rTMS −11.5 vs sham rTMS −10 HAM-D (24-item) scores over time: F = 36.5, d.f. = 189.0; p < 0.001). HAM-D (24-item) scores for an effect of treatment condition: (F = 3.3, d.f. = 49.0; p = 0.08) Time by treatment interaction (F = 0.9, d. f. = 189.0; p = 0.438)	Remission (≤10 on the HAM-D (24-item) and ≤ 60 % on the HAM-D (24-item) on 2 consecutive weeks) ITT sample rTMS 40 % vs sham rTMS 14.8 % (p < 0.05) PP sample rTMS 50 % vs sham rTMS 14.8 % (p < 0.05) Response (< 50 % on the HAM-D (24-item) on 2 consecutive weeks) ITT sample rTMS 44 % vs sham rTMS 18.5 % (p < 0.05) PP sample rTMS 55 % vs sham rTMS 18.5 % (p < 0.05)	Differences in the change in HAM-D (24-item) scores over time d = 0.873 Differences in the change in HAM-D (24-item) scores for an effect of treatment condition d = 0.519 Time by treatment interaction: d = 0.138	Change from baseline HAM-D: ⨁⨁◯◯
Blumberger 2022 Canada	Randomized non inferiority trial, 12 weeks	A rTMS vs A Sequential bilateral TBS	N = 172 (F=92; M=80) Mean age: 66.7 (6.04)	a) ≥ 60 years b) MDD (Mini-International Neuropsychiatric Interview) c) MADRS ≥ 18 d) no dementia (Short Blessed ≥ 10) ≥ 1 AD	Changes from baseline in MADRS scores: A rTMS −8.3 vs A TBS −9.9 (p-value = 0.0008) Changes from baseline in HRSD−17 scores: A rTMS −6.1 vs A TBS −7.1 (p = 0.0002)	Remission (≤ 10 on the MADRS): A rTMS 32.9 % vs A TBS 35.4 % (p 0.046) Response (≤ 50 % on the MADRS): A rTMS 32.9 % vs A TBS 44.3 % (p < 0.001) Remission (≤ 7 on the HAM-D (17 items)) A rTMS 27.2 % vs A TBS 33.8 % (p < 0.012) Response (≤ 50 % on the HAM-D (17 items)) A rTMS 29.6 %	Change in MADRS from baseline between 2 groups: d = 0.224 Change in HRSD−17 from baseline between 2 groups: d = 0.175	Change from baseline MADRS: ⨁⨁⨁◯

(*continued on next page*)

Table 1 (*continued*)

Author, year & country	Design	Arms and objectives	N, sex, Mean age	Inclusion criteria/ resistant depression definition	Change from baseline in depression rating	Outcomes response and/or remission	Absolute value of effect sizes	GRADE Score, certainty
						vs A TBS 41.9 % ($p < 0.001$)		
Stoppe 2006 Brasil	RCT, 4 weeks	S to RUL ECT vs S to BL ECT	N = 39 (F= 22; M=17) Mean age: 75.15	a) ≥ 60 years b) MDD (DSM-IV) c) MADRS≥ 20 d) no dementia ≥ 2 AD or ≥ 1 AD in severely ill patients, history of poor pharmacological response and good ECT results in previous episodes	Changes from baseline in MADRS scores: 4th session: RUL ECT −15.27 vs BL ECT −15.7 (U = 122.00, P = 0.06) 8th session: RUL ECT −24.28 vs BL ECT −21 (U = 118.00, P = 0.13) 12th session: RUL ECT −27.46 vs BL ECT −23.7 (U = 51.50, P = 0.06) 16th session: RUL ECT −24.1 vs BL ECT −27.8	Remission (≤ 10 on the MADRS): RUL ECT 88.2 % vs BL ECT 68.2 % (p = 0.25)	Not enough data provided to calculate effects sizes	Change from baseline MADRS: ⨁⨁◯◯
Reynolds 2010 USA	RCT, 16 weeks	IPT with escitalopram vs A DCM with escitalopram	N = 124 (W=85; M=39) Mean age: 72.3 (7.4)	a) ≥ 60 years b) MDD (DSM-IV) c) HAM-D (17 item) ≥ 15 d) mild cognitive impairment included dementia excluded (MMSE < 18 or with diagnosed dementia) Partial responders to one AD (escitalopram)	No data	Remission (≤ 7 on the HAM-D score for 3 weeks): IPT group: 58 % vs DCM group: 45 % (p = 0.14) in the ITT Response (three final scores on the HAM-D of 8–10.): IPT group: 82 % vs DCM group: 77 % in the ITT	Not enough data provided to calculate effects sizes	Remission (≤ 7 on the HAM-D score for 3 weeks): ⨁⨁⨁◯
Gebara 2019 USA	RCT, 8 weeks	immediate BBTI vs delayed BBTI (3-weeks wait-list control)	N = 11 (F=3; M=8) Mean age: 65 (4.1)	a) ≥ 60 years b) MDD (DSM−5) c) MADRS ≥ 15 d) Insomnia Severity Index (ISI) score ≥ 15. e) cognitive impairement excluded (MoCa <24) ≥ 1 AD	Changes from baseline in the PHQ−9 (minus the sleep) scores: Immediate BBTI: −3.8 (SE = 3.5) vs delayed BBTI: −2.4(SE = 2.2)	No data	Changes from baseline in the PHQ−9 (minus the sleep) between 2 groups: g = 0.196	Change from baseline PHQ−9: ⨁◯◯◯
Morimoto 2020 USA	DB, RCT, 4 weeks	nCCR designed to target CCD, active control not targeting CCD	N = 30 (F=19; M=11) Mean age: 73.5 years (7.8)	a) ≥ 60 years b) MDD (DSM-IV) c) MADRS > 15) d) no dementia (DSM-IV criteria for dementia) ≥ 1 AD	Changes from baseline in MADRS scores: nCCR: −12.1 (SD = 7.7) vs active control −6.6 (SD = 8.0). Treatment group X time interaction (F (1,61.8)= 11.37, p = 0.002)	Remission (<10 on the MADRS): nCCR group 58 % vs active control 8 % Response (≤50 % on MADRS): nCCR group 58 % vs active control 16 %	Changes from baseline in MADRS between 2 groups: d = 0.858	Change from baseline MADRS: ⨁◯◯◯

Abbreviations: A: Adding; AD: Antidepressant; BBTI: Brief Behavioral Treatment for Insomnia; BP: Bipolar Disorder; CCD: Cognitive Control Functions; CGI: Clinical Global Impression Scale; d: Cohen's d; DCM: Depression care management; DSM-5: Diagnostic and Statistical Manual of Mental Disorders, 5th Edition; ECT: Electroconvulsive Therapy; f: Cohen's f; F: Female; g: Hedges' g; GDS: Geriatric Depression Scale; HAM-D: Hamilton Depression Rating Scale; IDS-C30: Inventory of Depressive Symptomatology - Clinician Rated (30-item); ITT sample: intention-to-treat sample; IPT: interpersonal Psychotherapy; IV: Intravenous; M: Male; MDD: Major Depressive Disorder; MADRS: Montgomery-Åsberg Depression Rating Scale; MMSE: Mini-Mental State Examination; nCCR: Neuroplasticity based computerized cognitive remediation; PHQ-9: Patient Health Questionnaire-9; PP sample: per protocol sample; QIDS-SR-16: Quick Inventory of Depressive Symptomatology (16-item) (self-report); RCT: Randomized Controlled Trial; rTMS: repetitive transcranial magnetic stimulation; RUL: Right Unilateral; S: Switching; TBS: Theta Burst Stimulation; Ul: Unilateral; W: week **GRADE certainty ratings:** low: ⨁◯◯◯, very low: ⨁⨁◯◯, moderate: ⨁⨁⨁◯, high: ⨁⨁⨁⨁.

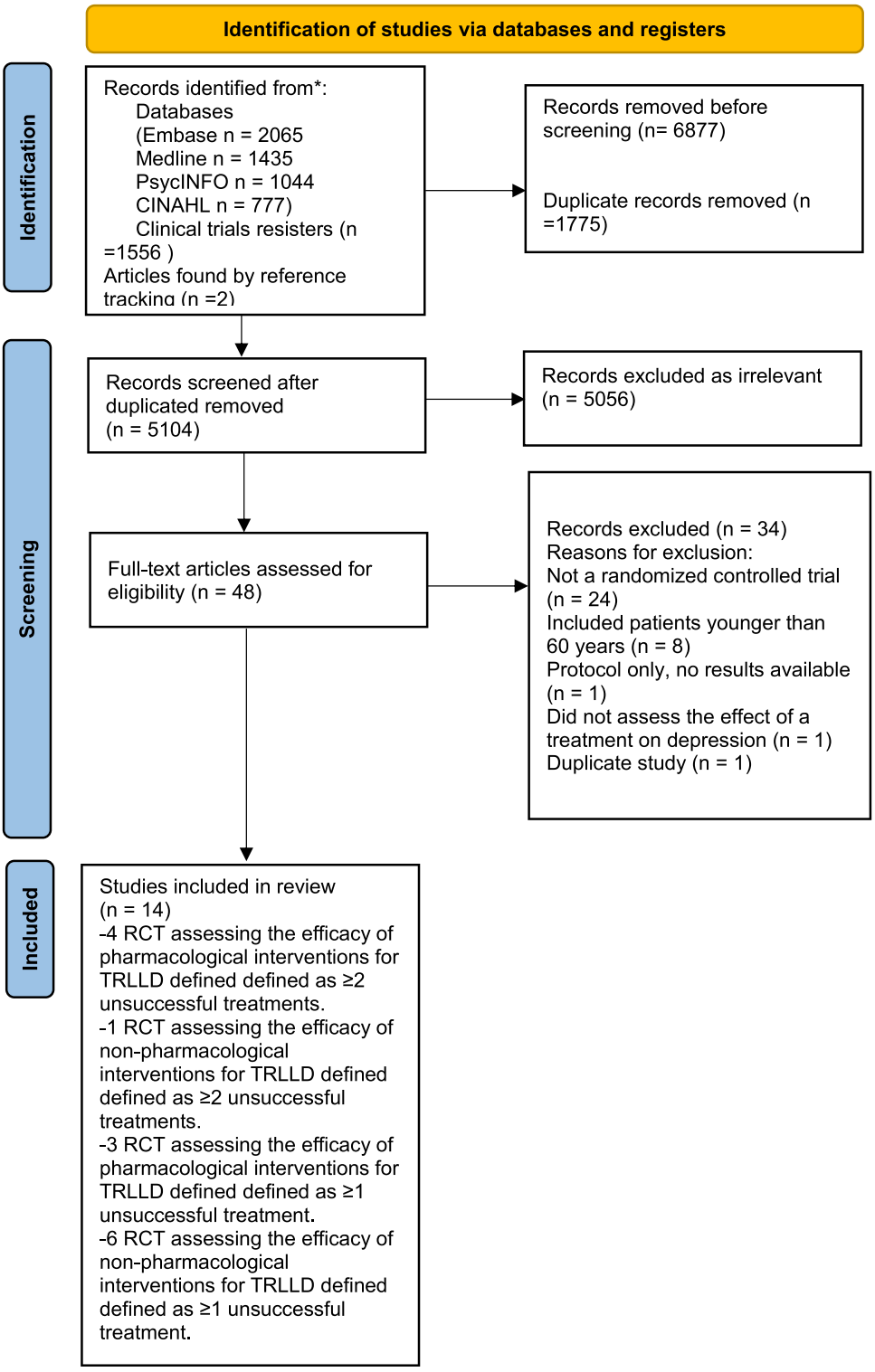

Fig. 1. Prisma Flowchart of the selection and identification of studies.

2.4. Effect size calculation

To facilitate comparison among the included trials, effect sizes were calculated based on Cohen's *d* (Lipsey and Wilson, 2001). The effect sizes however could not be aggregated due to the heterogeneity of studies.

3. Results

3.1. Search results

The search strategy yielded a total of 6877 articles (see Fig. 1 for details of the search process). After removal of duplicates and clearly irrelevant articles, we retrieved 48 full text records. Of these, 34 studies

were excluded for various reasons (see Table 1 of Excluded Studies Supplementary Online Material). A total of 14 studies were eligible for inclusion (Blumberger et al., 2022; Gebara et al., 2019; George et al., 2017; Kaster et al., 2018; Kok et al., 2007; Lenze et al., 2015; Lenze et al., 2023; Luzny, 2013; Mazeh et al., 2007; Morimoto et al., 2020; Ochs-Ross et al., 2020; Reynolds et al., 2010; Stoppe et al., 2006; Zhao et al., 2019) (see Table 1 for main study characteristics). Due to the heterogeneity of definitions of TRLLD, and interventions tested we were not able to perform a meta-analysis.

3.2. Description of studies

The mean age of participants was 69.3 years (taking into account only Step 1 of Lenze 2023), with a mean age range of 65 (Gebara et al., 2019) to 75.9 (Mazeh et al., 2007). Follow-up outcome timepoints varied from 7 days (George et al., 2017) to 16 weeks (Reynolds et al., 2010). All studies included 2 arms except for one, which had 2 steps with 3 arms in the first step (Lenze et al., 2023). Sample sizes ranged from 11 (Gebara et al., 2019) to 619 (in Step 1 (Lenze et al., 2023)).

Three studies evaluated effectiveness (Blumberger et al., 2022; Lenze et al., 2023; Luzny, 2013) and nine evaluated efficacy (Gebara et al., 2019; George et al., 2017; Kaster et al., 2018; Kok et al., 2007; Lenze et al., 2015; Mazeh et al., 2007; Morimoto et al., 2020; Ochs-Ross et al., 2020; Stoppe et al., 2006), two studies did not distinguish in their aims or methods if they were assessing effectiveness or efficacy (Reynolds et al., 2010; Zhao et al., 2019); for these two studies, following Gartlehner et al.'s recommendations (Gartlehner et al., 2006), one evaluated efficacy (Zhao et al., 2019) and the other effectiveness (Reynolds et al., 2010) (see Fig. 2).

3.3. Criteria for treatment resistant depression

All participants had a diagnosis of MDD according to DSM or ICD. Across all trials, participants were additionally required to have a minimum score on a specific scale on the PHQ-9, HAM, IDS-C30 or MADRS (see Table 1). Of the 14 studies, 5 studies included patients with LLD and two or more unsuccessful treatments (Lenze et al., 2023; Luzny, 2013; Mazeh et al., 2007; Ochs-Ross et al., 2020; Zhao et al., 2019). The remaining 9 studies included patients with LLD and only one unsuccessful treatment (Blumberger et al., 2022; Gebara et al., 2019; George et al., 2017; Kaster et al., 2018; Kok et al., 2007; Lenze et al., 2015; Morimoto et al., 2020; Reynolds et al., 2010; Stoppe et al., 2006).

3.4. Cognitive impairment or dementia

Most of the studies except two (George et al., 2017; Luzny, 2013) excluded people with dementia, cognitive impairment or cerebral organic disorders (see Table 1).

3.5. Treatments

3.5.1. Trials of pharmacological treatments

Two studies assessed the efficacy of switching to another antidepressant (paroxetine versus venlafaxine (Mazeh et al., 2007) or esketamine plus antidepressant versus placebo plus antidepressant (Ochs-Ross et al., 2020)), and 2 studies assessed the efficacy of adding a new psychotropic (aripiprazole versus placebo (Lenze et al., 2015); subcutaneous ketamine in separate sessions versus a single dose of midazolam randomly inserted in one of the first three sessions (George et al., 2017)). Three studies compared switching versus adding a new psychotropic (adding aripiprazole versus adding bupropion versus switching to bupropion (Lenze et al., 2023)); adding lithium versus switching to nortriptyline (Lenze et al., 2023) or adding lithium versus switching to phenelzine (Kok et al., 2007)). One study provided no information about the switching or adding strategy when reporting the effectiveness of intravenous (iv) citalopram versus electroconvulsive therapy (ECT) (Luzny, 2013).

3.5.2. Trials of non-pharmacological treatments

There were 7 trials of non-pharmacological interventions. Three trials examined effectiveness of adding rTMS (Blumberger et al., 2022; Kaster et al., 2018; Zhao et al., 2019). One study added rTMS versus not adding rTMS to an antidepressant (venlafaxine XR) (Zhao et al., 2019). One study added rTMS versus sham rTMS to stable doses of psychotropic medications (Kaster et al., 2018). The third study compared adding to pharmacotherapy rTMS versus sequential bilateral Theta Burst Stimulation (TBS) (Blumberger et al., 2022). One study assessed the efficacy of switching to bilateral (BL) versus right unilateral (RUL) ECT (Stoppe et al., 2006).

Two studies assessed the effectiveness or efficacy of psychotherapy

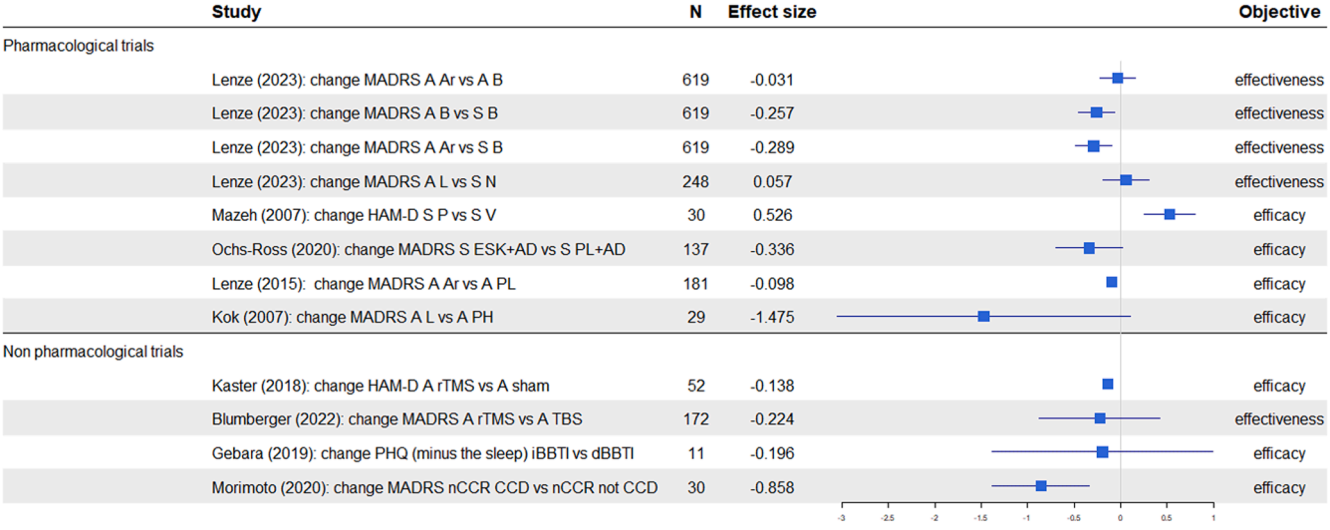

Fig. 2. Effect sizes across studies on change from baseline in depressive symptoms at post-treatment. **Abbreviations:** A: Adding; AD: Antidepressant; Ar: aripiprazole; B: bupropion; BBTI: Brief Behavioral Treatment for Insomnia; dBBTI: delayed Brief Behavioral Treatment for Insomnia; ESK: esketamine; HAM-D: Hamilton Depression Rating Scale; iBBTI: immediate Brief Behavioral Treatment for Insomnia; L: lithium; MADRS: Montgomery-Åsberg Depression Rating Scale; N: Nortriptyline; nCCR CCD: Neuroplasticity based computerized cognitive remediation targeting Cognitive Control Functions; nCCR not CCD: Neuroplasticity based computerized cognitive remediation not targeting Cognitive Control Functions; P: paroxetine; PH: phenelzine; PHQ-9.

(Gebara et al., 2019; Reynolds et al., 2010); one trial assessed the effectiveness of adding interpersonal psychotherapy (IPT) with escitalopram versus depression care management (DCM) with escitalopram (Reynolds et al., 2010); the other assessed the efficacy of immediate Brief Behavioral Treatment for Insomnia (BBTI) versus delayed BBTI (3-weeks wait-list control) (Gebara et al., 2019). We also identified one trial assessing the efficacy of cognitive remediation versus active control (Morimoto et al., 2020).

3.6. Primary outcome, change in depressive symptoms

All studies reported one or more measures of depressive symptoms as an outcome. Seven studies used MADRS, four the HAM-D with one study using the PHQ-9. Four studies used more than one depression measure (HAM-D, GDS or MADRS).

3.7. Secondary outcomes: Response or remission rates

Studies reported a dichotomous outcome of response and/or remission to treatment using different cut-offs in the MADRS or HAM-D (see Table 1).

Table 2
Side effects that occur in 5 % or more of the cases across all interventions and study arms.

Author, year & country	Arms	Side effets
Studies including patients with TRLLD (≥2 unsuccessful antidepressant treatments)		
Pharmacological trials		
Lenze 2023 USA	A Aripiprazole	Falls: 33 %, Dizziness: 17 %, Gastrointestinal distress: 13 %, Reduced salivation: 7 %, Tension, inner unrest, or anxiety: 14 %, Sleep disturbances: 18 %
	A Bupropion	Falls: 55 %, Dizziness: 20 %, Gastrointestinal distress: 17 %, Reduced salivation: 15 %, Tension, inner unrest, or anxiety: 10 %, Sleep disturbances: 9 %
	S Bupropion	Falls: 38 %, Dizziness: 20 %, Gastrointestinal distress: 18 %, Reduced salivation: 11 %, Tension, inner unrest, or anxiety: 14 %, Sleep disturbances: 16 %
	A Lithium	Falls: 47 %, Dizziness: 22 %, Gastrointestinal distress: 16 %, Reduced salivation: 10 %, Tension, inner unrest, or anxiety: 6 %, Sleep disturbances: 5 %
	S Nortriptyline	Falls: 38 %, Dizziness: 17 %, Gastrointestinal distress: 17 %, Reduced salivation: 42 %, Tension, inner unrest, or anxiety: 7 %, Sleep disturbances: 5 %
Mazeh 2007 Israel	S Venlafaxine	Nausea: 33 %, Dizziness: 26 %, Sedation: 26 %, Dry mouth: 20 %, Constipation: 20 %, Insomnia: 26 %
	S Paroxetine	Nausea: 26 %, Dizziness: 40 %, Sedation: 26 %, Dry mouth: 33 %, Constipation: 20 %, Insomnia: 26 %
Ochs-Ross 2020 USA	S Esketamine + AD	Dizziness: 21 %, Nausea: 18 %, Increased blood pressure: 12 %, Fatigue: 12 %, Headache: 12 %, Dissociation: 12 %, Vertigo: 11 %, Urinary tract infection: 8 %, Hypoesthesia oral: 7 %, Vomiting: 7 %, Dysgeusia: 6 %, Dysphoria: 6 %, Hypoesthesia: 6 %, Insomnia: 6 %, Paresthesia: 6 %
Lunzy 2013 Czech Republic	Intravenous Citalopram	Extrapyramidal parkinsonism: 8.3 %, Anxiety: 8.3 %
	Unilateral ECT	Memory impairment: 25 %
Non pharmacological trials		
Zhao 2019 China	A rTMS	No serious effects reported
Studies including patients with TRLLD (≥1 unsuccessful antidepressant treatments)		
Pharmacological trials		
Lenze 2015 USA	A Aripiprazole	Concentration difficulties: 9 %, Fatigue: 17 %, Sleepiness or sedation: 6 %, Tension/inner unrest: 15 %, Reduced sleep: 24 %, Increased dream activity: 27 %, Emotional indifference: 7.0 %, Tremor: 6 %, Akathisia: 12 %, Paraesthesia: 6 %, Headache: 12 %, Increased salivation: 10 %, Reduced salivation: 10 %, Nausea/vomiting: 12 %, Diarrhea: 9 %, Constipation: 8 %, Micturition disturbances: 12 %, Polyuria: 10 %, Orthostatic dizziness: 9 %, Photosensitivity: 6 %, Weight gain: 20 %, Diminished sexual desire: 14 %, Erectile dysfunction: 8 %, Orgastic Dysfunction: 6 %, Dry Vagina: 17 %
Kok 2007 The Netherlands	A Lithium	Anxiety: 7 %, Insomnia: 13 % Weakness/fatigue: 20 %, Increased appetite: 13 %, Headache: 13 %, Tremors: 80 %, Rigidity, stiffness: 7 %, Akathisia: 7 %, Dystonia: 13 %, Blurred vision: 13 %, Dry mouth: 33 %, Nausea, vomiting: 20 %, Diarrhea: 13 %, Constipation: 20 %, Syncope/dizziness: 27 %, Impaired urination: 20 %, Weight gain: 33 %
	S Phenelzine	Anxiety: 21 %, Agitation: 21 %, Insomnia: 57 %, Weakness/fatigue: 57 %, Memory impairment: 50 %, Increased appetite: 7 %, Headache: 7 %, Tremors: 21 %, Rigidity, stiffness: 21 %, Akathisia: 21 %, Dystonia: 7 %, Blurred vision: 28 %, Dry mouth: 57 %, Nausea, vomiting: 21 %, Diarrhea: 7 %, Constipation: 28 %, Syncope/dizziness: 43 %, Impaired urination: 21 %, Weight gain: 14 %
George 2017 Australia	A subcutaneous Ketamine	0.1 mg/kg: Flushing: 6 %, Dizziness: 6 %, Fatigue/Sleepiness/Poor Concentration/Feeling vage: 19 %, Dry mouth: 12 %, Blurred vision: 6 %, Restlessness: 12 %, Headache: 12 % 0.5 mg/kg: Dizziness: 44 %, Fatigue/Sleepiness/Poor Concentration/Feeling vage: 67 %, Paresthesia: 11 %, Blurred Vision: 44 %, Nausea: 11 %
Non pharmacological trials		
Kaster 2018 Canada	A rTMS	Headache after treatment: 56 %, Pain at stimulation site: 16 %
Blumberger, 2022 Canada	A Bilateral rTMS	Headache: 56 %, Nausea: 7 %, Dizziness: 19 %, Fatigue: 6 %, Anxiety/agitation: 8 %, Back/neck pain: 7 %
	A TBS	Headache: 54 %, Nausea: 8 %, Dizziness: 21 %, Fatigue: 6 %, Anxiety/agitation: 9 %, Back/neck pain: 11 %
Stoppe 2006 Brasil	Bilateral ECT	Post-ECT delirium responsible for treatment interruption: 18.2 %, Headache, Confusion/Delirium and Dizziness/Nausea: Not enough data provided to calculate the %
	Unilateral ECT	Dizziness/Nausea and Confusion/Delirium: Not enough data provided to calculate the %
Reynolds 2010 USA	A Escitalopram + IPT	Side effects were not considered as outcomes in the trial
	A Escitalopram + Care Management	Side effects were not considered as outcomes in the trial
Gebara 2019 USA	Behavioral Treatment for Insomnia	Side effects were not considered as outcomes in the trial
Morimoto 2017 USA	Computerized Cognitive Remediation	Side effects were not considered as outcomes in the trial

Abbreviations: A: Adding; AD: Antidepressant; ECT: Electroconvulsive Therapy; rTMS: repetitive transcranial magnetic stimulation; S: Switch- ing; TBS: Theta Burst Stimulation.

3.8. Ongoing trials

We did not identify any ongoing studies.

3.9. Risk of bias in included studies

Ratings of bias in included studies are presented in Table 2 of Supplementary Online Material. Risk of bias was detected predominantly in the randomization process and outcome measures. Two studies did not report details of the randomization process (Gebara et al., 2019; Luzny, 2013). In 6 studies, participants were not blinded to treatment (Blumberger et al., 2022; Gebara et al., 2019; Kok et al., 2007; Lenze et al., 2023; Reynolds et al., 2010; Zhao et al., 2019), and in two trials there were no blinded assessors (Mazeh et al., 2007; Zhao et al., 2019). Two studies did not provide information about blinding of assessors (Gebara et al., 2019), or the blinding process (Luzny, 2013). Given the small number of included studies, no assessment of publication bias was undertaken.

3.10. Effects of interventions

We did not identify any study reporting results of switching to placebo in monotherapy.

3.10.1. Studies assessing the effectiveness of switching to another antidepressant

The first study (n = 30) showed a statistically significant difference in changes from baseline in HAM-D scores when comparing paroxetine and venlafaxine, with more pronounced improvements in the venlafaxine group [F (1207) = 14.4, P < 0.0003] (Mazeh et al., 2007). The same study reported a response rate of 20 % when switching to either paroxetine or venlafaxine, and remission rates of 60 % when switching to venlafaxine and 33 % when switching to paroxetine (Mazeh et al., 2007).

The second study (n = 137) compared switching to esketamine and one antidepressant (duloxetine, escitalopram, sertraline, or venlafaxine XR) versus switching to placebo and one of the mentioned antidepressants (Ochs-Ross et al., 2020)). Although there was a difference in the change in MADRS between the two groups from baseline to day 28, with greater changes in the esketamine with antidepressant group, the results were not statistically significant (difference in LS mean MADRS −3.6 (p = 0.059)) (Ochs-Ross et al., 2020).

3.10.2. Studies assessing the effectiveness of adding a new psychotropic

In one study assessing the efficacy of aripiprazole versus placebo (n = 181), results showed that both groups showed improvement on depression (based on the HAM-D-17 score), but the aripiprazole group showed a significantly greater improvement (F[1170]=5.1, p = 0.03) (Lenze et al., 2015). The same study reported higher remission rates in the group receiving aripiprazole, with this group being two times more likely to achieve a treatment response (OR: 2.0 (1.1–3.7), p = 0.03) (Lenze et al., 2015).

The study comparing adding an ascending dose of subcutaneous ketamine versus adding a single dose of midazolam as a control treatment (n = 16), reported that the ketamine group receiving 0.4 mg/kg achieved greater improvements in depressive symptoms from baseline (8.7 points in MADRS) compared to the midazolam group (-1.1) (p < 0.001) (George et al., 2017). Both remission and response rates were higher in the ketamine compared to the midazolam group (George et al., 2017).

3.10.3. Studies assessing the effectiveness of switching versus adding a new psychotropic

We found three studies assessing the effectiveness of switching versus adding a new psychotropic. The first study included two steps; in the first step, patients (n = 619) received either an added dose of aripiprazole or bupropion or switched to bupropion (Lenze et al., 2023). In the second step, patients (n = 248) either received an added dose of lithium or switched to nortriptyline (Lenze et al., 2023).

In step 1, changes from baseline in MADRS scores were −7.60 (95 % CI, −9.20 to −5.99) in the aripiprazole-augmentation group, −7.23 (95 % CI, −8.86 to −5.59) in the bupropion-augmentation group, and −4.14 (95 % CI, −5.81 to −2.48) in the switch- to-bupropion group. In step 2, changes in MADRS scores were −4.63 (95 % CI, −6.78 to−2.49) in the lithium-augmentation group and −5.33 (95 % CI, −7.52 to −3.14) in the switch-to nortriptyline group. In Step 1, changes in depressive symptoms (MADRS) and remission rates for the augmentation groups (aripiprazole and bupropion) were statistically significant compared to the switch-to-bupropion group. The most significant reduction in MADRS scores from baseline and the highest remission rates were observed in the group adding aripiprazole, followed by the group adding bupropion.

In Step 2, neither the changes in depressive symptoms nor the remission rates between the lithium augmentation group and the switch-to-nortriptyline group differed to a statistically significant extent.

One trial assessed differences between adding lithium versus switching to phenelzine (n = 29). Both remission and response rates (MADRS and HAM-D) were significantly higher in the lithium augmentation group (Kok et al., 2007).

3.10.4. Studies assessing the effectiveness of rTMS

One of the studies (n = 88) showed that depression scores (HAMD) significantly decreased in the rTMS group compared with the control group at post-treatment (Zhao et al., 2019). Two studies assessed the efficacy of rTMS versus sham rTMS (n = 52) (Kaster et al., 2018) or versus sequential bilateral TBS (n = 172) (Blumberger et al., 2022). When comparing rTMS versus sham rTMS, the group receiving rTMS showed greater changes in depressive symptoms from baseline over time (F = 36.5, d.f. = 189.0; p < 0.001), but there was no significant differences in the change between the two groups (F = 0.9, d.f. = 189.0; p = 0.438); the group of patients receiving rTMS also had higher rates of response and remission (Kaster et al., 2018).

The third study compared the effectiveness of rTMS versus a sequential bilateral TBS treatment (n = 172) (Blumberger et al., 2022). The study demonstrated significant noninferiority of TBS compared to standard bilateral rTMS in depressive symptoms post-treatment, with an adjusted difference of 1.55 points in MADRS score change (95 % CI: −0.67). TBS also showed significant noninferiority in the MADRS response (44.3 % vs. 32.9 %, adjusted difference 11.4 %, 95 % CI: −1.1 %), within the noninferiority margin (δ < 15 %). Similarly, the MADRS remission rate for TBS was 35.4 % compared with 32.9 % for standard bilateral rTMS with an estimated adjusted difference of 2.5 % favoring TBS, with a lower 95 % CI of −9.7 %, smaller than the noninferiority δ< 10 %.

3.10.5. Studies assessing the effectiveness of ECT

One study compared the efficacy of BL versus RUL treatment with ECT (n = 39). Remission rates for RUL ECT (88.2 %) and BL ECT (68.2 %) were similar (p = 0.25) as were the reduction rates of depressive symptoms from baseline (Stoppe et al., 2006). The second study compared iv citalopram versus RUL ECT (n = 20); there were no statistically significant differences between the two groups in depressive symptoms post- treatment (Luzny, 2013).

3.10.6. Studies assessing the effectiveness of psychotherapy

Two studies assessed the effectiveness or efficacy of psychotherapy. The first added IPT versus a DCM intervention with escitalopram (n = 124). Remission rates for escitalopram with IPT and with DCM were similar across groups in intention-to-treat analyses (IPT vs. DCM: 58 [95 % CI: 46, 71] vs. 45 % [33,58]; p 1⁄4 0.14) (Reynolds et al., 2010).

In the second study (n = 11), assessing the therapeutic efficacy of

BBTI on symptoms of insomnia and depression, the treatment group experienced a 33 % rate of clinically significant positive change on the PHQ-9, compared to a 20 % rate in the delayed treatment group; however, this difference was not statistically significant (Gebara et al., 2019).

3.10.7. Studies assessing the effectiveness of cognitive remediation

In one study (n = 30) participants received neuroplasticity-based computerized cognitive remediation (nCCR), or an active control condition (Morimoto et al., 2020). The average reduction in depressive symptoms (MADRS) over 4 weeks was 12.1 points for the nCCR group, compared to 6.6 points in the control group. In the nCCR group, 58 % of participants achieved remission and response, while in the control group, only 8 % reached remission and 16 % responded. There was a significant interaction between group and time ($F(1, 61.8) = 11.37$, $p = 0.002$), indicating that the nCCR group experienced a greater decline in depressive symptoms over time compared to the control group.

3.11. Effect Sizes

The absolute values are presented in Table 1, while Fig. 2 illustrates effect sizes (Cohen's *d* and 95 % confidence intervals) reported or calculate based on data provided in each of the studies.

3.12. Side-effects

Side effects that occurred in more than 5 % of the treated cases are listed in Table 2.

All studies of pharmacological treatments reported side-effects affecting ≥ 5 % of the study sample. A total of 11 studies reported data on side-effects or adverse events associated with the interventions. The most commonly reported side-effects were tremor (in 80 % of patients with an added lithium intervention) (Kok et al., 2007), falls (in 47 % of an added lithium treatment or in 55 % when adding bupropion to an existing pharmacological medication) (Lenze et al., 2023) and insomnia (in 57 % of patients with switching to a phenelzine) (Kok et al., 2007). Fatigue was present in 57 % of patients switching to a phenelzine treatment (Kok et al., 2007) and 66 % of patients with 0.5 mg/kg of subcutaneous ketamine (George et al., 2017).

For non-pharmacological interventions, headache had a frequency of 56 % of patients with rTMS treatment in two studies (Blumberger et al., 2022; Kaster et al., 2018). Cognitive impairment was frequent in patients with ECT. One study found that participants who were randomized at BL ECT experienced greater cognitive impairment across all evaluations compared to the group that received RUL ECT. In this study, a neuropsychologist conducted a blinded assessment 2–4 days before the first ECT treatment and again 1 month after the final session. The difference was not significant after the 4th ECT (U = 47.00, P = 0.09) but reached significance after the 8th (U = 39.50, P = 0.03), the 12th (U = 29.00, P G 0.01), and 16th ECT sessions (U = 27.5, P G 0.01) (Stoppe et al., 2006). A small study (n = 20) reported memory impairment—defined in that trial as a decrease of 2 or more points on the Mini-Mental State Examination (MMSE)—in 25 % of the sample following RUL ECT treatment (Luzny, 2013). Post-ECT delirium was common in ECT treatment (Stoppe et al., 2006). The studies assessing the effectiveness of psychotherapy treatment did not report side effects as outcomes.

4. Discussion

In this review, we aimed to establish an up-to-date and accurate estimate of the effectiveness of both pharmacological and non-pharmacological treatments for older adults with resistant or refractory major depression. We encountered several challenges, the most important being the limited number of RCTs examining the effectiveness of antidepressant treatments in patients with TRLLD. The studies also varied greatly in terms of the definitions of TRD and the treatments evaluated, which meant that it was not possible to conduct a meta-analysis.

4.1. Studies assessing the effectiveness of pharmacological interventions

We identified three studies with relatively large sample sizes and low risk of bias, reporting on the effectiveness of aripiprazole, bupropion, lithium, nortriptyline and esketamine for the reduction of depressive symptoms in patients with TRLLD (Lenze et al., 2015; Lenze et al., 2023; Ochs-Ross et al., 2020). The first study showed that augmentation strategies with aripiprazole and bupropion resulted in statistically significant improvements in depressive symptoms and remission of depression rates compared to the group that switched to bupropion, based on a large sample (n = 619) (Lenze et al., 2023). The second step of this trial, reporting on a smaller sample size (n = 249), showed that switching to nortriptyline was associated with more important improvements from baseline in MADRS than adding lithium therapy; however, this difference did not reach statistical significance (Lenze et al., 2023).

The second study with a large sample size concluded that, when added to existing antidepressants, aripiprazole was significantly superior to placebo in treating depressive symptoms (Lenze et al., 2015). The effectiveness of aripiprazole as an augmentation strategy is in line with systematic reviews and meta-analyses assessing the effectiveness of TRD treatments in younger adults (Edwards et al., 2013; Nuñez et al., 2022; Zhou et al., 2015). The third study (n = 137) showed that switching to esketamine combined with an antidepressant was not associated with a statistically significant reduction of depressive symptoms when compared to switching to placebo plus an antidepressant (Ochs-Ross et al., 2020).

The main difference between our systematic review and that of Cooper et al. lies in the selection criteria. We included only randomized controlled trials (RCTs) with the initial aim of obtaining comparable results, whereas Cooper et al. additionally included primary research studies evaluating treatments for refractory depression, but not single case reports. Another key difference is that the review by Cooper et al. focused on older adults aged 55 and above. In contrast, to better define the population affected by LLD, we decided to include only studies involving participants aged 60 and older.

The remaining studies assessing the effectiveness of pharmacological interventions had small sample sizes, making their findings less reliable. One of the small-scale studies (n = 29), reported that augmenting treatment with lithium showed significantly greater response and remission rates compared to phenelzine (Kok et al., 2007). Another small study showed that doses of subcutaneous ketamine ≥ 0.2 mg/kg were significantly more efficacious than midazolam in treating depressive symptoms (George et al., 2017). Similarly, a small study reporting on 30 participants concluded that switching treatment to venlafaxine was significantly superior to switching to paroxetine for depressive symptoms (Mazeh et al., 2007). The results of these studies supporting the efficacy of ketamine and venlafaxine are interesting and indicate that larger studies are urgently needed to confirm these results.

4.2. Studies assessing the effectiveness of non-pharmacological interventions: rTMS

The effectiveness of rTMS was tested in three trials, with varying sample sizes and risk of bias. The study with the largest sample and lowest risk of bias showed that bilateral TBS was superior to standard bilateral rTMS in reducing depressive symptoms (Blumberger et al., 2022). The remaining two studies were smaller (Kaster et al., 2018; Zhao et al., 2019) and found that rTMS was effective compared to placebo or no treatment, with statistically significant changes in depressive symptoms at 4 weeks. Given that the evidence base for rTMS remains limited,

future large-scale studies with longer follow-ups and adhering to the conventionally used frequencies (10 Hz or 1 Hz) (Cappon et al., 2022) will be key to establishing the effectiveness of these treatments.

4.3. Studies assessing the effectiveness of non-pharmacological interventions: ECT

With respect to ECT, two studies with small sample sizes (Luzny, 2013; Stoppe et al., 2006) assessed effectiveness at post-treatment. One of these concluded that RUL ECT was as effective as BL ECT (Stoppe et al., 2006) in reducing depressive symptoms. The second study, which had high risk of bias in several domains, found that ECT was not more effective than iv citalopram in reducing symptoms of depression (Luzny, 2013). It is important to note that in this study participants with dementia were not specifically excluded. These findings indicate that the evidence base of the clinical effectiveness of ECT for TRLLD remains small, despite its widespread use.

4.4. Studies assessing the effectiveness of non-pharmacological interventions: psychotherapy and cognitive remediation interventions

We found insufficient evidence to derive meaningful conclusions from studies investigating the effectiveness of psychotherapy in TRLLD, highlighting the need for further research in this area. We identified only 2 studies that assessed the effectiveness of psychotherapy for the treatment of depressive symptoms in patients with TRLLD, defined as ≥ 1 unsuccessful treatments. The first study with a large sample size and low risk of bias found no differences on remission rates between IPT and DCM with escitalopram (Reynolds et al., 2010). The second study, which was of smaller size, concluded no significant differences between the immediate and delayed BBTI groups at post-treatment (Gebara et al., 2019). We also identified a small study reporting on a computerized cognitive remediation intervention which improved depressive symptoms compared to an active control condition. However, all studies so far remain small therefore any conclusions remain limited (Morimoto et al., 2020). Given the insufficient level of evidence for these interventions future large-scale studies focusing on psychological or cognitive remediation interventions are urgently needed.

4.5. Side effects

Based on the limited available data we can conclude that aripiprazole and ketamine (either in nasal spray form or subcutaneous ketamine HCl) appear to be the most well-tolerated (Lenze et al., 2023). Three studies assessed the side effects of rTMS; one of these studies concluded that there were no serious side effects associated with rTMS (Zhao et al., 2019). The remaining two studies reported that 56 % of patients experienced headaches after a rTMS treatment (Blumberger et al., 2022; Kaster et al., 2018). ECT was associated with a high risk of cognitive impairment and post seizure delirium (Luzny, 2013; Stoppe et al., 2006).

Further large-scale studies are needed to conduct a risk-benefit analysis of treatments for TRLLD. Studies including measurements of blood levels could also be particularly valuable, considering inter-individual variations in pharmacokinetics and pharmacodynamics (Hilmer, 2021). This is especially relevant for older patients, who often take multiple medications (Toh et al., 2023), some of which may act as enzyme inhibitors. Addressing these factors will enhance our understanding of treatment effectiveness and safety in older patients, helping to mitigate potential adverse effects.

4.6. Effect Size

This systematic review highlights that effect sizes reported across studies assessing changes in depressive symptoms from baseline to post-treatment were generally small. Adding lithium versus phenelzine showed a large effect size however, due to the small sample size, this effect was not found to be significant, while other pharmacological interventions were associated with small effects. Among non-pharmacological treatments, nCCR targeting CCD versus an active control showed a moderate to large effect size, the other non-pharmacological interventions had smaller effects. Adding aripiprazole was associated with a small effect size across different studies (vs. placebo or other treatments). However, these results should be interpreted with caution given the small sample sizes and variations in the populations studied.

4.7. Quality of the evidence

Risk of bias was unclear for multiple domains in several of the studies evaluated, with published information sometimes insufficient to determine risk of bias. Only the area of missing outcome data was judged as low risk, with most studies rated as high quality for this criterion. Several studies had uncertainties on the areas of random sequence generation, and allocation concealment, with some rated as high risk. Given that most studies to date remain small (George et al., 2017; Kaster et al., 2018; Kok et al., 2007; Stoppe et al., 2006), with risk of bias present across several domains, it will be important that future trials adhere to the highest standards of evaluation and reporting.

To assess the quality of the evidence, we employed the GRADE framework (Prime, 2024). The substantial heterogeneity in criteria for antidepressant resistance across studies, coupled with the absence of at least two studies with comparable comparators, limits grouping the studies under a common endpoint, such as change from baseline in depressive symptoms. Furthermore, not all studies assessed the same endpoint; for example, two studies (Luzny, 2013; Reynolds et al., 2010) evaluated remission and/or response rather than change to depressive symptoms from baseline.

Consequently, we conducted separate GRADE assessments for each study (or for individual steps within multi-phase studies). The results revealed considerable variability: five studies were rated as having very low certainty, four studies as low certainty, four studies as moderate certainty, with only one study—comprising two steps with distinct GRADE evaluations—achieving high certainty in both steps (see Table 1).

Overall, the certainty of evidence was rated as low or very low in the majority of studies. Future studies and evaluations of certainty of evidence are therefore likely to encounter the same challenges if the marked heterogeneity in pharmacological approaches and in the criteria used to define antidepressant resistance persists.

As mentioned earlier, we found a limited number of RCTs assessing the effectiveness of different interventions for TRLLD, with studies reporting on larger sample sizes urgently needed. However, this remains an important challenge in both research and clinical practice. Over the past decades, more than 500 RCTs have examined the effects of antidepressants or psychotherapies for depression, but only a few have specifically focused on TRD (Cuijpers et al., 2020). These trials often lack statistical power to determine who benefits most, limiting the reliability of individualized evidence. Varied outcome measures make it difficult to aggregate results, and most studies fail to assess long-term effects (Cuijpers et al., 2020).

There are also challenges that are unique to TRLLD. These relate to recruiting and retaining participants who are frail older people. Factors such as a lack of perceived benefit, poor health, and mobility problems often create significant barriers in recruiting older patients (Provencher et al., 2014). To address these issues, we propose international collaborations for the delivery of RCTs targeting TRLLD and leveraging registry datasets across nations to reflect "real-world" data. Additionally, addressing the heterogeneity of definitions of TRLLD, and standardizing outcomes are important objectives for future studies.

4.8. Strengths and limitations

This systematic review has some strengths as it represents the most updated synthesis since 2011, comprehensively covering the most relevant databases available to date. We included only RCTs, ensuring a high standard of evidence in our findings. Furthermore, multiple definitions of TRLLD were incorporated, allowing for a broader and more inclusive overview of the condition. However, there were several limitations, the most important being the small number of studies, and the variability in definitions and criteria of TRLLD used, which precluded a meta-analysis. This variability in definitions may have also led to the exclusion of some RCTs that otherwise might meet our inclusion criteria. The majority of studies included small sample sizes with only a few exceptions (Blumberger et al., 2022; Lenze et al., 2015; Lenze et al., 2023; Ochs-Ross et al., 2020; Reynolds et al., 2010; Zhao et al., 2019), which limits the robustness and generalizability of the findings. Another important limitation was the short duration of follow-up periods, with most studies reporting follow-up outcomes of only 4–7 weeks. This is an important limitation which presents significant challenges for understanding long-term outcomes of current treatments. We also found evidence of risk of bias across several studies, and different domains. Lastly, not all RCTs included were placebo-controlled, potentially impacting the consistency of comparisons.

4.9. Recommendations for future research

Currently, although guidelines do exist for TRD in the general adult population (American Psychological Association, 2019; Baba et al., 2022; Bennabi et al., 2019; Gebara et al., 2019; Rybak et al., 2021), there are currently no clinical guidelines for the treatment of TRLLD. Recently, recommendations have been published (Patrick et al., 2024; Steffens, 2024; Subramanian et al., 2023), however based on our review augmentation strategies with second-generation antipsychotics, lithium, or another antidepressant, as well as switching drug classes, are not based on high quality evidence. Non-pharmacological approaches like psychotherapy and lifestyle interventions which are valuable due to their low risk, are also urgently needed. Although ECT is widely used, our data show that it remains less studied, with limited clinical effectiveness data in older people with TRLLD, and preliminary evidence of a high risk of side effects.

A significant challenge for future research in the area is the lack of consensus regarding the definition of TRLLD, which necessitates an urgent review to advance research on the effectiveness of treatments for TRLLD. Similar efforts have already been undertaken by certain research groups studying TRD in older adults (Rybak et al., 2021).

5. Conclusion

Current evidence on the clinical effectiveness of both pharmacological and non-pharmacological treatments for TRLLD remains limited, with the quality of existing studies suboptimal. The absence of a standardized definition for TRLLD complicates the assessment of treatment effectiveness in clinical trials. Evidence from two adequately powered RCTs with a low risk of bias suggests that aripiprazole reduces depressive symptoms in patients with TRLLD, but this requires replication and longer term follow ups. Other treatments, including venlafaxine, rTMS, lithium, ketamine, and computerized cognitive remediation, show promise but require validation through larger, high-quality trials. There is an urgent need for well-designed RCTs to establish effective treatments and clinical guidelines for TRLLD.

Funding

Pozuelo Moyano, B. has received funding from the Department of Psychiatry at the Lausanne University Hospital, the Fondation Anna & André Livio Glauser and from the UNIL (Université de Lausanne) through the "Tremplin" grant from the "Bureau de l'Égalité" for academic advancement and research time. Howard, R., Reeves, S., and Orgeta, V. are supported by the UCLH NIHR Biomedical Research Centre.

Stewart, R., and Mueller, C. are part-funded by the NIHR Maudsley Biomedical Research Centre at the South London and Maudsley NHS Foundation Trust and King's College London. RS is also funded by i) the National Institute for Health Research (NIHR) Applied Research Collaboration South London (NIHR ARC South London) at King's College Hospital NHS Foundation Trust; ii) UKRI – Medical Research Council through the DATAMIND HDR UK Mental Health Data Hub (MRC reference: MR/W014386); iii) the UK Prevention Research Partnership (Violence, Health and Society; MR-VO49879/1), an initiative funded by UK Research and Innovation Councils, the Department of Health and Social Care (England) and the UK devolved administrations, and leading health research charities. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Young A.H., is supported by NIMH (USA); CIHR (Canada); NARSAD (USA); Stanley Medical Research Institute (USA); MRC (UK); Wellcome Trust (UK); Royal College of Physicians (Edinburgh); BMA (UK); UBC-VGH Foundation (Canada); WEDC (Canada); CCS Depression Research Fund (Canada); MSFHR (Canada); NIHR (UK); Janssen (UK) and EU Horizon 2020. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care. For the purposes of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Accepted Author Manuscript version arising from this submission.

CRediT authorship contribution statement

BPM: conceptualization, formal analysis, investigation, methodology, project administration, visualization, writing original draft and writing review and editing.
DGB: methodology, formal analysis and writing review and editing.
KJP: methodology, formal analysis and writing review and editing.
CM: conceptualization, writing review and editing.
AvG: conceptualization, writing review and editing.
PV: conceptualization, writing review and editing.
SRa: Formal analysis and writing review and editing.
RH: conceptualization, writing review and editing.
AY: writing review and editing.
RS: writing review and editing.
SRe: conceptualization, writing review and editing.
VO: conceptualization, formal analysis, investigation, methodology, project administration, visualization, writing original draft and writing review and editing.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests. Stewart, R. declares research support in the last 3 years from GSK and Takeda. Young, A.H. declares the following competing interests: i) Employed by King's College London; ii) Honorary Consultant South London and Maudsley NHS Foundation Trust (NHS UK); iii) Editor of Journal of Psychopharmacology and Deputy Editor, BJPsych Open; iv) Paid lectures and advisory boards for the following companies with drugs used in affective and related disorders: Flow Neuroscience, Novartis, Roche, Janssen, Takeda, Noema pharma, Compass, Astrazenaca, Boehringer Ingelheim, Eli Lilly, LivaNova, Lundbeck, Sunovion, Servier, Allegan, Bionomics, Sumitomo Dainippon Pharma, Sage, Neurocentrx, Otsuka; v) Principal Investigator in the Restore-Life VNS registry study funded by LivaNova.Principal Investigator on ESKETINTRD3004: "An Open-label, Long-term, Safety and Efficacy Study of Intranasal Esketamine in Treatment-resistant Depression."; vi) Principal Investigator on "The Effects of Psilocybin on Cognitive

Function in Healthy Participants"; vii) Principal Investigator on "The Safety and Efficacy of Psilocybin in Participants with Treatment-Resistant Depression (P-TRD); viii) "Principal Investigator on "A Double-Blind, Randomized, Parallel-Group Study with Quetiapine Extended Release as Comparator to Evaluate the Efficacy and Safety of Seltorexant 20 mg as Adjunctive Therapy to Antidepressants in Adult and Elderly Patients with Major Depressive Disorder with Insomnia Symptoms Who Have Responded Inadequately to Antidepressant Therapy.'' (Janssen); ix) Principal Investigator on "An Open-label, Long-term, Safety and Efficacy Study of Aticaprant as Adjunctive Therapy in Adult and Elderly Participants with Major Depressive Disorder (MDD).'' (Janssen); x) Principal Investigator on "A Randomized, Double-blind, Multicentre, Parallel-group, Placebo-controlled Study to Evaluate the Efficacy, Safety, and Tolerability of Aticaprant 10 mg as Adjunctive Therapy in Adult Participants with Major Depressive Disorder (MDD) with Moderate-to-severe Anhedonia and Inadequate Response to Current Antidepressant Therapy.'' xi) Principal Investigator on "A Study of Disease Characteristics and Real-life Standard of Care Effectiveness in Patients with Major Depressive Disorder (MDD) With Anhedonia and Inadequate Response to Current Antidepressant Therapy Including an SSRI or SNR.'' (Janssen); xii) UK Chief Investigator for Compass; COMP006 & COMP007 studies; xiii) UK Chief Investigator for Novartis MDD study MIJ821A12201. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.arr.2025.102710.

Appendix B. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.arr.2025.102710.

Data availability

Data will be made available on request.

References

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders, 5th ed, Washington, DC.

American Psychological Association, B.M.A.S.S.J.B.M.P.F.B.G.M.V.P.W.K.M.F.N.M.A., 2019. Clinical practice guideline for the treatment of depression across three age cohorts.

Baba, H., Kito, S., Nukariya, K., Takeshima, M., Fujise, N., Iga, J., Oshibuchi, H., Kawano, M., Kimura, M., Mizukami, K., Mimura, M., 2022. Guidelines for diagnosis and treatment of depression in older adults: A report from the Japanese Society of mood disorders. Psychiatry Clin. Neurosci. 76, 222–234.

Bennabi, D., Charpeaud, T., Yrondi, A., Genty, J.-B., Destouches, S., Lancrenon, S., Alaïli, N., Bellivier, F., Bougerol, T., Camus, V., 2019. Clinical guidelines for the management of treatment-resistant depression: French recommendations from experts, the French Association for Biological Psychiatry and Neuropsychopharmacology and the fondation FondaMental. BMC Psychiatry 19, 1–12.

Benson, C., Szukis, H., Sheehan, J.J., Alphs, L., Yuce, H., 2020. An Evaluation of the Clinical and Economic Burden Among Older Adult Medicare-Covered Beneficiaries With Treatment-Resistant Depression. Am. J. Geriatr. Psychiatry 28, 350–362.

Blaszczyk, A.T., Mathys, M., Le, J., 2023. A Review of Therapeutics for Treatment-Resistant Depression in the Older Adult. Drugs Aging 40, 785–813.

Blumberger, D.M., Mulsant, B.H., Thorpe, K.E., McClintock, S.M., Konstantinou, G.N., Lee, H.H., Nestor, S.M., Noda, Y., Rajji, T.K., Trevizol, A.P., Vila-Rodriguez, F., Daskalakis, Z.J., Downar, J., 2022. Effectiveness of Standard Sequential Bilateral Repetitive Transcranial Magnetic Stimulation vs Bilateral Theta Burst Stimulation in Older Adults with Depression: The FOUR-D Randomized Noninferiority Clinical Trial. JAMA Psychiatry 79, 1065–1073.

Buchalter, E.L.F., Oughli, H.A., Lenze, E.J., Dixon, D., Miller, J.P., Blumberger, D.M., Karp, J.F., Reynolds, C.F., 3rd, Mulsant, B.H., 2019. Predicting Remission in Late-Life Major Depression: A Clinical Algorithm Based Upon Past Treatment History. J. Clin. Psychiatry 80.

Canadian Coalition for Seniors' Mental Health, 2021. Guideline Update: Canadian Guidelines on Prevention, Assessment and Treatment of Depression Among Older Adults. Canadian Coalition for Seniors' Mental Health Markham, Ontario.

Cappon, D., den Boer, T., Jordan, C., Yu, W., Metzger, E., Pascual-Leone, A., 2022. Transcranial magnetic stimulation (TMS) for geriatric depression. Ageing Res Rev. 74.

Cooper, C., Katona, C., Lyketsos, K., Blazer, D., Brodaty, H., Rabins, P., De Mendonça Lima, C.A., Livingston, G., 2011. A systematic review of treatments for refractory depression in older people. Am. J. Psychiatry 168, 681–688.

Cuijpers, P., Stringaris, A., Wolpert, M., 2020. Treatment outcomes for depression: challenges and opportunities. Lancet Psychiatry 7, 925–927.

Edwards, S.J., Hamilton, V., Nherera, L., Trevor, N., 2013. Lithium or an atypical antipsychotic drug in the management of treatment-resistant depression: a systematic review and economic evaluation. Health Technol. Assess. 17, 1–190.

Gareri, P., De Fazio, P., De Sarro, G., 2002. Neuropharmacology of depression in aging and age-related diseases. Ageing Res Rev. 1, 113–134.

Gartlehner, G., Hansen, R.A., Nissman, D., Lohr, K.N., Carey, T.S., 2006. A simple and valid tool distinguished efficacy from effectiveness studies. J. Clin. Epidemiol. 59, 1040–1048.

Gaynes, B.N., Lux, L., Gartlehner, G., Asher, G., Forman-Hoffman, V., Green, J., Boland, E., Weber, R.P., Randolph, C., Bann, C., Coker-Schwimmer, E., Viswanathan, M., Lohr, K.N., 2020. Defining treatment-resistant depression. Depress Anxiety 37, 134–145.

Gebara, M.A., DiNapoli, E.A., Lederer, L.G., Bramoweth, A.D., Germain, A., Kasckow, J. W., Karp, J.F., 2019. Brief behavioral treatment for insomnia in older adults with late-life treatment-resistant depression and insomnia: a pilot study. Sleep. Biol. Rhythms 17, 287–295.

George, D., Gálvez, V., Martin, D., Kumar, D., Leyden, J., Hadzi-Pavlovic, D., Harper, S., Brodaty, H., Glue, P., Taylor, R., Mitchell, P.B., Loo, C.K., 2017. Pilot Randomized Controlled Trial of Titrated Subcutaneous Ketamine in Older Patients with Treatment-Resistant Depression. Am. J. Geriatr. Psychiatry 25, 1199–1209.

Godwin, M., Ruhland, L., Casson, I., MacDonald, S., Delva, D., Birtwhistle, R., Lam, M., Seguin, R., 2003. Pragmatic controlled clinical trials in primary care: the struggle between external and internal validity. BMC Med Res Method. 3, 1–7.

Gutsmiedl, K., Krause, M., Bighelli, I., Schneider-Thoma, J., Leucht, S., 2020. How well do elderly patients with major depressive disorder respond to antidepressants: a systematic review and single-group meta-analysis. BMC Psychiatry 20, 1–12.

Hamm, M.E., Karp, J.F., Lenard, E., Dawdani, A., Lavretsky, H., Lenze, E.J., Mulsant, B. H., Reynolds, C.F., Roose, S.P., Brown, P.J., 2022. "What else can we do?"-Provider perspectives on treatment-resistant depression in late life. J. Am. Geriatr. Soc. 70, 1190–1197.

Hilmer, S.N., 2021. Bridging geriatric medicine, clinical pharmacology and ageing biology to understand and improve outcomes of medicines in old age and frailty. Ageing Res Rev. 71, 101457.

Kaster, T.S., Daskalakis, Z.J., Noda, Y., Knyahnytska, Y., Downar, J., Rajji, T.K., Levkovitz, Y., Zangen, A., Butters, M.A., Mulsant, B.H., Blumberger, D.M., 2018. Efficacy, tolerability, and cognitive effects of deep transcranial magnetic stimulation for late-life depression: a prospective randomized controlled trial. Neuropsychopharmacology 43, 2231–2238.

Kern, D.M., Canuso, C.M., Daly, E., Johnson, J.C., Fu, D.J., Doherty, T., Blauer-Peterson, C., Cepeda, M.S., 2023. Suicide-specific mortality among patients with treatment-resistant major depressive disorder, major depressive disorder with prior suicidal ideation or suicide attempts, or major depressive disorder alone. Brain Behav. 13, e3171.

Kok, R.M., Vink, D., Heeren, T.J., Nolen, W.A., 2007. Lithium augmentation compared with phenelzine in treatment-resistant depression in the elderly: An open, randomized, controlled trial. J. Clin. Psychiatry 68, 1177–1185.

Lenze, E.J., Mulsant, B.H., Blumberger, D.M., Karp, J.F., Newcomer, J.W., Anderson, S.J., Dew, M.A., Butters, M.A., Stack, J.A., Begley, A.E., Reynolds, C.F., 2015. Efficacy, safety, and tolerability of augmentation pharmacotherapy with aripiprazole for treatment-resistant depression in late life: A randomised, double-blind, placebo-controlled trial. Lancet 386, 2404–2412.

Lenze, E.J., Mulsant, B.H., Roose, S.P., Lavretsky, H., Reynolds, C.F., Blumberger, D.M., Brown, P.J., Cristancho, P., Flint, A.J., Gebara, M.A., Gettinger, T.R., Lenard, E., Miller, J.P., Nicol, G.E., Oughli, H.A., Pham, V.T., Rollman, B.L., Yang, L., Karp, J.F., 2023. Antidepressant Augmentation versus Switch in Treatment-Resistant Geriatric Depression. N. Engl. J. Med 388, 1067–1079.

Lipsey, M.W., Wilson, D.B., 2001. Practical meta-analysis. Sage Publications, Thousand Oaks, CA.

Luzny, J., 2013. P. 2.f.009 Treating pharmacoresistant depression in psychogeriatry – should electroconvulsive therapy or intravenous seropram be used? Eur. Neuropsychopharmacol. 23, S400–S401.

Mazeh, D., Shahal, B., Aviv, A., Zemishlani, H., Barak, Y., 2007. A randomized, single-blind, comparison of venlafaxine with paroxetine in elderly patients suffering from resistant depression. Int Clin. Psychopharmacol. 22, 371–375.

McIntyre, R.S., Alsuwaidan, M., Baune, B.T., Berk, M., Demyttenaere, K., Goldberg, J.F., Gorwood, P., Ho, R., Kasper, S., Kennedy, S.H., Ly-Uson, J., Mansur, R.B., McAllister-Williams, R.H., Murrough, J.W., Nemeroff, C.B., Nierenberg, A.A., Rosenblat, J.D., Sanacora, G., Schatzberg, A.F., Shelton, R., Stahl, S.M., Trivedi, M.H., Vieta, E., Vinberg, M., Williams, N., Young, A.H., Maj, M., 2023. Treatment-resistant depression: definition, prevalence, detection, management, and investigational interventions. World Psychiatry 22, 394–412.

Morimoto, S.S., Altizer, R.A., Gunning, F.M., Hu, W., Liu, J., Cote, S.E., Nitis, J., Alexopoulos, G.S., 2020. Targeting Cognitive Control Deficits With Neuroplasticity-Based Computerized Cognitive Remediation in Patients With Geriatric Major

Depression: A Randomized, Double-Blind, Controlled Trial. Am. J. Geriatr. Psychiatry 28, 971–980.
Nuñez, N.A., Joseph, B., Pahwa, M., Kumar, R., Resendez, M.G., Prokop, L.J., Veldic, M., Seshadri, A., Biernacka, J.M., Frye, M.A., Wang, Z., Singh, B., 2022. Augmentation strategies for treatment resistant major depression: A systematic review and network meta-analysis. J. Affect Disord. 302, 385–400.
Ochs-Ross, R., Daly, E.J., Zhang, Y., Lane, R., Lim, P., Morrison, R.L., Hough, D., Manji, H., Drevets, W.C., Sanacora, G., Steffens, D.C., Adler, C., McShane, R., Gaillard, R., Wilkinson, S.T., Singh, J.B., 2020. Efficacy and Safety of Esketamine Nasal Spray Plus an Oral Antidepressant in Elderly Patients With Treatment-Resistant Depression—TRANSFORM-3. Am. J. Geriatr. Psychiatry 28, 121–141.
Patrick, R.E., Dickinson, R.A., Gentry, M.T., Kim, J.U., Oberlin, L.E., Park, S., Principe, J. L., Teixeira, A.L., Weisenbach, S.L., 2024. Treatment resistant late-life depression: A narrative review of psychosocial risk factors, non-pharmacological interventions, and the role of clinical phenotyping. J. Affect Disord. 356, 145–154.
Prime, M.Ua.E., 2024. GRADEpro GDT: GRADEpro Guidel. Dev. Tool. [Softw.].
Provencher, V., Mortenson, W.B., Tanguay-Garneau, L., Bélanger, K., Dagenais, M., 2014. Challenges and strategies pertaining to recruitment and retention of frail elderly in research studies: a systematic review. Arch. Gerontol. Geriatr. 59, 18–24.
Reynolds, C.F., 3rd, Dew, M.A., Martire, L.M., Miller, M.D., Cyranowski, J.M., Lenze, E., Whyte, E.M., Mulsant, B.H., Pollock, B.G., Karp, J.F., Gildengers, A., Szanto, K., Dombrovski, A.Y., Andreescu, C., Butters, M.A., Morse, J.Q., Houck, P.R., Bensasi, S., Mazumdar, S., Stack, J.A., Frank, E., 2010. Treating depression to remission in older adults: a controlled evaluation of combined escitalopram with interpersonal psychotherapy versus escitalopram with depression care management. Int J. Geriatr. Psychiatry 25, 1134–1141.
Riedel, M., Möller, H.-J., Obermeier, M., Schennach-Wolff, R., Bauer, M., Adli, M., Kronmüller, K., Nickel, T., Brieger, P., Laux, G., 2010. Response and remission criteria in major depression–a validation of current practice. J. Psychiatr. Res 44, 1063–1068.
Roose, S.P., Brown, P.J., 2022. What we Know, What We Do Not Know, and What We May Know Soon About Interventions for Treatment Resistant Depression in Late-Life? Am. J. Geriatr. Psychiatry 30, 557–559.
Rybak, Y.E., Lai, K.S.P., Ramasubbu, R., Vila-Rodriguez, F., Blumberger, D.M., Chan, P., Delva, N., Giacobbe, P., Gosselin, C., Kennedy, S.H., Iskandar, H., McInerney, S., Ravitz, P., Sharma, V., Zaretsky, A., Burhan, A.M., 2021. Treatment-resistant major depressive disorder: Canadian expert consensus on definition and assessment. Depress Anxiety 38, 456–467.
Soysal, P., Veronese, N., Thompson, T., Kahl, K.G., Fernandes, B.S., Prina, A.M., Solmi, M., Schofield, P., Koyanagi, A., Tseng, P.T., Lin, P.Y., Chu, C.S., Cosco, T.D., Cesari, M., Carvalho, A.F., Stubbs, B., 2017. Relationship between depression and frailty in older adults: A systematic review and meta-analysis. Ageing Res Rev. 36, 78–87.
Steffens, D.C., 2024. Treatment-Resistant Depression in Older Adults. N. Engl. J. Med 390, 630–639.
Stoppe, A., Louzã, M., Rosa, M., Gil, G., Rigonatti, S., 2006. Fixed high-dose electroconvulsive therapy in the elderly with depression: a double-blind, randomized comparison of efficacy and tolerability between unilateral and bilateral electrode placement. J. ect 22, 92–99.
Subramanian, S., Oughli, H.A., Gebara, M.A., Palanca, B.J.A., Lenze, E.J., 2023. Treatment-Resistant Late-Life Depression: A Review of Clinical Features, Neuropsychology, Neurobiology, and Treatment. Psychiatr. Clin. North Am. 46, 371–389.
Thase, M.E., Rush, A.J., 1997. When at first you don't succeed: sequential strategies for antidepressant nonresponders. J. Clin. Psychiatry 58 (13), 23–29.
Toh, J.J.Y., Zhang, H., Soh, Y.Y., Zhang, Z., Wu, X.V., 2023. Prevalence and health outcomes of polypharmacy and hyperpolypharmacy in older adults with frailty: a systematic review and meta-analysis. Ageing Res Rev. 83, 101811.
Tunis, S.R., Stryer, D.B., Clancy, C.M., 2003. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. Jama 290, 1624–1632.
World Health Organization, 2004. ICD-10: international statistical classification of diseases and related health problems: tenth revision. World Health Organization.
Zhao, X., Li, Y., Tian, Q., Zhu, B., Zhao, Z., 2019. Repetitive transcranial magnetic stimulation increases serum brain-derived neurotrophic factor and decreases interleukin-1β and tumor necrosis factor-α in elderly patients with refractory depression. J. Int Med Res 47, 1848–1855.
Zhou PhD, X., Keitner PhD, G.I., Qin, B.M.D., Ravindran PhD, A.V., Bauer PhD, M., Del Giovane PhD, C., Zhao PhD, J., Liu, Y.M.D., Fang PhD, Y., Zhang PhD, Y., Xie, P., 2015. Atypical Antipsychotic Augmentation for Treatment-Resistant Depression: A Systematic Review and Network Meta-Analysis. Int J. Neuropsychopharmacol. 18.