CADTH

TECHNOLOGY REVIEW: FOCUSED CRITICAL APPRAISAL

Drugs for Major Depressive Disorder

Service Line:	Technology Review
Issue:	24
Publication Date:	February 2020
Report Length:	19 Pages

CADTH

Authors: Mina Tadrous, Louis de Léséleuc

Cite As: *Drugs for Major Depressive Disorder*. Ottawa: CADTH; 2020 Feb. (CADTH technology review: focused critical appraisal; no. 24).

Acknowledgments: Mike Innes, Tarry Ahuja, Lauren Bresee

ISSN: 2369-7385 (online)

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada's health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.

Funding: CADTH receives funding from Canada's federal, provincial, and territorial governments, with the exception of Quebec.

CADTH

Table of Contents

Tables

Figures

Background

Major depressive disorder (MDD) is a common and recurrent mental health disorder. MDD is characterized by symptoms of persistent low mood; changes in appetite and sleep; fatigue; loss of motivation, interest, or pleasure; and feelings of worthlessness. It can also be associated with a substantial loss in productivity, reduced quality of life, and increased mortality from suicide.[1] Depression is the leading cause of disease disability and the third-leading cause of disease burden worldwide.[2,3] MDD is one of the more common mental health disorders, with one in 20 (5.4%) Canadians aged 15 years and older having reported symptoms that met the criteria for a mood disorder in the previous 12 months.[4] Further, one in eight adults (12.6%) identified symptoms that met the criteria for a mood disorder at some point during their lifetime.[4] Rates of depression are higher among women, occurring twice as often.

Multiple classes of drugs have been developed to treat depression, with many of them having differential effects on neurotransmitters. The majority of drugs to treat depression have targeted serotonin and norepinephrine. Examples of major antidepressant classes include tricyclic antidepressants (TCA), selective serotonin reuptake inhibitors, and serotonin norepinephrine reuptake inhibitors. Oral antidepressants are the most commonly prescribed drugs for mental health in Canada and some of the most widely prescribed drugs overall in the country.[5] According to Statistics Canada, about 9% of Canadians are dispensed antidepressants, higher than the Organisation for Economic Co-operation and Development average.[6] The management of MDD is regarded as highly personalized; providing access to a broad diversity of antidepressants may increase the likelihood of successfully treating MDD. In line with this principle, the Canadian Network for Mood and Anxiety Treatments (CANMAT) guidelines recommend selecting from a total of fifteen different antidepressants for initiating a first-line pharmacological treatment of major depression.[7] Importantly, remission rates of patients with depression given first-line pharmacotherapy is about 50%.[8] On average, it takes six to 12 weeks of therapy for patients to achieve remission.[9]

When new antidepressants become available in Canada, their comparative clinical effectiveness, safety, and cost-effectiveness relative to existing drugs must be established before they are reimbursed by public drug plans. Because of the large number of drugs indicated for the first-line treatment of MDD, a broad comparative review of antidepressant therapeutic attributes would help determine the relative value of existing and new antidepressants, and thus complement reviews focused on single new drugs. A comprehensive review performed by Cipriani et al. in 2018[10] was identified as potentially amenable to such purposes and will be the subject of a Focused Critical Appraisal by CADTH.

The objective of a CADTH Focused Critical Appraisal is to examine the methodology, scientific rigour, and clinical findings of a published study of importance.

Study Under Review[a]

Ciprani et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. The Lancet 2018[10]

Description of Study

Objective of Study

The primary aim of the study was to conduct an updated systematic review and network meta-analysis (NMA) to inform clinical practice by comparing different antidepressants for the acute treatment of adults with unipolar MDD.

Study Description

This study is a systematic review and NMA of double-blind, randomized controlled trial (RCT). The study authors compared the efficacy and acceptability of 21 antidepressants to either placebo or monotherapy with another antidepressant agent for the acute treatment of unipolar MDD in adults aged 18 years and older.

Methods

Literature Search

The authors searched a number of databases, including: “Cochrane Central Register of Controlled Trials, CINAHL, Embase, LILACS database, MEDLINE, MEDLINE In-Process, PsycINFO, AMED, the UK National Research Register, and PSYNDEX.” The time frame of the search was from the inception of each database until January 8, 2016, with no language restriction. In addition to the database search, they also conducted an expanded search of “published, unpublished, and ongoing RCTs in international trial registers, websites of drug approval agencies, and key scientific journals in the field.” As well, they contacted drug manufacturers and study authors to request unpublished information including missing data from included studies and unpublished pre- and post-market studies.

Table 1: Population, Interventions, Comparators, Outcomes, and Study Design Criteria for Study Inclusion

Criteria	Description
Population	Adults (≥18 years old) with a primary diagnosis of MDD based on “standard operationalised diagnostic criteria (Feighner criteria, Research Diagnostic Criteria, DSM-III, DSM-III-R, DSM-IV, DSM-5, and ICD-10)”
Interventions	Any of the following active antidepressants as oral monotherapy: • second-generation antidepressants with regulatory approval in the US, Europe, or Japan: agomelatine, bupropion, citalopram, desvenlafaxine, duloxetine, escitalopram, fluoxetine, fluvoxamine, levomilnacipran, milnacipran, mirtazapine, paroxetine, reboxetine, sertraline, venlafaxine, vilazodone, and vortioxetine • tricyclic antidepressants (included in the WHO Model List of Essential Medicines): amitriptyline and clomipramine

[a] Note to the reader: In this report, text enclosed between double quotation marks “” is taken verbatim from the published article under review.

Criteria	Description
	• trazodone and nefazodone
Comparators	Any active antidepressant monotherapy or placebo
Outcomes	*Primary Efficacy Outcome:* "response rate measured by the total number of patients who had a reduction of ≥ 50% of the total score on a standardized observer-rating scale for depression" *Primary Acceptability Outcome:* "treatment discontinuation measured by the proportion of patients who withdrew for any reason" *Secondary outcomes:* • Efficacy — defined as the end point score from the Hamilton Depression Rating Scale or Montgomery-Åsberg Depression Rating Scale • Remission — defined as the proportion of patients with remission of depressive symptoms • Tolerability — defined as the proportion of patients who discontinue due to an adverse event. Outcomes were measured at 8 weeks, if possible (range 4 to 12 weeks)
Study Design and Factors	• Double-blind RCTs • Included trials that allowed rescue medication (usually benzodiazepines or sedative hypnotic agents) for both the intervention and comparison groups • Only included doses within theraputic range • Excluded quasi-randomized trials, crossover trials, and cluster randomized trials • Excluded trials that were incomplete • Excluded trials with "20% or more of participants with bipolar disorder, psychotic depression, or treatment-resistant depression; or patients with a serious concomitant medical illness"
Language	No language restriction
Search Period	Database inception to January 8, 2016

MDD = major depressive disorder; RCT = randomized controlled trial.

Eligibility Criteria and Study Selection

Studies were eligible for inclusion if they were a double-blind RCT that enrolled patients with a primary diagnosis of MDD. The diagnosis had to be based on "standard operationalised diagnostic criteria" from the DSM-III, DSM-III-R, DSM-IV, DSM-5, research diagnostic criteria, ICD-10, or Feighner criteria. Data were only included for groups for whom the antidepressant was dosed within the established therapeutic range. Studies were excluded if they were quasi-randomized trials, crossover trials, were incomplete, or "included 20% or more of participants with bipolar disorder, psychotic depression, or treatment-resistant depression; or patients with a serious concomitant medical illness."

Data Extraction

Six pairs of investigators independently selected the studies for inclusion, reviewed materials, extracted the data, and completed the study quality assessments. Discrepancies "were resolved by consensus and arbitration by a panel of investigators within the review team."

Comparators

Comparators of interest were placebo and other currently available treatments:

- agomelatine[b]
- amitriptyline
- bupropion
- citalopram
- clomipramine
- desvenlafaxine
- duloxetine
- escitalopram
- fluoxetine
- fluvoxamine
- levomilnacipran
- milnacipran[a]
- mirtazapine
- nefazodone[a]
- paroxetine
- reboxetine[a]
- sertraline
- trazodone
- venlafaxine
- vilazodone
- vortioxetine

Outcomes

The primary outcomes were efficacy and acceptability. Efficacy was based on the response rate, which was defined as "the total number of patients who had a reduction of ≥ 50% of the total score on a standardized observer-rating scale for depression." If the response rate was not reported, it was calculated using a validated imputation method using the last observation carried forward in the context of a risk of bias assessment.[11] Acceptability was defined as the proportion of patients discontinuing treatment for any reason. The secondary outcomes of the study included the depression score at the end point, the rate of remission, and "the proportion of patients who dropped out early because of adverse events." When studies reported more than one standardized rating scale to measure depressive symptoms, the study authors "used a predefined hierarchy, based on psychometric properties and consistency of use across included trials." This hierarchy placed the Hamilton Depression Rating Scale (HAM-D) and the Montgomery-Åsberg Depression Rating Scale (MADRS) at the top of the hierarchy.[12] No specific HAM-D scoring was preferred (i.e., 17- or 24-item scale). Outcomes were reported at eight weeks, when possible. If information at eight weeks was not available, data ranging between four and 12 weeks was used with whatever data were closest to eight weeks. If time frames were of equal distance from eight weeks, the longer time frame was reported.

Quality Assessment of Included Studies

The risk of bias in the included studies was assessed using the *Cochrane Handbook for Systematic Reviews of Interventions*.[13] The study authors also assessed the confidence in the evidence that was used to calculate the primary outcome estimates with the Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework.[14] The risk of bias was reported for each component — specifically, sequence generation, allocation concealment, blinding of participants, blinding of therapist, blinding of assessors, selective reporting, and attrition bias. The risk of bias was categorized as low, moderate, and high using the following definition: "Studies were classified as having low risk of bias if

[b] Not available in Canada.

none of the domains above was rated as high risk of bias and three or less were rated as unclear risk; moderate if one was rated as high risk of bias or none was rated as high risk of bias but four or more were rated as unclear risk, and all other cases were assumed to pertain to high risk of bias." This information was planned to be used for sensitivity analyses.

Analytical Methods

The study authors conducted both pairwise and NMAs. They reported estimated summary odds ratios for dichotomous outcomes and standardized mean differences for continuous outcomes. The NMAs were conducted with group-level data using a random-effects model. The authors assumed that the amount of heterogeneity between all treatment comparisons was equal for the NMAs, and they assessed the amount of heterogeneity by comparing "the posterior distribution of the estimated heterogeneity variance with its predictive distribution." They assessed the transitivity assumption by comparing the distribution of variables that could act as potential effect modifiers. Effect modifiers included: year of publication, sponsorship, dosing schedule, probability of receiving placebo, baseline disease severity, multi-centre, dose ranges, and unpublished data. Funnel plots were used to assess if results differed based on the precision of the included trials. Lastly, they evaluated consistency in the network "using the design-by-treatment test and by separating direct evidence from indirect evidence."

For the two primary outcomes, "subgroup analyses and network meta-regression using study year, sponsorship, depressive severity at baseline, dosing schedule, study precision (i.e., small study effect), and novelty effect" were conducted to assess the robustness of results. Novelty effect was defined based on the comparator; if the agent was compared to placebo or an older agent, it was defined as novel. NMAs used "binomial likelihood for dichotomous outcomes, uninformative prior distributions for the treatment effects, and a minimally informative prior distribution for the common heterogeneity SD [standard deviation]." Models assumed uninformative priors for all meta-regression coefficients, and model convergence was evaluated using the Brooks and Gelman, and Gelman and Rubin, diagnostic and the visual inspection of three chains. The analysis was completed using OpenBUGS (version 3.2.2) and replicated in R (version 3.4.0). "Statistical evaluation of inconsistency and production of network graphs and result figures were done using the network and network graphs packages in Stata (version 14.2)." All code was shared as part of the protocol.

Results

The systematic review identified a total of 28,552 unique publications. Overall, 522 trials met the criteria for inclusion. All of the included trials were randomized, double-blind, parallel-group design clinical trials. They included studies conducted from 1979 to 2016 across 21 different antidepressants. Of the trials, 304 (58%) were placebo-controlled. The majority (83%) of studies were multi-centre studies, with 48% recruiting patients from North America. Overall, "46 (9%) of 522 trials were rated as high risk of bias, 380 (73%) trials as moderate, and 96 (18%) as low."

The 522 trials included a total of 116,447 participants: 87,052 patients were randomized to active treatment and 29,425 were randomly assigned placebo. The average study size was 224 patients, and the median duration of the included studies was eight weeks (interquartile range: 6 to 8) (Table 2). The mean patient age was 44 (SD 9) years and 62.3% of patients were women. A total of 464 (89%) studies evaluated baseline depression with the HAM-D 17-item, and the mean baseline score was 25.7 (SD 3.97), indicating moderate-to-severe depression.

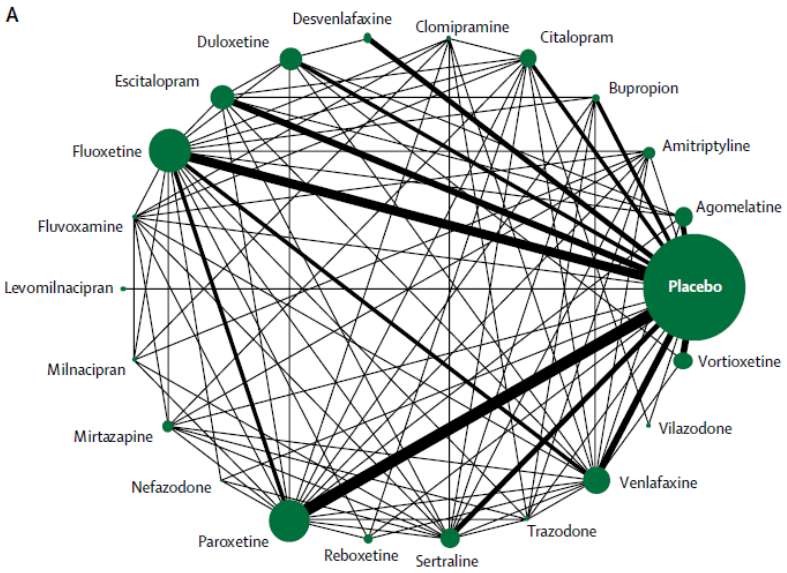
Table 2: Summary of Included Studies and Baseline Characteristics

Characteristics	
Number of studies	522
Total number of patients	116,447
Mean number of patients per study (SD)	224 (186)
Total number assigned to placebo	29,425 (25%)
Median duration of studies, weeks (IQR)	8 (6 to 8)
Multi-centre studies	391 (83%)
Proportion of outpatients	335 (77%)
Industry funding	409 (78%)
Unpublished information retrieved	274 (52%)
Mean age, years (SD)	44 years (9)
Proportion of females	62.3%

IQR = interquartile range; SD = standard deviation.

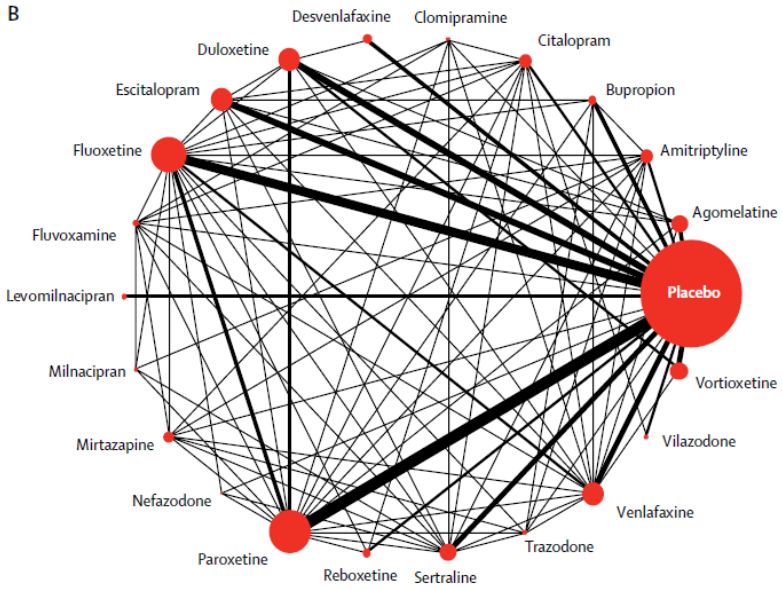
Reprinted (and modified) from Lancet, Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ, Ogawa Y et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. 2018 Apr 7;391(10128):1357-1366. https://doi.org/10.1016/S0140-6736(17)32802-7. Creative Commons Attribution License 4.0 http://creativecommons.org/licenses/by/4.0/

Figure 1: Network of Trials Included in the NMA for the Primary Efficacy Outcome



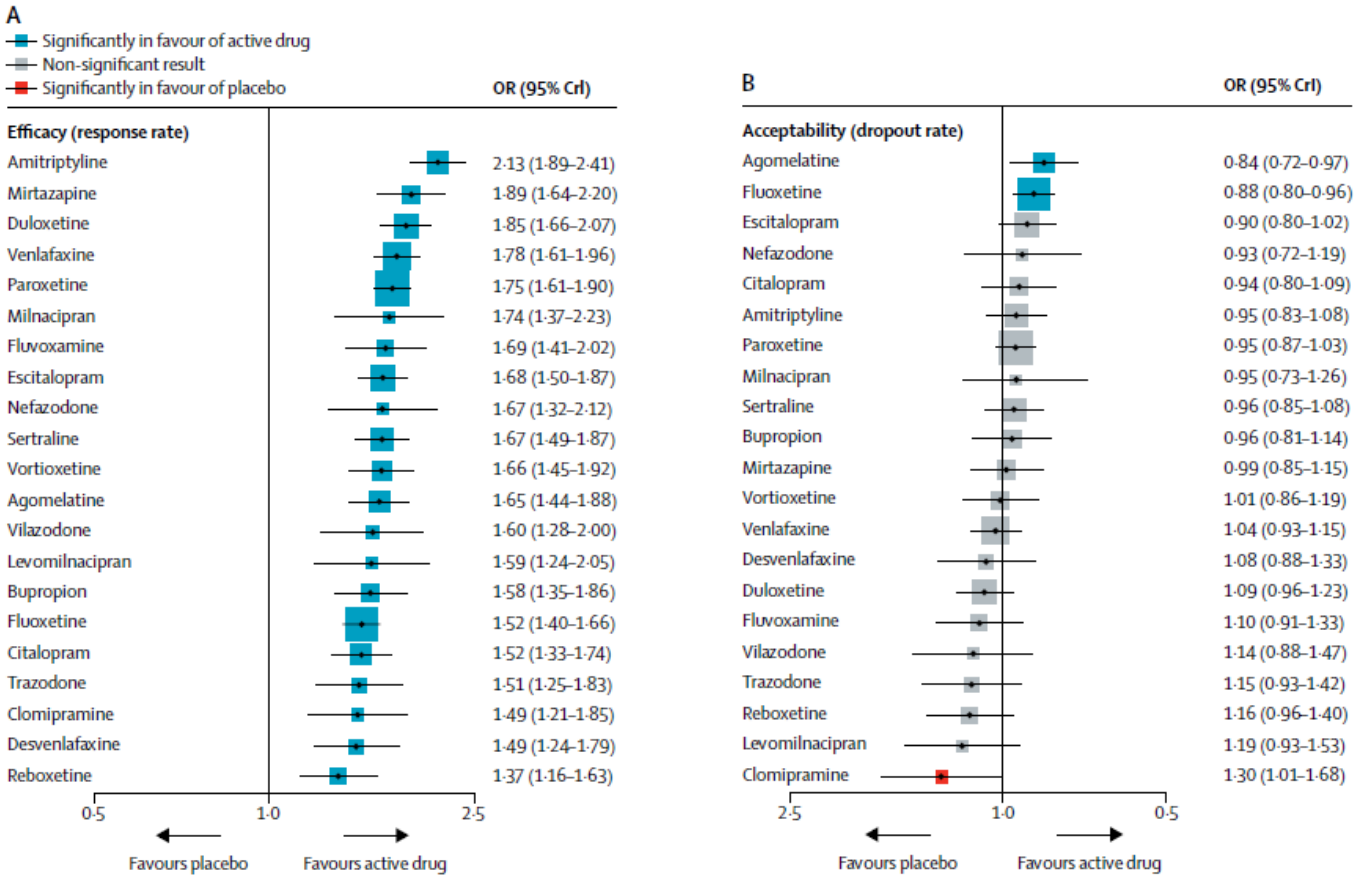
Reprinted from Lancet, Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ, Ogawa Y et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. 2018 Apr 7;391(10128):1357-1366. https://doi.org/10.1016/S0140-6736(17)32802-7. Creative Commons Attribution License 4.0 http://creativecommons.org/licenses/by/4.0/

Figure 2: Network of Trials Included in the NMA for the Primary Acceptability Outcome



Reprinted from Lancet, Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ, Ogawa Y et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. 2018 Apr 7;391(10128):1357-1366. https://doi.org/10.1016/S0140-6736(17)32802-7. Creative Commons Attribution License 4.0 http://creativecommons.org/licenses/by/4.0/

Figure 3: Forrest Plot of the Primary Efficacy (A) and Acceptability (B) Outcome Compared to Placebo



CrI = credible interval; OR = odds ratio.

Reprinted from Lancet, Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ, Ogawa Y et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. 2018 Apr 7;391(10128):1357-1366. https://doi.org/10.1016/S0140-6736(17)32802-7. Creative Commons Attribution License 4.0 http://creativecommons.org/licenses/by/4.0/

Figure 4: Relative Effect of Primary Outcomes of Efficacy and Acceptability (NMA of Head-to-Head Comparisons)

Efficacy (response rate) | Comparison | Acceptability (dropout rate)

Agom	0·72* (0·55–0·92)	0·80* (0·54–1·15)	0·89* (0·66–1·19)	0·57* (0·42–0·77)	0·62† (0·47–0·82)	0·97* (0·74–1·27)	0·85† (0·68–1·05)	0·69† (0·51–0·97)	0·79* (0·58–1·09)	0·81* (0·61–1·05)	0·70* (0·44–1·14)	0·81* (0·65–1·00)	0·53* (0·36–0·80)	0·86* (0·66–1·13)	0·69* (0·48–0·98)	0·74† (0·58–0·92)	1·24† (0·71–2·19)
0·96* (0·76–1·24)	**Amit**	1·10‡ (0·78–1·58)	1·23* (0·94–1·64)	0·79† (0·60–1·05)	0·87† (0·66–1·15)	1·35* (1·05–1·74)	1·18† (0·99–1·42)	0·97† (0·74–1·24)	1·10† (0·84–1·45)	1·12* (0·89–1·42)	0·98‡ (0·62–1·55)	1·12† (0·95–1·34)	0·74† (0·51–1·10)	1·20* (0·97–1·47)	0·96‡ (0·70–1·31)	1·02† (0·83–1·26)	1·72† (1·00–3·05)
0·87† (0·59–1·30)	0·91‡ (0·62–1·31)	**Bupr**	1·11‡ (0·76–1·67)	0·71† (0·49–1·07)	0·78† (0·53–1·18)	1·23* (0·84–1·80)	1·07‡ (0·76–1·50)	0·87‡ (0·59–1·30)	1·00‡ (0·66–1·49)	1·01† (0·70–1·47)	0·89‡ (0·51–1·54)	1·02‡ (0·73–1·43)	0·67† (0·42–1·08)	1·08‡ (0·75–1·56)	0·87‡ (0·57–1·30)	0·92‡ (0·66–1·30)	1·55† (0·85–2·94)
1·13* (0·88–1·47)	1·18* (0·93–1·49)	1·30† (0·88–1·93)	**Cita**	0·64† (0·47–0·87)	0·70* (0·51–0·95)	1·09* (0·85–1·42)	0·96* (0·76–1·21)	0·78* (0·57–1·06)	0·89* (0·64–1·23)	0·91† (0·68–1·21)	0·79‡ (0·49–1·32)	0·91* (0·71–1·17)	0·60† (0·41–0·87)	0·97‡ (0·74–1·25)	0·77* (0·53–1·13)	0·83† (0·64–1·07)	1·40† (0·78–2·48)
1·20* (0·91–1·59)	1·24† (0·98–1·58)	1·37† (0·93–2·04)	1·06* (0·82–1·38)	**Clom**	1·10† (0·80–1·51)	1·71* (1·27–2·29)	1·49† (1·16–1·90)	1·22† (0·88–1·67)	1·40† (1·00–1·92)	1·41* (1·05–1·91)	1·24‡ (0·76–2·00)	1·42† (1·12–1·79)	0·94‡ (0·62–1·41)	1·51† (1·15–1·96)	1·21† (0·83–1·73)	1·29† (0·99–1·67)	2·20† (1·22–3·90)
1·06* (0·82–1·37)	1·10† (0·84–1·42)	1·21† (0·81–1·81)	0·93* (0·71–1·22)	0·88† (0·66–1·18)	**Dulo**	1·56* (1·19–2·01)	1·37* (1·06–1·73)	1·12* (0·80–1·53)	1·28† (0·91–1·75)	1·30* (0·96–1·72)	1·13‡ (0·69–1·83)	1·30* (1·02–1·63)	0·86‡ (0·57–1·29)	1·38† (1·04–1·80)	1·10† (0·76–1·59)	1·18‡ (0·92–1·49)	1·99† (1·13–3·52)
0·90* (0·71–1·14)	0·93* (0·74–1·17)	1·03† (0·70–1·51)	0·79* (0·65–0·97)	0·75* (0·58–0·97)	0·85* (0·67–1·08)	**Escl**	0·87* (0·70–1·09)	0·71* (0·53–0·96)	0·81* (0·60–1·11)	0·83* (0·63–1·08)	0·72† (0·45–1·18)	0·83* (0·67–1·03)	0·55* (0·37–0·81)	0·88* (0·69–1·12)	0·70* (0·49–1·00)	0·75* (0·60–0·94)	1·27‡ (0·73–2·25)
1·20* (0·99–1·48)	1·25† (1·06–1·48)	1·38† (0·97–1·97)	1·06* (0·87–1·29)	1·00‡ (0·81–1·24)	1·14* (0·91–1·44)	1·34* (1·11–1·61)	**Fluo**	0·82* (0·64–1·04)	0·94* (0·72–1·20)	0·95* (0·77–1·16)	0·83† (0·54–1·30)	0·95* (0·83–1·09)	0·63† (0·44–0·90)	1·01† (0·84–1·21)	0·81* (0·60–1·09)	0·87† (0·74–1·01)	1·46† (0·85–2·53)
1·20* (0·91–1·61)	1·25† (0·99–1·59)	1·38† (0·93–2·07)	1·06* (0·82–1·39)	1·01‡ (0·76–1·32)	1·14† (0·85–1·54)	1·34* (1·03–1·75)	1·00* (0·80–1·25)	**Fluv**	1·14† (0·84–1·56)	1·16* (0·89–1·52)	1·01‡ (0·62–1·71)	1·16* (0·90–1·49)	0·77† (0·51–1·17)	1·23* (0·94–1·63)	0·99‡ (0·69–1·42)	1·06* (0·80–1·38)	1·78‡ (1·00–3·24)
1·07* (0·80–1·44)	1·11† (0·86–1·43)	1·23† (0·81–1·85)	0·94† (0·71–1·26)	0·89† (0·67–1·19)	1·01‡ (0·74–1·38)	1·19* (0·90–1·58)	0·89* (0·70–1·13)	0·89† (0·67–1·17)	**Miln**	1·02† (0·75–1·37)	0·88‡ (0·54–1·44)	1·02‡ (0·80–1·31)	0·67† (0·45–1·03)	1·08* (0·82–1·44)	0·86* (0·60–1·25)	0·93* (0·71–1·22)	1·56† (0·89–2·84)
0·93* (0·72–1·21)	0·97* (0·77–1·21)	1·07† (0·73–1·57)	0·82* (0·65–1·05)	0·78* (0·60–1·01)	0·88* (0·67–1·16)	1·04* (0·82–1·32)	0·78* (0·64–0·94)	0·78* (0·60–0·99)	0·87* (0·66–1·15)	**Mirt**	0·87† (0·55–1·41)	1·00* (0·82–1·23)	0·66* (0·45–0·99)	1·06* (0·84–1·35)	0·85* (0·62–1·18)	0·91* (0·73–1·13)	1·53† (0·89–2·72)
1·15† (0·76–1·76)	1·19† (0·80–1·78)	1·32‡ (0·80–2·20)	1·01‡ (0·67–1·54)	0·96‡ (0·63–1·45)	1·09‡ (0·71–1·68)	1·28* (0·86–1·94)	0·96‡ (0·66–1·40)	0·95‡ (0·63–1·46)	1·07‡ (0·70–1·67)	1·23* (0·82–1·86)	**Nefa**	1·15‡ (0·74–1·78)	0·75‡ (0·43–1·32)	1·23† (0·77–1·90)	0·98‡ (0·57–1·64)	1·04‡ (0·66–1·65)	1·76† (0·90–3·56)
1·01* (0·82–1·24)	1·05† (0·89–1·23)	1·16† (0·81–1·64)	0·89* (0·72–1·09)	0·84† (0·68–1·03)	0·95† (0·76–1·19)	1·12* (0·93–1·35)	0·84* (0·73–0·95)	0·84* (0·67–1·04)	0·94† (0·75–1·18)	1·08* (0·89–1·30)	0·88‡ (0·60–1·27)	**Paro**	0·66† (0·46–0·94)	1·06* (0·88–1·28)	0·85† (0·63–1·15)	0·91* (0·77–1·07)	1·53† (0·90–2·66)
1·44* (1·02–2·04)	1·50† (1·07–2·07)	1·65† (1·05–2·60)	1·27† (0·92–1·75)	1·20† (0·84–1·70)	1·36† (0·95–1·95)	1·60* (1·14–2·23)	1·20† (0·88–1·62)	1·20† (0·83–1·71)	1·35† (0·92–1·95)	1·54* (1·09–2·17)	1·25‡ (0·77–2·01)	1·43† (1·05–1·94)	**Rebo**	1·61† (1·09–2·34)	1·29† (0·81–2·01)	1·38† (0·94–1·99)	2·32† (1·24–4·41)
1·07* (0·85–1·37)	1·11* (0·92–1·35)	1·23† (0·85–1·79)	0·95† (0·76–1·18)	0·90† (0·71–1·13)	1·02‡ (0·79–1·32)	1·20* (0·97–1·48)	0·89‡ (0·76–1·05)	0·89† (0·70–1·13)	1·00† (0·77–1·30)	1·15* (0·93–1·43)	0·93‡ (0·63–1·37)	1·07* (0·90–1·26)	0·75† (0·54–1·04)	**Sert**	0·80* (0·58–1·11)	0·86* (0·70–1·05)	1·45† (0·84–2·54)
1·36* (0·99–1·87)	1·41† (1·06–1·86)	1·56† (1·04–2·31)	1·20* (0·88–1·63)	1·13† (0·83–1·54)	1·28† (0·92–1·79)	1·51* (1·12–2·04)	1·13† (0·87–1·46)	1·13† (0·82–1·55)	1·27* (0·91–1·76)	1·45* (1·09–1·94)	1·18‡ (0·75–1·84)	1·35* (1·04–1·75)	0·94‡ (0·64–1·39)	1·26† (0·95–1·67)	**Traz**	1·07‡ (0·77–1·47)	1·80† (0·98–3·38)
1·01* (0·82–1·26)	1·05† (0·87–1·27)	1·16† (0·82–1·65)	0·90† (0·72–1·10)	0·85† (0·67–1·06)	0·96† (0·77–1·21)	1·13* (0·93–1·37)	0·84† (0·73–0·97)	0·84* (0·66–1·07)	0·95* (0·73–1·23)	1·09* (0·89–1·33)	0·88‡ (0·59–1·30)	1·01† (0·86–1·17)	0·70† (0·51–0·97)	0·94* (0·78–1·13)	0·75† (0·57–0·98)	**Venl**	1·69† (1·01–2·86)
0·73‡ (0·42–1·26)	0·76‡ (0·44–1·29)	0·83‡ (0·45–1·54)	0·64† (0·37–1·11)	0·61† (0·35–1·05)	0·69† (0·40–1·20)	0·81‡ (0·47–1·39)	0·60† (0·36–1·02)	0·60† (0·34–1·05)	0·68† (0·39–1·20)	0·78‡ (0·45–1·34)	0·63† (0·33–1·19)	0·72† (0·43–1·22)	0·51† (0·28–0·92)	0·68† (0·39–1·16)	0·54† (0·30–0·95)	0·72† (0·43–1·19)	**Vort**

Agom = agomelatine; Amit = amitriptyline; Bupr = bupropion; Cita=citalopram; Clom = clomipramine; CrI=credible interval; Dulo = duloxetine; Esci = escitalopram; Fluo = fluoxetine; Fluv = fluvoxamine; Miln = milnacipran; Mirt = mirtazapine; Nefa = nefazodone; OR = odds ratio; Paro = paroxetine; Rebo =reboxetine; Sert = sertraline; Traz = trazodone; Venl = venlafaxine; Vort = vortioxetine.

"Drugs are reported in alphabetical order. Data are ORs (95% CrI) in the column-defining treatment compared with the row-defining treatment. For efficacy, ORs higher than 1 favour the column-defining treatment (ie, the first in alphabetical order). For acceptability, ORs lower than 1 favour the first drug in alphabetical order. To obtain ORs for comparisons in the opposite direction, reciprocals should be taken. Significant results are in bold and underscored. The certainty of the evidence (according to GRADE) was incorporated in this figure (appendix pp 231–65).

* Moderate quality of evidence. †Low quality of evidence. ‡ Very low quality of evidence."

Note: Includes 194 RCTs, and 34,196 patients.

Analysis Results

The network for the primary analyses included 432 RCTs comprised of 102,443 patients (Figure 3). The most commonly studied antidepressants were fluoxetine and paroxetine. Of the studies, 179 were head-to-head trials of active comparators. All antidepressants in the network had a placebo-controlled trial except for milnacipran. A minority of studies (n = 51) required imputation of response rates.

For the primary efficacy outcome, results of the NMA indicated that all treatments were more efficacious than placebo (Figure 3). For acceptability, only two drugs — agomelatine (OR: 0.84; 95% credible interval [CrI]: 0.72 to 0.97) and fluoxetine (OR: 0.88; 95% CrI: 0.80 to 0.96) — were more acceptable than placebo. In head-to-head comparisons from the NMA (Figure 4), a small number of comparisons were found to be statistically more efficacious or more acceptable than other antidepressants. Head-to-head comparisons were found to be statistically significant only for a minority of comparisons (19% of comparisons).

Secondary outcomes for response and remission also found that all treatments were more efficacious than placebo. The analysis of withdrawals due to adverse events found that all antidepressants had a greater frequency of dropouts compared with placebo, except for agomelatine (OR: 1.21; 95% CrI: 0.94 to 1.56). Meta-regression and subgroup analyses did not show any changes from the primary analysis. Preplanned sensitivity analysis limiting the network to only studies with a low risk of bias was not performed because of the small sample size (n = 39).

Critical Appraisal

Internal and External Validity

Internal

Ciprani et al. presented a transparent and robust synthesis of the evidence for the acute pharmacotherapies of MDD up to 2016. They conducted a comprehensive search of multiple databases over an extensive period of time. Overall, the methodology presented is in line with current methodological standards for systematic reviews (see Appendix 1). A number of steps were taken to ensure the inclusion of additional unpublished data, including a search of the grey literature. Screening of studies for eligibility occurred over multiple phases (titles/abstracts and full-texts) by multiple teams of reviewers working independently. Additionally, the authors have posted all of their data in hopes that full transparency may allow for improvement to their work. Importantly, the review was only conducted up to January 2016, potentially excluding recent evidence, especially for newer drugs.

Cipriani et al. also conducted a robust and full NMA. The analysis plan presented is in line with current methodological standards. They assessed and measured all the assumptions inherent in the NMA, including consistency and transitivity, and explored the impacts of these assumptions on results. There were some noted differences in results that highlight the heterogeneity of the data included. For example, there was greater variation in results for the head-to-head evidence than in the placebo-controlled evidence, highlighting the potential impact of publication bias and the high rates of placebo response in this therapeutic area. The authors noted that, although the results of individual studies varied greatly (as seen by the sensitivity analysis), the overall conclusions remained robust.

It is important to note that the studies included in the analyses were heterogeneous with respect to the populations included. For example, differences in ages, diagnostic criteria, previous treatments, and severity at baseline varied across studies. Additionally, due to the broad nature of this study, there was an inclusion of studies from as early as 1979, raising concerns that this may impact efficacy measures if outcome definitions and placebo response rates (for placebo-controlled studies) have changed over time. Importantly, meta-regressions and sensitivity analyses were conducted to attempt to control these factors and was not found to meaningfully change results. The authors reported that controlling for study

factors by meta-regression did not change results in a meaningful way. They explored the potential impact of study year and found minimal change to results. One important factor that was not discussed that may impact the analysis is concurrent non-pharmacological treatments. It appears that studies that allowed non-pharmacological treatments were categorized as placebo-controlled if no drug was given concurrently. This information was not collected or reported and could potentially present important variance across studies. Supportive medical management (e.g., sedative hypnotics and anxiolytics) and inclusion of psychotherapy may potentially introduce heterogeneity across studies. Lastly, head-to-head comparisons were more likely to be impacted by low-quality studies, based on results of risk of bias assessments, and thus some head-to-head comparisons must be interpreted carefully. Interestingly, the authors note a significant novelty affect in which newer drugs often performed better than older drugs.

External

Cipriani et al. only explored the acute use (eight weeks) of these treatments in adults. A study of UK antidepressant users found that nearly half of patients will be intermittent or chronic users of antidepressants.[15] Real-world studies of adherence to antidepressants have noted that 44% of users will discontinue at four weeks and more than half (52%) at 12 weeks.[16] This leaves a major gap in the knowledge of the long-term efficacy and acceptability with these agents. It is important to highlight that efficacy and tolerability are often quickly recognized after treatment initiation, and long-term use is strongly linked with initial response.[17] Lastly, as previously highlighted, there is no mention of concurrent psychotherapy or other non-pharmacological treatments in the study. This is important information as optimal treatment guidelines are developed and policies for the optimal use of these agents advanced.

A strength of this analysis is the inclusion of a broader evidence base. Importantly, this study does include drugs and dosages that are not available in Canada. It is noted that the study does not distinguish between formulations (i.e., immediate and extended-release). In addition to the well-known limitations associated with clinical trials, there are differences in the baseline characteristics of the included studies. They allowed for a broader inclusion of diagnostic tools unrealistic to what may occur globally in practice. Importantly, as is found with most clinical trials, the patients are younger and healthier than would be expected in real practice.[15,18-20] For example, the average age of patients in this study was 44 years. Thus, generalizing results to older patients may be limited. This is an important distinction with a high prevalence of antidepressants use among older individuals: recent US and Canadian data found that nearly 20% of those older than 60 years of age were treated with an antidepressant in the last month for a variety of indications.[15,21] Lastly, it is important to note that, in many of these trials, care was provided by psychiatrists, which may differ from the care given by family doctors who are the most common prescribers of antidepressants.[22-24]

Validity of Outcomes

Ciprani et al. reported two primary outcomes of efficacy (response rate) and acceptability (treatment discontinuations from any cause). These outcomes present a simple means to summarizing clinically relevant outcomes across a large evidence base. A 50% reduction in the HAM-D is the standard clinical response level.[25-27] This a strength of this analysis, as it allows the combination of various scales and maximizes the data used. The vast majority of studies did report HAM-D or MADRS scores. The authors selected the two most important

outcomes for decision-makers[12] and developed definitions to align a wide variation in outcomes definitions across a large number of tools. The primary outcomes reported are a robust means to assess across such a large number of studies that use various validated tools and measures. Importantly, in terms of safety, their analysis of acceptability and dropouts due to adverse events are a meaningful way to test across such a large number of drugs. The study does not report on some potentially relevant patient-centred outcomes such as quality of life or specific adverse events. This may limit the interpretability of the results for specific clinical concerns.

Overall Strengths and Limitations of the Study

The analysis is robust and thorough, and in line with current analytical methods and strategies. The overall quality of the study is high and meets all current standards as per the ISPOR assessment tool (see Appendix 1). The authors present a balance between ensuring usability of the information and ensuring the validity of the analyses conducted. Importantly, some of the solutions they created to combine various outcomes metrics (i.e., definitions of efficacy) helped maximize the evidence base included.

The study does have some important limitations worth noting when considering the applicability of the results. Firstly, the evidence base presented is drawn from clinical trials and thus may not be reflective of real-world practice and use. Important factors such as concurrent non-pharmacological therapies used in general practice and comorbidities were not explored and likely limited by their exclusion from RCTs. A large proportion of patients who are treated with antidepressants may have additional mental health comorbidities; they were not included in many of the studies analyzed. The second limitation, which is also acknowledged by Cipriani et al., is that some of the evidence included was of lower quality or did not have enough information reported to allow for analysis. The authors did plan to account for this through a variety of meta-regression and subgroup analyses. It is important to note that these sensitivity analyses did not change any of the findings. Lastly, the authors were not able to investigate important clinical and demographic modifiers of treatment response. This information would be useful in helping develop more nuanced clinical recommendations in selecting agents from among the 21 different antidepressants included. The authors rightfully suggest that future research must leverage patient-level data to explore these important factors.

Conclusions and Implications for Decision- or Policy-Making

This study concluded that all included antidepressants were more efficacious than placebo in their response rates, defined as the proportion of patients who had a ≥ 50% reduction in total score on a standardized observer-rating scale for depression. Patients who received agomelatine or fluoxetine were less likely to discontinue the medication for any reason compared with people who received placebo.

The findings presented herein are largely supportive of the existence of few differences between agents when all data were considered. The substantial variability between trials seems to be a common issue applying to all antidepressants and may be explained by the complexity of the condition, the subjectivity of scales used to measure effectiveness, varying patient perceptions and prior conditions, different study durations, and many other factors. The clinical significance of reported numerical differences remains to be determined. Differences may exist regarding the acceptability of agents, suggesting that multiple agents may need to be trialled by patients. Because of the generally moderate acceptability and effectiveness of antidepressants, it is widely argued that access to multiple agents would help optimize personalized treatment. Overall, these observations illuminate that all currently available treatments, regardless of novelty and price, are likely equal and can be used for patients with MDD based on their clinical attributes and personal preferences as part of a shared decision-making process.

It must be noted that these results are only applicable to adults and should not be generalized to children or adolescents. Additionally, although included in study populations, the findings may not fully hold true for older populations who are often not well-represented in clinical trials and have greater complexity. This study also highlights a gap in knowledge and important areas of future work that would leverage patient-specific characteristics to help optimize antidepressant selection.

As they stand, findings arising from this large body of evidence may help refine empirical clinical decision-making regarding the selection of initial antidepressants. However, observed differences are likely not important or robust enough to warrant any formal ranking that would inform specific reimbursement or procurement policies at the drug plan level, such as tiering or pricing. With evidence suggesting antidepressants are similarly effective and, as many antidepressants marketed in Canada are not publicly reimbursed in several jurisdictions (as of August 2019), efforts should be made to align current evidence with reimbursement policies in order to optimize access.

Consistency between the study under review and other similar studies was not explored in this focused report. While the current study was deemed of good quality and provides an exhaustive appraisal of available antidepressants, decision-makers may wish to further examine the findings of additional systematic reviews and network-meta-analyses with a similar scope conducted by other groups, as well as real-world evidence, to better understand the effectiveness and utilization of antidepressant treatments.

References

1. National Institute for Health and Care Excellence. Depression in adults: treatment and management [consultation draft]. 2019; https://www.nice.org.uk/guidance/gid-cgwave0725/documents/full-guideline-updated. Accessed 2019 Jul 31.
2. Institute for Health Metrics and Evaluation (IHME). Findings from the global burden of disease study 2017. Seattle (WA): IHME; 2018: http://www.healthdata.org/sites/default/files/files/policy_report/2019/GBD_2017_Booklet.pdf. Accessed 2019 Jul 31.
3. World Health Organization. Fact sheet: depression. 2018; https://www.who.int/news-room/fact-sheets/detail/depression. Accessed 2019 Jul 31.
4. Statistics Canada. Health at a glance: mental and substance use disorders in Canada. 2015; https://www150.statcan.gc.ca/n1/pub/82-624-x/2013001/article/11855-eng.htm. Accessed 2019 Jul 31.
5. Canadian Institute for Health Information. Prescribed drug spending in Canada, 2016: a focus on public drug program. Ottawa (ON): CIHI; 2016: https://secure.cihi.ca/free_products/Prescribed%20Drug%20Spending%20in%20Canada_2016_EN_web.pdf. Accessed 2019 Jul 31.
6. Statistics Canada. Prescription medication use by Canadians aged 6 to 79. 2014; https://www150.statcan.gc.ca/n1/pub/82-003-x/2014006/article/14032-eng.htm. Accessed 2019 Jul 31.
7. Kennedy SH, Lam RW, McIntyre RS, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 clinical guidelines for the management of adults with major depressive disorder: section 3. Pharmacological treatments. *Can J Psychiatry.* 2016;61(9):540-560.
8. Papakostas GI. Managing partial response or nonresponse: switching, augmentation, and combination strategies for major depressive disorder. *J Clin Psychiatry.* 2009;70 Suppl 6:16-25.
9. CAMH. Depression. 2019; https://www.camh.ca/en/health-info/mental-illness-and-addiction-index/depression. Accessed 2019 Jul 31.
10. Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *Lancet.* 2018;391(10128):1357-1366.
11. Furukawa TA, Cipriani A, Barbui C, Brambilla P, Watanabe N. Imputing response rates from means and standard deviations in meta-analyses. *Int Clin Psychopharmacol.* 2005;20(1):49-52.
12. Furukawa TA, Salanti G, Atkinson LZ, et al. Comparative efficacy and acceptability of first-generation and second-generation antidepressants in the acute treatment of major depression: protocol for a network meta-analysis. *BMJ open.* 2016;6(7):e010919.
13. Higgins JP, Altman DG, Gøtzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ.* 2011;343:d5928.
14. Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *J Clin Epidemiol.* 2011;64(4):380-382.
15. Moore M, Yuen HM, Dunn N, Mullee MA, Maskell J, Kendrick T. Explaining the rise in antidepressant prescribing: a descriptive study using the general practice research database. *BMJ.* 2009;339:b3999.
16. Vergouwen AC, Bakker A, Katon WJ, Verheij TJ, Koerselman F. Improving adherence to antidepressants: a systematic review of interventions. *J Clin Psychiatry.* 2003;64(12):1415-1420.
17. Stassen HH, Delini-Stula A, Angst J. Time course of improvement under antidepressant treatment: a survival-analytical approach. *Eur Neuropsychopharmacol.* 1993;3(2):127-135.
18. Beck CA, Patten SB, Williams JV, et al. Antidepressant utilization in Canada. *Soc Psychiatry Psychiatr Epidemiol.* 2005;40(10):799.
19. Patten SB, Williams JV, Lavorato DH, Wang JL, McDonald K, Bulloch AG. Descriptive epidemiology of major depressive disorder in Canada in 2012. *Can J Psychiatry.* 2015;60(1):23-30.
20. Raymond CB, Morgan SG, Caetano PA. Antidepressant utilization in British Columbia from 1996 to 2004: increasing prevalence but not incidence. *Psychiatr Serv.* 2007;58(1):79-84.
21. Wong J, Motulsky A, Eguale T, Buckeridge DL, Abrahamowicz M, Tamblyn R. Treatment indications for antidepressants prescribed in primary care in Quebec, Canada, 2006-2015. *JAMA.* 2016;315(20):2230-2232.
22. Andersson K, Melander A, Svensson C, Lind O, Nilsson JL. Repeat prescriptions: refill adherence in relation to patient and prescriber characteristics, reimbursement level and type of medication. *Eur J Public Health.* 2005;15(6):621-626.
23. Fairman KA, Drevets WC, Kreisman JJ, Teitelbaum F. Course of antidepressant treatment drug type, and prescriber's specialty. *Psychiatr Serv.* 1998;49(9):1180-1186.
24. Havard A, Straka P, Sara G, Lujic S, Tran DT, Jorm LR. Identifying patients using antidepressants for the treatment of depression: a predictive algorithm for use in pharmaceutical and medical claims data. *Pharmacoepidemiol Drug Saf.* 2019;28(3):354-361.
25. Hamilton Depression Rating Scale (HDRS). Gainesville (FL): University of Florida Med Consults College of Medicine; 2011: https://dcf.psychiatry.ufl.edu/files/2011/05/HAMILTON-DEPRESSION.pdf. Accessed 2019 Jul 31.
26. Lin CH, Chen CC, Wang FC, Lane HY. Percentage reduction of depression severity versus absolute severity after initial weeks of treatment to predict final response or remission. *Psychiatry Clin Neurosci.* 2013;67(4):265-272.
27. Roffman JL, Silverman BC, Stern T. Diagnostic rating scales and laboratory tests. In: Stern TA, Fricchione GL, Cassem NH, Jellinek M, Rosenbaum JF, eds. *Massachusetts General Hospital Handbook of General Hospital Psychiatry*. 6 ed. Philadelphia (PA): Saunders; 2010.

Appendix 1: Completed ISPOR Network Meta-analysis Assessment Questionnaire

	Question	Strength	Weakness
1	Is the population relevant?	**Yes**	No
2	Are any critical interventions missing?	**No**	Yes
3	Are any relevant outcomes missing?	No	**Yes**
4	Is the context (e.g., settings and circumstances) applicable to your population?	**Yes**	No
Credibility			
5	Did the researchers attempt to identify and include all relevant randomized controlled trials?	**Yes**	No
6	Do the trials for the interventions of interest form one connected network of randomized controlled trials?	**Yes**	No
7	Is it apparent that poor-quality studies were included, thereby leading to bias?	**No**	Yes
8	Is it likely that bias was induced by selective reporting of outcomes in the studies?	**No**	Yes
9	Are there systematic differences in treatment effect modifiers (i.e., baseline patient or study characteristics that impact the treatment effects) across the different treatment comparisons in the network?	No	**Yes**
10	If yes (i.e., there are such systematic differences in treatment effect modifiers), were these imbalances in effect modifiers across the different treatment comparisons identified prior to comparing individual study results?	**Yes**	No
11	Were statistical methods used that preserve within-study randomization (no naive comparisons)?	**Yes**	No
12	If both direct and indirect comparisons are available for pairwise contrasts (i.e., closed loops), was agreement in treatment effects (i.e., consistency) evaluated or discussed?	**Yes**	No
13	In the presence of consistency between direct and indirect comparisons, were both direct and indirect evidence included in the network meta-analysis?	Yes	**No**
14	With inconsistency or an imbalance in the distribution of treatment effect modifiers across the different types of comparisons in the network of trials, did the researchers attempt to minimize this bias with the analysis?	**Yes**	No

	Question	Strength	Weakness
15	Was a valid rationale provided for the use of random-effects or fixed-effects models?	**Yes**	No
16	If a random-effects model was used, were assumptions about heterogeneity explored or discussed?	**Yes**	No
17	If there are indications of heterogeneity, were subgroup analyses or meta-regression analysis with pre-specified covariates performed?	**Yes**	No
Reporting Quality and Transparency			
18	Is a graphical or tabular representation of the evidence network provided with information on the number of randomized controlled trials per direct comparison?	**Yes**	No
19	Are the individual study results reported?	**Yes**	No
20	Are all pairwise contrasts between Interventions, as obtained with the network meta-analysis, reported along with measures of uncertainty?	**Yes**	No
21	Are results of direct comparisons reported separately from results of the indirect comparisons or network meta-analysis?	**Yes**	No
22	Is a ranking of interventions provided given the reported treatment effects and its uncertainty by outcome?	**Yes**	No
23	Is the impact of important patient characteristics on treatment effects reported?	**Yes**	No
Conflict of Interest			
24	Were there any potential conflicts of interest?	Yes	**No**
25	If yes, were steps taken to address these?	**Yes**	No

Adapted from: Indirect Treatment Comparison/Network Meta-Analysis Study Questionnaire to Assess Relevance and Credibility to Inform Health Care Decision Making: An ISPOR-AMCP-NPC Good Practice Task Force Report Jansen, Jeroen P. et al. Value in Health, Volume 17, Issue 2, 157 – 173.