Depression module – evidence profile DEP2: Length of treatment with antidepressants for adults with depressive episode/disorder

WHO mhGAP guideline update: Mental Health Gap Action Programme (mhGAP) guideline for mental, neurological and substance use disorders

2023

World Health Organization

Contents

Mental Health Gap Action Programme (mhGAP) guideline for mental, neurological and substance use disorders, available at: https://www.who.int/publications/i/item/9789240084278

1. Background

Depression is a highly prevalent and recurrent mental disorder (Kessler, R. C., & Bromet, 2013). It has a great negative impact on the quality of life and functioning of the individuals, and it is associated with high societal and economic costs (Bloom et al., 2012; Ferrari et al., 2010). By 2030, depression is predicted to be one of the leading causes of disability and premature mortality worldwide (Mathers & Loncar, 2006). Reducing the depression burden by developing and scaling evidence-based interventions is now a major global priority (World Bank Group & World Health Organization, 2016).

Different types of antidepressants have been found to effectively reduce depressive symptomatology (Cipriani et al., 2018), and are currently recommended as a first line treatment for depression (Nathan & Gorman, 2015; Fletcher, Leaman, McSloy, & Leng, 2020; World Health Organization, 2016). The effects of antidepressants, however, are varied. Many patients do not improve or even experience deterioration (Thomas et al., 2013). Additionally, a long-standing concern is non-adherence to medications, which leads to symptom worsening, chronicity and increased suicidal rates (Ho et al., 2016).). Therefore, there is a need to further evaluate the short- and long-term balance between benefits and harms (Cipriani et al., 2018; Ioannidis, 2008). Moreover, previous research conducted in the UK suggested that half of the patients on antidepressants have been continuing this treatment for two years or more (Johnson et al., 2012; Petty et al., 2006), with similar increasing trends in long-term prescription in the US (Mojtabai & Olfson, 2014; Luo et al., 2020). Some side effects caused by antidepressants, including weight gain, sleep disturbance, and sexual dysfunction, could be increased with long-term use (Ferguson, 2001). Therefore, in the recommendation for antidepressant treatment, it is crucial to determine for how long the treatment should continue.

In the current report, we aimed to present the results of a systematic review of meta-analyses examining the association of treatment length with antidepressants efficacy and safety. Focusing on the most prescribed antidepressants, Tricyclic Antidepressants (TCAs) and Selective Serotonin Reuptake Inhibitors (SSRIs), we evaluated whether these pharmacotherapies were more effective and as safe as treatment as usual or pill placebo in adults with depressive disorders or elevated symptoms of depression, and for how long should the treatment continue. We reviewed the effects in a wide range of outcomes, including symptom reduction, suicide-related outcomes, adverse effects, and improvements in functioning.

2. Methodology

Evidence from recent meta-analyses covering the effectiveness and safety of antidepressant medications compared to treatment as usual or pill placebo for adults with depressive episode or disorders was summarized. With a focus on relapse prevention, the evaluation focused on for how long the treatments should continue following acute phase of treatment to prevent relapse.

2.1. PICO question

How long should treatment with antidepressants continue in adults with depressive episode/disorder?

Population (P): Adults with depressive episode/disorder and/or elevated depressive symptoms
Intervention (I): Antidepressant medicines: TCAs, SSRIs
Comparator (C): Placebo, treatment as usual
Outcomes (O):

List critical outcomes:

- **Critical outcome 1:** Reduction of symptoms
- **Critical outcome 2:** Adverse effects
- **Critical outcome 3:** Suicide-related outcomes
- **Critical outcome 4:** Improvement in quality of life and functioning
- **Critical outcome 5:** Relapse

List important outcomes:

None specified

2.2. Search strategy

Existing systematic reviews were identified by conducting searches in the following bibliographic databases:

- PubMed
- PsycINFO
- Embase
- Cochrane reviews
- Global Index Medicus

The search strings were designed in collaboration with a Medical Information Specialist at the Vrije Universiteit Amsterdam. We designed the search strings by combining blocks with free and index terms indicative for 1) Depression (*Type of Participants*), 2) Antidepressants (TCA and SSRIs) (*Types of interventions*), and 3) terms related to systematic reviews and meta-analyses (*Type of studies*). The search strings for PubMed can be accessed in Appendix I. In line with the WHO guideline methodology, indicating that evidence obtained for the development of guidelines should be as recent as possible (World Health Organization, 2014), therefore, the period of the searches covered from 1 January 2019 until 31 January 2022. No restrictions were applied for language.

2.3. Data collection and analysis

As the first stage in selecting relevant studies, records retrieved from the bibliographic databases were assessed for eligibility by examining their titles and abstracts, based on the inclusion and exclusion criteria developed a priori. Studies were included if they were (i) Systematic reviews of randomized controlled trials (RCTs). (ii) Had adult participants (>18 years) with a primary diagnosis of depression as established by a diagnostic interview or elevated symptoms of depression according to cut off scores on self-report scales. (iii) Evaluated the effectiveness or safety of SSRIs or TCAs compared to pill placebo/ treatment as usual (iv) Reported outcomes regarding mental health symptoms, adverse effects, quality of life and functioning and suicide related outcomes. We excluded studies that had participants with secondary depression (due to medical conditions/illness, trauma, etc), bipolar disorder, psychotic depression, and treatment resistant depression. The full text of articles found to be potentially relevant based on their titles and abstracts were retrieved and examined considering the same inclusion criteria in the second stage of study selection. Data from eligible studies were extracted into pre-defined templates that include the general characteristics of the study, population, intervention, comparator and outcomes. When there was an overlap between studies (i.e. they evaluated the same antidepressant medications, in similar target populations, and reported the same outcomes), we selected the meta-analysis based on the following criteria and in the following order: (i) Recency (more recent publication covering a more recent search period) (ii) number of included RCTs, (iii) broadness of the review (covering multiple antidepressants and groups of antidepressants compared to pill placebo and/or treatment as usual, with a wide range of outcomes) (iv) AMSTAR ratings.

Two reviewers (AA and MC/CM) independently assessed the eligibility of the studies identified and extracted data from study reports. Discrepancies between the reviewers were resolved through discussions. The search strategy and results reporting the databases searched, the strategy used to search each database, the total number of citations retrieved from each database, and the reasons for excluding some publications after reviewing the full text have been documented. The flow of articles throughout the search and up to the final cohort of included studies is shown in Figure 1, which includes the number of excluded articles and the reasons for any exclusions at the full-text screening stage.

2.4. Selection and coding of identified records

Rayyan and Endnote were used for the management of references. Rayyan was used during the first two stages of the project, involving the selection of studies based on titles, abstracts, and full texts. Endnote was used to store the references and pdfs of the included studies for the remaining stages of the project. Data extraction was conducted in excel files with a predefined format which was designed by the involved reviewers. A wide range of study level data regarding date of searches, target population characteristics, type of intervention and control, average length of interventions, total number of participants, mean age, proportion of women and risk of bias were extracted. All data was collected by two independent reviewers and discrepancies were resolved through discussions.

2.5. Quality assessment

The quality of the included systematic reviewers was assessed with the **AMSTAR quality appraisal tool 2**. Two independent researchers (AA and MC/CM) applied the AMSTAR-2 checklist to the included studies, and any disagreements were discussed with a third researcher.

The certainty of the evidence was assessed using **GRADE** (Grading of Recommendations, Assessment, Development and Evaluations). When available, we extracted the GRADE assessments from the meta-analysis. When the GRADE assessment was not available, we assessed it ourselves examining the following criteria:

- **Risk of bias (RoB):** We extracted the RoB ratings from the individual studies included in the meta-analyses (when available). We calculated the percentage of trials rated at low, high, and unclear risk of bias. Based on this information, and in order to take consistent decisions across the available evidence, we rated the RoB GRADE item using a decision tree. This decision tree can be accessed in Appendix II.
- **Inconsistency:** We judged inconsistency by examining heterogeneity statistics: I^2, which indicates the percentage of heterogeneity between effect sizes, and its 95% confidence interval (95% CI). When the 95% CI of the I^2 is not reported, we computed it and used it in our judgements. We judged inconsistency as serious when I^2 was over 75% and its 95% CI substantially overlaps with the category of considerable heterogeneity (above 75%). Substantial overlap was estimated with the median of the 95% CI. If the 95% CI was not available or could not be calculated, we rated it as serious if heterogeneity was larger than 50% (category of substantial heterogeneity). If I^2 was not reported and could not be calculated, we rated it as serious.
- **Indirectness**: Direct evidence was derived from research that directly compares the interventions which we are interested in, delivered to the participants in which we are interested, and that measures the outcomes important to patients. We rated for each particular comparison how indirect the reviewed evidence was in terms of population, intervention, and outcomes.
- **Imprecision**: We rated this item based on a standard power calculation ($\alpha = 0.05$ and $\beta = 0.20$) for detecting an effect size of 0.2, which requires a sample size of 400 participants in total. We judged as serious for all analyses that included less than 400 participants. Analyses including less than 100 participants was rated as very serious. A rating of serious was given when the number of participants included in the analyses was not available.

  Other considerations: For this item we explored publication bias. We rated it as serious if there was evidence for publication bias in the meta-analyses, based on statistical tests. However, we did not downgrade the evidence if a meta-analysis did not investigate it.

2.6. Analysis of subgroups or subsets

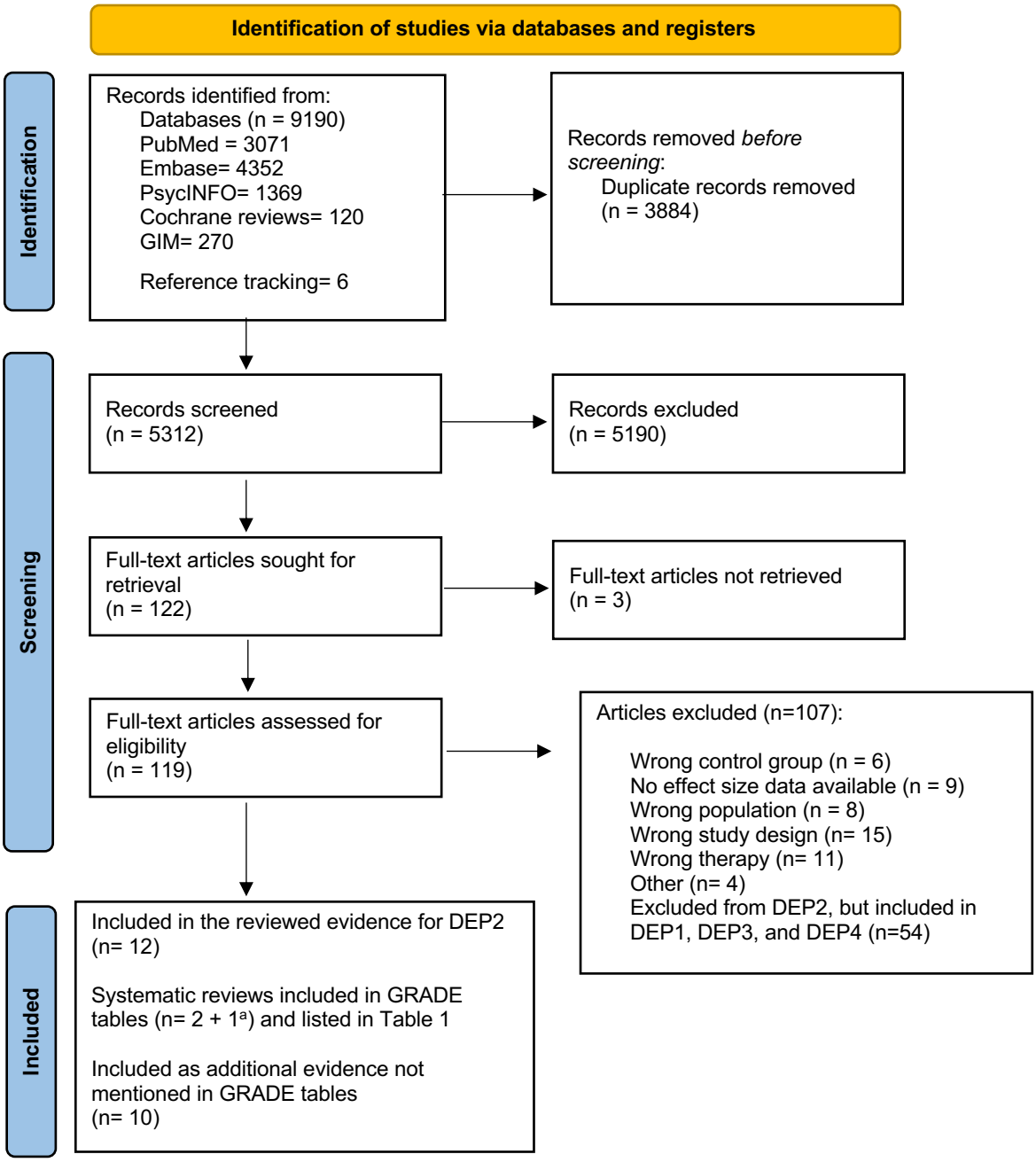
Since we reviewed existing systematic reviews, we considered the subgroups or subsets that were available in the included meta-analyses. The subgroups of interest were:

- Time of follow up (e.g., 0-3 months, 3-6 months, 6-9 months, 9-12 months and 1 year)
- Dose setting of antidepressant medication (e.g., fixed dose setting and flexible dose setting)

3. Results

3.1. Systematic reviews and/or studies identified by the search process

Figure 1: PRISMA 2020 flow diagram for systematic review of reviews which includes searches of databases and registers only



GIM: Global Index Medicus

a. One RCT was identified through other sources. The data was pooled with a selected review to update the meta-analysis.

3.2. Included In GRADE tables/footnotes

KATO, M., HORI, H., INOUE, T., IGA, J., IWATA, M., INAGAKI, T., SHINOHARA, K., IMAI, H., MURATA, A., MISHIMA, K. & TAJIKA, A. 2021. Discontinuation of antidepressants after remission with antidepressant medication in major depressive disorder: a systematic review and meta-analysis. Mol Psychiatry, 26, 118-133.

LEWIS, G., MARSTON, L., DUFFY, L., FREEMANTLE, N., GILBODY, S., HUNTER, R., KENDRICK, T., KESSLER, D., MANGIN, D., KING, M., LANHAM, P., MOORE, M., NAZARETH, I., WILES, N., BACON, F., BIRD, M., BRABYN, S., BURNS, A., CLARKE, C. S., HUNT, A., PERVIN, J. & LEWIS, G. 2021. Maintenance or Discontinuation of Antidepressants in Primary Care. *N Engl J Med,* 385**,** 1257-1267.

ZHOU, D., LV, Z., SHI, L., ZHOU, X., LIN, Q., CHEN, X., WAN, L., LI, Y., RAN, L., HUANG, Y., WANG, G., LI, D., WANG, W., LIU, C. & KUANG, L. 2020. Effects of antidepressant medicines on preventing relapse of unipolar depression: a pooled analysis of parametric survival curves. *Psychological medicine***,** 1-9.

3.3. Excluded from GRADE tables/footnotes

BALDESSARINI, R. J., LAU, W. K., SIM, J., SUM, M. Y. & SIM, K. 2015. Duration of initial antidepressant treatment and subsequent relapse of major depression. *J Clin Psychopharmacol,* 35**,** 75-6.

BRAUN, C., ADAMS, A., RINK, L., BSCHOR, T., KUHR, K. & BAETHGE, C. 2020. In search of a dose-response relationship in SSRIs-a systematic review, meta-analysis, and network meta-analysis. Acta Psychiatr Scand, 142, 430-442.

CHENG, Q., HUANG, J., XU, L., LI, Y., LI, H., SHEN, Y., ZHENG, Q. & LI, L. 2020. Analysis of Time-Course, Dose-Effect, and Influencing Factors of Antidepressants in the Treatment of Acute Adult Patients With Major Depression. *Int J Neuropsychopharmacol,* 23**,** 76-87.

FURUKAWA, T. A., CIPRIANI, A., COWEN, P. J., LEUCHT, S., EGGER, M. & SALANTI, G. 2019. Optimal Dose of Selective Serotonin Reuptake Inhibitors, Venlafaxine, and Mirtazapine in Major Depression: A Systematic Review and Dose-Response Meta-Analysis. *Focus (Am Psychiatr Publ),* 18**,** 211-219.

FURUKAWA, T. A., SALANTI, G., COWEN, P. J., LEUCHT, S. & CIPRIANI, A. 2020. No benefit from flexible titration above minimum licensed dose in prescribing antidepressants for major depression: systematic review. *Acta Psychiatr Scand,* 141**,** 401-409.

FURUKAWA, T. A., SHINOHARA, K., SAHKER, E., KARYOTAKI, E., MIGUEL, C., CIHAROVA, M., BOCKTING, C. L. H., BREEDVELT, J. J. F., TAJIKA, A., IMAI, H., OSTINELLI, E. G., SAKATA, M., TOYOMOTO, R., KISHIMOTO, S., ITO, M., FURUKAWA, Y., CIPRIANI, A., HOLLON, S. D. & CUIJPERS, P. 2021. Initial treatment choices to achieve sustained response in major depression: a systematic review and network meta-analysis. World Psychiatry, 20, 387-396.

HENGARTNER, M. P. 2020. How effective are antidepressants for depression over the long term? A critical review of relapse prevention trials and the issue of withdrawal confounding. *Therapeutic Advances in Psychopharmacology,* 10.

HENSSLER, J., HEINZ, A., BRANDT, L. & BSCHOR, T. 2019. Antidepressant Withdrawal and Rebound Phenomena. *Dtsch Arztebl Int,* 116**,** 355-361.

HOLPER, L. 2020. Optimal doses of antidepressants in dependence on age: Combined covariate actions in Baye*sian network meta*-anal**y**sis. EClinicalMedicine, 18, 100219.

MACHMUTOW, K., MEISTER, R., JANSEN, A., KRISTON, L., WATZKE, B., HÄRTER, M. C. & LIEBHERZ, S. 2019. Comparative effectiveness of continuation and maintenance treatments for persistent depressive disorder in adults. Cochrane Database Syst Rev, 5, Cd012855.

SIM, K., LAU, W. K., SIM, J., SUM, M. Y. & BALDESSARINI, R. J. 2015. Prevention of Relapse and Recurrence in Adults with Major Depressive Disorder: Systematic Review and Meta-Analyses of Controlled Trials. *Int J Neuropsychopharmacol,* 19.

SØRENSEN, A., JUHL JØRGENSEN, K. & MUNKHOLM, K. 2022. Clinical practice guideline recommendations on tapering and discontinuing antidepressants for depression: a systematic review. *Ther Adv Psychopharmacol,* 12**,** 20451253211067656.

THOM, R. P., ALEX, ER, J. L., BARON, D., GARAKANI, A., GROSS, L., PINE, J. H., RADHAKRISHNAN, R., SLABY, A. & SUMNER, C. R. 2021. Selective Serotonin Reuptake Inhibitors: How Long Is Long Enough? Journal of Psychiatric Practice, 27, 361-371.

VAN LEEUWEN, E., VAN DRIEL, M. L., HOROWITZ, M. A., KENDRICK, T., DONALD, M., DE SUTTER, A. I., ROBERTSON, L. & CHRISTIAENS, T. 2021. Approaches for discontinuation versus continuation of long-term antidepressant use for depressive and anxiety disorders in adults. Cochrane Database Syst Rev, 4, Cd013495.

Table 1: PICO Table

Serial Number	Intervention/ Comparison	Outcomes	Systematic reviews (Name, Year)	Justification/Explanation for systematic review
1	**Pharmacotherapy compared to pill placebo in adults with depressive disorders**	Reduction in mental health symptoms	-	No available recent meta-analytic evidence on this outcome (N/A)
		Adverse effects	-	N/A
		Suicide related outcomes	-	N/A
		Improvement in QAL and Functioning	-	N/A
		Relapse	Zhou et al., 2020	Most recent high-quality meta-analysis available on relapse related outcomes at different follow-up times in pharmacotherapy compared to placebo in adults with a diagnosis of depression
		Relapse	Kato et al., 2020 + Lewis et al., 2021	Most recent high-quality meta-analysis available on the relapse rate after 6 months post remission in pharmacotherapy compared to pill placebo on depressive symptoms in adults with a diagnosis of depression
2	**Fixed antidepressant dose compared to pill placebo in adults with depressive disorders**	Reduction in mental health symptoms	-	N/A
		Adverse effects	-	N/A
		Suicide related outcomes	-	N/A
		Improvement in QAL and Functioning	-	N/A
		Relapse	Kato et al., 2020	Most recent high-quality meta-analysis available on the relapse rate in pharmacotherapy with a fixed antidepressant dose setting compared to pill placebo on depressive symptoms in adults with a diagnosis of depression

Serial Number	Intervention/ Comparison	Outcomes	Systematic reviews (Name, Year)	Justification/Explanation for systematic review
3	**Flexible antidepressant dose compared to pill placebo in adults with depressive disorders**	Reduction in mental health symptoms		N/A
		Adverse effects		N/A
		Suicide related outcomes		N/A
		Improvement in QAL and Functioning		N/A
		Relapse	Kato et al., 2020	Most recent high-quality meta-analysis available on the relapse rate in pharmacotherapy with a flexible antidepressant dose setting compared to pill placebo on depressive symptoms in adults with a diagnosis of depression

3.4. Narrative description of studies that contributed to GRADE analysis

Kato et al., 2021: A significant clinical issue encountered after a successful acute major depressive disorder (MDD) treatment is the relapse of depressive symptoms. Although continuing maintenance therapy with antidepressants is generally recommended, there is no established protocol on whether or not it is necessary to prescribe the antidepressant used to achieve remission. In this meta-analysis, the risk of relapse and treatment failure when either continuing with the same drug used to achieved remission or switching to a placebo was assessed in several clinically significant subgroups. The pooled odds ratio (OR) (±95% confidence intervals (CI)) was calculated using a random effects model. Across 40 studies (n = 8890), the relapse rate was significantly lower in the antidepressant group than the placebo group by about 20% (OR = 0.38, CI: 0.33-0.43, p < 0.00001; 20.9% vs 39.7%). The difference in the relapse rate between the antidepressant and placebo groups was greater for tricyclics (25.3%; OR = 0.30, CI: 0.17-0.50, p < 0.00001), SSRIs (21.8%; OR = 0.33, CI: 0.28-0.38, p < 0.00001), and other newer agents (16.0%; OR = 0.44, CI: 0.36-0.54, p < 0.00001) in that order, while the effect size of acceptability was greater for SSRIs than for other antidepressants. A flexible dose schedule (OR = 0.30, CI: 0.23-0.48, p < 0.00001) had a greater effect size than a fixed dose (OR = 0.41, CI: 0.36-0.48, p < 0.00001) in comparison to placebo. Even in studies assigned after continuous treatment for more than 6 months after remission, the continued use of antidepressants had a lower relapse rate than the use of a placebo (OR = 0.40, CI: 0.29-0.55, p < 0.00001; 20.2% vs 37.2%). The difference in relapse rate was similar from a maintenance period of 6 months (OR = 0.41, CI: 0.35-0.48, p < 0.00001; 19.6% vs 37.6%) to over 1 year (OR = 0.35, CI: 0.29-0.41, p < 0.00001; 19.9% vs 39.8%). The all-cause dropout of antidepressant and placebo groups was 43% and 58%, respectively, (OR = 0.47, CI: 0.40-0.55, p < 0.00001). The tolerability rate was ~4% for both groups. The rate of relapse (OR = 0.32, CI: 0.18-0.64, p = 0.0010, 41.0% vs 66.7%) and all-cause dropout among adolescents was higher than in adults. To prevent relapse and treatment failure, maintenance therapy, and careful attention for at least 6 months after remission is recommended. SSRIs are well-balanced agents, and flexible dose adjustments are more effective for relapse prevention.

Lewis et al., 2021: BACKGROUND: Patients with depression who are treated in primary care practices may receive antidepressants for prolonged periods. Data are limited on the effects of maintaining or discontinuing antidepressant therapy in this setting. METHODS: We conducted a randomized, double-blind trial involving adults who were being treated in 150 general practices in the United Kingdom. All the patients had a history of at least two depressive episodes or had been taking antidepressants for 2 years or longer and felt well enough to consider stopping antidepressants. Patients who had received citalopram, fluoxetine, sertraline, or mirtazapine were randomly assigned in a 1:1 ratio to maintain their current antidepressant therapy (maintenance group) or to taper and discontinue such therapy with the use of matching placebo (discontinuation group). The primary outcome was the first relapse of depression during the 52-week trial period, as evaluated in a time-to- event analysis. Secondary outcomes were depressive and anxiety symptoms, physical and withdrawal symptoms, quality of life, time to stopping an antidepressant or placebo, and global mood ratings. RESULTS: A total of 1466 patients underwent screening. Of these patients, 478 were enrolled in the trial (238 in the maintenance group and 240 in the discontinuation group). The average age of the patients was 54 years; 73% were women. Adherence to the trial assignment was 70% in the maintenance group and 52% in the discontinuation group. By 52 weeks, relapse occurred in 92 of 238 patients (39%) in the maintenance group and in 135 of 240 (56%) in the discontinuation group (hazard ratio, 2.06; 95% confidence interval, 1.56 to 2.70; P<0.001). Secondary outcomes were generally in the same direction as the primary outcome.

Patients in the dis- continuation group had more symptoms of depression, anxiety, and withdrawal than those in the maintenance group. CONCLUSIONS: Among patients in primary care practices who felt well enough to discontinue antidepressant therapy, those who were assigned to stop their medication had a higher risk of relapse of depression by 52 weeks than those who were assigned to maintain their current therapy. (Funded by the National Institute for Health Re- search; ANTLER ISRCTN number, ISRCTN15969819.)

Zhou et al., 2020: BACKGROUND: Major depressive disorder is characterized by a high risk of relapse. We aimed to compare the prophylactic effects of different antidepressant medicines (ADMs). METHODS: PubMed, Cochrane Central Register of Controlled Trials, Embase and the Web of Science were searched on 4 July 2019. A pooled analysis of parametric survival curves was performed using a Bayesian framework. The main outcomes were hazard ratios (HRs), relapse-free survival and mean relapse-free months. RESULTS: Forty randomized controlled trials were included. The 1-year relapse-free survival for ADM (76%) was significantly better than that for placebo (56%). Most of the relapse difference (86.5%) occurred in the first 6 months. Most HRs were not constant over time. Compared with placebo, the HRs of several drugs (vilazodone, nefazodone, quetiapine, mirtazapine, amitriptyline, fluvoxamine, hypericum extract and tianeptine) became closer to 1 over time and crossed the invalid line (HR = 1) before 12 months. The HRs of paroxetine, desvenlafaxine and bupropion approached 1 over time, but they remained superior to the placebo within 1 year. Other anti- depressants (selegiline, vortioxetine, levomilnacipran, fluoxetine, agomelatine, citalopram, sertraline, venlafaxine, duloxetine, milnacipran, reboxetine, phenelzine and gepirone) were continuously superior to placebo over time. Proof of benefit after 6 months of follow-up was not established partially because of small differences between the drug and placebo after 6 months. Almost all studies used an 'enriched' randomized discontinuation design, which may explain the high relapse rates in the first 6 months after randomization. CONCLUSIONS: The superiority of ADM v. placebo was mainly attributed to the difference in relapse rates that occurred in the first 6 months. Our analysis provided evidence that the prophylactic efficacy was not constant over time. A beneficial effect was observed, but the prevention of new episodes after 6 months was questionable. These findings may have implications for clinical practice.

3.5. Grading the Evidence

GRADE Table 1: Pharmacotherapy compared to pill placebo in adults with depressive disorders

Author(s): Arpana Amarnath, Marketa Ciharova, Clara Miguel
Question: Pharmacotherapy compared to pill placebo in adults with depression
Population: General Adult [a]
Reference List: Zhou et al., 2020; Kato et al., 2020; Lewis et al., 2021 [b]

Certainty assessment								Effect	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	№ of patients	Absolute (95% CI)		
Reduction in mental health symptoms – Not available										
-	-	-	-	-	-	-	-	-	-	CRITICAL
Adverse effects - Not available										
-	-	-	-	-	-	-	-	-	-	CRITICAL
Suicide related outcomes - Not available										
-	-	-	-	-	-	-	-	-	-	CRITICAL
Improvement in QAL and Functioning - Not available										
-	-	-	-	-	-	-	-	-	-	CRITICAL

Certainty assessment								Effect	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	№ of patients	Absolute (95% CI)		
Relapse – 1 year relapse rate – Zhou, 2020										
NR	RCT	serious [c]	serious [d]	serious [e]	serious [f]	none	NR	24% ADM 44% placebo	⨁◯◯◯ VERY LOW	CRITICAL
Relapse – mean relapse free months at 1 year – Zhou, 2020										
NR	RCT	serious [c]	serious [d]	serious [e]	serious [f]	none	NR	10 months ADM 8 months placebo	⨁◯◯◯ VERY LOW	CRITICAL
Relapse – mean difference in relapse rate at 0-3 months – Zhou, 2020										
NR	RCT	serious [c]	serious [d]	serious [e]	serious [f]	none	NR	63.9%	⨁◯◯◯ VERY LOW	CRITICAL
Relapse – mean difference in relapse rate at 3-6 months – Zhou, 2020										
NR	RCT	serious [c]	serious [d]	serious [e]	serious [f]	none	NR	22.6%	⨁◯◯◯ VERY LOW	CRITICAL
Relapse – mean difference in relapse rate at 6-9 months – Zhou, 2020										
NR	RCT	serious [c]	serious [d]	serious [e]	serious [f]	none	NR	9.3%	⨁◯◯◯ VERY LOW	CRITICAL

Certainty assessment								Effect	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	№ of patients	Absolute (95% CI)		
Relapse – mean difference in relapse rate at 9-12 months – Zhou, 2020										
NR	RCT	serious [c]	serious [d]	serious [e]	serious [f]	none	NR	4.2%	⊕◯◯◯ VERY LOW	CRITICAL
Relapse – Relapse rate (6 months) [g] **–** Kato, 2020 + Lewis, 2021										
8	RCT	serious [h]	not serious	serious [e]	not serious	none	1347	OR **0.43** CI 0.33 to 0.56	⊕⊕◯◯ LOW	CRITICAL

ADM: Antidepressant medication (pharmacotherapy) **CI:** Confidence interval; **OR:** Odds Ratio; **RCTs**: Randomized Controlled Trials; **SMD:** Standard Mean Difference; **QAL:** Quality of life

Interpretation of outcomes:
1 year relapse – the relapse rate at 1 year was lower for pharmacotherapy compared to placebo
Mean relapse free months – the average of relapse free months over a year was higher in the pharmacotherapy groups than in the placebo groups
Mean difference in relapse rate – most of the difference in 1 year relapse rates between pharmacotherapy and placebo occurred in the first 3 months (63.9%) and first 6 months (85.5 %). The difference in relapse free rates becomes much smaller after 6 months.
Relapse rate – Below 1 favors pharmacotherapy; above 1 favors pill placebo

Explanations:

a. Adults (>18 years) with a diagnosis of MDD who achieved remission or response after acute phase of treatment or recovery after continuation phase of treatment.

b. The individual studies included in Kato et al., 2020 were pooled with one additional RCT identified via reference tracking (Lewis et al., 2021). The RCT was included since it provided follow-up of 52 weeks, while the available meta-analysis provided outcomes only until 6 months

c. The risk of bias was aggregated for the entire meta-analyses. It has been rated as serious because the number of high-risk studies was above 25%.

d. The I^2 was not reported and could not be calculated.

e. Indirectness has been downgraded because no meta-analysis answering the exact research question has been conducted, but several studies provide indications of answers to the research questions indirectly. The presented evidence is based on these studies.

f. The number of participants included in the analyses was not available and this seriously affects the certainty of evidence

g. Relapse rate was calculated after 6 months of continuation treatment after remission

h. Vast majority of the included studies (>60%) have an unclear risk of bias and this seriously affects the certainty of evidence.

GRADE Table 2: Fixed antidepressant dose compared to pill placebo in adults with depressive disorders

Author(s): Arpana Amarnath, Marketa Ciharova, Clara Miguel
Question: Pharmacotherapy compared to pill placebo in adults with depression [a]
Population: General Adult
Reference List: Kato et al., 2020

Certainty assessment								Effect	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	№ of patients	Absolute (95% CI)		
Reduction in mental health symptoms - Not available										
-	-	-	-	-	-	-	-	-	-	CRITICAL
Adverse effects - Not available										
-	-	-	-	-	-	-	-	-	-	CRITICAL
Suicide related outcomes - Not available										
-	-	-	-	-	-	-	-	-	-	CRITICAL
Improvement in QAL and Functioning - Not available										
-	-	-	-	-	-	-	-	-	-	CRITICAL
Relapse – Relapse rate (9 months) [b] **–** Kato, 2020										

Certainty assessment								Effect	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	№ of patients	Absolute (95% CI)		
27	RCT	serious [c]	not serious	serious [d]	not serious	none	7042	OR **0.41** CI 0.36 to 0.48	⨁⨁◯◯ LOW	CRITICAL

CI: Confidence interval; **OR:** Odds Ratio; **RCTs**: Randomized Controlled Trials; **QAL:** Quality of life

Interpretation of outcomes:
Relapse rate – Below 1 favors pharmacotherapy; above 1 favors pill placebo

Explanations:

a. Adults (>18 years) with a diagnosis of depression who responded to an active drug in an acute treatment phase.
b. Relapse rate was calculated after 9 months of continuation treatment after remission
c. Vast majority of the included studies (>60%) have an unclear risk of bias and this seriously affects the certainty of evidence.
d. Indirectness has been downgraded because no meta-analysis answering the exact research question has been conducted, but several studies provide indications of answers to the research questions indirectly. The presented evidence is based on these studies

GRADE Table 3: Flexible antidepressant dose compared to pill placebo in adults with depressive disorders

Author(s): Arpana Amarnath, Marketa Ciharova, Clara Miguel
Question: Pharmacotherapy compared to pill placebo in adults with depression [a]
Population: General Adult
Reference List: Kato et al., 2020

Certainty assessment								Effect	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	№ of patients	Absolute (95% CI)		
Reduction in mental health symptoms - Not available										
-	-	-	-	-	-	-	-	-	-	CRITICAL
Adverse effects - Not available										
-	-	-	-	-	-	-	-	-	-	CRITICAL
Suicide related outcomes - Not available										
-	-	-	-	-	-	-	-	-	-	CRITICAL
Improvement in QAL and Functioning - Not available										
-	-	-	-	-	-	-	-	-	-	CRITICAL
Relapse – Relapse rate (12 months) [b] **–** Kato, 2020 + Lewis, 2021										

Certainty assessment								Effect	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	№ of patients	Absolute (95% CI)		
13	RCT	serious [c]	not serious	serious [d]	not serious	none	1857	OR **0.30** CI 0.23 to 0.48	⨁⨁◯◯ LOW	CRITICAL

CI: Confidence interval; **OR:** Odds Ratio; **RCTs**: Randomized Controlled Trials; **QAL:** Quality of life

Interpretation of outcomes:
Relapse rate – Below 1 favors pharmacotherapy; above 1 favors pill placebo

Explanations:

a. Adults (>18 years) with a diagnosis of depression who responded to an active drug in an acute treatment phase.
b. Relapse rate was calculated after 1 year of continuation treatment after remission
c. Vast majority of the included studies (>60%) have an unclear risk of bias and this seriously affects the certainty of evidence.
d. Indirectness has been downgraded because no meta-analysis answering the exact research question has been conducted, but several studies provide indications of answers to the research questions indirectly. The presented evidence is based on these studies

3.6. Additional evidence not mentioned in GRADE tables

Baldessarini et al., 2015: Background: The efficacy, limitations, and methods of studying antidepressant treatment continued beyond initial weeks of acute major depression remain incompletely resolved. Aims: For subjects treated in controlled trials for acute depression, we analyzed the relationship of relapse risk within 12 months of rerandomizing to placebo versus duration of initial treatment and putative stabilization. Methods: With data from placebo arms of 45 relevant controlled trials identified in recent, systematic reviews were pooled and analyzed by regression modeling. Results: There was a strong inverse correlation of shorter initial treatment and greater relapse risk after rerandomizing to placebo treatment, best fit to a power function ($P \leq 0.003$); relapse risk differed by 11.4-fold, declining sharply as initial treatment continued for 16 to 20 weeks or more. Conclusions: Discontinuation of antidepressant treatment for major depressive episodes at times less than 6 months was associated with rising risks after randomization to continuation with placebo. This relationship requires critical consideration in both clinical management of depressed patients and the design and interpretation of treatment discontinuation trials.

Braun et al., 2020: OBJECTIVE: Recent meta-analyses on dose-response relationships of SSRIs are largely based on indirect evidence. We analyzed RCTs directly comparing different SSRI doses. METHOD: Systematic literature search for RCTs. Two raters independently screened articles and extracted data. Across SSRIs, doses defined as low, medium, and high doses, based on drug manufacturers' product monographs, were analyzed in pairwise random-effects meta-analyses and in a sensitivity network meta-analysis with regard to differences in antidepressive efficacy (primary outcome). We also analyzed all direct comparisons of different dosages of specific SSRIs. (Prospero CRD42018081031). RESULTS: Out of 5333 articles screened, we included 33. Comparisons of dosage groups (low, medium, and high) resulted in only small and clinically non-significant differences for SSRIs as a group, the strongest relating to medium vs low doses (SMD: -0.15 [95%-CI: -0.28; -0.01) and not sustained in a sensitivity analysis. Among different doses of specific SSRIs, no statistically significant trend emerged for efficacy at higher doses, but 60 mg/day fluoxetine are statistically significantly inferior to 20 mg/day. Paroxetine results are inconclusive: 10 mg/day are inferior to higher doses, but 30 and 40 mg/day are inferior to 20 mg/day. Meaningful effects cannot be ruled out for certain drugs and dosages, often investigated in only one trial. Dropout rates increase with dose-particularly due to side effects. Network meta-analyses supported our findings. CONCLUSIONS: There is no conclusive level I or level II evidence of a clinically meaningful dose-response relationship of SSRIs as a group or of single substances. High SSRI doses are not recommended as routine treatment.

Cheng et al., 2019: OBJECTIVE: Model-based meta-analysis was used to describe the time-course and dose-effect relationships of antidepressants and also simultaneously investigate the impact of various factors on drug efficacy. METHODS: This study is a reanalysis of a published network meta-analysis. Only placebo-controlled trials were included in this study. The change rate in depression rating scale scores from baseline was used as an efficacy indicator because a continuous variable is more likely to reflect subtle differences in efficacy between drugs. RESULTS: A total 230 studies containing 64 346 patients were included in the analysis. The results showed that the number of study sites (single or multi-center) and the type of setting (inpatient or noninpatient) are important factors affecting the efficacy of antidepressants. After deducting the placebo effect, the maximum pure drug efficacy value of inpatients was 18.4% higher than that of noninpatients, and maximum pure drug efficacy value of single-center trials was 10.2% higher than that of multi-central trials. Amitriptyline showed the highest drug efficacy. The remaining 18 antidepressants were comparable or had little difference. Within the approved dose range, no significant dose-response relationship was observed. However, the time-course relationship is

obvious for all antidepressants. In terms of safety, with the exception of amitriptyline, the dropout rate due to adverse events of other drugs was not more than 10% higher than that of the placebo group. CONCLUSION: The number of study sites and the type of setting are significant impact factors for the efficacy of antidepressants. Except for amitriptyline, the other 18 antidepressants have little difference in efficacy and safety.

Furukawa et al., 2019: BACKGROUND: Depression is the single largest contributor to non-fatal health loss worldwide. Second-generation antidepressants are the first-line option for pharmacological management of depression. Optimising their use is crucial in reducing the burden of depression; however, debate about their dose dependency and their optimal target dose is ongoing. We have aimed to summarise the currently available best evidence to inform this clinical question. METHODS: We did a systematic review and dose-response meta-analysis of double-blind, randomised controlled trials that examined fixed doses of five selective serotonin reuptake inhibitors (SSRIs; citalopram, escitalopram, fluoxetine, paroxetine, and sertraline), venlafaxine, or mirtazapine in the acute treatment of adults (aged 18 years or older) with major depression, identified from the Cochrane Central Register of Controlled Trials, CINAHL, Embase, LILACS, MEDLINE, PsycINFO, AMED, PSYNDEX, websites of drug licensing agencies and pharmaceutical companies, and trial registries. We imposed no language restrictions, and the search was updated until Jan 8, 2016. Doses of SSRIs were converted to fluoxetine equivalents. Trials of antidepressants for patients with depression and a serious concomitant physical illness were excluded. The main outcomes were efficacy (treatment response defined as 50% or greater reduction in depression severity), tolerability (dropouts due to adverse effects), and acceptability (dropouts for any reasons), all after a median of 8 weeks of treatment (range 4-12 weeks). We used a random-effects, dose-response meta-analysis model with flexible splines for SSRIs, venlafaxine, and mirtazapine. FINDINGS: 28 554 records were identified through our search (24 524 published and 4030 unpublished records). 561 published and 121 unpublished full-text records were assessed for eligibility, and 77 studies were included (19 364 participants; mean age 42·5 years, SD 11·0; 7156 [60·9%] of 11 749 reported were women). For SSRIs (99 treatment groups), the dose-efficacy curve showed a gradual increase up to doses between 20 mg and 40 mg fluoxetine equivalents, and a flat to decreasing trend through the higher licensed doses up to 80 mg fluoxetine equivalents. Dropouts due to adverse effects increased steeply through the examined range. The relationship between the dose and dropouts for any reason indicated optimal acceptability for the SSRIs in the lower licensed range between 20 mg and 40 mg fluoxetine equivalents. Venlafaxine (16 treatment groups) had an initially increasing dose-efficacy relationship up to around 75-150 mg, followed by a more modest increase, whereas for mirtazapine (11 treatment groups) efficacy increased up to a dose of about 30 mg and then decreased. Both venlafaxine and mirtazapine showed optimal acceptability in the lower range of their licensed dose. These results were robust to several sensitivity analyses. INTERPRETATION: For the most commonly used second-generation antidepressants, the lower range of the licensed dose achieves the optimal balance between efficacy, tolerability, and acceptability in the acute treatment of major depression. FUNDING: Japan Society for the Promotion of Science, Swiss National Science Foundation, and National Institute for Health Research.

Furukawa et al., 2020: BACKGROUND: In fixed-dose antidepressant trials, the lower range of the licensed dose achieves the optimal balance between efficacy and tolerability. Whether flexible upward titration while side-effects permit provides additional benefits is unknown. METHODS: We did a systematic review of placebo-controlled randomized trials that examined selective serotonin reuptake inhibitors (SSRIs), venlafaxine or mirtazapine in the acute treatment of major depression. Our primary outcome was response, defined as 50% or greater reduction in depression severity. Secondary outcomes included drop-outs due to adverse effects and drop-outs for any reason. We conducted random-effects meta-analyses to calculate the ratios of odds ratios (RORs) between trials comparing the flexible dose titrating above the minimum licensed

dose against placebo and those comparing the fixed minimum licensed dose against placebo. RESULTS: We included 123 published and unpublished randomized controlled trials (29 420 participants). There was no evidence supporting efficacy of the flexible dosing over the fixed low dose of SSRIs (ROR 0.96, 95% CI: 0.73 to 1.25), venlafaxine (1.24, 0.96 to 1.60) or mirtazapine (0.77, 0.33 to 1.78). No important differences were noted for tolerability or for any subgroup analyses except the superior efficacy of venlafaxine flexible dosing between 75 and 150 mg over the fixed 75 mg (1.30, 1.02 to 1.65). CONCLUSION: There was no evidence to support added value in terms of efficacy, tolerability or acceptability of flexibly titrating up the dosage over the minimum licensed dose of SSRIs or mirtazapine. For venlafaxine, increased efficacy can be expected by flexibly titrating up to 150 mg.

Furukawa et al., 2021: Major depression is often a relapsing disorder. It is therefore important to start its treatment with therapies that maximize the chance of not only getting the patients well but also keeping them well. We examined the associations between initial treatments and sustained response by conducting a network meta-analysis of randomized controlled trials (RCTs) in which adult patients with major depression were randomized to acute treatment with a psychotherapy (PSY), a protocolized antidepressant pharmacotherapy (PHA), their combination (COM), standard treatment in primary or secondary care (STD), or pill placebo, and were then followed up through a maintenance phase. By design, acute phase treatment could be continued into the maintenance phase, switched to another treatment or followed by discretionary treatment. We included 81 RCTs, with 13,722 participants. Sustained response was defined as responding to the acute treatment and subsequently having no depressive relapse through the maintenance phase (mean duration: 42.2±16.2 weeks, range 24-104 weeks). We extracted the data reported at the time point closest to 12 months. COM resulted in more sustained response than PHA, both when these treatments were continued into the maintenance phase (OR=2.52, 95% CI: 1.66-3.85) and when they were followed by discretionary treatment (OR=1.80, 95% CI: 1.21-2.67). The same applied to COM in comparison with STD (OR=2.90, 95% CI: 1.68-5.01 when COM was continued into the maintenance phase; OR=1.97, 95% CI: 1.51-2.58 when COM was followed by discretionary treatment). PSY also kept the patients well more often than PHA, both when these treatments were continued into the maintenance phase (OR=1.53, 95% CI: 1.00-2.35) and when they were followed by discretionary treatment (OR=1.66, 95% CI: 1.13-2.44). The same applied to PSY compared with STD (OR=1.76, 95% CI: 0.97-3.21 when PSY was continued into the maintenance phase; OR=1.83, 95% CI: 1.20-2.78 when PSY was followed by discretionary treatment). Given the average sustained response rate of 29% on STD, the advantages of PSY or COM over PHA or STD translated into risk differences ranging from 12 to 16 percentage points. We conclude that PSY and COM have more enduring effects than PHA. Clinical guidelines on the initial treatment choice for depression may need to be updated accordingly.

Hengartner, 2020: The aim of this article is to discuss the validity of relapse prevention trials and the issue of withdrawal confounding in these trials. Recommendations for long-term antidepressant treatment are based almost exclusively on discontinuation trials. In these relapse prevention trials, participants with remitted depression are randomised either to have the antidepressant abruptly discontinued and replaced by inert placebo or to continue active treatment. The drug–placebo difference in relapse rates at the end of the maintenance phase is then interpreted as a prophylactic drug effect. These trials consistently produce remarkable benefits for maintenance treatment. However, the internal validity of this trial protocol is compromised, as research has shown that abruptly stopping antidepressants can cause severe withdrawal reactions that lead to (or manifest as) depression relapses. That is, there is substantial withdrawal confounding in discontinuation trials, which renders their findings uninterpretable. It is not clear to what degree the drug–placebo separation in relapse prevention (discontinuation) trials is due to withdrawal reactions, but various estimations suggest that it is presumably the majority. A review of findings based on other methodologies, including real-world long-term

effectiveness trials like STAR*D and various naturalistic cohort studies, do not indicate that antidepressants have considerable prophylactic effects. As absence of evidence does not imply evidence of absence, no definitive conclusions can be drawn from the literature. To enable a thorough risk–benefit evaluation, real-world effectiveness trials should not only focus on relapse prevention, but also assess antidepressants' long-term effects on social functioning and quality of life. Thus far, reliable long-term data on these outcome domains are lacking.

Henssler et al., 2019: Background: Antidepressants are among the most commonly prescribed drugs worldwide. They are often discontinued, frequently without the knowledge of the prescribing physician. It is, therefore, important for physicians to be aware of the withdrawal and rebound phenomena that may arise, in order to prevent these phenomena, treat them when necessary, and counsel patients appropriately. Methods: This review is based on a comprehensive, structured literature search on antidepressant withdrawal phenomena that we carried out in the CENTRAL, PubMed (Medline), and Embase databases. We classified the relevant publications and reports by their methodological quality. Results: Out of a total of 2287 hits, there were 40 controlled trials, 38 cohort studies and retrospective analyses, and 271 case reports that met the inclusion criteria. Withdrawal manifestations are usually mild and self-limiting; common ones include dizziness, headache, sleep disturbances, and mood swings. More serious or prolonged manifestations rarely arise. There is an increased risk with MAO inhibitors, tricyclic antidepressants, venlafaxine, and paroxetine; on the other hand, for agomelatine and fluoxetine, abrupt discontinuation seems to be unproblematic. There is also some evidence of rebound phenomena, i.e., of higher relapse rates or especially severe relapses of depression after the discontinuation of an antidepressant. Conclusion: A robust evidence base now indicates that there can be acute withdrawal phenomena when antidepressants are discontinued. Putative rebound phenomena have not been adequately studied to date. It is recommended that antidepressants should be tapered off over a period of more than four weeks.

Holper, 2020: Background: The meta-analysis by Furukawa et al. (The Lancet Psychiatry 2019, 6(7)) reported optimal doses for antidepressants in adult major depressive disorder (MDD). The present reanalysis aimed to adjust optimal doses in dependence on age. Methods: Analysis was based on the same dataset by Cipriani et al. (The Lancet 2018, 391(10128)) comparing 21 antidepressants in MDD. Random-effects Bayesian network meta-analysis was implemented to estimate the combined covariate action using restricted cubic splines (RCS). Balanced treatment recommendations were derived for the outcomes efficacy (response), acceptability (dropouts for any reason), and tolerability (dropouts due to adverse events). Findings: The combined covariate action of dose and age suggested agomelatine and escitalopram as the best-balanced antidepressants in terms of efficacy and tolerability that may be escalated until 40 and 60 mg/day fluoxetine equivalents (mg/day (FE)), respectively, for ages 30-65 years. Desvenlafaxine, duloxetine, fluoxetine, milnacipran, and vortioxetine may be escalated until 20-40 mg/day (FE) , whereas bupropion, citalopram, mirtazapine, paroxetine, and venlafaxine may not be given in doses > 20 mg/day (FE) . Amitriptyline, clomipramine, fluvoxamine, levomilnacipran, reboxetine, sertraline, and trazodone revealed no relevant balanced benefits and may therefore not be recommended for antidepressant treatment. None of the antidepressants was observed to provide balanced benefits in patients >70 years because of adverse events exceeding efficacy. Interpretation: Findings suggest that the combined covariate action of dose and age provides a better basis for judging antidepressant clinical benefits than considering dose or age separately and may thus inform decision makers to accurately guide antidepressant dosing recommendations in MDD.

Machmutow et al., 2019: BACKGROUND: Persistent depressive disorder (PDD) is defined as a depressive disorder with a minimum illness duration of two years, including four diagnostic subgroups (dysthymia, chronic major depression, recurrent major depression with incomplete

remission between episodes, and double depression). Persistent forms of depression represent a substantial proportion of depressive disorders, with a lifetime prevalence ranging from 3% to 6% in the Western world. Growing evidence indicates that PDD responds well to several acute interventions, such as combined psychological and pharmacological treatments. Yet, given the high rates of relapse and recurrences of depression following response to acute treatment, long-term continuation and maintenance therapy are of great importance. To date, there has been no evidence synthesis available on continuation and maintenance treatments of PDDs. OBJECTIVES: To assess the effects of pharmacological and psychological (either alone or combined) continuation and maintenance treatments for persistent depressive disorder, in comparison with each other, placebo (drug/attention placebo/non-specific treatment control), and treatment as usual (TAU). Continuation treatments are defined as treatments given to currently remitted people (remission is defined as depressive symptoms dropping below case level) or to people who previously responded to an antidepressant treatment. Maintenance therapy is given during recovery (which is defined as remission lasting longer than six months). SEARCH METHODS: We searched Ovid MEDLINE (1950-), Embase (1974-), PsycINFO (1967-) and the Cochrane Central Register of Controlled Trials (CENTRAL) to 28 September 2018. An earlier search of these databases was also conducted for RCTs via the Cochrane Common Mental Disorders Controlled Trial Register (CCMD-CTR) (all years to 11 Dec 2015). In addition, we searched grey literature resources as well as the international trial registers ClinicalTrials.gov and ICTRP to 28 September 2018. We screened reference lists of included studies and contacted the first author of all included studies. SELECTION CRITERIA: We included randomized (RCTs) and non-randomized controlled trials (NRCTs) in adults with formally diagnosed PDD, receiving pharmacological, psychological, or combined continuation and maintenance interventions. DATA COLLECTION AND ANALYSIS: Two review authors independently selected studies and extracted and analysed data. The primary efficacy outcome was relapse/recurrence rate of depression. The primary acceptance outcome was dropout due to any reason other than relapse/recurrence. We performed random-effects meta-analyses using risk ratios (RR) for dichotomous outcomes and mean differences (MD) for continuous outcomes, with 95% confidence intervals (CI). MAIN RESULTS: We included 10 studies (seven RCTs, three NRCTs) involving 840 participants in this review, from which five studies investigated continuation treatments and five studies investigated maintenance treatments. Overall, the included studies were at low-to-moderate risk of bias. For the three NRCTs, the most common source of risk of bias was selection of reported results. For the seven RCTs, the most common sources of risk of bias was non-blinding of outcome assessment and other bias (especially conflict of interest due to pharmaceutical sponsoring). Pharmacological continuation and maintenance therapies the most common comparison was antidepressant medication versus tablet placebo (five studies). Participants taking antidepressant medication were probably less likely to relapse or to experience a recurrent episode compared to participants in the placebo group at the end of the intervention (13.9% versus 33.8%, RR 0.41, 95% CI 0.21 to 0.79; participants = 383; studies = 4; I^2 = 54%, moderate quality evidence). Overall dropout rates may be similar between participants in the medication and placebo group (23.0% versus 25.5%, RR 0.90, 95% CI 0.39 to 2.11; RCTs = 4; participants = 386; I^2 = 64%, low quality evidence). However, sensitivity analyses showed that the primary outcome (rate of relapse/recurrence) showed no evidence of a difference between groups when only including studies with low risk of bias. None of the studies compared pharmacological or psychological treatments versus TAU. Psychological continuation and maintenance therapies One study compared psychological therapies versus attention placebo/non-specific control. One study compared psychotherapy with medication. The results of the studies including psychotherapy might indicate that continued or maintained psychotherapy could be a useful intervention compared to no treatment or antidepressant medication. However, the body of evidence for these comparisons was too small and uncertain to draw any high-quality conclusions. Combined psychological and pharmacological continuation and maintenance therapies Three studies compared combined psychological and pharmacological therapies with pharmacological therapies alone. One study compared combined

psychological and pharmacological therapies with psychotherapeutic therapies alone. However, the body of evidence for these comparisons was too small and uncertain to draw any high-quality conclusions Comparison of different antidepressant medications Two studies reported data on the direct comparison of two antidepressants. However, the body of evidence for this comparison was too small and uncertain to draw any high-quality conclusions. AUTHORS' CONCLUSIONS: Currently, it is uncertain whether continued or maintained pharmacotherapy (or both) with the reviewed antidepressant agents is a robust treatment for preventing relapse and recurrence in people with PDD, due to moderate or high risk of bias as well as clinical heterogeneity in the analysed studies. For all other comparisons, the body of evidence was too small to draw any final conclusions, although continued or maintained psychotherapy might be effective compared to no treatment. There is need for more high-quality trials of psychological interventions. Further studies should address health-related quality of life and adverse events more precisely, as well as assessing follow-up data.

Sim et al., 2016: Background: Findings of substantial remaining morbidity in treated major depressive disorder (MDD) led us to review controlled trials of treatments aimed at preventing early relapses or later recurrences in adults diagnosed with MDD to summarize available data and to guide further research. Methods: Reports (n = 97) were identified through systematic, computerized literature searching up to February 2015. Treatment versus control outcomes were summarized by random-effects meta-analyses. Results: In 45 reports of 72 trials (n = 14 450 subjects) lasting 33.4 weeks, antidepressants were more effective than placebos in preventing relapses (response rates [RR] = 1.90, confidence interval [CI]: 1.73–2.08; NNT = 4.4; $p< 0.0001$). In 35 reports of 37 trials (n = 7253) lasting 27.0 months, antidepressants were effective in preventing recurrences (RR = 2.03, CI 1.80–2.28; NNT = 3.8; $p < 0.0001$), with minor differences among drug types. In 17 reports of 22 trials (n = 1 969) lasting 23.7 months, psychosocial interventions yielded inconsistent or inconclusive results. Conclusions: Despite evidence of the efficacy of drug treatment compared to placebos or other controls, the findings further underscore the substantial, unresolved morbidity in treated MDD patients and strongly encourage further evaluations of specific, improved individual and combination therapies (pharmacological and psychological) conducted over longer times, as well as identifying clinical predictors of positive or unfavorable responses and of intolerability of long-term treatments in MDD.

Sørensen et al., 2022: BACKGROUND: Tapering and discontinuing antidepressants are important aspects of the management of patients with depression and should therefore be considered in clinical practice guidelines. OBJECTIVES: We aimed to assess the extent and content, and appraise the quality, of guidance on tapering and discontinuing antidepressants in major clinical practice guidelines on depression. METHODS: Systematic review of clinical practice guidelines on depression issued by national health authorities and major national or international professional organisations in the United Kingdom, the United States, Canada, Australia, Singapore, Ireland and New Zealand (PROSPERO CRD42020220682). We searched PubMed, 14 guideline registries and the websites of relevant organisations (last search 25 May 2021). The clinical practice guidelines were assessed for recommendations and information relevant to tapering and discontinuing antidepressants. The quality of the clinical practice guidelines as they pertained to tapering and discontinuation was assessed using the AGREE II tool. RESULTS: Of the 21 included clinical practice guidelines, 15 (71%) recommended that antidepressants are tapered gradually or slowly, but none provided guidance on dose reductions, how to distinguish withdrawal symptoms from relapse or how to manage withdrawal symptoms. Psychological challenges were not addressed in any clinical practice guideline, and the treatment algorithms and flow charts did not include discontinuation. The quality of the clinical practice guidelines was overall low. CONCLUSION: Current major clinical practice guidelines provide little support for clinicians wishing to help patients discontinue or taper antidepressants in terms of mitigating and managing withdrawal symptoms. Patients who have deteriorated upon following current guidance on tapering and

discontinuing antidepressants thus cannot be concluded to have experienced a relapse. Better guidance requires better randomised trials investigating interventions for discontinuing or tapering antidepressants.

Thom et al., 2021: Selective serotonin reuptake inhibitors (SSRIs) are among the most commonly prescribed medications. They are among the first-line medications for several chronic or relapsing-remitting psychiatric conditions, including major depressive disorder and anxiety disorders. The advantages of SSRI use include ease of titration and their tolerability and safety profile. Guidelines for the short-term use of SSRIs are widely available, but there is no well-organized guidance on how and whether to maintain a patient on SSRIs for the long-term. In this article, we discuss the benefits and possible adverse consequences of long-term SSRI use, as well as clinical practice considerations when using SSRIs chronically. The major benefit of long-term SSRI use is relapse prevention. The current literature suggests that the general health risks of long-term SSRI use are low; however, further research, particularly in special populations including youth and the elderly, is needed. Long-term SSRI use increases the risk of tachyphylaxis and discontinuation syndrome. Recognizing that many patients may remain on SSRIs for many years, there are several factors that prescribers should consider if they choose to use an SSRI when initiating treatment and during long-term monitoring. The decision to continue or to discontinue an SSRI should be an active one, involving both the patient and prescriber, and should be revisited periodically. Patients who remain on SSRIs for the long-term should also have periodic monitoring to reassess the risk-benefit ratio of remaining on the SSRI, as well as to assess the safety, tolerability, and efficacy of the medication.

Van Leeuwen et al., 2021: BACKGROUND: Depression and anxiety are the most frequent indication for which antidepressants are prescribed. Long-term antidepressant use is driving much of the internationally observed rise in antidepressant consumption. Surveys of antidepressant users suggest that 30% to 50% of long-term antidepressant prescriptions had no evidence-based indication. Unnecessary use of antidepressants puts people at risk of adverse events. However, high-certainty evidence is lacking regarding the effectiveness and safety of approaches to discontinuing long-term antidepressants. OBJECTIVES: To assess the effectiveness and safety of approaches for discontinuation versus continuation of long-term antidepressant use for depressive and anxiety disorders in adults. SEARCH METHODS: We searched all databases for randomised controlled trials (RCTs) until January 2020. SELECTION CRITERIA: We included RCTs comparing approaches to discontinuation with continuation of antidepressants (or usual care) for people with depression or anxiety who are prescribed antidepressants for at least six months. Interventions included discontinuation alone (abrupt or taper), discontinuation with psychological therapy support, and discontinuation with minimal intervention. Primary outcomes were successful discontinuation rate, relapse (as defined by authors of the original study), withdrawal symptoms, and adverse events. Secondary outcomes were depressive symptoms, anxiety symptoms, quality of life, social and occupational functioning, and severity of illness. DATA COLLECTION AND ANALYSIS: We used standard methodological procedures as expected by Cochrane. MAIN RESULTS: We included 33 studies involving 4995 participants. Nearly all studies were conducted in a specialist mental healthcare service and included participants with recurrent depression (i.e. two or more episodes of depression prior to discontinuation). All included trials were at high risk of bias. The main limitation of the review is bias due to confounding withdrawal symptoms with symptoms of relapse of depression. Withdrawal symptoms (such as low mood, dizziness) may have an effect on almost every outcome including adverse events, quality of life, social functioning, and severity of illness. Abrupt discontinuation Thirteen studies reported abrupt discontinuation of antidepressant. Very low-certainty evidence suggests that abrupt discontinuation without psychological support may increase risk of relapse (hazard ratio (HR) 2.09, 95% confidence interval (CI) 1.59 to 2.74; 1373 participants, 10 studies) and there is insufficient evidence of its effect on adverse events (odds ratio (OR) 1.11, 95% CI 0.62 to 1.99;

1012 participants, 7 studies; I^2 = 37%) compared to continuation of antidepressants, without specific assessment of withdrawal symptoms. Evidence about the effects of abrupt discontinuation on withdrawal symptoms (1 study) is very uncertain. None of these studies included successful discontinuation rate as a primary endpoint. Discontinuation by "taper" Eighteen studies examined discontinuation by "tapering" (one week or longer). Most tapering regimens lasted four weeks or less. Very low-certainty evidence suggests that "tapered" discontinuation may lead to higher risk of relapse (HR 2.97, 95% CI 2.24 to 3.93; 1546 participants, 13 studies) with no or little difference in adverse events (OR 1.06, 95% CI 0.82 to 1.38; 1479 participants, 7 studies; I^2 = 0%) compared to continuation of antidepressants, without specific assessment of withdrawal symptoms. Evidence about the effects of discontinuation on withdrawal symptoms (1 study) is very uncertain. Discontinuation with psychological support Four studies reported discontinuation with psychological support. Very low-certainty evidence suggests that initiation of preventive cognitive therapy (PCT), or MBCT, combined with "tapering" may result in successful discontinuation rates of 40% to 75% in the discontinuation group (690 participants, 3 studies). Data from control groups in these studies were requested but are not yet available. Low-certainty evidence suggests that discontinuation combined with psychological intervention may result in no or little effect on relapse (HR 0.89, 95% CI 0.66 to 1.19; 690 participants, 3 studies) compared to continuation of antidepressants. Withdrawal symptoms were not measured. Pooling data on adverse events was not possible due to insufficient information (3 studies). Discontinuation with minimal intervention Low-certainty evidence from one study suggests that a letter to the general practitioner (GP) to review antidepressant treatment may result in no or little effect on successful discontinuation rate compared to usual care (6% versus 8%; 146 participants, 1 study) or on relapse (relapse rate 26% vs 13%; 146 participants, 1 study). No data on withdrawal symptoms nor adverse events were provided. None of the studies used low-intensity psychological interventions such as online support or a changed pharmaceutical formulation that allows tapering with low doses over several months. Insufficient data were available for the majority of people taking antidepressants in the community (i.e. those with only one or no prior episode of depression), for people aged 65 years and older, and for people taking antidepressants for anxiety. AUTHORS' CONCLUSIONS: Currently, relatively few studies have focused on approaches to discontinuation of long-term antidepressants. We cannot make any firm conclusions about effects and safety of the approaches studied to date. The true effect and safety are likely to be substantially different from the data presented due to assessment of relapse of depression that is confounded by withdrawal symptoms. All other outcomes are confounded with withdrawal symptoms. Most tapering regimens were limited to four weeks or less. In the studies with rapid tapering schemes the risk of withdrawal symptoms may be similar to studies using abrupt discontinuation which may influence the effectiveness of the interventions. Nearly all data come from people with recurrent depression. There is an urgent need for trials that adequately address withdrawal confounding bias, and carefully distinguish relapse from withdrawal symptoms. Future studies should report key outcomes such as successful discontinuation rate and should include populations with one or no prior depression episodes in primary care, older people, and people taking antidepressants for anxiety and use tapering schemes longer than 4 weeks.

4. From Evidence to Recommendations

4.1. Summary of findings

Table 3: Summary of findings table

GRADE Table	Source	Outcome	Specific Outcome	Number of Studies	Effects	Certainty of Evidence
GRADE Table 1: Pharmacotherapy compared to pill placebo in adults with depressive disorders	Zhou et al., 2020; Kato et al., 2020 + Lewis et al, 2021	Reduction in mental health symptoms	-	-	-	N/A
		Adverse effects	-	-	-	N/A
		Suicide related outcomes	-	-	-	N/A
		Improvement in quality of life and functioning	-	-	-	N/A
		Relapse	1 year relapse rate	NR	24% ADM 44% placebo	⊕◯◯◯ VERY LOW
			Mean relapse free months at 1 year	NR	10 months ADM 8 months placebo	⊕◯◯◯ VERY LOW
			Mean difference in relapse rate at 0-3 months	NR	63.9%	⊕◯◯◯ VERY LOW
			Mean difference in relapse rate at 3-6 months	NR	22.6%	⊕◯◯◯ VERY LOW
			Mean difference in relapse rate at 6-9 months	NR	9.3%	⊕◯◯◯ VERY LOW
			Mean difference in relapse rate at 9-12 months	NR	4.2%	⊕◯◯◯ VERY LOW
		Relapse	Relapse rate after 6 months post remission	8	OR **0.43** CI 0.33 to 0.56	⊕⊕◯◯ LOW
GRADE Table 2:	Kato et al., 2020	Reduction in mental health symptoms	-	-	-	N/A

GRADE Table	Source	Outcome	Specific Outcome	Number of Studies	Effects	Certainty of Evidence
Fixed antidepressant dose compared to pill placebo in adults with depressive disorders		Adverse effects	-	-	-	N/A
		Suicide related outcomes	-	-	-	N/A
		Improvement in quality of life and functioning	-	-	-	N/A
		Relapse	Relapse rate after 9 months post remission	27	OR **0.41** CI 0.36 to 0.48	⊕⊕◯◯ LOW
GRADE Table 3: Flexible antidepressant dose compared to pill placebo in adults with depressive disorders	Kato et al., 2020	Reduction in mental health symptoms	-	-	-	N/A
		Adverse effects	-	-	-	N/A
		Suicide related outcomes	-	-	-	N/A
		Improvement in quality of life and functioning	-	-	-	N/A
		Relapse	Relapse rate after 12 months post remission	13	OR **0.30** CI 0.23 to 0.48	⊕⊕◯◯ LOW

4.2 Evidence to decision

Table 4: Evidence to decision table
Please note * indicates evidence from overarching qualitative review by Gronholm et al, 2023.

CRITERIA, QUESTIONS		JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
Priority of the problem	Is the problem a priority? The more serious a problem is, the more likely it is that an option that addresses the problem should be a priority (e.g., diseases that are fatal or disabling are likely to be a higher priority than diseases that only cause minor distress). The more people who are affected, the more likely it is that an option that addresses the problem should be a priority.			
	• Are the consequences of the problem serious (that is, severe or important in terms of the potential benefits or savings)? • Is the problem urgent? • Is it a recognised priority (such as based on a political or policy decision)? [Not relevant when an individual patient perspective is taken]	☐ No ☐ Probably no ☐ Probably yes ☒ **Yes** ☐ Varies ☐ Don't know	• By 2030, depression is predicted to be one of the leading causes of disability and premature mortality worldwide. Reducing the depression burden by developing and scaling up evidence-based interventions is therefore a major global priority. • Previous research suggests that half of the patients on antidepressants have been continuing this treatment for two years or more. However, side effects have been known to increase with long term use. Therefore, it is crucial to determine for how long the treatment should continue.	

CRITERIA, QUESTIONS	JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
Desirable Effects	How substantial are the desirable anticipated effects? The larger the benefit, the more likely it is that an option should be recommended.		
• Judgments for each outcome for which there is a desirable effect • How substantial (large) are the desirable anticipated effects (including health and other benefits) of the option (taking into account the severity or importance of the desirable consequences and the number of people affected)?	☐ Trivial ☒ **Small** ☐ Moderate ☐ Large ☐ Varies ☐ Don't know	• The relapse rate at 1 year was lower for pharmacotherapy compared to placebo • The average of relapse free months over a year was higher in the pharmacotherapy groups than in the placebo groups • Most of the difference in 1 year relapse rates between pharmacotherapy and placebo occurred in the first 3 months (63.9%) and first 6 months (85.5 %). The difference in relapse free rates becomes much smaller after 6 months. • Both fixed and flexible pharmacotherapy dose settings were associated with lower relapse rates compared to pill placebo and 9 and 12 months respectively.	

CRITERIA, QUESTIONS	JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
Undesirable Effects	How substantial are the undesirable anticipated effects? The greater the harm, the less likely it is that an option should be recommended.		
• Judgments for each outcome for which there is an undesirable effect • How substantial (large) are the undesirable anticipated effects (including harms to health and other harms) of the option (taking into account the severity or importance of the adverse effects and the number of people affected)?	☐ Large ☐ Moderate ☐ Small ☐ Trivial ☐ Varies ☒ **Don't know**	There was no direct evidence to evaluate the effects of the length of treatment on undesirable effects • In terms of acceptability, the pooled OR of 32 studies including 7146 subjects was 0.47 (95% CI, 0.40–0.55, $Z = 9.50$ $p < 0.00001$), favoring antidepressant continuation over placebo. The RD of the rate of acceptability between antidepressant (43.3%) and placebo (58.2%) groups was 0.17 (95% CI, 0.14–0.20, $Z = 10.68$ $p < 0.00001$) and NNT was 7. • For the rate of tolerability, pooled ORs of 28 studies with 6897 subjects was 1.15 (95% CI, 0.79–1.67, $Z = 0.72$ $p = 0.47$) and RD was 0.01 (95% CI, −0.01 to 0.02, $Z = 1.03$ $p = 0.30$) without significant differences between antidepressant (4.1%) and placebo (3.9%) groups	

CRITERIA, QUESTIONS	JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
Certainty of evidence	What is the overall certainty of the evidence of effects? The less certain the evidence is for critical outcomes (those that are driving a recommendation), the less likely that an option should be recommended (or the more important it is likely to be to conduct a pilot study or impact evaluation, if it is recommended).		
• What is the overall certainty of this evidence of effects, across all of the outcomes that are critical to making a decision? • See GRADE guidance regarding detailed judgments about the quality of evidence or certainty in estimates of effects	☒ **Very low** ☐ Low ☐ Moderate ☐ High ☐ No included studies	• The certainty of evidence for the mean difference in relapse rates between antidepressant medications (ADM) and pill placebo from 0 to 9 months, and relapse free months at 1 year post remission, response or recovery is **very low**. • The certainty of evidence is **low** for relapse rates at 6, months post remission between ADM and pill placebo. • There is **low** evidence for relapse rates at 9 and 12 months for ADM compared to pill placebo in fixed and flexible dose settings respectively.	

CRITERIA, QUESTIONS	JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
Values	Is there important uncertainty about or variability in how much people value the main outcomes? The more likely it is that differences in values would lead to different decisions, the less likely it is that there will be a consensus that an option is a priority (or the more important it is likely to be to obtain evidence of the values of those affected by the option). Values in this context refer to the relative importance of the outcomes of interest (how much people value each of those outcomes). These values are sometimes called 'utility values'.		
• Is there important uncertainty about how much people value each of the main outcomes? • Is there important variability in how much people value each of the main outcomes?	☐ Important uncertainty or variability ☐ Possibly important uncertainty or variability ☒ **Probably no important uncertainty or variability** ☐ No important uncertainty or variability	• There was no direct evidence to evaluate values and preferences of people • *Overall, the studies highlighted importance and recognition of importance of mental health interventions and the outcomes of those interventions on people's mental health and wellbeing. • The value could be limited by certain factors and barriers present in the health systems. For instance, low awareness, poor funding and poor political buy-in, or other social barriers (Badu *et al.* 2018; Padmanathan & De Silva 2013; Sarkar *et al.* 2021; Verhey *et al.* 2020). • Social networks or raising awareness can facilitate adoption and recognition of mental health issues and the perceived value of the interventions (Amaral *et al.* 2018; Brooke-Sumner *et al.* 2015; Dickson & Bangpan 2018; Verhey *et al.* 2020).	

CRITERIA, QUESTIONS		JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
Balance of effects	Does the balance between desirable and undesirable effects favor the intervention or the comparison? The larger the desirable effects in relation to the undesirable effects, taking into account the values of those affected (i.e. the relative value they attach to the desirable and undesirable outcomes) the more likely it is that an option should be recommended.			
	• Judgments regarding each of the four preceding criteria • To what extent do the following considerations influence the balance between the desirable and undesirable effects: - How much less people value outcomes that are in the future compared to outcomes that occur now (their discount rates)? - People's attitudes towards undesirable effects (how risk averse they are)? - People's attitudes towards desirable effects (how risk seeking they are)?	☐ Favors the comparison ☐ Probably favors the comparison ☐ Does not favor either the intervention or the comparison ☒ **Probably favors the intervention** ☐ Favors the intervention ☐ Varies ☐ Don't know	• The balance of the effects probably favours ADM continuation over pill placebo. Nevertheless, more research is needed to draw firm conclusions on the balance between favourable and unfavourable outcomes of ADM vs pill placebo.	

CRITERIA, QUESTIONS		JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
Resources required	How large are the resource requirements (costs)? The greater the cost, the less likely it is that an option should be a priority. Conversely, the greater the savings, the more likely it is that an option should be a priority.			
	• How large is the difference in each item of resource use for which <u>fewer</u> resources are required? • How large is the difference in each item of resource use for which <u>more</u> resources are required? • How large an investment of resources would the option require or save?	☐ Large costs ☐ Moderate costs ☐ Negligible costs and savings ☐ Moderate savings ☐ Large savings ☒ **Varies** ☐ Don't know	There was no direct evidence to evaluate resource requirements	Additional information: • This varies depending on the type of antidepressant and treatment duration

CRITERIA, QUESTIONS		JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
Certainty of evidence of required resources	What is the certainty of the evidence of resource requirements (costs)?			
	• Have all-important items of resource use that may differ between the options being considered been identified? • How certain is the evidence of differences in resource use between the options being considered (see GRADE guidance regarding detailed judgments about the quality of evidence or certainty in estimates)? • How certain is the cost of the items of resource use that differ between the options being considered? • Is there important variability in the cost of the items of resource use that differ between the options being considered?	☐ Very low ☐ Low ☐ Moderate ☐ High ☒ **No included studies**	There was no direct evidence to evaluate resource requirements	

CRITERIA, QUESTIONS		JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
Cost effectiveness	Does the cost-effectiveness of the intervention favor the intervention or the comparison? The greater the cost per unit of benefit, the less likely it is that an option should be a priority.			
	• Judgments regarding each of the six preceding criteria • Is the cost effectiveness ratio sensitive to one-way sensitivity analyses? • Is the cost effectiveness ratio sensitive to multivariable sensitivity analysis? • Is the economic evaluation on which the cost effectiveness estimate is based reliable? • Is the economic evaluation on which the cost effectiveness estimate is based applicable to the setting(s) of interest?	☐ Favors the comparison ☐ Probably favors the comparison ☐ Does not favor either the intervention or the comparison ☐ Probably favors the intervention ☐ Favors the intervention ☐ Varies ☒ **No included studies**	No reviews examining cost effectiveness identified	

CRITERIA, QUESTIONS	JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
Health equity, equality and non-discrimination	What would be the impact on health equity, equality and non-discrimination? (WHO INTEGRATE) Health equity and equality reflect a concerted and sustained effort to improve health for individuals across all populations, and to reduce avoidable systematic differences in how health and its determinants are distributed. Equality is linked to the legal principle of non-discrimination, which is designed to ensure that individuals or population groups do not experience discrimination on the basis of their sex, age, ethnicity, culture or language, sexual orientation or gender identity, disability status, education, socioeconomic status, place of residence or any other characteristics. All recommendations should be in accordance with universal human rights standards and principles. The greater the likelihood that the intervention increases health equity and/or equality and that it reduces discrimination against any particular group, the greater the likelihood of a general recommendation in favor of this intervention.			
• How are the condition and its determinants distributed across different population groups? Is the intervention likely to reduce or increase existing health inequalities and/or health inequities? Does the intervention prioritise and/or aid those furthest behind? • How are the benefits and harms of the intervention distributed across the population? Who carries the burden (e.g. all), who benefits (e.g. a very small sub-group)? • How affordable is the intervention for individuals, workplaces or communities? • How accessible - in terms of physical as well as informational access - is the intervention across different population groups? • Is there any suitable alternative to addressing the condition, does the intervention represent the only available option? Is this option proportionate to the need, and will it be subject to periodic review?	☐ Reduced ☐ Probably reduced ☐ Probably no impact ☐ Probably increased ☐ Increased ☒ **Varies** ☐ Don't know	There was no direct evidence to evaluate health equity, equality and non-discrimination. *The review noted considerations for ensuring MNS interventions are equitable, equally available and non-discriminatory: • Accessibility, physical/practical considerations • time & travel constraints. • Accessibility, informational barriers • Affordability - medication and treatment costs These factors may be exacerbated for certain groups: • People with low education/literacy - e.g. written instructions, psychoeducation materials • Women - travel restrictions, stronger stigma/shame, caregiving responsibilities Low resource settings - affordability/cost considerations exacerbated	

CRITERIA, QUESTIONS	JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
Feasibility	Is the intervention feasible to implement? The less feasible (capable of being accomplished or brought about) an option is, the less likely it is that it should be recommended (i.e. the more barriers there are that would be difficult to overcome).		
• Can the option be accomplished or brought about? • Is the intervention or option sustainable? • Are there important barriers that are likely to limit the feasibility of implementing the intervention (option) or require consideration when implementing it?	☐ No ☐ Probably no ☒ **Probably yes** ☐ Yes ☐ Varies ☐ Don't know	There was no direct evidence to evaluate feasibility. *Included reviews considered feasibility, and how this can be enhanced • Acceptability of interventions for stakeholders - requires increased engagement with specialist staff, increased visibility of the task-sharing workforce within health facilities, perception of usefulness by providers and service users (e.g. via positive feedback), context-specific interventions, standardised implementation steps for simpler decision-making and delivery • Health worker workload, competency- requires training, refreshers, supervision; networking with others in same role. • Availability of a task-sharing workforce • Availability of caregivers • Participant education and literacy requires verbal explanations/tasks. • Logistical issues such as e.g. mobile populations, affordability of travel to receive care, lack of private space. • Limited resources/mental health budget Sustainability considerations: • Training and supervision • Integrating into routine clinical practice	

CRITERIA, QUESTIONS		JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
			Provider type (e.g. formally employed lay health workers vs. volunteers)	
Human rights and sociocultural acceptability	Is the intervention aligned with human rights principles and socio-culturally acceptable? (WHO INTEGRATE) This criterion encompasses two distinct constructs: The first refers to an intervention's compliance with universal human rights standards and other considerations laid out in international human rights law beyond the right to health (as the right to health provides the basis of other criteria and sub-criteria in this framework). The second, sociocultural acceptability, is highly time-specific and context-specific and reflects the extent to which those implementing or benefiting from an intervention as well as other relevant stakeholder groups consider it to be appropriate, based on anticipated or experienced cognitive and emotional responses to the intervention. The greater the sociocultural acceptability of an intervention to all or most relevant stakeholders, the greater the likelihood of a general recommendation in favor of this intervention.			
	• Is the intervention in accordance with universal human rights standards and principles? • Is the intervention socio-culturally acceptable to patients/beneficiaries as well as to those implementing it? To which extent do patients/beneficiaries value different non-health outcomes? • Is the intervention socio-culturally acceptable to the public and other relevant stakeholder groups? Is the intervention sensitive to sex, age, ethnicity, culture or language, sexual orientation or gender identity, disability status, education, socio-economic status, place of residence or any other relevant characteristics? • How does the intervention affect an individual's, population group's or organization's autonomy, i.e. their ability to make a competent, informed and voluntary decision?	☐ No ☐ Probably no ☐ Probably yes ☐ Yes ☒ **Varies** ☐ Don't know	There was no direct evidence to evaluate alignment with human rights principle and socio-cultural acceptability. *The review noted a number of considerations which would impact the right to health and access to healthcare. E.g. stigma and discrimination and lack of confidentiality could affect the help-seeking among service users. • The importance of socio-cultural acceptability of MNS interventions was clearly expressed. Pre-intervention considerations that take into account cultural and social aspects improve the acceptability of implemented interventions. • When interventions were perceived as appropriate for the culture and target group, the content and medium of the intervention received more positive feedback from service users and caregivers Also, considerations of age,	

CRITERIA, QUESTIONS	JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
• How intrusive is the intervention, ranging from low intrusiveness (e.g. providing information) to intermediate intrusiveness (e.g. guiding choices) to high intrusiveness (e.g. restricting or eliminating choices)? Where applicable, are high intrusiveness and/or impacts on the privacy and dignity of concerned stakeholders justified?		sex and language have been highlighted as important to acceptability and accessibility. Mitigating steps to improve sociocultural acceptability include: • To train health workers in non-judgemental care • Integrate preventative mental health awareness messages to reduce the stigma • Train acceptable counsellors for the local settings and target groups Facilitate the use of indigenous/ local phrases and terms to increase acceptability, accessibility and fidelity.	

4.3. Summary of judgements

Table 5: Summary of judgements

This provides a snapshot of the evidence to decision table.

Priority of the problem	- Don't know	- Varies		- No	- Probably No	- Probably Yes	✓ **Yes**
Desirable effects	- Don't know	- Varies		- Trivial	✓ **Small**	- Moderate	- Large
Undesirable effects	✓ **Don't know**	- Varies		- Large	- Moderate	- Small	- Trivial
Certainty of the evidence	- No included studies			✓ **Very low**	- Low	- Moderate	- High
Values				- Important uncertainty or variability	- Possibly important uncertainty or variability	✓ **Probably no important uncertainty or variability**	- No important uncertainty or variability
Balance of effects	- Don't know	- Varies	- Favors comparison	- Probably favors comparison	- Does not favor either	✓ **Probably favors intervention**	- Favours intervention
Resources required	- Don't know	✓ **Varies**	- Large costs	- Moderate costs	- Negligible costs or savings	- Moderate savings	- Large savings
Certainty of the evidence on required resources	✓ **No included studies**			- Very low	- Low	- Moderate	- High
Cost–effectiveness	✓ **No included studies**	- Varies	- Favors comparison	- Probably favors comparison	- Does not favor either	- Probably favors intervention	- Favors intervention
Equity, equality and non-discrimination	- Don't know	✓ **Varies**	- Reduced	Probably reduced	- Probably no impact	- Probably increased	- Increased
Feasibility	- Don't know	- Varies		- No	- Probably No	✓ **Probably Yes**	- Yes
Human rights and socio-cultural acceptability	- Don't know	✓ **Varies**		- No	- Probably No	- Probably Yes	- Yes

✓ Indicates category selected, - Indicates category not selected

5. References

Amaral CE, Onocko-Campos R, de Oliveira PRS, Pereira MB, Ricci ÉC, Pequeno ML, Emerich B, Dos Santos RC, Thornicroft G. Systematic review of pathways to mental health care in Brazil: narrative synthesis of quantitative and qualitative studies. Int J Ment Health Syst. 2018 Oct 31;12:65. doi: 10.1186/s13033-018-0237-8. PMID: 30450125; PMCID: PMC6208112.

Badu, E., O'Brien, A. P., & Mitchell, R. (2018). An integrative review of potential enablers and barriers to accessing mental health services in Ghana. Health research policy and systems, 16, 1-19.

Bloom, D. E., Cafiero, E., Jané-Llopis, E., Abrahams-Gessel, S., Bloom, L. R., Fathima, S., ... & O'Farrell, D. (2012). The global economic burden of noncommunicable diseases (No. 8712). Program on the Global Demography of Aging.

Brooke-Sumner C, Petersen I, Asher L, Mall S, Egbe CO, Lund C. Systematic review of feasibility and acceptability of psychosocial interventions for schizophrenia in low and middle income countries. BMC Psychiatry. 2015 Feb 12;15:19. doi: 10.1186/s12888-015-0400-6. PMID: 25886524; PMCID: PMC4382830

Cipriani, A., Furukawa, T.A., Salanti, G., Chaimani, A., Atkinson, L.Z., Ogawa, Y., Leucht, S., Ruhe, H.G., Turner, E.H., Higgins, J.P. and Egger, M., 2018. 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. Focus, 16(4), pp.420-429.

Dickson K, Bangpan M. What are the barriers to, and facilitators of, implementing and receiving MHPSS programmes delivered to populations affected by humanitarian emergencies? A qualitative evidence synthesis. Glob Ment Health (Camb). 2018 Jun 1;5:e21. doi: 10.1017/gmh.2018.12. PMID: 29997893; PMCID: PMC6036649.

Ferguson, J.M., 2001. SSRI antidepressant medications: adverse effects and tolerability. Primary care companion to the Journal of clinical psychiatry, 3(1), p.22.

Ferrari, A. J., Charlson, F. J., Norman, R. E., Patten, S. B., Freedman, G., Murray, C. J., ... & Whiteford, H. A. (2013). Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. PLoS medicine, 10(11), e1001547.

Fletcher, R., Leaman, J., McSloy, A. and Leng, G., 2020. NICE Update NICE public health guidance update.

Gronholm, P., Makhmud, A., Barbui, C., Brohan, E. & Chowdhary, N. (2023). Qualitative evidence regarding the experience of receiving and providing care for mental health conditions in non-specialist settings in low-and middle-income countries: A systematic review of reviews. BMJ Mental Health, 26:e300755.

Ho, S. C., Chong, H. Y., Chaiyakunapruk, N., Tangiisuran, B., & Jacob, S. A. (2016). Clinical and economic impact of non-adherence to antidepressants in major depressive disorder: a systematic review. Journal of affective disorders, 193, 1-10.

Ioannidis JP. Effectiveness of antidepressants: an evidence myth constructed from a thousand randomized trials? Philos Ethics Humanit Med 2008; 3: 14

Johnson CF, Macdonal HJ, Atkinson P, et al. Reviewing long-term antidepressants can reduce drug burden: a prospective observational cohort study. Br J Gen Pract 2012;62:773–9.

Kessler, R. C., & Bromet, E. J. (2013). The epidemiology of depression across cultures. Annual review of public health, 34, 119-138.

Luo, Y., Kataoka, Y., Ostinelli, E.G., Cipriani, A. and Furukawa, T.A., 2020. National prescription patterns of antidepressants in the treatment of adults with major depression in the US between 1996 and 2015: a population representative survey based analysis. Frontiers in psychiatry, 11, p.35.

Mathers, C. D., & Loncar, D. (2006). Projections of global mortality and burden of disease from 2002 to 2030. PLoS medicine, 3(11), e442.

Mojtabai R, Olfson M. National trends in long-term use of antidepressant medications: Results from the US National Health and Nutrition Examination Survey. J Clin Psychiatry 2014;75:169–77.

Nathan, P.E. and Gorman, J.M. eds., 2015. A guide to treatments that work. Oxford University Press.

Padmanathan P, De Silva MJ. The acceptability and feasibility of task-sharing for mental healthcare in low and middle income countries: a systematic review. Soc Sci Med. 2013 Nov;97:82-6. doi: 10.1016/j.socscimed.2013.08.004. Epub 2013 Aug 24. PMID: 24161092.

Petty DR, House A, Knapp P, et al. Prevalence, duration and indications for prescribing of antidepressants in primary care. Age Ageing 2006;35:523–6.

Sarkar, S., Taylor, A., Dutta, P., Bajaj, M., Nash, J., Ravola, M., Ievleva, S., Llyod, C., Ola, P., Jenkins, B. and Sengupta, B., 2021. Health disparity and COVID-19—A retrospective analysis. Health Science Reports, 4(3), p.e345.

Thomas L, Kessler D, Campbell J, et al. Prevalence of treatment-resistant depression in primary care: cross-sectional data. Br J Gen Pract. 2013;63(617):e852-e858. doi:10.3399/bjgp13X675430

Verhey, I.J., Ryan, G.K., Scherer, N. et al. Implementation outcomes of cognitive behavioural therapy delivered by non-specialists for common mental disorders and substance-use disorders in low- and middle-income countries: a systematic review. Int J Ment Health Syst 14, 40 (2020). https://doi.org/10.1186/s13033-020-00372-9

World Bank Group, WHO (2016). Report of Proceedings of Event “Out of the Shadows: Making Mental Health a Global Development Priority.” Washington, DC: World Bank Group, 2016. http://www.worldbank.org/en/topic/health/brief/mental-health. (Accessed May 2019).

World Health Organization, 2016. mhGAP intervention guide for mental, neurological and substance use disorders in non-specialized health settings: mental health Gap Action Programme (mhGAP). World Health Organization.

Appendix I: Search terms used to identify systematic reviews

PubMed

1# Depression

"Depression"[Mesh] OR "Depressive Disorder"[Mesh] OR “depress*”[tiab] OR “dysthymi*”[tiab] OR “mood disorder*”[tiab] OR “affective disorder*”[tiab] OR “dysphoric disorder*”[tiab]

2# Antidepressants

"Antidepressive Agents"[Mesh:NoExp] OR "Serotonin Uptake Inhibitors"[Mesh] OR "Antidepressive Agents, Tricyclic" [Mesh] OR "Fluoxetine"[Mesh] OR "Citalopram"[Mesh] OR "Sertraline"[Mesh] OR "Nortriptyline"[Mesh] OR "Antidepressive Agents" [Pharmacological Action] OR "Serotonin Uptake Inhibitors" [Pharmacological Action] OR "Antidepressive Agents, Tricyclic" [Pharmacological Action] OR "antidepressiv*"[tiab] OR "anti-depressiv*"[tiab] OR antidepressant*[tiab] OR "anti-depressant*"[tiab] OR thymoleptic*[tiab] OR thymoanaleptic*[tiab] OR "Serotonin Reuptake Inhibitor*"[tiab] OR "Serotonin Re-uptake Inhibitor*"[tiab] OR "Serotonin uptake Inhibitor*"[tiab] OR “serotonin specific reuptake inhibitor*”[tiab] OR “serotonin specific re-uptake inhibitor*”[tiab] OR SSRI*[tiab] OR TCA[tiab] OR TCAs[tiab] OR alaproclate [tiab] OR Citalopram [tiab] OR Celexa [tiab] OR Cipramil [tiab] OR Escitalopram [tiab] OR Lexapro [tiab] OR Cipralex [tiab] OR Fluoxetine [tiab] OR Prozac [tiab] OR Sarafem [tiab] OR Fluvoxamine [tiab] OR Luvox [tiab] OR Faverin [tiab] OR Paroxetine [tiab] OR Paxil [tiab] OR Seroxat [tiab] OR Sertraline [tiab] OR Zoloft [tiab] OR Lustral [tiab] OR Vilazodone [tiab] OR Viibryd [tiab] OR femoxetine [tiab] OR indalpine [tiab] OR Zimeldine [tiab] OR Amitriptyline [tiab] OR Elavil [tiab] OR Endep [tiab] OR Amitriptylinoxide [tiab] OR Amioxid [tiab] OR Ambivalon [tiab] OR Equilibrin [tiab] OR Clomipramine [tiab] OR Anafranil [tiab] OR Desipramine [tiab] OR Norpramin [tiab] OR Pertofrane [tiab] OR Dibenzepin [tiab] OR Noveril [tiab] OR Victoril [tiab] OR Dimetacrine [tiab] OR Istonil [tiab] OR Dosulepin [tiab] OR Prothiaden [tiab] OR Doxepin [tiab] OR Adapin [tiab] OR Sinequan [tiab] OR Imipramine [tiab] OR Tofranil [tiab] OR Lofepramine [tiab] OR Lomont [tiab] OR Gamanil [tiab] OR Melitracen [tiab] OR Dixeran [tiab] OR Melixeran [tiab] OR Trausabun [tiab] OR Nitroxazepine [tiab] OR Sintamil [tiab] OR Nortriptyline [tiab] OR Pamelor [tiab] OR Aventyl [tiab] OR Noxiptiline [tiab] OR Agedal [tiab] OR Elronon [tiab] OR Nogedal [tiab] OR Opipramol [tiab] OR Insidon [tiab] OR Pipofezine [tiab] OR Azafen [tiab] OR Azaphen [tiab] OR Protriptyline [tiab] OR Vivactil [tiab] OR Trimipramine [tiab] OR Surmontil [tiab] OR Amoxapine [tiab] OR Asendin [tiab] OR cericlamine [tiab] OR dapoxetine [tiab] OR ifoxetine [tiab] OR litoxetine [tiab] OR lubazodone [tiab] OR moxifetin [tiab] OR nomelidine [tiab] OR norcitalopram [tiab] OR norfluoxetine [tiab] OR seproxetine [tiab] OR norsertraline [tiab] OR omiloxetine [tiab]

3# SR + MA filter

("Meta-Analysis" [Publication Type] OR "Meta-Analysis as Topic"[Mesh] OR metaanaly*[tiab] OR meta-analy*[tiab] or metanaly*[tiab] OR "Systematic Review" [Publication Type] OR systematic[sb] OR meta-analysis[Filter] OR systematicreview[Filter] OR "Cochrane Database Syst Rev"[Journal] or prisma[tiab] OR

“preferred reporting items”[tiab] OR prospero[tiab] OR ((systemati*[ti] OR umbrella[ti] OR “structured literature”[ti]) AND (review[ti] OR overview[ti])) OR “systematic review”[tiab] OR “umbrella review”[tiab] OR “structured literature review”[tiab] OR “systematic qualitative review”[tiab] OR “systematic quantitative review”[tiab] OR “systematic search and review”[tiab] OR “systematized review”[tiab] OR “systematised review”[tiab] OR “systemic review”[tiab] OR “systematic literature review”[tiab] OR “systematic integrative literature review”[tiab] OR “systematically review”[tiab] OR “scoping literature review”[tiab] OR “scoping review”[tiab] OR “systematic critical review”[tiab] OR “systematic integrative review”[tiab] OR “systematic evidence review”[tiab] OR “systematic integrative literature review”[tiab] OR “systematic mixed studies review”[tiab] OR “systematized literature review”[tiab] OR “systematic overview”[tiab] OR “Systematic narrative review”[tiab] OR “narrative review”[tiab] OR metasynthes*[tiab] OR meta-synthes*[tiab]) NOT ("Comment" [Publication Type] OR "Letter" [Publication Type] OR "Editorial" [Publication Type] OR (("Animals"[Mesh] OR "Models, Animal"[Mesh]) NOT "Humans"[Mesh]))

 # Timeframe
2019-2022

Appendix II: Decision Tree used to evaluate ROB GRADE item

Figure 2: Developed tree for the assessment of the risk of bias item in GRADE