

Adverse health events associated with antipsychotic use: An umbrella review

Rachel Tsz Suet Chow
Kellogg College

A thesis submitted in fulfilment of the requirements for the degree:

Master of Science by Research in Psychiatry

Trinity Term, 2021

Word Count: 21,919

Acknowledgements

I would like to express my deepest gratitude to everyone who contributed to my work presented in this thesis. I am especially grateful to my supervisor Professor Seena Fazel and Dr. Tapio Paljärvi, for their consistent support and guidance throughout the year. Many thanks to Dr. Stella Botchway for her constructive feedback on my work. I feel very honoured to be part of the Forensic Psychiatry team and look forward to future collaborations. I would also like to thank Professor Phil Cowen and Professor Rebecca Park for supporting my Transfer of Status viva. Thanks to Professor Phil Burnet for his many useful advice on my progress.

Special thanks to David Chow, Beverley Chan and Kirsten Chow, for their love and support as always. Thank you to my friends in Oxford for making my time here so enjoyable and all the late-night conversations. I also personally thank my friends in London and Manchester for their encouragement and being so supportive.

I disclose no conflict of interest or external source of funding for this umbrella review.

Table of Contents

Acknowledgements ..... *i*
Table of Contents ..... *ii*
List of Abbreviations ..... *iv*
Abstract ..... *v*
Chapter 1: Introduction *1*
1.1 The Use of Antipsychotic Medications in Treatment **1**
1.2 Adverse Drug Events Associated with Antipsychotics **3**
1.3 The Importance of Umbrella Reviews in Pharmacovigilance **5**
1.4 Previous Umbrella Reviews on Antipsychotic Medications' Adverse Effects **8**
1.5 Aims of the Current Umbrella Review **10**
Chapter 2: Methods *10*
2.1 Data Sources **10**
2.2 Study Screening and Eligibility Criteria **11**
2.3 Data Extraction and AMSTAR Critical Appraisal **13**
2.4 Data Analysis **15**
2.5 Assessment of Evidence Quality and Credibility **17**
2.6 Subgroup Analyses **21**
2.7 Classification and Criteria for Adverse Outcomes **21**
Chapter 3: Results *24*
3.1 Study Characteristics **24**
3.2 Overview of Effect Size Data **35**
3.2.1 Circulatory System Outcomes 36
3.2.2 Endocrinological and Metabolic Outcomes 38
3.2.3 Injury 40
3.2.4 Mortality 40
3.2.5 Musculoskeletal and Movement-Related Outcomes 41
3.2.6 Neuropsychiatric Outcomes 43
3.2.7 Other Outcomes 45
3.2.8 Adverse Outcome Categories 47
3.3 Assessment of Heterogeneity **51**
3.4 Subgroup Analyses **52**
3.5 AMSTAR 2 Quality Assessment **55**
3.6 Prediction Intervals **58**
3.7 Small Study Effects **58**
3.8 Excess Significance Bias **60**
3.9 Summary of Quality Assessments **64**
Chapter 4: Discussion *67*

4.1 Antipsychotic-Associated Adverse Outcomes **67**
4.2 Quality Assessment **69**
4.3 Comparison with Previous Umbrella Review **74**
4.4 Susceptibility to Adverse Drug Outcomes **75**
4.5 Strengths and Limitations of This Umbrella Review **77**
4.6 Implications **81**
4.7 Conclusion **83**
References ***85***
Appendix A: PRISMA Harms Checklist ***98***
Appendix B: Search Strategy for PubMed ***100***
Appendix C: Screening Template ***101***
Appendix D: AMSTAR 2 Critical Appraisal Instrument ***103***
Appendix E: Formulae Used in Meta-Analysis ***105***
Appendix F: Outcomes Excluded from Analyses with Reasons ***107***
Appendix G: Potentially Relevant But Excluded Studies ***113***
Appendix H: Individual Effect Sizes Reported in Meta-Analyses ***117***
Appendix I: Distribution of RORs ***126***
Appendix J: Strength Assessment of Risk Ratios and Odds Ratios ***127***
Supplementary References ***128***

List of Abbreviations

AD	Antidepressant
AMSTAR	A Measurement Tool to Assess Systematic Reviews
BAP	British Association for Pharmacology
BMI	Body mass index
CI	Confidence interval
CTCAE	Common Terminology Criteria for Adverse Events
DoTs	Dose, Time and Susceptibility
DSM	Diagnostic and Statistical Manual
ESS	Excess statistical significance
FGA	First-generation antipsychotics
GRADE	Grading of Recommendations, Assessment, Development and Evaluations
HR	Hazards ratio
ICD	International Classification of Diseases
MDD	Major depressive disorder
MHRA	Medicines and Healthcare products Regulatory Agency
NICE	National Institute for Health and Care Excellence
OCD	Obsessive compulsive disorder
OFC	Olanzapine-fluoxetine combination
OR	Odds ratio
PI	Prediction interval
PRISMA	Preferred reporting items for systematic reviews and meta-analyses
PTSD	Post-traumatic stress disorder
RCT	Randomised controlled trial
ROR	Ratio of odds ratios
RR	Risk ratio
SGA	Second-generation antipsychotics
SMI	Serious mental illnesses
URC	Umbrella Review Criteria
WHO	World Health Organisation

Abstract

Title: Adverse health events associated with antipsychotic use: An umbrella review
Name: Rachel Tsz Suet Chow
Degree: Master of Science by Research in Psychiatry
Affiliations: Department of Psychiatry & Kellogg College, University of Oxford
Term: Trinity Term, 2021

INTRODUCTION: Antipsychotic medications are prescribed in nearly 800,000 patients in England. Monitoring the adverse effects of antipsychotics is essential in the prevention of harm to patients. Numerous systematic reviews have found evidence for antipsychotics' adverse effects, but few attempts have been made to assess the robustness of these evidence. This umbrella review aimed to grade and summarise the evidence for all adverse health events associated with antipsychotic use in the community settings, and assess the potential factors moderating the antipsychotic-adverse event associations, through the comparison across outcomes, quality assessments and identification of the current literature gaps.

METHODS: A systematic search was conducted on 6 databases (PubMed, Embase, PsycINFO, Scopus, CINAHL and Cochrane Library of Systematic Reviews). The risk of each adverse outcome was estimated with RRs or ORs. Subgroup analyses were conducted on the distinct classes and types of antipsychotics, age, diagnoses and study designs. The quality of the evidence of reported associations was examined via assessments on the extent of heterogeneity, 95% prediction intervals, small study effects and excess significance bias in the underlying review literature. The methodological quality of included meta-analyses was assessed with the AMSTAR 2 tool.

RESULTS: Twenty-seven meta-analyses of RCTs or observational studies were identified. The included meta-analyses of RCTs and observational studies reported statistically significant effect sizes for 26 adverse outcomes. Meta-analyses of RCTs found antipsychotics to increase the risk of endocrinological and metabolic, neuropsychiatric and movement-related outcomes. Meta-analyses on observational data found significant associations between antipsychotic use and injuries, circulatory system outcomes and mortality. Evidence was of variable quality, given the heterogenous findings, asymmetry in the funnel plots, and the diverse AMSTAR 2 ratings, prediction intervals and excess significance ratios. The most robust adverse outcomes were gait abnormality (RR 2.74, 95% CI 1.66 to 4.53) and hypertonia (RR 2.91, 95% CI 1.31 to 6.47). Other outcomes were supported by evidence of low-to-moderate quality. Subgroup analyses found atypical antipsychotics to be associated with the highest risk of adverse outcomes, compared to typical antipsychotics. Within SGAs, olanzapine was associated with a higher risk of adverse outcomes than quetiapine and risperidone. In terms of patient variables, older adults aged over 65 and patients with dementia were more likely to experience antipsychotic-associated adverse outcomes than adults and patients with schizophrenia.

CONCLUSIONS: Significant associations were found between antipsychotic use and a wide range of body functions. No conclusive evidence was found for how pharmacological and clinical factors impact on the occurrence of adverse outcomes due to the non-statistically significant difference between subgroups. The inconsistent quality of current evidence demonstrated the need for better conducted studies following reporting guidelines in evaluating the harm profiles of antipsychotics.

Chapter 1: Introduction

1.1 The Use of Antipsychotic Medications in Treatment

Antipsychotic pharmacotherapy has been considered the first-line treatment for psychotic disorders and is also utilised in the treatment of other SMIs such as bipolar disorder and MDD. Prescribed in 798 thousand identified mental health patients in England (NHS, 2020), antipsychotic medications are categorised into the first-generation (typical/ neuroleptics; FGAs) and second-generation (atypical; SGAs) classes. Currently, twelve FGAs and eight SGAs are licensed for use in the UK (MHRA, 2005). The most commonly prescribed FGAs in UK primary care are haloperidol, chlorpromazine and trifluoperazine, whereas olanzapine, risperidone and quetiapine are the most commonly prescribed SGAs (Marston et al., 2014).

FGAs serve as a strong dopamine antagonist by blocking dopaminergic receptors and transmission in the nigrostriatal, mesolimbic and tuberoinfundibular pathways (Habibi, 2010; Li et al., 2016; Lind et al., 2005). FGAs' D_2 antagonism is considered their hallmark pharmacology and has been found to correlate with their clinical potency in treating positive symptoms associated with schizophrenia (i.e., hallucinations and delusions) (Li et al., 2016). Nevertheless, the fact that they operate by reducing the dopamine overactivity means little benefits are offered to managing negative symptoms such as emotional and cognitive deficits, speech difficulties and social withdrawal, which are associated with deficiency in dopaminergic activity in the mesocortical pathway (Bishara & Taylor, 2008). Concerns as such brought about the emergence of SGAs, which significantly outperform FGAs in their effectiveness against negative symptoms. Referred to as serotonin-dopamine receptor antagonists, most SGAs' mechanism of action is characterised by their potent blockade of $5HT_{2A}$ receptors and relatively weaker antagonism on D_2 receptors (Patil & Schwartz, 2018). Such $5HT_{2A}$ antagonism is thought to underlie their efficacy in improving negative symptoms

or cognitive impairment among patients as it increases the release of dopamine in the prefrontal cortex (Miyamoto et al., 2012). The distinction between SGAs' affinity for $5HT_{2A}$ and D_2 receptors is also considered a basis for their reduced resultant extrapyramidal symptoms compared to FGAs, given the enhancement of dopaminergic transmission via the blockades of $5HT_{2A}$ receptors within the nigrostriatal pathway (Kusumi et al., 2015).

Antipsychotic pharmacotherapy for major depressive disorder is most likely prescribed as an augmentation of antidepressants (e.g., serotonin reuptake inhibitor). While antidepressant monotherapy has been widely regarded as the first-line intervention for depression of all degrees of severity, only around one-third of patients prescribed with antidepressants experienced remission and less than half of them responded to the therapy (Rush et al. 2006; Trivedi et al., 2006). Over the past decade, there has been an accumulation of evidence on the effectiveness of antipsychotic augmentation strategies for such phenomenon termed "refractory" or "treatment-resistant" depression (Nelson & Papakostas, 2009; Wang et al., 2016). In fact, the NICE clinical guideline recommends aripiprazole, quetiapine, risperidone and olanzapine to be the first-line augmentation treatments for depression (NICE, 2009). However, a meta-analysis of RCTs found that while aripiprazole, quetiapine and risperidone significant increased response rates in treatment-resistant patients, no significant difference between olanzapine-fluoxetine combination and placebo was found (Spielmans et al., 2013). The authors highlighted the generally small effect sizes for all efficacy outcomes and the substantial evidence of antipsychotic-associated adverse events. They concluded that cautions must be taken when advising the use of antipsychotics in the treatment of treatment-resistant depression. Such evidence indicates the need for research comprehensively assessing the efficacy-safety profiles of antipsychotic augmentation therapy.

1.2 Adverse Drug Events Associated with Antipsychotics

Adverse drug events are defined as any injury resulting from the administration of a drug, within which those signified by a causal relationship between the drug and harm are referred to as adverse drug reaction, that is any unintended and noxious response to a drug at normal doses (Nebeker et al., 2004; WHO, 2005). Not only could adverse drug events obscure the significance of antipsychotic prescription in the clinical practice, they substantially contribute to patients' non-adherence to treatment, resulting in worsening symptoms (Garcia et al., 2016). Some adverse events could severely impair patients' quality of life or even be fatal (Chawla & Kumar, 2017; Vermeulen et al., 2017).

FGAs' effectiveness is largely limited by extrapyramidal disorders, cognitive dulling and hyperprolactinaemia. While their effectiveness is considered optimal when they antagonise 72% of D_2 receptors in the brain, an occupancy of 75 to 80% onwards triggers acute extrapyramidal symptoms (Chokhawala & Stevens, 2021). FGA's tight binding on D_2 receptors in the mesolimbic pathway underlies their efficacy in reducing positive symptoms, while their long-term occupancy of the D_2 receptors in the nigrostriatal and tuberoinfundibular pathways fosters extrapyramidal symptoms and hyperprolactinaemia respectively (Leppien et al., 2019; Lind et al., 2005). The adverse effect profiles of FGAs could be examined according to their potency. The classification of FGAs in terms of potency is self-explanatory: those with high affinity for D_2 pathways are categorised to be of 'high-potency' (e.g., fluphenazine, haloperidol, pimozide, trifluoperazine), followed by intermediate- (e.g., perphenazine, thiothixene) and low-potency with less affinity for D_2 receptors and relatively more affinity for alpha-adrenergic and H_1-histaminic pathways (e.g., chlorpromazine, promazine, thioridazine) (Gardner et al., 2005). High-potency FGAs are commonly associated with extrapyramidal disorders, such as Parkinsonism, given the depletion of dopamine in the extrapyramidal system of the brain (Weiden, 2007).

Although superior to the FGAs in reducing the incidence of extrapyramidal disorders (Divac et al., 2014), SGAs are associated with long-term metabolic side effects such as weight gain, hyperprolactinaemia and heightened cholesterol levels, and more rarely with serious medical events such as blood and cardiovascular disorders, as indicated by evidence from case-control and cohort studies on patients of schizophrenia (Correll et al., 2015; Muench & Hamer, 2010). Moreover, research found patients with SMI are at a two-to-three times-higher risk for premature death compared to the general population (John et al., 2018; Ringen et al., 2014). While the pharmacokinetics of antipsychotics could undeniably narrow such mortality gap, cross-sectional studies found that their resultant adverse drug events–particularly cardiovascular disease–ironically serve as a major factor of mortality in the SMI or schizophrenia (Vancampfort et al., 2016). Likewise, the use of SGA augmentation may do more harm than good. NICE cautioned that the reports of side effects and dropouts were more prevalent in patients prescribed with SGA augmentative therapy than those on antidepressant monotherapy (Taylor et al., 2020). A meta-analysis also identified a fourfold-higher rate of discontinuation due to adverse events associated with SGAs as an augmentative therapy, compared to placebo (Nelson & Papakostas, 2009).

Proposed by Aronson and Ferner (2003), the DoTs classification system allows the categorisation of adverse drug events into several types of reactions in terms of the medication's dosage, duration between administration and adverse events, and one's vulnerability to adverse events (Calderón-Ospina & Bustamante-Rojas, 2010). In terms of dose-response relationship, adverse events are classified as either supratherapeutic reactions (i.e., at doses beyond recommended dose), collateral reactions (i.e., within recommended dose) or hypersensitivity reactions (i.e., at doses below recommended dose). Adverse events could also be time-independent or time-dependent reactions, occurring fast (i.e., on administration), early, intermediately, late or delayed. Susceptibility, as a category, encompasses a variety of

patient-level factors, including predisposition, age, gender, physiology, pre-existing diseases and the presence of exogenous substance. The DoTs classification system highlights the relevance of individual variables in the development of adverse events in antipsychotic users and thus the need for investigation via informational sources such as clinical and epidemiological studies (MHRA, n.d.). Scattered evidence highlighted the role of multiple clinical (i.e., age, gender, comorbid physical and psychiatric illnesses) and pharmacological (i.e., types of antipsychotic, duration and dose) factors in exacerbating the risk of ventricular arrhythmia, sudden cardiac death, dyslipidemia, neutropenia and agranulocytosis, hypotension, seizure and sialorrhea (Wu, et al., 2015; Yoshida & Takeuchi, 2021). Inconsistent evidence was also found for the effect of age and sex in neurological adverse outcomes and tardive dyskinesia (Straus et al., 2004; Tenback et al., 2009). Syntheses and appraisal of these findings are required for valid conclusions to be made.

In contrast to the well-documented efficacy profiles of antipsychotics, the broad array of antipsychotic-associated adverse events has been inconsistently reported. Such inconsistencies likely originated from the deficient documentation of events and the variations in the assessment measures for adverse events and reporting methods, precluding meaningful comparisons of adverse effects across different antipsychotics (Pope et al., 2018). Considering the serious implications of antipsychotics' adverse effects, researchers should acknowledge the strategic decisions made throughout their collection and assessment of adverse event data, so that a standardised framework could be established to improve the methodology and accuracy of findings in future research (Loke et al., 2007).

1.3 The Importance of Umbrella Reviews in Pharmacovigilance

A fundamental principle of prescribing pharmacological treatments is to inform patients of potential harms to aid cost-benefit analysis in clinical decision-making.

Pharmacovigilance is the science and activities related to the detection, understanding, assessment and prevention of adverse drug events (WHO, 2002). This could be achieved through multiple means of detection, such as automated surveillance of patient treatment data, which is considered highly important and complementary with voluntary reporting (Aspden et al., 2004). An example of voluntary reporting is the Yellow Card Scheme, developed by the UK government, to reinforce pharmacovigilance in England: public reports of adverse drug reactions associated with medications are collected for evaluation, prompting effective recognition and prevention of foreseeable harms to patients (MHRA, 2019).

Observational studies offer insights into the automated surveillance of clinical data as they often analyse data obtained from chart reviews or hospital records. They provide findings reflective of real-life clinical practices owing to their representative samples and longer follow-ups, thus are widely applicable in the scientific research on serious adverse events (Golder et al., 2011). Nevertheless, inconsistently reported data and unrecognised confounding variables pose threats on their credibility and prevent the inferences of causality.

RCTs are regarded as the gold standard design for establishing causality in pharmacological treatments. In fact, researchers suggest RCTs to triumph on the hierarchy of evidence addressing clinical questions (Barton, 2000). RCTs are less prone to bias and capture clinically relevant effect sizes for most acutely occurring adverse events (Solmi et al., 2018). Still, unlike observational studies, their findings could be criticised for being underpowered in detecting differences between treatment arms and lacking external validity for their susceptibility to the exclusion of patients with significant psychiatric or medical comorbidities and those who may be more treatment-resistant or vulnerable to adverse drug events (Phillips et al., 2019; Suzuki et al., 2019).

While a meta-analysis found little systematic difference between the adverse effect estimates generated from meta-analyses of RCTs and meta-analyses of observational studies,

the authors suggested that systematic reviews of a broad range of study designs would be more preferable in building a complete picture of harm profiles, as opposed to those restricting to a single design (Golder et al., 2011). Given the complementary benefits and flaws of the observational and RCT designs in studying the adverse effects of medications, findings of reviews are expected to be the most conclusive when studies of both designs are included.

Numerous systematic reviews and meta-analyses have been conducted to aggregate findings on the adverse effects of antipsychotics from RCTs or observational studies. Systematic reviews involve standardised inclusion criteria and search strategies for the identification of all relevant studies on a given area, and as one of its components, meta-analyses quantitatively synthesise and compare outcomes across studies to provide precise effects denoted by statistical significance (Grant & Booth, 2009). While useful in addressing clinical research questions, the accumulation of systematic reviews and meta-analyses overtime could hinder clinical decision-making and result in research wastes, owing to their potentially flawed but unevaluated methodology and redundancy in their included studies (Correll et al., 2017). Furthermore, the time and resources required to draw effective conclusions from these fragmented evidence are simply incongruent to the clinical reality.

To address such knowledge gap, umbrella reviews could be conducted to offer rapid syntheses of evidence by pooling and appraising the data of all systematic reviews on antipsychotic-associated adverse events (Aromataris et al., 2015). An umbrella review is a systematic review of systematic reviews or meta-analysis of meta-analyses. It enables potential assessments on the consistency and reason of findings and the conclusions made in multiple reviews, which are crucial in research on the safety of psychotropic medications (Solmi et al., 2018). Ideally, the consistency in harm effect estimates from both observational studies and RCTs could be juxtaposed and have their credibility and validity fortified using an umbrella review.

Like systematic reviews and meta-analyses, the PRISMA checklist has been recommended as a tool to inform the reporting of umbrella reviews (Papageorgiou & Biondi-Zoccai, 2016). The PRISMA statement contains a 27-item checklist of components deemed essential for documentation in a systematic review or meta-analysis encompassing seven areas: title, abstract, introduction, methods, results, discussion and funding. Reviewers are encouraged to publish a protocol for their study in registries such as PROSPERO and to provide rationale and meaning for each item on the checklist with examples and justifications. In view of the generally inadequate reporting of adverse events-related outcomes in clinical reviews, the PRISMA harms checklist was developed to account for the characteristics of reviews reporting adverse events, either as a primary or secondary outcome (Zorzela et al., 2016). Adapted from the original PRISMA statement which focuses primarily on the reporting of treatment efficacy, PRISMA harms consists of at least four extension items deemed compulsory in systematic reviews investigating harms, regardless of whether they are analysed in association with efficacy. These additional items address the explicitness of rationales and meanings in studying harms in the review (item 1), how zero events are handled (item 14), the assessment and definition of harms by reviewers (item 18), and the synthesis of results to ensure inferences of causality (item 21). Each systematic review studying adverse events should report the items in both the PRISMA statement and PRISMA harms.

1.4 Previous Umbrella Reviews on Antipsychotic Medications' Adverse Effects

An umbrella review was conducted on the risk of physical diseases associated with antipsychotics, antidepressants and mood stabilisers (Correll et al., 2015), which highlighted antipsychotics' higher propensity for a range of nutritional and metabolic, endocrine, cardiovascular, respiratory, musculoskeletal, hematologic and neurological diseases, such as weight gain, diabetes mellitus, hypertension, stroke, pneumonia, constipation, movement

disorders and seizure, relative to mood stabilisers and antidepressants. The review has presented a preliminary but valuable attempt in summarising the scattered evidence on the adverse outcomes associated with different psychiatric medications. Nevertheless, its methodology seemed to be less established, likely because the development of uniform approaches in conducting umbrella reviews did not flourish until 2015. This review has applied a relatively vague inclusion criteria, which identified articles reporting a wide range of effect metrics, meaning quantitative comparisons between the risk of different adverse events were not possible. Furthermore, the review has not evaluated the quality of the included evidence.

The only umbrella review identified to date centralising evidence on antipsychotics' adverse effects conducted meta-analyses on 6 systematic reviews of 58 observational studies and found that antipsychotic use increased the risk of 6 life-threatening health events: hip fracture, venous thromboembolism, stroke, myocardial infarction, pneumonia and sudden cardiac death (Papola et al., 2019). Weak-to-moderate effect sizes of $1.45<OR<2.24$ were found for these outcomes, with the largest being the risk of sudden cardiac death. Comprehensive quality assessments were carried out, with the strength of each association graded using the URC and GRADE framework. Nevertheless, the evidence generally had low quality, demonstrating the need for credibility assessments on the evidence for other adverse effects of antipsychotics as well. The fact that their scope was limited to six adverse medical events and only observational studies were included means findings from this review could be criticised for their low representativeness. More common and acute adverse events such as weight gain and sedation, which tend to be investigated using RCTs, remained uncaptured. A major factor in the action of antipsychotics are the classes or individual types of antipsychotics. Nevertheless, they conducted no subgroup analyses on the types of antipsychotics and the study designs involved in their meta-analysis to evaluate the adverse effects of distinct factors and how they might contribute to the high heterogeneity exhibited in their effect sizes.

1.5 Aims of the Current Umbrella Review

My objective was to expand the scope of my search to include all adverse outcomes reported in the current literature base—unrestricted to life-threatening medical events—and systematic reviews of both experimental and observational studies. This umbrella review aimed to: (i) systematically grade and give an overview of the current evidence for the adverse health outcomes associated with antipsychotic use in the community settings, (ii) compare the antipsychotic-associated risk of adverse health outcomes across adverse outcome categories, and finally, (iii) assess the factors which may affect the occurrence of adverse health outcomes.

Chapter 2: Methods

This umbrella review followed the PRISMA harms reporting guideline (see Appendix A). The protocol for this study was registered on PROSPERO (CRD42020223706) on 14th December 2020.

2.1 Data Sources

From inception to 19th December 2020, a systematic search was conducted on the following databases: PubMed, Embase, PsycINFO, Scopus, CINAHL and Cochrane Library of Systematic Reviews. Emails were sent to authors to retrieve papers from journals without open or institutional access.

Efficacy studies of antipsychotics often report adverse drug events–such information is, however, seldom present in their title or abstract. I therefore refrained from using generic and specific adverse effect terms, owing to the generally poor reporting of adverse drug outcomes in published reviews, and instead aimed to also identify efficacy studies by searching individual

medication names in titles or abstracts to increase search sensitivity. Furthermore, the authors of the "Adverse effects" chapter in the Cochrane Handbook evaluated different search approaches for systematic reviews in an overview and found that using subheadings is the most efficient search method to retrieve results of adverse effects, given the resultant optimal sensitivity and precision (Golder et al., 2019). Taking into account the above factors, my search strategy incorporated the names of individual antipsychotic medications (e.g., *fluphenazine; haloperidol; pimozide; thiothixene; aripiprazole; clozapine; olanzapine; quetiapine; risperidone; amisulpride*) and search strings comprising of database-specific indexing terms (e.g., MeSH terms) attached to adverse drug events-related subheadings (e.g., *antipsychotic agents/ae, to, po* for PubMed; and *neuroleptic agent/ae, to, tm, po* for Embase). Study design filters developed by the National Library of Medicine and BMJ were incorporated in the search strategy to identify systematic reviews and meta-analyses. See Appendix B for the full search strategy for PubMed.

2.2 Study Screening and Eligibility Criteria

Studies were examined for eligibility based on predetermined inclusion criteria. I included peer-reviewed and published systematic reviews or meta-analyses, which: (i) are of randomised controlled trials and observational studies with a cohort, case-control, nested case-control or cross-sectional design reporting the associations between adverse health outcomes and any of the 29 antipsychotic medications of interest: chlorpromazine, fluphenazine, haloperidol, pimozide, prochlorperazine, trifluoperazine, benperidol, flupentixol, levomepromazine, pericyazine, promazine, sulpiride, zuclopenthixol, acepromazine, droperidol, loxapine, trifluprомazine, perphenazine, thioridazine, molindone, thiothixene, aripiprazole, clozapine, olanzapine, quetiapine, risperidone, amisulpride, paliperidone and ziprasidone; (ii) examined antipsychotic prescription of either monopharmacy or

polypharmacy; (iii) examined adverse health outcomes associated with antipsychotic use in human populations of any age with any psychiatric illness or medical condition; (iv) reported adverse health outcomes with study-level data which allow the calculation of ORs and RRs (e.g., in terms of prevalence or effect estimates); and (v) measured adverse health outcomes as the primary or secondary outcome. To retrieve all published meta-analyses, no restriction was applied on the year range of publication. My antipsychotic medications of interest were determined based on the British National Formulary, United States Pharmacopeia-National Formulary, and 'Medicines used in psychotic disorders' from the WHO Model List of Essential Medicines (2019) to ensure the universality of my findings. Only English articles were included.

Excluded studies include: (i) articles with methodologies other than systematic review and meta-analysis, (ii) meta-analyses of quasi-randomised controlled trials; (iii) meta-analyses examining in-utero exposure to antipsychotics and the subsequent adverse events experienced by infants; (iv) meta-analyses which did not specify specific adverse outcomes (i.e., "any adverse effects"); (v) meta-analyses primarily investigating antipsychotic overdose. As I aimed to study the adverse health events in the community setting, meta-analyses which did not conduct a stratified analysis for settings and included at least 30% of primary studies with primarily institutionalised samples (i.e., inpatients, residents in care facilities) were excluded. Meta-analyses on surgical or palliative care patients were also excluded.

With reference to a screening template, titles and abstracts were first screened to eliminate clearly irrelevant articles (Appendix C). To proceed, title and abstracts must fulfil all the inclusion criteria, whereas indication of a single exclusion criterion would result in their exclusion. For those marked 'unclear' for any criterion, their full texts were retrieved for inspection. The full texts of potentially relevant studies identified during initial screening were then subsequently retrieved and screened. For each antipsychotic-adverse event association reported in reviews of RCTs on the same antipsychotic medication and population, the meta-

analysis with the highest quality was selected when multiple meta-analyses reporting effect sizes for the same outcome were identified to avoid data duplication. When there was a partial overlap or no overlap in the antipsychotic medication and population of these meta-analyses, the largest meta-analysis with the largest number of trials was selected. In circumstances where more than one meta-analysis included the same number of trials, the meta-analysis with more participants was included. Given the more inclusive inclusion criteria of meta-analyses of observational studies, the most recent meta-analysis was selected. When one or more meta-analyses were published in the same year, the meta-analysis with the largest number of studies was selected.

2.3 Data Extraction and AMSTAR Critical Appraisal

Data were extracted and transcribed onto a systematised data extraction form. For each included systematic review and meta-analysis, I extracted: the author(s); publication year; design of primary studies; number of databases searched; year range of review; number of primary studies; study location; settings; number of participants; participants' age; participants' diagnoses; type(s) of antipsychotics; adverse health events; effect sizes with 95% confidence intervals; and for reviews of observational studies, the number of cases exposed to the adverse outcome and the number of cases without exposure to the adverse outcome. All uncertainties and discrepancies were resolved through discussion or consultation with supervisors. A second extractor will be recruited for potential publication.

The methodological quality of included reviews was critically appraised using the 16-item AMSTAR 2 tool (Table S1, Appendix D). Each item was assessed by 3 options, "yes", "partial yes" or "no", in terms of 10 domains, within which 7 were regarded as the 'critical domains' impactful on the reviews' overall validity (items 2, 4, 7, 9, 11, 13 and 15) (Shea et al., 2017). These options were dichotomised into positive ("yes" and "partial yes") and negative

("no") results. Rather than the heavy emphasis on the score system in the previous version of the AMSTAR tool (Shea et al., 2007), AMSTAR 2 was used to determine the overall confidence in the meta-analyses' findings with ratings of "high", "moderate", "low" and "critically low" according to the number of critical and non-critical methodological flaws in each meta-analysis (Table S2, Appendix D). A negative result in a critical item indicates one critical weakness. Given the respective natures of RCTs and non-randomised studies of interventions, different criteria are applied on meta-analyses of RCTs and observational studies. For instance, item 9 highlights the need for appropriate techniques in assessing risk of bias, for which meta-analyses of RCTs must have assessed risk of bias in terms of random allocation, the blinding of patients and clinicians, allocation sequence, and the presence of selective reporting, while meta-analyses of observational studies must have assessed risk of bias in terms of confounding, selection bias, methods to ascertain treatment exposures and outcomes, and the presence of selective reporting (Shea et al., 2017).

In the AMSTAR 2 checklist, for a systematic review to score 'partially yes' or 'yes' in item 7, it must include a list of all potentially relevant full texts excluded from the analysis. Although useful, the inclusion of a full list of references may not be practical for all authors (Hailes et al., 2017). Therefore, this criterion was modified that a summary of excluded studies, in the form of PRISMA flowchart or description of reasons in the results section for instance, would suffice. Furthermore, item 15 specifies the requirement of investigation on publication bias via graphical or statistical tests such as funnel plots. However, the authors of several meta-analyses stated that while attempts were made to investigate small study effects graphically, the small number of included trials could diminish the meaningfulness of such investigation and hence no funnel plots were computed. In fact, the AMSTAR 2 guidance document suggests that the importance of publication bias investigation could depend on the context and setting (Shea et al., 2017). Therefore, this criterion was revised to require only active attempts to

identify publication bias and if graphical or statistical measures were not possible or appropriate, justification is required.

2.4 Data Analysis

Subsequent to the extraction of effect sizes, I conducted meta-analyses on the effect estimate and 95% confidence intervals for each antipsychotic-related adverse health event with the Dersimonian-Laird random-effects model. To aid summarisation, if a single meta-analysis reported multiple effect sizes respectively for different types of antipsychotics, levels of doses and/ or variables of patient characteristics, these effect sizes were pooled to devise an overall summary effect estimate to be presented in the main meta-analysis sections. The respective effect sizes and information on individual antipsychotics, dose-response relationships, duration and patient characteristics were discussed and investigated via subgroup analyses. See Appendix E for all formulae used in my meta-analysis.

The risks of adverse health events associated with antipsychotic use were expressed in terms of ORs and RRs. RRs are the ratio of the probability of an event happening in the exposed group versus that in the unexposed group. ORs compare the odds (i.e., probability) of an event happening in the exposed group against the odds of it happening in the unexposed group (Szumilas, 2010; Tenny & Hoffman, 2021). Appendix J presents the criteria established by researchers in assessing the direction and strength of RRs and ORs. Still, conclusions made on the strength of antipsychotic-adverse event associations may be misleading if only the magnitude of effect sizes was taken into account. Therefore, these criteria were only used as a guide for assessing the direction of effects.

ORs has been regarded as the most appropriate measure in estimating the risk of rare events (prevalence of <10%). Contrarily, owing to ORs' potential for overestimating the risk of common events (Balasubramanian et al., 2015), all effect sizes were converted into RRs

where possible for meta-analyses reporting outcomes from RCTs. When meta-analyses of RCTs reported the number of participants with and without adverse events from the antipsychotic-exposed and antipsychotic-unexposed groups and effect sizes other than RRs, I re-estimated their effect sizes in terms of RRs (see formula 1).

Direct conversion from ORs into RRs (or RRs into ORs) requires baseline prevalence and could be carried out using this formula:

$$RR = \frac{OR}{1 - p + p(OR)}$$

where p = baseline risk of adverse event/ risk of adverse event in the control group

However, since baseline risk data is often unavailable, I have refrained from directly converting all effect sizes into RRs or ORs.

Under the circumstance where the adverse event is rare, ORs and RRs are approximately equivalent and could be used interchangeably. Therefore, all the ORs and RRs for adverse outcomes reported in meta-analyses of observational studies were combined in a single forest plot. Contrarily, since RCTs report more common events, adverse events reported in RRs and ORs by meta-analyses of RCTs were presented separately on different forest plots.

Some meta-analyses provided no corresponding p-values for their reported effect estimates and confidence intervals. To obtain information on the statistical significance of each adverse outcome in this umbrella review, the effect sizes and confidence intervals of the concerned outcomes (extrapyramidal disorder, recurrence of psychotic symptoms, sexual dysfunction, hypersomnia, falls, all-cause death, cardiac death and sudden death) were log-transformed and applied in formula 3 to estimate their p-values (Altman & Bland, 2011).

For each overlapping adverse outcome reported in meta-analyses of either RCTs or observational studies, I compared their effect sizes by computing a ROR (Golder et al., 2011). A ROR of 1.0 indicates complete consensus or no discrepancy between the two study designs. Effect estimates extracted from RCTs were converted to ORs, where applicable, using formula

4 to aid calculation. Furthermore, adopted from a previous umbrella review of both study designs (He et al., 2018), a X^2 Q test was conducted to investigate the difference beyond chance (p<0.10) between their respective effect estimates.

All meta-analyses were conducted in Stata-IC version 16.0 using the metan commands and Microsoft Excel 2020.

2.5 Assessment of Evidence Quality and Credibility

Complementary to the quality appraisal using AMSTAR 2, statistical tests for heterogeneity, excess significance bias, prediction intervals and small-study effects were conducted to further assess the quality of the included reviews.

Heterogeneity of the effect estimates across included reviews was assessed with the I^2 statistic, which computes the proportion of observed dispersion due to heterogeneity rather than chance, as a percentage out of 100%. The cut-offs for low, moderate and high heterogeneity were 25%, 50% and 75% respectively (Huedo-Medina et al., 2006; Ioannidis et al., 2007). Among the meta-analyses of RCTs, 28 out of 35 adverse events had their heterogeneity reported. All adverse events extracted from the included meta-analyses of observational studies had their heterogeneity reported. For meta-analyses which reported the heterogeneity of their primary studies with the Cochran's Q statistic, Q was converted into I^2 with formula 5. The heterogeneity statistics of two adverse outcomes (extrapyramidal disorder and recurrence of psychotic symptoms) were converted from Q to I^2.

The 95% prediction interval for each summary effect estimate was estimated to account for between-study heterogeneity, as an expression for the uncertainty in the expected effect in future studies investigating the adverse health events associated with antipsychotic use (IntHout et al., 2016). It represents the region in which there is a 95% chance that the potential adverse effect in a prospective study on AP medication would lie within. A prediction interval

with a null value of no difference (i.e., 1 for RR and OR) included signifies weak evidence (Spineli & Pandis, 2020). Heterogeneity statistics (i.e., τ^2) were extracted for the calculation of 95% prediction interval using formula 6 in Appendix G (Higgins et al., 2009; Inthout et al., 2016).

I considered using the Egger's regression test to evaluate the presence of small-study effects among the included meta-analyses, given their universality for all types of outcomes. Nevertheless, it assumes no association between effect sizes and precision under the null hypothesis of publication bias (Egger et al., 1997), which although applicable for continuous data under normality (e.g., sample mean), is less plausible for discrete data constituting binary outcomes such as OR and RR (Peters et al., 2006, 2007; Rücker et al., 2008). The variance estimates of log(OR) and log(RR) are statistically dependent on the estimated log(OR) and log(RR), thus inducing asymmetry in the data points of the funnel plot even when small-study effects are absent, if studies are small and heterogenous (Lin & Chu et al., 2018). This means the Egger's test is highly prone to nominal type I error when substantial heterogeneity exists between studies (Debray et al., 2017). Unlike the rank-correlation or regression tests developed earlier which rely heavily on variance (Begg & Mazumdar, 1994; Egger et al., 1997; Peter et al., 2006; Harbord et al., 2006), the arcsine test applies arcsine and variance-stabilising transformations on the estimates and variance estimators of log(OR) and log(RR) to obtain treatment effects nearly independent of their variance (Rücker et al., 2008). Given the high heterogeneity among my reported adverse outcomes, the arcsine-Thompson test was used in the current umbrella review, owing to its superior power and independence from the variance estimators compared to other tests of publication bias and better performance under high heterogeneity relative to other arcsine test counterparts, arcsine-Rank and arcsine-Egger (Rücker et al., 2008; Thompson & Sharp, 1999). The arsine-Thompson test is also recommended for application in the investigation of rare and common events (Jin et al., 2014).

The association between the arcsine difference (i.e., difference of arcsine-transformed observed risks) and their estimated variance is then tested, in which a summary effect less conservative than p <0.10 indicates publication bias (Rücker et al., 2008). I also computed the conventionally used contour-enhanced funnel plot, where the log(RR) or log(OR) of each adverse event and the standard error of log(RR) or log(OR) lie on the x-axis and y-axis respectively (Sterne & Egger, 2001), to investigate the presence of discrepancies between the symmetry patterns reported in the arcsine-Thompson test and the funnel plot. Ideally, data points in the funnel plot should constitute a roughly symmetrical, inverted funnel shape.

To detect excess significance bias, the Ioannidis' test was used to examine whether the observed number of original studies with nominally significant (p <0.05) results (O) was larger than the expected number of original studies with nominally significant results (E) at $\alpha = 0.05$ (Ioannidis & Trikalinos, 2007). For each adverse outcome in this umbrella review, I computed the ratio of the summary RR or OR of a meta-analysis to the RR or OR of the largest component study in the same meta-analysis, with the former and latter signifying O and E respectively (Kavvoura et al., 2008). This was computed under the assumption that the effect size provided by the largest component study in each meta-analysis was the closest to the true effect and the most precise. For a high-quality review or meta-analysis, a ratio close to 1 is anticipated, as the observed and expected numbers of ORs are similar. In contrast, the occurrence of O>E is likely a result of selective reporting and biased analysis – in this case, the O:E ratio would be larger than 1, serving as an indicator of excess significance bias.

The URC have been utilised in previous umbrella reviews to grade each reported association connoted by a statistically significant (i.e., $p \leq 0.05$) summary effect size as 'convincing' (Class I), 'highly suggestive' (Class II), 'suggestive' (Class III) or 'weak' (Class IV) in terms of strength, sample size and the presence of biases (Aromataris et al., 2015; Papola et al., 2019; Solmi et al., 2020). For instance, convincing evidence should have: (i) a sample of

>1000 cases; (ii) highly significant effect summary estimates ($p < 10^{-6}$) in the random-effects calculations; (iii) 95% prediction intervals excluding the null value; (iv) low ($I^2 < 25\%$) or moderate ($25\% \leq I^2 < 50\%$) heterogeneity; and (v) no evidence for small-study effects (*p*-value from egger's test) and excessive significance bias (ESS ratio). Evidence was classified as non-significant if association had a *p*-value larger than 0.05. The fact that the cut-offs are continuous could hinder interpretation and hence challenge the determination of credibility. For example, whether the inclusion of 1000 cases suggest definite credibility is controversial–999 cases are not substantially different from 1000 cases (Papatheodorou, 2019). The inclusion of <1000 cases prevents an evidence from being classified as any grade above class IV (weak), even when other criteria are met. As a quality assessment tool, the URC has not been validated and no study, to my knowledge, was conducted to evaluate its credibility. That said, the URC presents a standardised framework and stringent criteria for the important qualities indicative of high quality evidence (Fusar-Poli & Radua, 2018). Therefore, while the associations reported in the current umbrella review were not graded with credibility ratings from the URC, results from the quality assessments will be presented in a summarisation table established based on the URC to aid informative judgments made by readers.

Since the URC highlights the use of random-effects model in meta-analysis to account for potential heterogeneity, I made sure the effect sizes for adverse outcomes with statistical significance were estimated using the random-effects model.

The statistical tests for quality assessments were conducted using the metabias and confunnel commands in Stata-IC version 16.0, and the metagen and metabias commands in RStudio version 1.4.1106 (arcsine-Thompson test).

2.6 Subgroup Analyses

Subgroup analyses allow investigations on how a pooled treatment effect estimate varies across different subsets of samples (Richardson et al., 2019). By dividing all data included in a meta-analysis into subgroups according to trial or patient characteristics, with a meta-analysis subsequently performed on each subset, sources of heterogeneity in the adverse events among patients with different clinically relevant characteristics could be identified. In the current umbrella review, included meta-analyses which reported effect sizes from observational studies were assessed for potential subgroup analyses. Sufficient effect size data were available for the following subgroups: class of APs, individual APs, patients' age groups, diagnosis and study designs.

2.7 Classification and Criteria for Adverse Outcomes

The CTCAE v.5.0 was adapted to categorise the meta-analysed adverse health outcomes (US Department of Health and Human Services, 2020). Eligible meta-analyses provided effect size data for over 150 adverse health outcomes (see Appendix F for all excluded adverse events with reasons). Given the laborious analyses required for such large volume of adverse events, several criteria were used to constrict the range of reported adverse outcomes. First, outcomes with data obtained from less than 3 primary studies were excluded (e.g., clinically significant weight loss, toxicity) as the estimation of heterogeneity in meta-analyses is often difficult with a small number of studies and inaccuracy in such could bias effect sizes and overly narrow confidence intervals (von Hippel, 2015). To further narrow down the reported outcomes, specific medical conditions or diagnoses (e.g., recurrence of psychotic symptoms) and their synonymous or associated symptoms (e.g., psychosis, delusions, hallucinations and schizophrenia) were collectively considered as the same outcome. Furthermore, to maximise the clinical relevance of this umbrella review, I excluded outcomes

characterised by clinical observations, which do not require medical attention (e.g., increased or decreased appetite, increased or decreased duration of sleep). Outcomes which are not standalone diagnoses but overlapping symptoms of multiple medical conditions (e.g., dizziness, headache, upper respiratory infection, pain and gastrointestinal symptoms such as vomiting, dyspepsia, toothache, constipation and xerostomia) were excluded. Dr. Stella Botchway, Researcher from the Forensic Psychiatry research group at the University of Oxford, was consulted to resolve uncertainties in the clinical relevance and inclusion of extracted adverse outcomes.

For metabolic outcomes, only clinically significant binary outcomes (e.g., with or without weight gain) were included, while laboratory investigations and continuous outcomes such as mean change in weight or cholesterol level were excluded as such findings could be deemed less clinically meaningful. Eligible meta-analyses reported clinically significant weight gain with varying cut-offs of 5%, 7% and 10%. The conventional definition of clinically significant weight change is still in question and varies between 3%, 5% and 7% from one medical literature to another (Keeny et al., 2013). Stevens et al. (2006) highlighted the distinction between weight maintenance and clinically significant weight change by proposing their respective cut-off values to be 3% and 5%. Therefore, in this umbrella review, adverse outcomes of weight gain of at least 5% from the baseline weight were classified collectively as clinically significant weight gain (Kolotkin et al., 2009; Wadden et al., 2011).

Ultimately, 42 adverse events supported by sufficient number of studies were deemed clinically relevant and included in this umbrella review (Table 1).

Table 1. Adverse events included in this umbrella review, as classified with categories adapted from CTCAE v.5.0

Classifications	**Adverse events**
Circulatory system disorder	Hypotension, hypertension, myocardial infarction, pulmonary embolism, stroke, venous thromboembolism

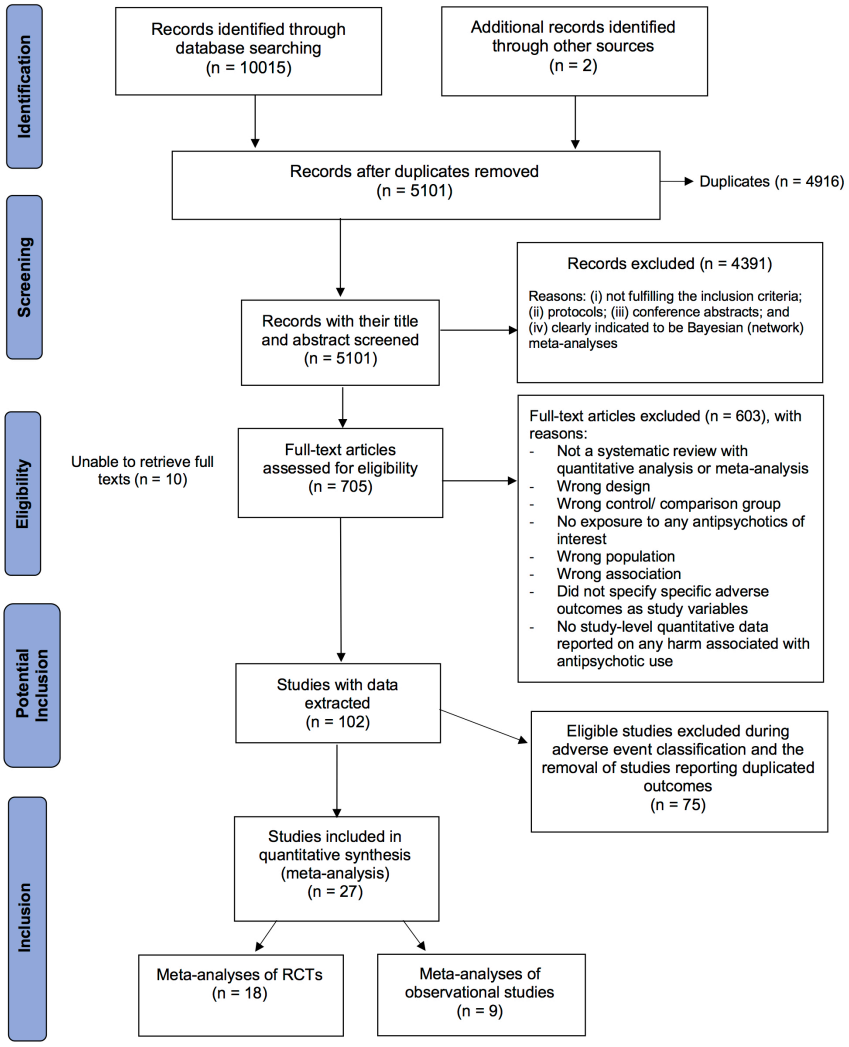
Endocrinological and metabolic outcomes	Clinically significant HbA1C levels, clinically significant weight gain, diabetes mellitus, electrocardiogram abnormalities, endocrine abnormalities (except diabetes), gestational diabetes mellitus, metabolic syndrome
Injury	Falls, fracture, hip fracture, injury
Mortality	All-cause death, cardiac death, sudden death
Musculoskeletal and movement-related outcomes	Extrapyramidal disorder, gait abnormality, musculoskeletal pain
Neuropsychiatric outcomes	Agitation, anxiety, completed suicide, depression, hypersomnia, hypertonia, hyperkinesia, insomnia, recurrence of psychotic symptoms, suicidal ideation, suicide attempt, sedation, somnolence
Other outcomes	Blurred vision, menstrual disorder, peripheral oedema, pneumonia, sexual dysfunction, rash

Chapter 3: Results

3.1 Study Characteristics

In total, the literature search yielded 10,015 references. Following de-duplication, 705 full texts were scrutinised out of the 5,101 papers from initial screening (Figure 1). Full-text screening identified 102 eligible meta-analyses. After narrowing down the scope of reported adverse outcomes and the removal of studies reporting overlapping outcomes, 75 studies were excluded (Appendix G). Ultimately, this umbrella review included 27 meta-analyses with 40,589,364 participants and over 820 associations from primary studies. Among the included studies, 18 and 9 meta-analyses provided data for 35 and 13 adverse health events from RCTs and observational studies respectively.

Figure 1. PRISMA flowchart of the selection process for included systematic reviews and meta-analysis



The included meta-analyses of RCTs reported data from over 567 associations with sample sizes ranging from 517 to 29,049, and 168,482 participants in total (Table 2). All but two included meta-analyses of RCTs compared adverse health events between antipsychotic monotherapy and placebo. The remaining two meta-analyses reported the adverse effects of antipsychotics as an augmentation to antidepressant pharmacotherapy in treatment-resistant depression, with one investigating the adverse events associated with olanzapine by comparing OFC with fluoxetine adjunct to placebo, and the other with aripiprazole by comparing

aripiprazole-antidepressant augmentation therapy and antidepressant plus placebo (Luan et al., 2017; Luan et al., 2018). Study participants were diagnosed with a wide range of psychiatric disorders, encompassing schizophrenia, schizoaffective disorder, bipolar disorder, major depressive disorder, treatment-resistant depression, dementia, autism, Tourette syndrome and miscellaneous conditions involved in off-label antipsychotic use.

The included meta-analyses of observational studies yielded over 40,420,882 participants, with sample sizes ranging from 11,789 to 28,729,856 (1 review unreported) (Table 3). All meta-analyses compared the adverse health events between antipsychotic-exposed individuals and antipsychotic-unexposed controls in 253 associations. Primary studies mostly had case-control and cohort designs, with some being self-controlled case series, case-crossover or cross-sectional studies.

Table 2. Characteristics of all included meta-analyses of RCTs

Author	**Year**	**Age**	**Diagnoses or population characteristics**	**Settings**	**Antipsychotic(s) / Comparator** **Dose; Duration**	**Adverse event**	***k***	***N***
Bai et al.	2020	≥18 years	DSM-III or ICD-9 or above Bipolar disorder	Unclear	Aripiprazole Asenapine Cariprazine Haloperidol Lurasidone Olanzapine Paliperidone Quetiapine Risperidone Ziprasidone/ Placebo 0.25-800 mg/d 12.5-50 mg/2wks 3-104 weeks	Clinically significant weight gain (≥7%)	54	17,478
						Somnolence	58	17,284
Defruyt et al.	2011	≥18 years	Bipolar disorder with a depressive episode	Mixed (most outpatient)	SGAs (Aripiprazole, Quetiapine, Olanzapine)/ Placebo	Recurrence of psychotic symptoms	12	4,619
Demyttenaere et al.	2019	≥18 years	Schizophrenia, Schizoaffective disorder, Bipolar disorder or MDD	Unclear	SGAs/ Placebo	Extrapyramidal disorder	48	19,854
Derry & Moore	2007	Depressive episodes: *M* = 36-42 years; Mania or mixed	Bipolar I or II disorder with depressive, mania or mixed episodes	Unclear	Haloperidol Olanzapine Quetiapine Risperidone	Depression	5	1,040

		episodes: *M* = 35-43 years			Ziprasidone/ Placebo			
Harrington et al.	2010	Adults and older adults	Acute Schizophrenia, Schizoaffective disorder or Bipolar I disorder	Unclear	Paliperidone/ Placebo 1.5-15 mg/d ≤6 weeks	Sexual dysfunction	9	2,688
Hay et al.	2015	≥18 years	DSM-IV or DSM-IV-TR Schizophrenia or Schizoaffective disorder	Mixed	Asenapine/ Placebo ≤12 weeks	Clinically significant HbA1C level	3	641
Lao et al.	2016	≥18 years (Mean range: 35.5-46.4)	Schizophrenia, Bipolar I mania and depression or MDD	Unclear	Cariprazine/ Placebo 0.75-12 mg/d 3-8 weeks	Hypotension	7	3,207
						Suicidal ideation	6	2,900
Luan et al.	2017	Adults	Treatment-resistant depression	Unclear	Olanzapine @OFC/ Fluoxetine+ Placebo 6-18 mg/d	Hypersomnia	*	*
						Peripheral oedema	*	*
Luan et al.	2018	Adults	Treatment-resistant depression	Unclear	Aripiprazole (plus AD)/ AD+Placebo 2-20 mg/d 4-12 weeks	Blurred vision	8	*
Ma et al.	2014	40-97 years	Dementia	Unclear	SGAs (Aripiprazole, Olanzapine, Quetiapine and Risperidone)/ Placebo	Gait abnormality	16	2,845
						Injury	10	3,370

					6-26 weeks			
Maglione et al.	2011	18-64 years; ≥65 years	Off-label antipsychotics users with all diagnoses (Adults: Anxiety, Eating disorders, Depression, OCD, PTSD, Personality disorders, Substance abuse…) (Elderlies: Dementia)	Mixed	Olanzapine Quetiapine Risperidone/ Placebo	Diabetes mellitus	8	4,046
		18-64 years			Olanzapine Quetiapine Risperidone Ziprasidone/ Placebo	Endocrine abnormalities except diabetes	5	911
Maher et al.	2011	18-64 years	Off-label antipsychotic users with all diagnoses (Anxiety, Eating disorders, Depression, OCD, PTSD, Personality disorders, Substance abuse…)	Unclear	Aripiprazole Olanzapine Quetiapine Risperidone Ziprasidone/ Placebo	Fatigue	30	8,613
		18-64 years; ≥65 years	Off-label antipsychotic users with all diagnoses (Adults: Anxiety, Eating disorders, Depression, OCD, PTSD, Personality disorders, Substance abuse…) (Elderlies: Dementia)		Aripiprazole Olanzapine Quetiapine Risperidone/ Placebo	Sedation	71	15,760
		≥65 years	Off-label antipsychotic users with Dementia			Stroke	11	3381
Nussbaum & Stroup	2008	≥18 years	DSM-IV Schizophrenia	Outpatient or initially started in hospital and then outpatient	Paliperidone/ Placebo ≤12 weeks	Hyperkinesia	4	1,360
						Hypertonia	5	1,225

Nussbaum & Stroup	2012	≥18 years	DSM-IV Schizophrenia	Outpatient or initially started in hospital and then outpatient	Paliperidone/ Placebo *Med* = 13 weeks	Anxiety	5	2,180
						Hypertension	5	2,180
						Menstrual disorder	3	762
						Musculoskeletal pain	5	2,180
						Pneumonia	1	517
						Rash	5	2,180
						Suicide attempt	3	2,178
Polcwiartek et al.	2015	6-95 years	Schizophrenia, Schizoaffective disorder, Bipolar disorder, Tourette syndrome, Autism, Alzheimer's disease	Unclear	Aripiprazole/ Placebo	Electrocardiogram abnormalities	18	6,106
Romeo et al.	2018	All age groups	DSM-III, DSM-IV, DSM-V MDD or Bipolar depression with depressive or mixed mania episode	Unclear	Aripiprazole Cariprazine/ Placebo >5 mg/d (ARI) >2.5 mg/d (CAR)	Agitation	14	*
						Insomnia	24	*
Rotella et al.	2020	All age groups	Schizophrenia or Bipolar disorder	Unclear	SGAs/ Placebo ≥52 weeks	Myocardial infarction	2	600
					FGA (Fluphenazine) SGAs (Aripiprazole, Paliperidone, Quetiapine, Risperidone) / Placebo ≥52 weeks	Cardiac death	5	1,051

Schneider-Thoma et al.	2018	All age groups (most 18-65, followed by children and adolescents, and older adults)	Any diagnosis (e.g., Schizophrenia, Bipolar disorder, MDD)	Unclear	AP/ Placebo 1.5 hours-104 weeks	All-cause death	91	29,049
						Completed suicide	18	6,749

* Information unavailable; *k* = the number of primary studies

Table 3. Characteristics of all included meta-analyses of observational studies

Author	Year	Definition and determination of adverse event	Population characteristics	Study designs	Total *N*	Adverse event	*k*
Kucukgoncu et al.	2020	ICD, chart reviews, livery or health records, patient reports from the GP or study records	Diagnosed with Schizophrenia, Schizoaffective disorder, Bipolar disorder, Depression, OCD, PTSD, Anxiety, Alcohol/ substance use disorder, Personality disorder, psychotic episodes or other non-affective psychosis	Case-control and cohort studies	1,683,312	Gestational diabetes mellitus	4
Liu et al.	2021	ICD, autopsy records, hospital medical records, objective tests	Any diagnosis	Case-control and cohort studies	28,729,856	Pulmonary embolism	4
		ICD, post-mortem records, hospital medical records, objective tests			2,767,005	Venous thromboembolism	21
Mortensen et al.	2020	*	M_{age} = 71 years	Case-control and cohort studies	306,713	Hip fracture	8
Nosé et al.	2015	ICD, medical records, radiographical findings for aspiration pneumonia, community-acquired pneumonia and pneumonia	Age: 15-65 years and ≥65 years	Case-control and nested case-control studies and 1 self-controlled case series	28,726	Pneumonia	6

Papola et al.	2018	Self-report during interview, ICD and record from healthcare insurer database,	M_{age} = 55.4 years Diagnosed with Dementia, Parkinson's Disease, Schizophrenia and miscellaneous conditions	Case-control, prospective and retrospective cohort, case-crossover studies	1,753,724	Fracture	7
Seppala et al.	2018	Medical records, recurrent falls in the past 3, 11 or 12 months, recall	≥60 years old Community subgroup	Case-control, cohort and cross-sectional studies	*	Falls	4
Vancampfort et al.	2015	ATP-III, ATP-III-A, IDF or WHO Metabolic syndrome diagnosis	M_{age} = 41.3 years Diagnosed with DSM-IV or ICD-10 Schizophrenia or a related psychotic disorder, Bipolar disorder or MDD	Prospective and retrospective cohort and cross-sectional studies	13,165	Metabolic syndrome	143
Yang et al.	2018	*	Age = 10-100 years Any diagnosis (Dementia, Alzheimer's disease, Schizophrenia, Parkinson's disease…) Follow-up: 30 days - 16 years	Case-control and cohort studies	2,500,546	All-cause death	34
					409,294	Cardiac death	7
					11,789	Sudden death	3
Zivkovic et al.	2019	Formal diagnosis, hospitalisation, self-report	M_{age} (Dementia) = 78-81 M_{age} (Psychiatric) = 43-57 M_{age} (General population) = 45-85 Diagnosed with Dementia, Schizophrenia, Depression, Mood disorder or not specified	Case-control and cohort studies	130,145	Myocardial infarction	7
		Hospitalisation for stroke or transient ischaemic attack, death from cerebrovascular event, first diagnosis			2,086,619	Stroke	15

			Follow-up: 1 week - 12 years				

* Information unavailable; *k* = the number of primary studies or associations

3.2 Overview of Effect Size Data

Among meta-analyses of RCTs, 14 out of 37 adverse outcomes had statistically significant RRs or ORs (p<0.05). The largest and smallest effect sizes reported in RRs were peripheral oedema (RR 10.16, 95% CI 4.61 to 22.38) and sexual dysfunction (RR 1.01, 95% CI 1.00, 1.02) respectively. The largest OR was for extrapyramidal disorder (OR 2.43, 95% CI 1.91 to 3.10) and the smallest for insomnia (OR 1.66, 95% CI 1.35 to 2.04). Non-significant associations were found for the following 22 outcomes: myocardial infarction, diabetes mellitus, injury, clinically significant HbA1c level, electrocardiogram abnormalities, musculoskeletal pain, hyperkinesia, stroke, anxiety, depression, recurrence of psychotic symptoms, suicide attempt, suicidal ideation, completed suicide, menstrual disorder, sexual dysfunction, pneumonia, rash, hypertension, hypotension, and all-cause death.

Effect estimates for 13 adverse outcomes were extracted from meta-analyses of observational studies. All associations were statistically significant (p<0.05). The largest effect size was for metabolic syndrome (OR 4.97, 95% CI 4.03 to 6.13) and the smallest for fracture (OR 1.17, 95% CI 1.04 to 1.31). The distribution of ROR for overlapping outcomes reported in both RCTs and observational studies can be found in Appendix I.

Several meta-analyses reported multiple effect sizes, respective to the type of individual antipsychotic medications, levels of doses and/ or patient variables, for a single outcome, but did not provide a summary effect size. Nine adverse outcomes were extracted from these meta-analyses: diabetes mellitus, endocrine abnormalities, fatigue, clinically significant weight gain, insomnia, sedation, somnolence, stroke, agitation, depression, and metabolic syndrome. To aid summarisation, I estimated an overall effect size by pooling these individual effect sizes together for each outcome. Ideally, when no summary estimate is provided, effect size data from component primary studies should be extracted and re-meta-analysed to obtain an overall estimate and heterogeneity statistics. Nevertheless, no effect size data from primary studies

was available in these reviews so the computation of summary effect estimates could only be carried out at review-level for these outcomes. Some meta-analyses of RCTs presented separate effect sizes respective to different antipsychotics but also reported summary effect sizes (myocardial infarction, injury, extrapyramidal disorder, gait abnormality, recurrence of psychotic symptoms and all-cause death). These component associations demonstrated potential variations in the risk of adverse outcomes among different types of antipsychotics, dose level, duration and age groups. See Appendix H for the individual associations presented in forest plots.

The effect size and confidence intervals for sexual dysfunction were converted from attributable risk to RR (see formula 2 in Appendix E). Among the included meta-analyses of observational studies, four outcomes (gestational diabetes mellitus, all-cause death, cardiac death and sudden death) were reported in RRs. Separate effect sizes for stroke and myocardial infarction were reported respective to the case-control and cohort study designs in HRs and ORs (Zivkovic et al., 2019). For pneumonia, separate ORs were also presented respective to case-control and cohort study designs (Nosé et al., 2015). However, as the effect size data from each primary study was reported, these were pooled together using the random-effects model to estimate a summary OR for each outcome (Table 4).

3.2.1 Circulatory System Outcomes

Myocardial infarction was found to be significantly associated with antipsychotic use in the meta-analysis of observational studies, but not in RCTs (Figure S1) (Table 4). All observational primary studies reporting observational study data were adjusted for age and sex, and most adjusted also for cardiovascular comorbidities. The summary effect size extracted from the meta-analysis of RCTs was much larger than that from observational studies (*ROR* = 2.61). Nevertheless, the OR extracted from the meta-analysis of RCTs had wide confidence

intervals and hence did not reach statistical significance. Results from the X^2 Q test indicated that no significant difference beyond chance existed between the effect sizes from RCTs and observational studies ($p = 0.41$).

A single meta-analysis of observational studies provided effect sizes for pulmonary embolism and venous thromboembolism; both were statistically significant (Table 4). The extracted effect sizes for both outcomes were obtained by pooling both adjusted and unadjusted estimates from primary studies in the original meta-analysis. Adjusted confounders included age, sex, race, settings, weight, BMI, comorbidities, socioeconomic status, etc.

Meta-analysis of observational studies showed that antipsychotic exposure was significantly associated with stroke. All included studies were adjusted for patients' age and sex. A similar effect size for stroke was found in the meta-analysis of RCTs (*ROR* = 1.12, $p = 0.71$), although it was statistically non-significant given the null value in its confidence interval (Figure S13).

The RRs for hypertension and hypotension were not statistically significant (Table 4).

Table 4. Summary effect size for the association between antipsychotic use and circulatory system outcomes reported in meta-analyses of RCTs and observational studies

Adverse outcome	Effect metric	Effect size	95% CI	*p*	I^2 (%)	95% PI
RCTs						
Hypertension	RR	1.33	(0.52, 3.38)	0.55	0.0	(0.52, 3.39)
Hypotension	RR	0.93	(0.76, 1.13)	0.45	0.0	*
Myocardial infarction	OR	3.08	(0.32, 29.73)	0.336	*	*
Stroke[△]	RR	1.38	(0.78, 2.46)	0.269	0.0[†]	(0.78, 2.45)[†]
Observational studies						
Myocardial infarction	OR	1.18	(1.03, 1.36)	0.020	82.6	(0.79, 1.77)
Pulmonary embolism	OR	3.68	(1.23, 11.05)	0.020	90.0	(0.12, 109.87)
Stroke	OR	1.24	(1.06, 1.46)	0.0071	84.0	(0.64, 2.39)
Venous thromboembolism	OR	1.55	(1.36, 1.76)	<0.001	85.0	(0.92, 2.62)

* Unreported or could not be determined, [△]Effect size obtained by pooling multiple associations together on review level, [†]Reflects the heterogeneity and resultant prediction intervals of studies on systematic review level rather than primary study level

3.2.2 Endocrinological and Metabolic Outcomes

SGAs significantly increased the risk of endocrine abnormalities (except diabetes) in off-label antipsychotic users (Table 5). Individual RRs for olanzapine, quetiapine, risperidone and ziprasidone ranged from 1.95 to 5.00, compared to placebo, within which the largest estimate was for ziprasidone (RR = 5.00). Nevertheless, the extent of increase in the risk of endocrine abnormalities by ziprasidone was uncertain, given its wide confidence interval containing the null value (95% CI 0.25 to 99.95) (Figure S2). No significant association was found for diabetes mellitus (Figure S3).

Effect size data for clinically significant weight gain was available for a variety of antipsychotic medications, with different doses and durations. These individual effect sizes were pooled to calculate a summary RR, which showed that the use of antipsychotics significantly increased the risk of clinically significant weight gain in patients with bipolar disorder (Table 5). Among individual effect sizes, the largest RRs were found for olanzapine at 5-20 mg/d administered across 6 to 8 weeks (RR 36.56, 95% CI 11.72 to 114.04), asenapine at 10-20 mg/d administered across 3 weeks and quetiapine at 400-800 mg/d administered across 3 weeks. The smallest RR, which indicated a decreased risk of clinically significant weight gain, was found for aripiprazole at 15-30 mg/d administered in 3 weeks (RR 0.57, 95% CI 0.30, 1.08). The same medication at the same dose level administered for 26 weeks yielded a much larger RR of 16.05, its confidence interval, however, was wide and included the null value (95% CI 0.94 to 274.71). To summarise, the RRs for asenapine, cariprazine, lurasidone, paliperidone, olanzapine, quetiapine, risperidone and ziprasidone all indicated an increased risk of clinically significant weight gain. Aripiprazole increased the risk under longer duration but decreased the risk under short duration. Haloperidol decreased the risk of clinically significant

weight gain (RR 0.84, 95% CI 0.38 to 1.83). See Figure S6 for the detailed effect size data of all associations.

The RR for clinically significant HbA1C levels (RR 2.24 CI 1.06 to 4.75) was originally estimated using a fixed-effects model by Hay et al. (2015). To reaffirm its significance, the Dersimonian-Laird random-effects model was applied to re-estimate its summary effect estimates. Re-estimation showed that the use of asenapine use resulted in no significant difference in the risk of clinically significant HbA1C levels among patients with schizophrenia.

A significant OR for metabolic syndrome was estimated by pooling the separate ORs for aripiprazole, clozapine, olanzapine, quetiapine, risperidone and FGAs together (Table 5). Among these individual APs, clozapine had the largest effect size (OR 7.81, 95% CI 6.02 to 10.22), followed by olanzapine, quetiapine, FGAs, risperidone, amisulpride and aripiprazole (OR 3.25, 95% CI 2.36 to 4.49) (Figure S7).

Among pregnant women diagnosed with a wide range of psychiatric disorders, antipsychotics exposure was significantly associated with the risk of gestational diabetes mellitus. The effect estimate was adjusted for participants' BMI, smoking and additional smoking use (Table 5).

Table 5. Summary effect sizes of the association between antipsychotic use and endocrinological and metabolic outcomes reported in the meta-analysis of RCTs

Adverse outcome	Effect metric	Effect size	95% CI	*p*	I^2 (%)	95% PI
RCTs						
Clinically significant weight gain (≥7%)△	RR	2.82	(2.07, 3.84)	<0.001	73.1†	(0.69, 11.49) †
Clinically significant HbA1C level	RR	2.24	(0.62, 8.10)	0.220	44.3	(0.063, 79.36)
Diabetes mellitus△	RR	1.63	(0.84, 3.19)	0.152	11.4†	(0.49, 5.39) †
Endocrine abnormalities, except diabetes△	RR	2.21	(1.28, 3.81)	0.004	0.0†	(0.67, 7.29) †
Observational studies						
Gestational diabetes mellitus	RR	1.30	(1.02, 1.66)	0.032	14.4	(0.48, 3.56)

Metabolic syndrome[△]	OR	4.97	(4.03, 6.13)	<0.001	71.4[†]	(2.68, 9.22)[†]

[△] Effect size obtained by pooling multiple associations together on review level, [†]Reflects the heterogeneity and resultant prediction intervals of studies on systematic review level rather than primary study level

3.2.3 Injury

All effect sizes for injuries reported in meta-analyses of observational studies were statistically significant (Table 6). The OR for falls was adjusted for age, sex, sociodemographic and clinical factors in community-living antipsychotic users aged over 60. The meta-analysis reporting the effect estimate for fracture only included primary studies which adjusted for at least age, sex and clinical factors such as comorbidities, BMI, alcohol consumption and concomitant medications. No significant association was found for injury (Figure S5).

Table 6. Summary effect size for the association between antipsychotic use and injury reported in meta-analyses of RCTs and observational studies

Adverse outcome	Effect metric	Effect size	95% CI	*p*	I^2 (%)	95% PI
RCTs						
Injury	RR	0.98	(0.86, 1.12)	1.194	0.0	*
Observational studies						
Falls	OR	2.30	(1.24, 4.26)	0.008	0.0	*
Fracture	OR	1.17	(1.04, 1.31)	0.009	84.0	(0.89, 1.53)
Hip fracture	OR	1.85	(1.36, 2.51)	<0.001	94.0	*

* Unreported or could not be determined

3.2.4 Mortality

All outcomes extracted from meta-analyses of RCTs had confidence intervals including the null value (Table 7). See Figures S17 and S18 for the individual associations of all-cause death and cardiac death.

Larger and statistically significant effect sizes for death outcomes were found in the included meta-analysis of observational studies, which mostly included primary studies

adjusted for age, sex, race, comorbidities, and concomitant medication use. Results indicated that antipsychotic-exposure was associated with an increased risk of all-cause death, cardiac death and sudden death, compared to non-exposure (Table 7). The highest risk was associated with sudden death.

A direct comparison between the ORs and RRs for all-cause death and cardiac death extracted from the included meta-analyses of RCTs and observational studies was deemed appropriate. A moderate-size and close-to-one ROR was found for all-cause death (0.79, p = 0.23) and cardiac death (0.93, p = 0.93) respectively, with no significant differences beyond chance found between the two study designs in the X^2 Q test. Such patterns indicated little discrepancies to exist between the RCTs and observational study designs.

Table 7. Summary effect sizes of the association between antipsychotic use and mortality reported in meta-analyses of RCTs and observational studies

Adverse outcome	Effect metric	Effect size	95% CI	*p*	I^2 (%)	95% PI
RCTs						
All-cause death	OR	1.19	(0.93, 1.53)	0.172	0.0	(0.93, 1.53)
Cardiac death[△]	OR	1.96	(0.46, 8.28)	0.360	0.0[†]	(0.46, 8.28)[†]
Observational studies						
All-cause death	RR	1.50	(1.12, 1.99)	0.006	99.3	*
Cardiac death	RR	2.10	(1.29, 3.42)	0.003	76.7	*
Sudden death	RR	3.70	(2.68, 5.12)	<0.001	0	*

*Unreported or could not be determined, [△]Effect size obtained by pooling multiple associations together on review level, [†]Reflects the heterogeneity and resultant prediction intervals of studies on systematic review level rather than primary study level

3.2.5 Musculoskeletal and Movement-Related Outcomes

Meta-analyses of RCTs provided sufficient data for estimating the risk of multiple neurological adverse outcomes (Table 8). Among patients with schizophrenia, major depressive disorder, schizoaffective disorder and bipolar disorders, the use of antidepressants significantly increased the risk of extrapyramidal symptoms relative to placebo. Individually,

asenapine, lurasidone and cariprazine all increased the odds of extrapyramidal symptoms, with cariprazine having the largest effect size (OR 4.35, 95% CI 2.80, 6.75) (Figure S11).

The summary RRs for hypertonia (RR 3.24 CI 1.92 to 6.63) and gait abnormalities (RR 3.57 CI 2.28 to 5.60) were originally estimated with a fixed-effects model (Ma et al., 2014; Nussbaum & Stroup, 2008). A re-estimation of these effect sizes was carried out using the Dersimonian-Laird random-effects model. Re-estimation resulted in wider confidence intervals and reduced statistical significance for these outcomes, although they still remain statistically significant (Table 8). The meta-analysis showed that the risk of gait abnormalities was significantly increased with the use of antipsychotics in patients with dementia. Authors also reported the individual effect sizes for aripiprazole, olanzapine, quetiapine and risperidone, which all increased the risk of gait abnormalities (Figure S12). Aripiprazole had the largest RR of 5.29 (95% CI 0.71, 39.47), but uncertainties existed in such increase as its confidence interval contained the null value.

Among patients with schizophrenia, the use of paliperidone decreased the risk of musculoskeletal pain, although the difference was statistically non-significant (Table 8).

Table 8. Summary effect sizes of the association between antipsychotic use and musculoskeletal and movement-related outcomes reported in meta-analyses of RCTs

Adverse outcome	Effect metric	Effect size	95% CI	*p*	I^2 (%)	95% PI
Extrapyramidal disorder	OR	2.43	(1.91, 3.10)	<0.001	54.0	*
Gait abnormalities	RR	2.74	(1.66, 4.53)	<0.001	0.0	(1.52, 4.95)
Hyperkinesia	RR	1.67	(0.97, 2.89)	0.06	0.0	(0.97, 2.88)
Hypertonia	RR	2.91	(1.31, 6.47)	0.009	0.0	(0.80, 10.65)
Musculoskeletal pain	RR	0.80	(0.47, 1.37)	0.41	0.0	(0.47, 1.37)

* Unreported or could not be determined, △Effect size obtained by pooling multiple associations together on review level, †Reflects the heterogeneity and resultant prediction intervals of studies on systematic review level rather than primary study level

3.2.6 Neuropsychiatric Outcomes

Only meta-analyses of RCTs provided effect size data for psychiatric adverse outcomes. The use of antipsychotics significantly increased the risk of agitation, sedation, and sleep outcomes, including hypersomnia, insomnia, and somnolence (Table 9).

There were opposing dose-response relationships for the risk of agitation associated with aripiprazole and cariprazine (Figure S14). As the dose of aripiprazole increased from 5-15 mg/d to over 25 mg/d, there was a decrease in effect size from a OR of 3.63 (95% CI 1.36 to 9.69) to 1.04 (95% CI 0.72 to 1.50). In contrast, the OR for cariprazine increased from 1.80 (95% CI 0.89 to 3.62) to 3.41 (95% CI 0.75 to 15.42) as its dose level increased from 2.5-5 mg/d to >5 mg/d.

Associations with sedation were stratified based on individual antipsychotic medications (aripiprazole, olanzapine, quetiapine and risperidone) and patients' age and population (older adults with dementia and adults) (Figure S9). In general, larger RRs were found for older adults with dementia than adults without dementia. Across these populations, quetiapine yielded the largest effect sizes (RR 3.69, 95% CI 2.26 to 6.03 for older adults with dementia and RR 2.89, 95% CI 2.62 to 3.20 for adults), followed by olanzapine, aripiprazole and risperidone. Data for ziprasidone was available only for adults and an increased risk was found (RR 2.45, 95% CI 1.57 to 3.81). The pooled summary RR indicated a statistically significant increased risk of sedation by these antipsychotics.

Among the three sleep outcomes, the risk of hypersomnia was larger than that of somnolence. No comparison could be made between insomnia and other sleep outcomes as it was reported in a different effect metric (i.e., OR).

Aggregation of the individual ORs for aripiprazole and cariprazine under different dose levels found significantly increased odds of insomnia in patients with major depressive disorder

and bipolar disorder. In terms of component associations, aripiprazole had ORs ranging between 1.90 and 1.99, whereas the ORs for cariprazine ranged between 1.38 and 1.58 (Figure S8). Furthermore, olanzapine in olanzapine-fluoxetine combination treatment also significantly increased the risk of hypersomnia.

Separate RRs were pooled to estimate a summary effect size for somnolence respective to individual antipsychotics as treatment for bipolar disorder under different duration (Figure S10). Antipsychotic use significantly increased the risk of somnolence. Among patients who used antipsychotic in acute treatment for bipolar mania, all antipsychotic administrations increased the risk of somnolence, with the exception of risperidone long-acting injection. Quetiapine extended release administered in 3 weeks had the largest effect size (RR 4.79, 95% CI 1.75 to 13.10), followed by haloperidol (RR 4.48, 95% CI 1.39 to 14.33) and cariprazine (RR 4.39, 95% CI 0.96 to 19.97). Likewise, quetiapine extended-release provided the largest RR for somnolence as an acute treatment for bipolar depression (RR 6.80, 95% CI 4.47 to 10.33), followed by ziprasidone (RR 4.30, 95% CI 2.43 to 7.61) and quetiapine immediate release at 300 mg/d (RR 3.58, 95% CI 2.57 to 4.99). Aripiprazole, cariprazine and lurasidone had the smallest RRs ranging between 1.31 and 1.83. For the use of antipsychotics as bipolar maintenance treatment, paliperidone had the largest effect size (RR 10.85, 95% CI 0.81 to 194.53), while that of olanzapine at 5-20 mg/d (104 weeks) (RR 5.29, 95% CI 0.22 to 128.30) and at 10 mg/d (RR 4.12, 95% CI 1.42 to 12.01) were the second and third largest respectively. Quetiapine, olanzapine at 5-20 mg/d (48 weeks) and risperidone long-acting injectable at 25-60 mg/2 weeks increased the risk of somnolence, while aripiprazole and asenapine decreased such risk. Risperidone long-acting injectable at 12.5-60 mg/2 weeks had no association with the occurrence of somnolence (RR 0.97, 95% CI 0.06, 15.33).

No significant association was found for depression and the recurrence of psychotic symptoms (Figures S15 and S16).

Table 9. Summary effect size for the association between antipsychotic use and neuropsychiatric outcomes reported in meta-analyses of RCTs

Adverse outcome	Effect metric	Effect size	95% CI	*p*	I^2 (%)	95% PI
Agitation△	OR	2.01	(1.14, 3.56)	0.016	62.3†	(0.33, 12.37) †
Anxiety	RR	0.74	(0.52, 1.07)	0.11	0.0	(0.52, 1.06)
Completed suicide	OR	1.15	(0.47, 2.81)	0.772	0.0	(0.47, 2.81)
Depression△	RR	1.08	(0.73, 1.60)	0.707	0.0†	(0.73, 1.60) †
Hypersomnia	RR	3.57	(1.92, 6.63)	<0.001	*	*
Insomnia△	OR	1.66	(1.35, 2.04)	<0.001	0.0†	(1.35, 2.04) †
Recurrence of psychotic symptoms	RR	0.73	(0.48, 1.11)	0.141	40.4	*
Sedation△	RR	2.48	(2.12, 2.90)	<0.001	67.1†	(1.55, 3.95) †
Somnolence△	RR	2.81	(2.34, 3.37)	<0.001	49.3†	(1.44, 5.47) †
Suicide attempt	RR	1.27	(0.26, 6.35)	0.77	0.0	(0.26, 6.28)
Suicidal ideation	RR	0.91	(0.65, 1.27)	0.57	0.0	(0.65, 1.27)

* Unreported or could not be determined, △Effect size obtained by pooling multiple associations together on review level, †Reflects the heterogeneity and resultant prediction intervals of studies on systematic review level rather than primary study level

3.2.7 Other Outcomes

A highly significant summary effect size for fatigue was estimated from the individual RRs for five individual SGAs, ranging from 0.85 for risperidone (95% CI 0.34 to 2.11) to 10.62 for ziprasidone (95% CI 0.62 to 181.22) in off-label antipsychotic users. However, the ziprasidone-associated risk of fatigue was statistically non-significant, given its wide confidence interval with null value. Aripiprazole, olanzapine and quetiapine had summary RRs of 2.59, 1.90 and 2.56 respectively. See Figure S4 for the detailed data of all individual associations on fatigue.

Compared to fluoxetine plus placebo, olanzapine in olanzapine-fluoxetine combination significantly increased the risk of peripheral oedema in patients with treatment-resistant depression (Table 10). Nevertheless, as the number of trials and participants in this association was unreported in the meta-analysis, its effect size and *p* value should be interpreted with caution.

Aripiprazole, as an augmentation of antidepressant treatment, significantly increased the risk of blurred vision in patients with treatment-resistant depression, compared to those treated with antidepressant plus placebo (Table 10).

No association was found between antipsychotic use and sexual dysfunction and menstrual disorder (Table 10). The confidence interval of sexual dysfunction included the null value, meaning its statistical significance should be interpreted with caution.

Meta-analysis of observational studies found a highly significant OR for pneumonia associated with antipsychotic use in patients with all age groups (Table 10). Most primary observational studies were adjusted for potential confounders, such as age, sex, physical health factors (e.g., smoking status, history of cardiovascular disease), comorbidity, hospital admissions, drug use, etc. The meta-analysis of observational studies yielded a much larger effect size for pneumonia than the re-estimated effect size of RCTs (OR 0.32, 95% CI 0.02 to 5.22) (*ROR* = 0.17), although no significant difference beyond chance was found between the two study designs in the X^2 Q test (p = 0.22). Nevertheless, the RCT data for pneumonia was extracted from a single trial with a sample of only 517 participants. Therefore, this effect size was excluded from subsequent quality assessments as examining its heterogeneity, publication bias and excess significance ratio could be meaningless.

Schizophrenia patients who received paliperidone showed an increased risk of developing rash compared to placebo, although the difference was not clinically significant (Table 10).

Table 10. Summary effect size for the association between antipsychotic use and other outcomes reported in meta-analyses of RCTs

Adverse outcome	Effect metric	Effect size	95% CI	*p*	I^2 (%)	95% PI
RCTs						
Blurred vision	RR	4.05	(1.68, 9.75)	0.002	*	*
Electrocardiogram abnormalities	RR	0.85	(0.51, 1.43)	0.54	*	*

Fatigue[△]	RR	2.17	(1.61, 2.93)	<0.001	49.0[†]	(0.92, 5.14)[†]
Menstrual disorder	RR	1.59	(0.32, 8.04)	0.57	0.0	(0.32, 7.98)
Pneumonia	RR	0.33	(0.02, 5.17)	0.43	n/a	*
Peripheral oedema	RR	10.16	(4.61, 22.38)	<0.001	*	*
Rash	RR	1.38	(0.53, 3.57)	0.52	0.0	(0.53, 3.58)
Sexual dysfunction	RR	1.01	(1.00, 1.02)	0.049	0.0	*
Observational studies						
Pneumonia	OR	1.84	(1.62, 2.09)	<0.001	27.0	(1.38, 2.45)

* Unreported or could not be determined, [△]Effect size obtained by pooling multiple associations together on review level, [†]Reflects the heterogeneity and resultant prediction intervals of studies on systematic review level rather than primary study level, n/a: not applicable since data was extracted from only one RCT

3.2.8 Adverse Outcome Categories

Figures 2 and 3 show the RRs and ORs for the selected outcomes in terms of the adverse outcome categories I adapted from CTCAE v.5. Only categories comprising of at least two adverse outcomes are discussed here.

The sample sizes of the included meta-analyses of RCTs were generally small. For instance, the effect sizes of pneumonia, menstrual disorder and myocardial infarction were extracted from a maximum of 3 studies with less than 1000 participants. These contribute to imprecision and hence their wide confidence intervals. To maximise the meaningfulness of my comparison between adverse outcome categories, only meta-analyses of RCTs with more than 10 primary studies were pooled and displayed on the forest plots.

RRs extracted from meta-analyses of RCTs were not presented in a forest plot, as 4 out of 6 categories had only one outcome and there was substantial heterogeneity within the remaining 2 categories. The pooled RR across adverse outcome categories indicated a significant increase in the risk of adverse outcomes associated with antipsychotic use (RR 1.69, 95% CI 1.15 to 2.48, $p = 0.008$). However, high heterogeneity was found in the effect size ($I^2 = 94.1\%$). In terms of individual categories, antipsychotic use increased the risk of neuropsychiatric outcomes (RR 1.79, 95% CI 1.03 to 3.12, $p = 0.039$, $I^2 = 94.1\%$). Other outcomes had a RR indicative of an increase in risk, but its confidence interval contained the

null value, and no significant result was found for such increase (RR 1.39, 95% CI 0.56 to 3.48, $p = 0.480$, $I^2 = 89.5\%$).

For adverse outcomes reported in ORs, antipsychotic use significantly increased the risk of musculoskeletal and movement-related outcomes ($p<0.0005$) and neuropsychiatric outcomes ($p<0.0005$) (Figure 2). No significant increase in the risk of mortality was found ($p = 0.161$). No heterogeneity was found within the individual estimates within the neuropsychiatric and mortality outcome categories. Pooling of the effects sizes from these outcome categories found antipsychotic use to significantly increase the risk of adverse outcomes ($p = 0.001$), although high heterogeneity was found in the summary effect estimate ($I^2 = 77.0\%$).

Aggregation of effect sizes extracted from meta-analyses of observational studies found antipsychotic exposure to be significantly associated with an increase in the odds of all adverse outcome categories, except endocrinological and metabolic outcomes ($p = 0.163$) (Figure 3). Among the categories with statistically significant ORs, the largest pooled effect size was found for mortality ($p = 0.007$), followed by injury ($p = 0.027$) and circulatory system outcomes ($p = 0.002$). Nevertheless, high heterogeneity was found for all three categories, with I^2 ranging between 77.8% to 88.1%, thus their summary effect estimates must be interpreted with caution. Similar to the pooled RRs and ORs from meta-analyses of RCTs, a significantly increase but heterogeneous risk of adverse outcomes was found to associate with antipsychotic use ($p = <0.0005$, $I^2 = 94.6\%$).

Figure 2. ORs for adverse health outcomes associated with antipsychotic use found in meta-analyses of RCTs, grouped by adverse outcome categories

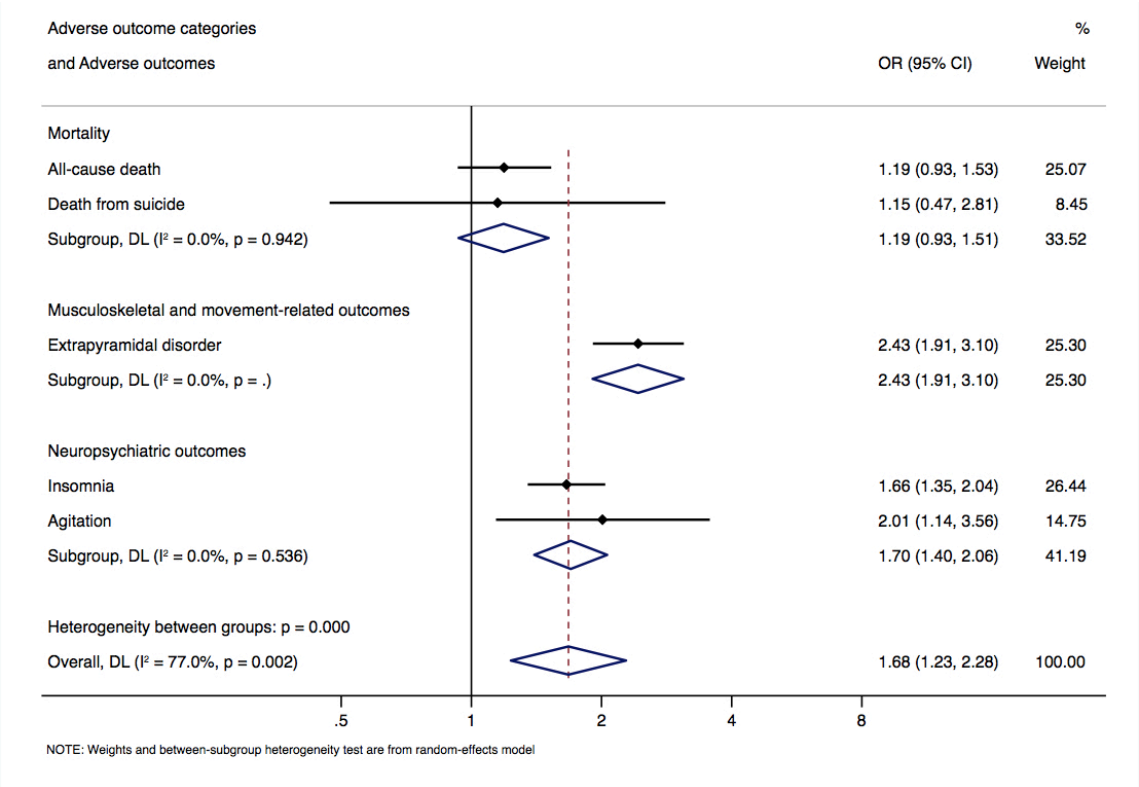
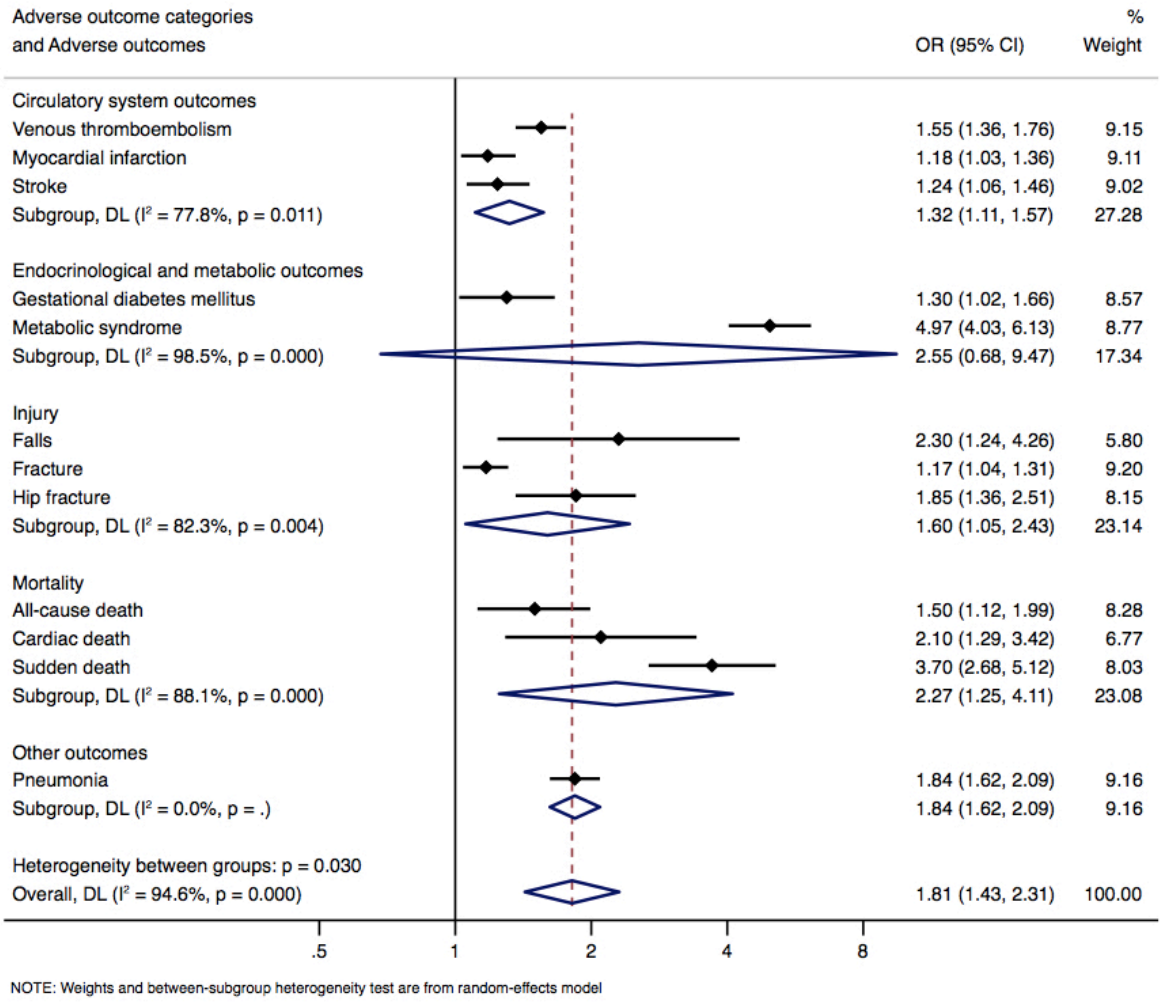


Figure 3. ORs and RRs (as OR equivalence) of adverse health outcomes associated with antipsychotic use found in meta-analyses of observational studies, grouped by adverse outcome categories



3.3 Assessment of Heterogeneity

Completed lists of heterogeneity statistics can be found in Tables 4-10 from section 3.2.1 to 3.2.7. For the summary estimates of adverse outcomes reported in meta-analyses of RCTs, the effects of 20 outcomes exhibited no heterogeneity, with an I^2 of 0% (endocrine abnormalities except diabetes, injury, musculoskeletal pain, gait abnormalities, hyperkinesia, hypertonia, insomnia, stroke, agitation, anxiety, depression, suicide attempt, suicidal ideation, completed suicide, menstrual disorder, sexual dysfunction, rash, hypertension, hypotension, all-cause death and cardiac death). Diabetes mellitus had heterogeneity higher than 0% but did not meet the threshold for low heterogeneity. Low heterogeneity ($25\% < I^2 < 50\%$) was found among diabetes mellitus, fatigue, clinically significant HbA1C level, somnolence and recurrence of psychotic symptoms. Moderate heterogeneity ($50\% < I^2 < 75\%$) existed in the remaining 5 adverse outcomes: clinically significant weight gain, extrapyramidal disorder, sedation and agitation. Six adverse outcomes had neither I^2 nor Q reported (myocardial infarction, blurred vision, peripheral oedema, electrocardiogram abnormalities and hypersomnia). No heterogeneity statistics was reported for pneumonia as its effect estimate was obtained from one primary study. Caution must be taken when interpreting heterogeneity statistics as outcomes with small number of primary studies or smaller average sample (N <80) often yield low heterogeneity and this does not necessarily indicate the absence of heterogeneity (Huedo-Medina et al., 2006; von Hippel, 2015).

Adverse outcomes reported in meta-analyses of observational studies generally exhibited high heterogeneity, with 3 outcomes having an I^2 over 90% (pulmonary embolism, hip fracture and all-cause death); and 4 outcomes with an I^2 over 80% (venous thromboembolism, myocardial infarction, fracture and stroke). High heterogeneity was also found for cardiac death. Metabolic syndrome had moderate heterogeneity. No or low heterogeneity was found in falls, gestational diabetes mellitus, pneumonia and sudden death.

3.4 Subgroup Analyses

Because most summary effect estimates of outcomes reported in the included meta-analyses of observational studies had high heterogeneity, subgroup analyses were conducted to explore the potential effect of individual factors in the occurrence of adverse outcomes. Subgroup analyses based on the following variables were possible: classes of antipsychotics, individual antipsychotics, age, diagnoses and observational study designs.

Three meta-analyses reported separate effect estimates to conduct a subgroup analysis for the two classes of antipsychotics (i.e., SGAs and FGAs) for five adverse outcomes: all-cause death, cardiac death, pneumonia, sudden death and venous thromboembolism. Thirty-eight primary studies provided data on FGAs and 44 primary studies provided data on SGAs. Despite having larger ORs in three out of five outcomes (all-cause death, cardiac death and sudden death) compared to SGAs (OR 1.70, 95% CI 1.44 to 2.01), the overall effect estimate reported for FGAs was smaller (OR 1.66, 95% CI 1.48 to 1.85). No significant difference was found between their summary ORs ($p = 0.787$). Within the classes, sudden death (OR 2.60, 95% CI 1.01, 6.74) and venous thromboembolism (OR 1.47, 95% CI 1.21 to 1.78) were associated with the largest and smallest ORs for FGAs, whereas the odds of pneumonia (OR 1.98, 95% CI 1.67 to 2.35) and all-cause death (OR 1.31, 95% CI 1.00 to 1.72) were the largest and smallest in the SGA group.

As for individual antipsychotic medications, subgroup analysis was possible given the effect sizes for metabolic syndrome, pneumonia and venous thromboembolism provided separately for olanzapine, quetiapine and risperidone. The effect estimate for risperidone was pooled from the largest number of studies ($k = 37$), followed by olanzapine ($k = 29$) and quetiapine ($k = 24$). Overall, the odds of adverse outcomes associated with olanzapine (OR 2.80, 1.18 to 6.65) triumphed over that of quetiapine (OR 2.58, 95% CI 1.10 to 6.05) and risperidone (OR 2.21, 95% CI 1.04 to 4.68), for which olanzapine had a large OR, followed

by quetiapine and risperidone with moderate ORs. No significance difference in the overall ORs for adverse events was found across the subgroups ($p = 0.916$). In terms of individual associations, all outcomes had the largest effect sizes for olanzapine. Quetiapine provided larger ORs for metabolic syndrome (OR 5.14, 95% CI 3.75 to 7.07) and pneumonia (OR 1.78, 95% CI 1.39 to 2.28) than risperidone (OR 4.57, 95% CI 3.48 to 6.03 and OR 1.44, 95% CI 1.12 to 1.85), whereas risperidone was associated with larger odds of venous thromboembolism (OR 1.63, 95% CI 1.16 to 2.31) than quetiapine (OR 2.58, 95% CI 1.10 to 6.05).

Comparison of ORs across age groups was possible for participants below the age of 65 years and participants over the age of 65, for which only effect size data of all-cause death and pneumonia was available. Results from the subgroup analysis indicate that antipsychotic users aged 65 or above had a larger effect size for all-cause death (OR 1.59, 95% CI 1.13 to 2.23) and a larger overall OR for adverse outcomes (OR 1.80, 95% CI 1.53 to 2.11) than those younger than 65 (all-cause death: OR 1.09, 95% CI 0.67 to 1.78; overall: OR 1.50, 95% CI 0.89 to 2.50), while the odds of developing pneumonia was the same (OR 1.86, 95% CI 1.53 to 2.20 or 2.25) in the two age groups. No significance difference was found in the overall ORs between the two age groups ($p = 0.505$). The disparate odds of all-cause death in younger and older individuals likely underlie the contrasting heterogeneity between the two age subgroups (below 65: $I^2 = 74.9\%$; over 65: $I^2 = 0.0\%$). Nevertheless, the larger overall OR for adverse outcomes found for the older age group could be attributed to the significantly larger number of primary studies ($k = 25$) with extractable data for ≥65-years olds compared to <65-year olds ($k = 9$), thus these pooled effect estimates should be deemed provisional only.

Sufficient data was obtained for the comparison of ORs between antipsychotic users with dementia ($k = 10$) and antipsychotic users with schizophrenia ($k = 9$). The ORs for both

all-cause death and venous thromboembolism were slightly larger in patients with dementia (OR 1.24, 95% CI 0.85 to 1.80 and OR 1.44, 95% CI 0.99 to 2.08) than that with schizophrenia (OR 1.07 95% CI 0.58 to 1.96 and OR 1.34, 95% CI 1.15 to 1.55). The test for inter-group heterogeneity suggested that such difference was non-significant ($p = 0.944$). The same pattern was illustrated in the overall effect estimates for the two groups (dementia: OR 1.34, 95% CI 1.03 to 1.74; schizophrenia: OR 1.32, 95% CI 1.14 to 1.53), with no heterogeneity found within neither group. Small effect sizes were found for both the dementia and schizophrenia groups.

Case-control studies ($k = 39$) provided larger ORs for four out of five outcomes (all-cause death, myocardial infarction, pneumonia and venous thromboembolism) and had a larger overall OR for adverse outcomes than cohort studies ($k = 40$) (case-control studies: OR 1.55, 95% CI 1.23 to 1.95; cohort studies: OR 1.48, 95% CI 1.32 to 1.64), although this difference was found to be non-significant ($p = 0.717$). Nevertheless, moderate heterogeneity exists within the ORs extracted from case-control studies ($I^2 = 69.9\%$), whereas no heterogeneity was found in the cohort studies subgroup. It should, however, be noted that the data for pneumonia was provided by only one cohort study, relative to five case-control studies.

Given the numerous sub-associations presented according to varied pharmacological and patient variables, conducting subgroup analyses using data from RCTs could be pragmatically difficult. For example, the RRs for clinically significant weight gain were calculated respective to different levels of doses and durations. Therefore, no subgroup analysis was conducted using data from meta-analyses of RCTs.

3.5 AMSTAR 2 Quality Assessment

Confidence assessment based on the AMSTAR 2 criteria suggests that the meta-analyses included in this umbrella review were of variable quality: 4 out of 27 included meta-analyses were rated as having high quality, 3 received a moderate confidence rating, 10 had a low-quality rating and 10 were of critically low quality (Table 11). Over half of the included meta-analyses (15 out of 27) contain no explicit statement indicating their establishment of a written protocol prior to the commencement of their research (i.e., Q2), which is regarded as a critical weakness, thus preventing their quality from being rated to be of any level above low or critically low. Insufficient attempt in addressing or taking into account the impact of the variable risk of bias of their included studies on the findings (i.e., Q12 and Q13) also resulted in ratings of low or critically low in 10 meta-analyses. Furthermore, almost all meta-analyses gave no consideration on or justification for their inclusion of either RCTs or observational studies. All but one meta-analyses acknowledged their source(s) of funding and declared potential conflicts of interest.

In terms of the adverse outcomes reported in meta-analyses of RCTs, ten out of 34 outcomes (hypertonia, musculoskeletal pain, hyperkinesia, anxiety, menstrual disorder, rash, hypertension, all-cause death, suicide attempt and completed suicide) had a high AMSTAR 2 quality rating, one outcome (clinically significant HbA1c level) had moderate quality, 14 outcomes (endocrine abnormalities except diabetes, fatigue, gait abnormalities, insomnia, sedation, agitation, myocardial infarction, diabetes mellitus, injury, stroke, recurrence of psychotic symptoms, suicidal ideation, hypotension and cardiac death) had low quality and 9 outcomes (blurred vision, peripheral oedema, clinically significant weight gain, extrapyramidal disorder, hypersomnia, somnolence, sexual dysfunction, electrocardiogram abnormalities and depression) were of critically low quality.

Out of the 13 adverse outcomes extracted from meta-analyses of observational studies, fracture had a high AMSTAR 2 quality rating. Three (hip fracture, pulmonary embolism and venous thromboembolism) and 5 (gestational diabetes mellitus, pneumonia, all-cause death, cardiac death, and sudden death) adverse outcomes were of moderate and low quality respectively. Finally, four (falls, metabolic syndrome, myocardial infarction and stroke) outcomes had a critically low quality rating.

Table 11. Overall and individual items' confidence ratings in the AMSTAR 2 checklist

Author (Year)	**Design**	**Q1**	***Q2***	**Q3**	***Q4***	**Q5**	**Q6**	***Q7***	**Q8**	***Q9***	**Q10**	***Q11***	**Q12**	***Q13***	**Q14**	***Q15***	**Q16**	**Rating**
Bai et al. (2020)	RCT	Y	N	N	Y	UC	UC	Y	Y	N	N	N	N	N	N	N	Y	Critically low
Defruyt et al. (2011)	RCT	Y	N	N	Y	Y	Y	Y	Y	PY	N	Y	Y	Y	Y	PY	Y	Low
Demyttenaere et al. (2019)	RCT	Y	N	N	Y	UC	UC	Y	Y	N	N	Y	N	N	Y	Y	Y	Critically low
Derry & Moore (2007)	RCT	Y	N	N	PY	UC	UC	Y	Y	PY	Y	Y	N	PY	Y	N	Y	Critically low
Harrington et al. (2010)	RCT	Y	N	N	Y	UC	UC	PY	Y	N	N	Y	N	N	Y	N	Y	Critically low
Hay et al. (2015)	RCT	Y	Y	N	Y	Y	Y	PY	PY	Y	Y	Y	N	PY	N	Y	Y	Moderate
Lao et al. (2016)	RCT	Y	Y	N	Y	UC	UC	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Low
Luan et al. (2017)	RCT	Y	N	N	Y	Y	Y	Y	PY	Y	N	Y	N	N	Y	Y	Y	Critically low
Luan et al. (2018)	RCT	Y	N	N	Y	UC	Y	Y	PY	Y	N	Y	N	N	Y	Y	Y	Critically low
Ma et al. (2014)	RCT	Y	N	N	Y	UC	UC	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Low
Maglione et al. (2011)	RCT	Y	PY	Y	Y	Y	Y	PY	Y	Y	PY	Y	N	N	N	Y	Y	Low
Maher et al. (2011)	RCT	Y	Y	Y	Y	Y	Y	Y	Y	Y	PY	Y	N	N	N	Y	Y	Low
Nussbaum & Stroup (2008)	RCT	Y	Y	N	Y	Y	Y	PY	Y	Y	Y	Y	Y	Y	Y	Y	Y	High

Nussbaum & Stroup (2012)	RCT	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	High
Polcwiartek et al. (2015)	RCT	Y	N	N	Y	Y	UC	Y	Y	N	N	Y	N	N	Y	N	Y	Critically low
Romeo et al. (2018)	RCT	Y	N	N	PY	UC	Y	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Low
Rotella et al. (2020)	RCT	Y	PY	N	Y	Y	Y	Y	Y	Y	N	Y	N	N	N	Y	Y	Low
Schneider-Thoma et al. (2018)	RCT	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	High
Kucukgoncu et al. (2020)	OBS	Y	N	Y	Y	UC	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Low
Liu et al. (2021)	OBS	Y	Y	N	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Moderate
Mortensen et al. (2020)	OBS	Y	Y	N	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	PY	Y	Y	Moderate
Nose et al. (2015)	OBS	Y	Y	Y	Y	Y	Y	N	Y	Y	N	Y	Y	Y	Y	Y	Y	Low
Papola et al. (2018)	OBS	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	High
Seppala et al. (2018)	OBS	Y	N	Y	Y	Y	Y	Y	Y	Y	N	Y	N	N	Y	N	Y	Critically low
Vancampfort et al. (2013)	OBS	Y	N	Y	Y	Y	UC	Y	UC	PY	UC	Y	N	N	Y	Y	N	Critically low
Yang et al. (2018)	OBS	Y	N	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Low
Zivkovic et al. (2019)	OBS	Y	N	PY	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Critically low

Italicised questions were regarded as critical items;
RCT = meta-analyses of randomised controlled trials; OBS = meta-analyses of observational studies;
'Y' = yes, 'PY' = partial yes, 'N' = no, 'UC' = unclear

3.6 Prediction Intervals

Sufficient data was available for the calculation of 95% prediction intervals for 24 outcomes extracted from meta-analyses of RCTs. Among these outcomes, gait abnormalities, insomnia, sedation and somnolence had prediction intervals with a lower limit above the null value of 1, indicating a 95% certainty that future observations would find a positive antipsychotic-associated risk of these outcomes. In contrast, the prediction intervals of the remaining 19 outcomes included lower limits below the null value, suggesting that there is a 95% chance that future observations would find null effect sizes for these outcomes.

Among the adverse outcomes extracted from meta-analyses of observational studies, the calculation of prediction intervals was possible for 8 outcomes (pulmonary embolism, venous thromboembolism, myocardial infarction, metabolic syndrome, stroke, gestational diabetes mellitus and pneumonia). Two of these outcomes (metabolic syndrome and pneumonia) had prediction intervals rejecting the null hypothesis as their lower limits exclude the null value, whereas the remaining outcomes had prediction intervals with lower limits below 1.

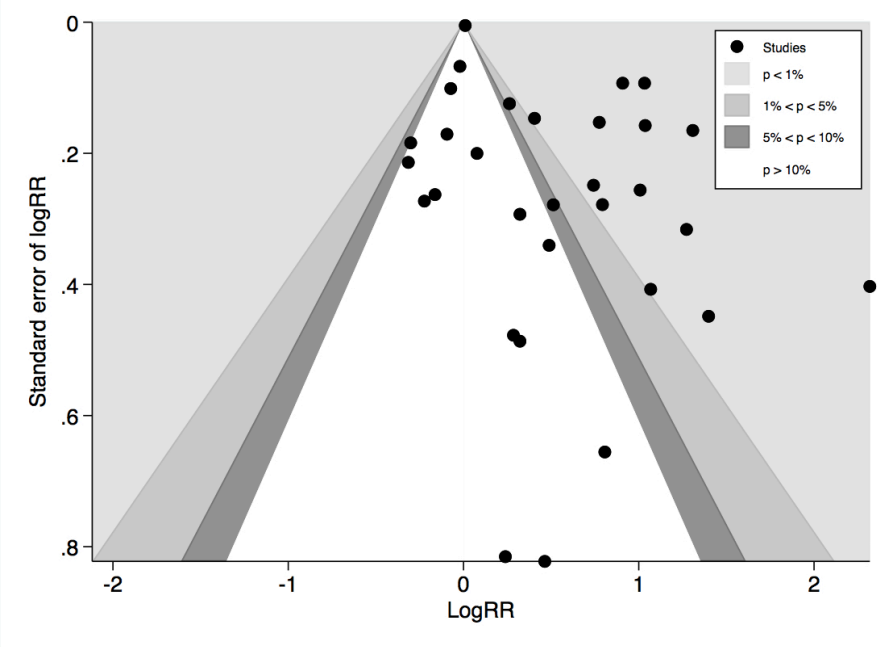
3.7 Small Study Effects

Publication bias was assessed by detecting the presence of small study effects using the arcsine-Thompson test. As adverse outcomes were reported in terms of both RRs and ORs, two respective funnel plots and arcsine tests were computed and performed for these effect sizes.

For adverse outcomes reported in RRs, an asymmetrical contour-enhanced funnel plot was generated (Figure 4). Data points at the top (i.e., larger reviews with lower standard error) scatter widely from the middle to the right and those indicative of higher standard error at the bottom do not seem to scatter evenly across the left- and right-hand side of the funnel. Data lying outside the limit lines also signify considerable heterogeneity in some reviews. However,

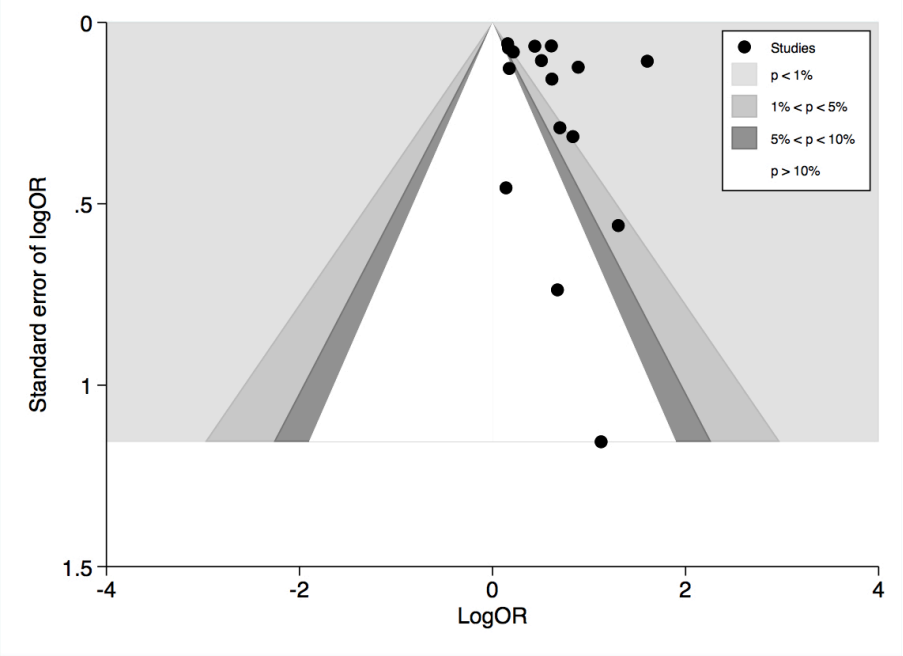
the arcsine-Thompson test found no significant results indicating asymmetry (p = 0.1730). Whether such asymmetrical pattern was caused by publication bias, however, is unclear because although close-to-half of the data points lie in areas of high statistical significance, more than half of the data points also lie in areas of low statistical significance.

Figure 4. Contour-enhanced funnel plot for evidence of small study effects in adverse outcomes reported in RRs



Likewise, no significant evidence for asymmetricity in the size and precision was found among the ORs extracted from included meta-analyses by the arcsine-Thompson test (p = 0.3973), despite the seemingly asymmetrical pattern illustrated in the contour-funnel plot of log(OR) and standard error. The data points of larger meta-analyses with lower standard error cluster relatively closely to each other, and majority if not all the data points lie on the right with missing studies on the left (Figure 5).

Figure 5. Contour-enhanced funnel plot for evidence of small study effects in adverse outcomes reported in ORs



Caution, however, must still be taken when interpreting the results generated from the arcsine-Thompson tests as the pre-arcsine-transformed log(RR) and log(OR) were estimated with a generic inverse variance model (with Dersimonian-Laird estimator) using the metagen function on RStudio–estimation using the metabin command was not possible because only the log-effect sizes and standard errors (calculated from 95% confidence intervals), rather than the number of exposed and unexposed participants with or without outcomes, were available.

3.8 Excess Significance Bias

ESS ratios (i.e., ratio of the summary OR of a meta-analysis to the OR of the largest component study in the same meta-analysis) were computed for 18 and 8 adverse outcomes reported in the meta-analyses of RCTs and observational studies respectively (figures 6 and 7; table 12).

Among the adverse outcomes reported in meta-analyses of RCTs, an ESS ratio greater than 1 was found for injury, clinically significant HbA1C levels, musculoskeletal pain, rash, hypotension, all-cause death and death from suicide. This indicates that for approximately 39% of the adverse outcomes, a summary effect estimate larger than the effect size of the largest primary study (i.e., true effect) was reported by the meta-analysis, thus suggesting evidence for excess significance bias. No evidence was found for the excess of studies with significant results in myocardial infarction, electrocardiogram abnormalities, gait abnormalities, hyperkinesia, hypertonia, anxiety, suicide attempt, suicidal ideation, menstrual disorder and hypertension as they had ESS ratios equal to or smaller than 1. ESS ratios ranging from 0.50 to 3.15 were found for the adverse outcomes reported in meta-analyses of observational studies, among which myocardial infarction, pulmonary embolism and pneumonia demonstrated apparent excess significant bias, whereas gestational diabetes, venous thromboembolism, fracture, stroke and all-cause death had ESS ratios equals to or below 1, indicating the absence of excess significance bias.

The calculation of the ESS ratio for the remaining 21 outcomes was not possible due to insufficient effect size data provided on their largest component studies.

Figure 6. The relationship between the reported summary effect size and the effect size quoted in the largest primary study in the meta-analysis for each adverse outcome extracted from RCTs

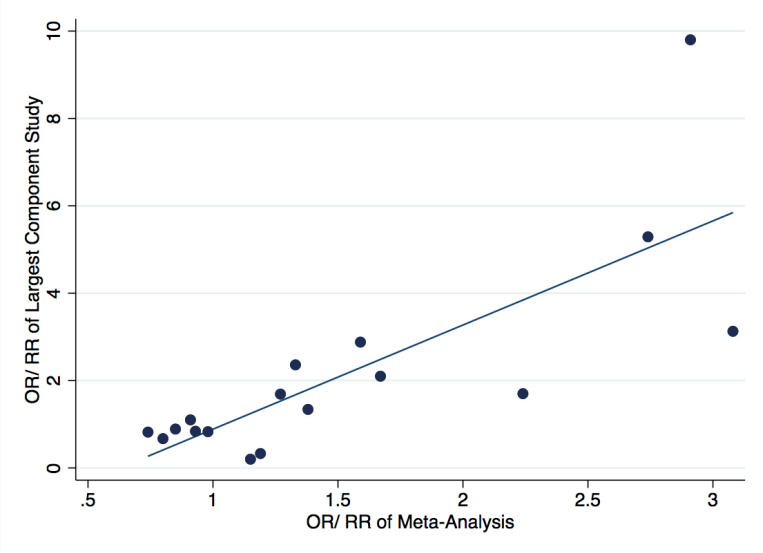


Figure 7. The relationship between the reported summary effect size and the effect size quoted in the largest primary study in the meta-analysis for each adverse outcome extracted from observational studies

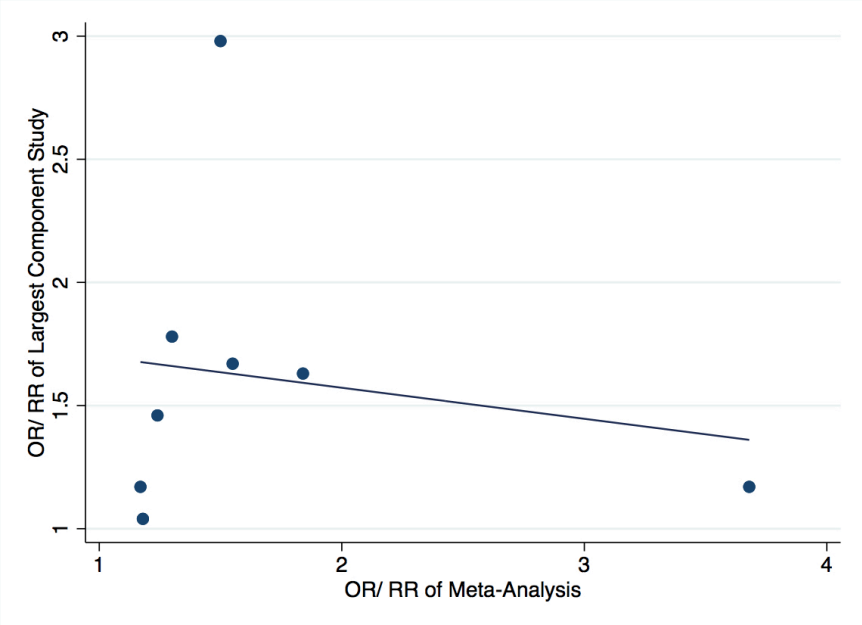


Table 12. Summary effect size and effect size of the largest primary study in the meta-analysis and the ESS ratio for each adverse outcome

Adverse outcome	ES(meta-analysis) (O)	ES(largest primary study) (E)	ESS (O/E) ratio
RCTs			
Myocardial infarction	3.08 (0.32, 29.73)	3.13 (0.13, 77.35)	0.98

Injury	0.98 (0.86, 1.12)	0.83 (0.65, 1.05)	1.18
Clinically significant HbA1c level	2.24 (0.62, 8.10)	1.70 (0.18, 16.17)	1.32
Electrocardiogram abnormalities	0.85 (0.51, 1.43)	0.89 (0.08, 9.81)	0.96
Musculoskeletal pain	0.80 (0.47, 1.37)	0.67 (0.06, 7.36)	1.19
Gait abnormalities	2.74 (1.66, 4.53)	5.29 (0.71, 39.47)	0.52
Hyperkinesia	1.67 (0.97, 2.89)	2.10 (0.75, 5.92)	0.80
Hypertonia	2.91 (1.31, 6.47)	9.80 (0.59, 163.02)	0.30
Anxiety	0.74 (0.52, 1.07)	0.82 (0.42, 1.63)	0.90
Suicide attempt	1.27 (0.26, 6.35)	1.69 (0.08, 35.11)	0.75
Suicidal ideation	0.91 (0.65, 1.27)	1.10 (0.66, 1.85)	0.82
Menstrual disorder	1.59 (0.32, 8.04)	2.88 (0.30, 27.14)	0.55
Rash	1.38 (0.53, 3.57)	1.34 (0.29, 6.27)	1.03
Hypertension	1.33 (0.52, 3.38)	2.36 (0.12, 45.49)	0.56
Hypotension	0.93 (0.76, 1.13)	0.84 (0.55, 1.29)	1.11
All-cause death	1.19 (0.93, 1.53)	0.33 (0.01, 8.06)	3.61
Death from suicide	1.15 (0.47, 2.81)	0.20 (0.01, 4.24)	5.75
Observational studies			
Pneumonia	1.84 (1.62, 2.09)	1.64 (1.40, 1.92)	1.13
Gestational diabetes	1.30 (1.02, 1.66)	1.78 (1.05, 3.03)	0.73
Pulmonary embolism	3.68 (1.23, 11.05)	1.17 (1.13, 1.21)	3.15
Venous thromboembolism	1.55 (1.36, 1.76)	1.67 (1.42, 1.98)	0.93
Fracture	1.17 (1.04, 1.31)	1.17 (1.13, 1.21)	1.00
Myocardial infarction	1.18 (1.03, 1.36)	1.04 (0.82, 1.33)	1.13
All-cause death	1.50 (1.12, 1.99)	2.98 (2.92, 3.04)	0.50
Stroke	1.24 (1.06, 1.46)	1.46 (1.22, 1.75)	0.85

3.9 Summary of Quality Assessments

An overall score of 0 to 5, from low (<2.5) to moderate (2.5 - 3) to high (4 - 5), was given for each reported adverse outcome according to the following criteria: (i) whether their 95% prediction intervals rejects the null hypothesis (yes = 1, no = 0); (ii) statistical significance ($p \leq 0.05$ = 1, $p > 0.05$ = 0); (iii) level of heterogeneity (low = 1, moderate = 0.5, high = 0); (iv) AMSTAR 2 rating (high = 1, moderate = 0.5, low = 0, critically low = 0) and (v) presence of excess significance bias (no = 1, yes = 0) (Tables 13 and 14). All outcomes had an overall score of at least 1. Among the 33 outcomes scored below 2.5, those extracted from meta-analyses of RCTs had prediction intervals with the null value, and low AMSTAR 2 ratings, showed excess significance bias and/ or did not reach statistical significance, whereas outcomes extracted from meta-analyses of observational studies mostly suffered from the problems of high heterogeneity, low AMSTAR 2 ratings and/ or having prediction intervals including the null value. Thirteen adverse outcomes had a moderate overall quality score. Two outcomes had an overall score of 4, thus were of high quality. No outcome had a score of 5.

Table 13. Summary table of results from the quality assessments on adverse outcomes extracted from meta-analyses of RCTs

Adverse outcome	No. of cases with the outcome	95% PI excludes null value?	*p*	I^2	AMSTAR 2 rating	Presence of ESS	Overall score
Significant associations							
Endocrine abnormalities except diabetes	39	No	4.00×10^{-3}	Low	Low	*	2
Blurred vision	*	*	2.00×10^{-3}	*	Critically low	*	1
Fatigue	468	No	4.28×10^{-7}	Low	Low	*	2

Peripheral oedema	*	*	1.70×10^{-8}	*	Critically low	*	1
Clinically significant weight gain	881	No	4.09×10^{-11}	Moderate	Critically low	*	1.5
Extrapyramidal disorder	*	*	2.70×10^{-12}	Moderate	Critically low	*	1.5
Gait abnormalities	117	Yes	9.46×10^{-5}	Low	Low	No	4
Hypersomnia	*	*	6.60×10^{-5}	*	Critically low	*	1
Hypertonia	46	No	9.0×10^{-3}	Low	High	No	4
Insomnia	*	Yes	1.57×10^{-6}	Low	Low	*	3
Sedation	2879	Yes	2.66×10^{-28}	Moderate	Low	*	2.5
Somnolence	1449	Yes	1.67×10^{-29}	Low	Critically low	*	3
Agitation	*	No	1.60×10^{-2}	Moderate	Low	*	1.5
Sexual dysfunction	*	*	4.90×10^{-2}	Low	Critically low	*	2
Non-significant associations							
Myocardial infarction	*	*	0.336	*	Low	No	1
Diabetes mellitus	41	No	0.152	Low	Low	*	1
Injury	469	*	1.194	Low	Low	Yes	1
Clinically significant HbA1C level	21	No	0.220	Low	Moderate	Yes	2
Electrocardiogram abnormalities	44	*	0.54	*	Critically low	No	1
Musculoskeletal pain	28	No	0.41	Low	High	Yes	2
Hyperkinesia	69	No	0.06	Low	High	No	3
Stroke	35	No	0.269	Low	Low	*	1
Anxiety	74	No	0.11	Low	High	No	3
Depression	59	No	0.707	Low	Critically low	*	1
Recurrence of psychotic symptoms	88	*	0.141	Low	Low	*	1
Suicide attempt	4	No	0.77	Low	High	No	3
Suicidal ideation	97	No	0.57	Low	Low	No	2
Menstrual disorder	4	No	0.57	Low	High	No	3
Rash	17	No	0.52	Low	High	Yes	2
Hypertension	15	No	0.55	Low	High	No	3
Hypotension	250	*	0.45	Low	Low	Yes	1
All-cause death	207	No	0.172	Low	High	Yes	2
Cardiac death	*	No	0.360	Low	Low	*	1
Completed suicide	13	No	0.772	Low	High	Yes	2

*Review provided insufficient information for determination

Table 14. Summary table of results from the quality assessments on adverse outcomes extracted from meta-analyses of observational studies

Adverse outcome	No. of cases with the outcome	95% PI excludes null value?	*p*	I^2	AMSTAR 2 rating	Presence of ESS	Overall score
All-cause death	*	*	6.00×10^{-3}	High	Low	No	2
Cardiac death	*	*	3.00×10^{-3}	High	Low	*	1
Falls	*	*	8.00×10^{-3}	Low	Critically low	*	2
Fracture	221	No	9.00×10^{-3}	High	High	No	3
Gestational diabetes mellitus	140	No	3.20×10^{-2}	Low	Low	No	3
Hip fracture	*	*	9.48×10^{-5}	High	Moderate	*	1.5
Metabolic syndrome	*	Yes	8.87×10^{-51}	Moderate	Critically low	*	2.5
Myocardial infarction	1,489 (2 NR)	No	2.00×10^{-2}	High	Critically low	Yes	1
Pneumonia	4906	Yes	7.09×10^{-21}	Low	Low	Yes	3
Pulmonary embolism	451,286	No	2.00×10^{-2}	High	Moderate	Yes	1.5
Venous thromboembolism	103,427	No	8.02×10^{-11}	High	Moderate	No	2.5
Stroke	1,324 (5 NR)	No	7.10×10^{-3}	High	Critically low	No	2
Sudden death	*	*	1.56×10^{-14}	Low	Low	*	2

*Reviews provided insufficient information for determination; (number NR) indicates the number of primary studies with unreported number of cases

Chapter 4: Discussion

By combining evidence from meta-analyses of both RCTs and observational studies, the current umbrella review aimed to systematically assess the adverse health outcomes associated with typical and atypical antipsychotics. To my knowledge, this is the first umbrella review grading the evidence of antipsychotic-associated adverse outcomes from meta-analyses of both RCTs and observational studies. Therefore, this review will add to the existing literature by providing a more comprehensive assessment of the various adverse outcomes associated with antipsychotics.

4.1 Antipsychotic-Associated Adverse Outcomes

Most adverse outcome data identified using my inclusion criteria were for SGAs, with effect sizes reported for aripiprazole, asenapine, amisulpride, cariprazine, clozapine, lurasidone, paliperidone, olanzapine, quetiapine, risperidone, and ziprasidone. The abundance of findings on these SGAs was anticipated given their rapidly growing popularity in recent years (Claassen & Park, 2021; Hirsch et al., 2018). Effect size data indicated associations between antipsychotics and impaired mental, cardiovascular, cerebrovascular, metabolic, endocrine, sensory, respiratory, and movement-related outcomes.

The meta-analysis on 18 meta-analyses of RCTs found antipsychotics to increase the risk or odds of 25 out of 35 adverse health outcomes in individuals prescribed with antipsychotic medications, as compared with those treated with placebo. The increase in risk or odds was statistically significant in 14 outcomes. None of the RRs suggesting decreased risks were statistically significant. Overall, SGAs increased the risk of sleep outcomes and such increase was statistically significant. In terms of other neuropsychiatric outcomes, no association was found between antipsychotic use with anxiety, depression, recurrence of

psychotic symptoms and suicidality outcomes. However, there was a two-fold greater risk of agitation in those prescribed with aripiprazole or cariprazine, compared to patients prescribed with placebo. Despite the 'extrapyramidal symptoms-free' label pharmaceutical companies commonly associate SGAs with, this umbrella review found SGAs to significantly increase the risk of movement-related outcomes. This is in line with the findings of a previous overview, which has found SGAs to increase the prevalence of movement disorders in patients with schizophrenia or psychosis (Martino et al., 2018). Motor side effects are considered a result of SGAs' excessive blockade of D_2 receptors (Glazer, 2000; Werner & Coveñas, 2017). On the other hand, the meta-analyses of observational studies found significant associations between antipsychotics and serious adverse health outcomes, including stroke and myocardial infarction, injury, and death.

While the effect sizes for mortality outcomes reported in meta-analyses of RCTs and meta-analyses of observational studies were similar, no statistically significant result was found for the effect sizes of RCTs. The different length of follow-up of the two study designs may have led to the discrepancies in the statistical significance (Kimachi et al., 2021). For instance, the meta-analysis of RCTs had follow-ups ranging from 1.5 hours to 104 weeks (Schneider-Thoma et al., 2018), whereas the follow-ups in the meta-analysis of observational studies spanned from 1 week to 12 years (Yang et al., 2018). This means that mortality may be a delayed adverse outcome of antipsychotic use, highlighting RCTs' failure to capture adverse outcomes occurring possibly years after the first exposure. Among the mortality outcomes, comparison across the RCT and observational study designs was possible for all-cause death, which indicated no significant difference beyond chance.

Both meta-analyses of RCTs and meta-analyses of observational studies found evidence for the association between antipsychotic use and increased risks or odds of metabolic outcomes. This finding is consistent with previous reviews, which found SGAs to be associated

with a wide range of metabolic side effects (Tschoner et al., 2007). As an antagonist, SGAs' high affinity for H_1-histamine receptors has been the long-acknowledged mechanism underlying metabolic adverse outcomes (Nasrallah, 2008).

Attempts were made to identify potential pharmacological heterogeneity in observational data via subgroup analyses on the classes and individual types of antipsychotics, which, however, produced no informative findings. The two subgroup analyses found SGAs to have larger ORs compared to FGAs, and olanzapine to have the largest ORs, followed by quetiapine and risperidone. The differences between subgroups were not statistically significant and the confidence intervals of outcomes overlapped. Furthermore, high heterogeneity was also found within the SGA and all individual antipsychotic subgroups.

Although effect sizes of different magnitudes were found, no meaningful comparisons could be made across individual adverse outcomes. For instance, comparing clinically significant HbA1C level and fatigue would have been problematic because the confidence intervals included the null value, which indicated uncertainty about the true effect size. This also illustrated why I refrained from interpreting the effect sizes in terms of strength – concluding associations to be 'strong' simply because they have large effect sizes, when substantial variability exists among treatment effects, could be misleading. No meaningful interpretation could be made on the comparative risks or odds of adverse outcomes across the adverse outcome categories either, due to the substantial heterogeneity existed within each category.

4.2 Quality Assessment

The evidence for antipsychotic-associated adverse outcomes was generally of low quality Most associations scored below 2.5 in the quality assessment out of a total score of 5. High-to-low quality evidence was found for the outcomes reported in meta-analyses of RCTs.

Movement-related outcomes, i.e., gait abnormalities (RR 2.74, 95% CI 1.66 to 4.53) and hypertonia (RR 2.91, 95% CI 1.31 to 6.47) were supported by high quality evidence. Sleep outcomes generally had moderate quality evidence, except hypersomnia which scored 1. The associations reported in meta-analyses of observational studies had moderate-to-low quality. Moderate quality evidence was found for the association between antipsychotic use and injury, metabolic and circulatory outcomes.

Majority of the meta-analyses included in this umbrella review had low or critically low quality according to AMSTAR 2 criteria. Meta-analyses without pre-registered protocols are especially prone to the selective reporting (Ayorinde et al., 2020). The lack of adequate discussion of meta-analyses on the risk of bias threatened the credibility of the reported associations. Meta-analyses which accounted for the risk of bias in primary studies assessed bias using the Cochrane Risk of Bias (RoB) tool, Newcastle-Ottawa Scale and SIGN checklist. Most meta-analyses included studies with low risk of bias. A minority of meta-analyses reported high risk of selective reporting, with other sources of bias being the inadequate matching of participant variables, funding from pharmaceutical companies and high non-response rate (Nussbaum & Stroup, 2008; Papola et al., 2018; Schneider-Thoma et al., 2018). While meta-analyses helped evaluating the consistency among findings from primary studies, the fundamental flaws in primary studies' methodology would still have posed threats to the validity of this umbrella review, if they were left unaddressed with inadequate assessment on the risk of bias, which was the case for almost half of the included meta-analyses.

Effect sizes had substantial heterogeneity, which most possibly stemmed from the variable pharmacological variables (i.e., types of antipsychotics, doses and duration) and confounders in observational studies. For observational data, carefully adjusted effect estimates were available for falls and gestational diabetes mellitus only. The meta-analyses of most other outcomes reported effect estimates pooled from studies with and without adjustment

for confounders, while the meta-analysis for metabolic syndrome reported no details on adjustment. Unadjusted confounders such as patients' age, sex and concomitant medication use may have given rise to the clinical heterogeneity observed in the analysed outcomes. For instance, lithium is sometimes prescribed concomitantly with antipsychotics for the treatment of mania among patients with psychotic disorders (Scherket al., 2007). Unlike antipsychotics, lithium has been found to have either no or a protective effect against stroke in population-based studies, which may potentially contribute to the heterogenous results for the risk of stroke (Chen et al., 2019; Lan et al., 2015). On the other hand, tricyclics and serotonin reuptake inhibitors, which are often augmented by antipsychotics for treatment-resistant depression and bipolar disorder (Fleurence et al., 2009), were found to associate with an increased risk of hip fracture (Dragioti et al., 2019). The exact effect of antipsychotics on the risk of injury outcomes could not be determined, given the effect of concomitant antidepressant use as a confounder. SMIs themselves are also associated with an increased risk of myocardial infarction and stroke (Correll et al., 2017; Zivkovic et al., 2019). Therefore, the extent to which antipsychotics manifested circulatory system outcomes in patients diagnosed with SMIs was unclear.

Misdiagnosis or misclassification of adverse outcomes may have affected results. Varied means of documentation of adverse health events were used among the included observational studies. Some primary studies used standardised diagnostic criteria from the ATP, ICD, IDF or WHO, while others obtained outcome information from medical or post-mortem records. Furthermore, the cases of some circulatory system outcomes and injury were determined via self-reports, meaning effect size data for such were especially vulnerable to self-report bias. Together with the wide range of age groups and psychiatric diagnoses among participants, the inconsistent definition of adverse outcomes contributed to the substantial heterogeneity found among observational data, which should be addressed with subgroup analyses to isolate the influence of individual factors. Nevertheless, the scattered identification

of patients' baseline clinical characteristics in majority of the meta-analyses provided little input for subgroup analyses.

Case-control studies identify participants who developed the concerned outcomes before comparing their previous exposure to aetiological determinants with those unaffected by the concerned outcomes (Coggon et al., 2003). This means case-control studies are prone to selection bias as the identification of a control group with characteristics fully matched with those from the case group is difficult. These diverse patient characteristics may be reflected in the considerable heterogeneity among the case-control subgroup in the stratified analysis on study designs, which nevertheless found no significant difference between case-control and cohort studies. Furthermore, most included meta-analyses indicated the utilisation of administration databases in primary studies, so selection bias might be minimal.

RCT data were not bias-free either. For instance, the quantity of data collected in RCTs depended wholly on participants' adherence to the antipsychotic treatments. This is especially problematic when patients had co-morbid mental and physical conditions or substance use disorders which potentially underlie nonadherence in schizophrenia (Haddad et al., 2014), as patients who refused antipsychotic administration would have been excluded. Likewise, Maglione et al. (2011) and Maher et al. (2011) included antipsychotic off-label users with substance use disorder, meaning the findings from these meta-analyses may not be applicable on non-compliant antipsychotic users.

Owing to their narrow inclusion criteria and restrictive recruitment processes, RCTs are especially prone to selection bias. Their inadequate representativeness of the clinical population could dramatically underestimate the heterogeneity in their reported outcomes (Kravitz et al., 2004). While no or generally low heterogeneity was found among effect sizes extracted from RCTs, ecologically limited findings could seriously undermine the quality of their evidence. Furthermore, adverse events are often the secondary outcome in meta-analyses

of RCTs (e.g., Nussbaum and Stroup (2012)), meaning the quality of investigation on such may have fallen short compared to the primary outcome (i.e., efficacy).

A potentially important concern on the meta-analyses' quality was the asymmetry found in the assessment for publication bias. There were apparent discrepancies between the graphical and statistical results from the funnel plots and arcsine-Thompson tests, thus implying the role of variance between the included meta-analyses in contributing to the asymmetry in the funnel plots. For instance, high heterogeneity was exhibited in over half (61.5%) of the effect estimates reported in meta-analyses of observational studies. As most of the outcomes reported in ORs were extracted from meta-analyses observational studies, the substantial between-study heterogeneity may explain such asymmetry. Still, heterogeneity may not fully account for the asymmetrical pattern displayed among RRs, which were mainly reported in meta-analyses of RCTs. In fact, majority of the outcomes extracted from RCTs were of low heterogeneity. An alternative explanation for the asymmetry could be the constituting small-scaled RCTs reporting predominantly larger effect sizes. This could be substantiated by the little number of cases with the adverse event for each outcome. For example, large and positively trending RRs for menstrual disorder, endocrine abnormalities except diabetes and hypertonia were extracted from trials with only 4, 39 and 46 cases respectively.

Another measure of variability was 95% prediction intervals, which account for the range within which the risk of each adverse outcome in future studies or occurrences would lie (Riley et al., 2011). Circulatory system, neuropsychiatric and movement-related outcomes had prediction intervals without the null value, indicating that antipsychotic use would lead to an increased risk of these outcomes in 95% of future studies.

Among meta-analyses which provided sufficient information for the calculation of ESS ratios, 39% (7/18) and 50% (4/8) of outcomes showed excess significance bias, as their

observed summary effect size exceeded the expected effect size from the largest study, which was assumed to be the closest to the true effect. Taking into account the variable confounders and potential small study effects discussed above, it is probable that the substantial between-study heterogeneity, inadequate number of cases and/ or even publication bias played a part in the excess of positive effect sizes (Tsilidis et al., 2012).

4.3 Comparison with Previous Umbrella Review

For pneumonia, hip fracture, venous thromboembolism, stroke, sudden cardiac death and myocardial infarction, the range of my effect sizes ($1.18<OR<2.10$) resembled those reported by Papola (2019), who found ORs between 1.45 and 2.24. Discrepancies existed between individual effect sizes, with the most prominent being that of myocardial infarction, which had the smallest and largest effect sizes respectively in the current and Papola et al.'s reviews. I found smaller effect sizes for stroke and sudden cardiac death, and a larger effect size for hip fracture.

A direct comparison between the quality of the included evidence with that of Papola et al. was not possible, given the different criteria we applied on quality assessments. Papola et al. concluded the evidence for pneumonia to be of the highest grade (i.e., convincing) according to the URC, while it was given a moderate rating in the current umbrella review even though our data were extracted from the same meta-analysis. This could be mainly attributed to the fact that the URC does not account for the AMSTAR 2, which suggests the evidence for pneumonia to be of low quality. AMSTAR 2 addresses the methodological flaws underlying risks of bias which tend to be overlooked in many systematic reviews, unlike the URC (Lorenz et al., 2019). Moreover, when assessing excess significance bias, I computed the ESS ratios by dividing the effect size of the included meta-analysis by that of the largest study included in the same meta-analysis, while Papola et al. assessed the observed and expected number of

primary studies with statistically significant effect sizes–this may explain why my assessment on excess significance bias generated positive results for pneumonia when Papola et al. did not.

Compared with the meta-analyses included in Papola et al.'s umbrella review, the meta-analyses reporting the same outcomes in this review were generally of lower quality in terms of AMSTAR 2 ratings. For instance, I assigned a moderate, critically low and low confidence rating to hip fracture, stroke and cardiac death respectively, whereas high, low and moderate ratings were given respectively for these outcomes in Papola et al.'s review, although the association of venous thromboembolism was supported by higher quality evidence in the current review. Furthermore, with the exception of venous thromboembolism, all overlapped outcomes in Papola et al.'s review had higher statistical significance than this review. This could be explained by the larger sample sizes of the meta-analyses for all outcomes except stroke included in Papola et al.'s review than the current review.

Nevertheless, effect size data of the adverse outcomes reported in this review were less heterogeneous than that in Papola et al.'s review, which found I^2 predominantly above 90% and as high as 99%. Another barrier to effectively comparing the quality of the current and previous evidence was the missing information from my included meta-analyses. Quite a few meta-analyses, in particular that of hip fracture and cardiac death, failed to report data from primary studies and τ^2 statistics, which prevents the calculation of 95% prediction intervals and ESS ratios. This means that only tentative conclusions could be made on the comparative certainty or confidence for these outcomes in the current and Papola et al.'s umbrella reviews.

4.4 Susceptibility to Adverse Drug Outcomes

Multiple subgroup analyses were conducted given the separate effect sizes reported, respective to different clinical and patient variables, by the included meta-analyses of observational studies. The respective ORs of subgroups indicated that antipsychotic users aged

above 65 or diagnosed with dementia were more at risk of adverse events than those below 65 or diagnosed with schizophrenia. Like pharmacological variables, no statistical difference was found between the age and diagnosis subgroups. The failure to detect differences may have stemmed from the inadequate number of adverse events for the tests of effect. Researchers suggested that, for clinical research, a much larger sample size is required to detect meaningful differences between subgroups than that between an overall treatment and null effects (Brookes et al., 2004; Burke et al., 2015). The fact that there were only 9 and 10 studies in the schizophrenia and dementia groups respectively means their effect estimates likely lacked statistical power, which could result in false negatives. A large-scale meta-analysis of RCTs pointed out that observational studies' investigations on antipsychotic-associated physical adverse events are multivariate and may not be informative in terms of causality (Schneider-Thoma et al., 2019). Nevertheless, their findings from over 300 RCTs resembled that of my subgroup analysis, for which older patients (>65 years) showed the highest risk of somatic adverse events, relative to adults and children or adolescents (Schneider-Thoma et al., 2019).

The intertwining relationship between age and diagnosis serves as a major caveat of interpreting the current findings on individual patient characteristics. In particular, older individuals are more prone to dementia, which has been found to elevate susceptibility to antipsychotic-associated mortality (Liperoti et al., 2009). The pharmacological interactions between antipsychotics and cholinesterase inhibitors, prescribed for the treatment of dementia, also contribute to serious psychiatric, respiratory and vascular extra-cardiac adverse drug reactions (Pariente et al., 2010). In fact, most research on antipsychotic-associated adverse events in older adults were conducted on dementia patients (Kate et al., 2015). The subgroup analyses and reported individual associations failed to reveal the respective extent to which patients with and without dementia were prone to adverse health outcomes. Furthermore, no

subgroup analysis was possible on children or adolescents, which have been found to be more at risk of antipsychotics' adverse effects than adults (Pillay et al., 2018).

Overall, results from subgroup analyses suggested that patients' age and psychiatric diagnosis do not modify their susceptibility to adverse events associated with antipsychotics. However, the notable within-group heterogeneity and small samples considerably reduced the validity of the test of interactions, meaning the interpretation of age and diagnosis' potential impact on heterogeneity in adverse event susceptibility was uncertain. Still, my attempted subgroup analyses explored multiple directions of evidence and could therefore be regarded as a strength of this umbrella review, considering the lack of comprehensive investigations on individual variables potentially impactful on drug-response relationships in the current literature.

4.5 Strengths and Limitations of This Umbrella Review

The utilisation of subheadings and MeSH terms maximised the sensitivity of my search and helped identifying a wide range of adverse outcomes. Comprehensive quality assessments were conducted to grade the included evidence to inform clinical decision-making. I made strategic choices to resolve the pragmatic difficulties arisen from the inconsistently reported data by modifying criteria for data analysis and quality assessments.

However, my approach in pooling individual associations reported respective to clinical and patient factors may be inappropriate given the vast amount of heterogeneous data retrieved. The fact that the component meta-analyses presented these estimates in individual associations in the first place may very well suggest that they were meant to be interpreted individually. Considering the varying pharmacology, it may not be sensible to regard SGAs as a homogenous entity (Leucht et al., 2009). While simplifying the complex mechanisms and downplaying discrepant evidence is useful in establishing consistency in the literature, ill-

informed clinical decision-making could manifest serious harm to patients. The extent to which discrete but related effect sizes should be meta-analysed is still debateable in the current literature. Still, I synthesised these individual associations purely as a means of summarisation and provided all component effect sizes which, along with findings from quality assessments, should offer adequate rooms for evaluation during informed decision-making.

I assumed the ORs and HRs of circulatory system outcomes reported respectively for case-control and cohort studies to be equivalent and meta-analysed them to estimate a summary OR for each individual outcome. HR estimates the likelihood that one suffers from an adverse event at a given point in time and measures risks that may change continuously (Hernàn, 2010), while OR and RR do not take into account the timing during which the adverse event occurred (George et al., 2020). Due to HR's different nature from that of OR and RR, pooling them together may have given rise to the high heterogeneity found in these outcomes. Still, the Cochrane Handbook for Systematic Reviews of Intervention suggests that the synthesis of case-control and cohort data may take place when studies address similar research questions and the random-effects model is used, meaning my combination of such could be deemed acceptable (Higgins et al., 2021).

I included two meta-analyses of RCTs focusing solely on patients with treatment-resistant depression, who may be regarded as atypical relative to non-treatment-resistant patients (Luan et al., 2017; Luan et al., 2018). For instance, antipsychotic augmentations were found to increase the risk of neuropsychiatric and other outcomes. Treatment-resistant patients' atypicality could be expressed in neurological terms. For example, they were found to have significantly reduced grey matter volume in their putamen while non-treatment resistant (i.e., major depressive disorder) patients were not (Klok et al., 2019). There was a possibility that the putamen atrophy in treatment-resistant patients have heightened the risk of hypersomnia, which is associated with decreased cerebral blood flow in the default mode network (i.e., in

the putamen) (Pomares et al., 2019). The antipsychotic-associated risk of hypersomnia might be overestimated in this review as it may have been lower in non-treatment resistant patients. As no subgroup analysis was possible, whether the predisposed differences in physiology between treatment-resistant and non-treatment resistant patients affected their susceptibility to adverse events was unknown. These presented a caveat to generalising the findings from treatment-resistant patients on the non-treatment-resistant populations.

Bias may have existed in my classification of extracted adverse outcomes. The definitions of adverse events in previous literature have been broad and discordant. For example, the terms 'asthenia' and 'fatigue' are often used interchangeably in the current literature. However, some regard asthenia and fatigue as separate events and suggested that the former refers to a subjective sensation of energy loss, while the latter is a symptom of tiredness resulted from effort (Noble, 2020). I treated them as a single outcome during adverse event classification and ultimately the selection of studies. Although doing this could avoid redundancy if authors of meta-analyses were using these terms interchangeably, there may have been loss of data if they were describing separate events. Nevertheless, little information was given by authors on how their outcomes were defined in majority of included meta-analyses. Subsuming outcomes of similar nature but give different experience (e.g., extrapyramidal disorder included akathisia, tardive dyskinesia, dystonia, etc.) also decreased the specificity of my findings.

Unlike previous umbrella reviews on psychiatric medications' adverse outcomes (Dragioti et al., 2018; Papola et al., 2019), I extracted summary effect sizes from included meta-analyses rather than re-estimated the summary estimates by aggregating the associations from primary studies, which could lower the consistency in my effect size data and prediction intervals. I was also not able to assess publication bias individually for each meta-analysis, since the computation of arcsine's p-value requires the effect sizes from primary studies. Due to the insufficient data on primary studies provided by many included meta-analyses, I

calculated the ESS ratios under the simplest assumption in terms of effect sizes. Conventionally, E is obtained by calculating the sum of the statistical power estimate for each component study in the meta-analysis, which is computed with an algorithm based on the non-central *t* distribution (Tsilidis et al., 2012). The effect size-based ESS ratios may be flawed as effect sizes tend to be overestimated when the statistical power of the studies is low, and even more so when results are highly significant (Curran-Everett, 2017).

Unlike previous umbrella reviews (Dragioti et al., 2019; Papola et al., 2019), the GRADE framework was not used to appraise the quality of meta-analyses in the current review. The GRADE tool addresses the risk of bias, inconsistency, indirectness, imprecision, publication bias and confounders in given evidence, and was regarded as a good alternative for the URC (Schlesinger et al., 2019). Its emphasis on the risk of bias makes it a universally acclaimed assessment tool for the methodological quality of systematic reviews (Zähringer et al., 2020). Again, the lack of accessibility to primary study data and the scope of the current project precluded the assessment on risk of bias using the Cochrane Risk of Bias 2 tool and small study effects for each meta-analysis (Higgins et al., 2021).

The application of the conventional threshold, i.e., $\alpha = 0.05$, in null hypothesis testing has long been criticised (Szucs and Ioannidis, 2017). A 0.005 threshold may mitigate the risk of false positives, especially for potentially underpowered RCTs (Ioannidis, 2018; Wayatt et al., 2018). In the current review, an association was given a score of 1 as long as it had a p-value below 0.05. For instance, the confidence interval for sexual dysfunction contained the null value of 1.00 but its p-value suggested that the effect was significant ($p = 0.049$). This may have led to my overestimation of the true effect. Nonetheless, there still seems to be no consensus on the definite threshold indicating statistical power (Solmi et al., 2018), with researchers criticising the overemphasis on p-values in evidence-based medicine and some even proposing to abandon the use of such (Evans et al., 2021; Amrhein, et al., 2019). These

thresholds may merely be arbitrary cut-offs. Along with heterogeneity statistics and prediction intervals, my presentation of effect sizes and confidence intervals already allowed meaningful interpretation of findings (Ranstam, 2012).

4.6 Implications

This umbrella review contributes to the pharmacovigilance of antipsychotics and shed light on the inconsistent quality of evidence in the current literature, which serves as a major barrier to applying evidence-based medicine into clinical practice. Underestimating the risk of adverse outcomes associated with psychiatric medications could compromise treatment efficacy or even cause serious harm to patients, while overestimating such risk could deny patients potentially effective pharmacological interventions.

The identification of life-threatening adverse drug events is no doubt of crucial importance. Still, this should not take away the merit of research on more common adverse outcomes. One may assume metabolic outcomes to be less life-threatening than cardiovascular events and hence overlook the importance of managing such. Nevertheless, metabolic side effects are strong mediators of physical comorbidities or even death (De Hert et al., 2012). Clinicians should monitor the metabolic functioning in patients taking antipsychotics, particularly SGAs, to ensure preventive measures are in place when comorbidity with serious adverse outcomes is foreseen (Nasrallah, 2007).

Adverse events tend to be investigated as an outcome secondary to efficacy in clinical trials, and the design of analysis on the harm profiles of antipsychotics is generally poorer, relative to investigations on antipsychotics' efficacy. Treatment success in antipsychotic pharmacotherapy requires long-term administration to mitigate symptoms and prevent relapse (Higashi et al., 2013), but the burden of metabolic and sleep outcomes was found to be associated with nonadherence (DiBonaventura et al., 2012). Efficacy and safety should not be

viewed as separate entities: an antipsychotic regarded to have prominent therapeutic benefits should not have life-threatening adverse effects, and same the other way round, an antipsychotic with serious adverse effects is not effective.

It would be sensible to assume that the differing pharmacokinetics of antipsychotics account for distinctions in adverse effects between the individual classes or types of antipsychotics and dose-response relationships. Likewise, different individual variables such as age and diagnosis may play a role in patients' susceptibility to adverse outcomes. Nevertheless, the low-quality data from primary studies precluded interpretations of the relative effect of these factors on the occurrence of adverse outcomes. This important literature gap remained unaddressed in the current umbrella review. Furthermore, my findings may not be informative for clinicians in low-income countries where FGAs are still more commonly used than SGAs (Teferra et al., 2013).

The reporting of harms in systematic reviews is generally poor, at least for antipsychotics. Researchers should develop the practice of registering protocols to improve transparency of methodologies and ensure reliable reporting. The use of validated research guidelines such as the AMSTAR 2 and PRISMA harms checklists could help researchers establish a clear rationale for the structure and scope of their reviews. My umbrella review serves as an initial evaluation on the wide range of antipsychotic-associated events reported in the current evidence base and provides rationales for future reviews in developing a more narrowly focused scope to assess specific events which may be clinically significant. Still, findings from this review were just the tip of the iceberg. Despite my effort in synthesising a wide range of health outcomes, the risk of many other outcomes is still unknown. Regular review on the rapidly accumulating systematic reviews is required.

An important safety concern for psychiatric medications is their propensity for suicidal ideation and behaviour in users. Researchers suggested that antipsychotics may indirectly

induce pro-suicidal neurological activity in patients subjected to neuropsychiatric and/ or movement-related outcomes, such as extrapyramidal effects (Mihanović et al., 2010). The fact that no significant evidence was found between antipsychotic use and increased suicidality in the current umbrella review may be reassuring.

Disentangling the complex mechanisms underlying the distinctions in adverse effects between individual antipsychotics is challenging. The intertwining relationship between the variables involved in antipsychotic administration serves as a major problem in research. Substantial heterogeneity among findings seems to be unavoidable if an overall landscape was to be captured for all antipsychotics. It may be more beneficial to comprehensively study each antipsychotic individually, rather than to simplify antipsychotics and aggregate their complex multivariate relationships into a single entity. Future systematic reviews or meta-analyses could address the substantial heterogeneity in results by conducting subgroup analyses or meta-regression.

4.7 Conclusion

Principal findings of this umbrella review substantiated previous evidence indicating the positive link between antipsychotic use and circulatory system, neuropsychiatric, movement-related, endocrinological and metabolic outcomes, injuries, and mortality. Nevertheless, the quality of evidence was generally low. Findings from the comparisons of effect sizes across individual factors affecting the occurrence of adverse events were inconclusive, due to the imprecision in the direction of adverse effects and heterogeneity in the baseline patient and pharmacological characteristics. This review should not be interpretated as a dispute over the use of antipsychotics but a call for higher quality research following methodological guidelines such as the AMSTAR 2 or PRISMA Harms and addressing the risk of bias and inconsistent reporting of data. Until more active efforts are invested into improving

the quality of primary studies, only tentative conclusions can be made on the adverse effects of antipsychotics.

References

Altman, D. G., & Bland, J. M. (2011). How to obtain the P value from a confidence interval. *BMJ (Clinical research ed.)*, *343*, d2304. https://doi.org/10.1136/bmj.d2304

Amrhein, V., Greenland, S., & McShane, B. (2019). Scientists rise up against statistical significance. *Nature*, *567*(7748), 305–307. https://doi.org/10.1038/d41586-019-00857-9

Aromataris, E., Fernandez, R., Godfrey, C. M., Holly, C., Khalil, H., & Tungpunkom, P. (2015). Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. *International journal of evidence-based healthcare*, *13*(3), 132–140. https://doi.org/10.1097/XEB.0000000000000055

Aronson, J. K., & Ferner, R. E. (2003). Joining the DoTS: new approach to classifying adverse drug reactions. *BMJ (Clinical research ed.)*, *327*(7425), 1222–1225. https://doi.org/10.1136/bmj.327.7425.1222

Institute of Medicine (US) Committee on Data Standards for Patient Safety, Aspden, P., Corrigan, J. M., Wolcott, J., & Erickson, S. M. (Eds.). (2004). *Patient Safety: Achieving a New Standard for Care*. National Academies Press (US).

Ayorinde, A. A., Williams, I., Mannion, R., Song, F., Skrybant, M., Lilford, R. J., & Chen, Y. F. (2020). Assessment of publication bias and outcome reporting bias in systematic reviews of health services and delivery research: A meta-epidemiological study. *PloS one*, *15*(1), e0227580. https://doi.org/10.1371/journal.pone.0227580

Bai, Y., Yang, H., Chen, G., & Gao, K. (2020). Acceptability of Acute and Maintenance Pharmacotherapy of Bipolar Disorder: A Systematic Review of Randomized, Double-Blind, Placebo-Controlled Clinical Trials. *Journal of clinical psychopharmacology*, *40*(2), 167–179. https://doi.org/10.1097/JCP.0000000000001169

Balasubramanian, H., Ananthan, A., Rao, S., & Patole, S. (2015). Odds ratio vs risk ratio in randomized controlled trials. *Postgraduate medicine*, *127*(4), 359–367. https://doi.org/10.1080/00325481.2015.1022494

Barton S. (2000). Which clinical studies provide the best evidence? The best RCT still trumps the best observational study. *BMJ (Clinical research ed.)*, *321*(7256), 255–256. https://doi.org/10.1136/bmj.321.7256.255

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, *50*(4), 1088–1101.

Bentler, S. E., Liu, L., Obrizan, M., Cook, E. A., Wright, K. B., Geweke, J. F., Chrischilles, E. A., Pavlik, C. E., Wallace, R. B., Ohsfeldt, R. L., Jones, M. P., Rosenthal, G. E., & Wolinsky, F. D. (2009). The aftermath of hip fracture: discharge placement, functional status change, and mortality. *American journal of epidemiology*, *170*(10), 1290–1299. https://doi.org/10.1093/aje/kwp266

Bishara, D., & Taylor, D. (2008). Upcoming agents for the treatment of schizophrenia: mechanism of action, efficacy and tolerability. *Drugs*, *68*(16), 2269–2292. https://doi.org/10.2165/0003495-200868160-00002

Brookes, S. T., Whitely, E., Egger, M., Smith, G. D., Mulheran, P. A., & Peters, T. J. (2004). Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. *Journal of clinical epidemiology*, *57*(3), 229–236. https://doi.org/10.1016/j.jclinepi.2003.08.009

Burke, J. F., Sussman, J. B., Kent, D. M., & Hayward, R. A. (2015). Three simple rules to ensure reasonably credible subgroup analyses. *BMJ (Clinical research ed.)*, *351*, h5651. https://doi.org/10.1136/bmj.h5651

Calderón-Ospina, C., & Bustamante-Rojas, C. (2010). The DoTS classification is a useful way to classify adverse drug reactions: a preliminary study in hospitalized patients. *The*

International journal of pharmacy practice, *18*(4), 230–235. https://doi.org/10.1111/j.2042-7174.2010.00039.x

Chawla, S., & Kumar, S. (2017). Adverse Drug Reactions and their Impact on Quality of Life in Patients on Antipsychotic Therapy at a Tertiary Care Center in Delhi. *Indian journal of psychological medicine*, *39*(3), 293–298. https://doi.org/10.4103/0253-7176.207332

Chen, P. H., Tsai, S. Y., Pan, C. H., Chang, C. K., Su, S. S., Chen, C. C., & Kuo, C. J. (2019). Mood stabilisers and risk of stroke in bipolar disorder. *The British journal of psychiatry : the journal of mental science*, *215*(1), 409–414. https://doi.org/10.1192/bjp.2018.203

Chokhawala, K., & Stevens, L. (2020). Antipsychotic medications. *StatPearls [Internet]*.

Claassen, J. N., & Park, J. S. (2021). Examining the dispensing patterns of antipsychotics in Australia from 2006 to 2018 - A pharmacoepidemiology study. *Research in social & administrative pharmacy : RSAP*, *17*(6), 1159–1165. https://doi.org/10.1016/j.sapharm.2020.09.001

Coggon, D., Barker, D., & Rose, G. (2009). *Epidemiology for the Uninitiated*. John Wiley & Sons.

Correll, C. U., Detraux, J., De Lepeleire, J., & De Hert, M. (2015). Effects of antipsychotics, antidepressants and mood stabilizers on risk for physical diseases in people with schizophrenia, depression and bipolar disorder. *World psychiatry : official journal of the World Psychiatric Association (WPA)*, *14*(2), 119–136. https://doi.org/10.1002/wps.20204

Correll, C. U., Rubio, J. M., Inczedy-Farkas, G., Birnbaum, M. L., Kane, J. M., & Leucht, S. (2017). Efficacy of 42 Pharmacologic Cotreatment Strategies Added to Antipsychotic Monotherapy in Schizophrenia: Systematic Overview and Quality Appraisal of the Meta-analytic Evidence. *JAMA psychiatry*, *74*(7), 675–684. https://doi.org/10.1001/jamapsychiatry.2017.0624

Curran-Everett D. (2017). Explorations in statistics: the assumption of normality. *Advances in physiology education*, *41*(3), 449–453. https://doi.org/10.1152/advan.00064.2017

De Fruyt, J., Deschepper, E., Audenaert, K., Constant, E., Floris, M., Pitchot, W., Sienaert, P., Souery, D., & Claes, S. (2012). Second generation antipsychotics in the treatment of bipolar depression: a systematic review and meta-analysis. *Journal of psychopharmacology (Oxford, England)*, *26*(5), 603–617. https://doi.org/10.1177/0269881111408461

De Hert, M., Detraux, J., van Winkel, R., Yu, W., & Correll, C. U. (2011). Metabolic and cardiovascular adverse effects associated with antipsychotic drugs. *Nature reviews. Endocrinology*, *8*(2), 114–126. https://doi.org/10.1038/nrendo.2011.156

Debray, T., Moons, K., & Riley, R. D. (2018). Detecting small-study effects and funnel plot asymmetry in meta-analysis of survival data: A comparison of new and existing tests. *Research synthesis methods*, *9*(1), 41–50. https://doi.org/10.1002/jrsm.1266

Demyttenaere, K., Detraux, J., Racagni, G., & Vansteelandt, K. (2019). Medication-Induced Akathisia with Newly Approved Antipsychotics in Patients with a Severe Mental Illness: A Systematic Review and Meta-Analysis. *CNS drugs*, *33*(6), 549–566. https://doi.org/10.1007/s40263-019-00625-3

Derry, S., & Moore, R. A. (2007). Atypical antipsychotics in bipolar disorder: systematic review of randomised trials. *BMC psychiatry*, *7*, 40. https://doi.org/10.1186/1471-244X-7-40

Dibonaventura, M., Gabriel, S., Dupclay, L., Gupta, S., & Kim, E. (2012). A patient perspective of the impact of medication side effects on adherence: results of a cross-sectional nationwide survey of patients with schizophrenia. *BMC psychiatry*, *12*, 20. https://doi.org/10.1186/1471-244X-12-20

Divac, N., Prostran, M., Jakovcevski, I., & Cerovac, N. (2014). Second-generation antipsychotics and extrapyramidal adverse effects. *BioMed research international*, *2014*, 656370. https://doi.org/10.1155/2014/656370

Dragioti, E., Solmi, M., Favaro, A., Fusar-Poli, P., Dazzan, P., Thompson, T., Stubbs, B., Firth, J., Fornaro, M., Tsartsalis, D., Carvalho, A. F., Vieta, E., McGuire, P., Young, A. H., Shin, J. I.,

Correll, C. U., & Evangelou, E. (2019). Association of Antidepressant Use With Adverse Health Outcomes: A Systematic Umbrella Review. *JAMA psychiatry*, *76*(12), 1241–1255. https://doi.org/10.1001/jamapsychiatry.2019.2859

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ (Clinical research ed.)*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Evans, S., Anderson, J. M., Johnson, A. L., Checketts, J. X., Scott, J., Middlemist, K., Fishbeck, K., & Vassar, M. (2021). The Potential Effect of Lowering the Threshold of Statistical Significance From P < .05 to P < .005 in Orthopaedic Sports Medicine. *Arthroscopy : the journal of arthroscopic & related surgery : official publication of the Arthroscopy Association of North America and the International Arthroscopy Association*, *37*(4), 1068–1074. https://doi.org/10.1016/j.arthro.2020.11.041

Feng, Z., Hirdes, J. P., Smith, T. F., Finne-Soveri, H., Chi, I., Du Pasquier, J. N., Gilgen, R., Ikegami, N., & Mor, V. (2009). Use of physical restraints and antipsychotic medications in nursing homes: a cross-national study. *International journal of geriatric psychiatry*, *24*(10), 1110–1118. https://doi.org/10.1002/gps.2232

Fleurence, R., Williamson, R., Jing, Y., Kim, E., Tran, Q. V., Pikalov, A. S., & Thase, M. E. (2009). A systematic review of augmentation strategies for patients with major depressive disorder. *Psychopharmacology bulletin*, *42*(3), 57–90.

Fusar-Poli, P., & Radua, J. (2018). Ten simple rules for conducting umbrella reviews. *Evidence-based mental health*, *21*(3), 95–100. https://doi.org/10.1136/ebmental-2018-300014

García, S., Martínez-Cengotitabengoa, M., López-Zurbano, S., Zorrilla, I., López, P., Vieta, E., & González-Pinto, A. (2016). Adherence to Antipsychotic Medication in Bipolar Disorder and Schizophrenic Patients: A Systematic Review. *Journal of clinical psychopharmacology*, *36*(4), 355–371. https://doi.org/10.1097/JCP.0000000000000523

George, A., Stead, T. S., & Ganti, L. (2020). What's the Risk: Differentiating Risk Ratios, Odds Ratios, and Hazard Ratios?. *Cureus*, *12*(8), e10047. https://doi.org/10.7759/cureus.10047

Golder, S., Loke, Y. K., & Bland, M. (2011). Meta-analyses of adverse effects data derived from randomised controlled trials as compared to observational studies: methodological overview. *PLoS medicine*, *8*(5), e1001026. https://doi.org/10.1371/journal.pmed.1001026

Golder, S., Peryer, G., & Loke, Y. K. (2019). Overview: comprehensive and carefully constructed strategies are required when conducting searches for adverse effects data. *Journal of clinical epidemiology*, *113*, 36–43. https://doi.org/10.1016/j.jclinepi.2019.05.019

Graff-Guerrero, A., Rajji, T. K., Mulsant, B. H., Nakajima, S., Caravaggio, F., Suzuki, T., Uchida, H., Gerretsen, P., Mar, W., Pollock, B. G., & Mamo, D. C. (2015). Evaluation of Antipsychotic Dose Reduction in Late-Life Schizophrenia: A Prospective Dopamine D2/3 Receptor Occupancy Study. *JAMA psychiatry*, *72*(9), 927–934. https://doi.org/10.1001/jamapsychiatry.2015.0891

Grant, M. J., & Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. *Health information and libraries journal*, *26*(2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Habibi, M. (2010). Dopamine Receptors. In K. Kompoliti, & L. Verhagen (Eds.), *Encyclopedia of movement disorders (Vol. 1)*. Academic Press.

Haddad, P. M., Brain, C., & Scott, J. (2014). Nonadherence with antipsychotic medication in schizophrenia: challenges and management strategies. *Patient related outcome measures*, *5*, 43–62. https://doi.org/10.2147/PROM.S42735

Hägg, S., Jönsson, A. K., & Spigset, O. (2009). Risk of venous thromboembolism due to antipsychotic drug therapy. *Expert opinion on drug safety*, *8*(5), 537–547. https://doi.org/10.1517/14740330903117271

Hailes, H. P., Yu, R., Danese, A., & Fazel, S. (2019). Long-term outcomes of childhood sexual abuse: an umbrella review. *The lancet. Psychiatry*, *6*(10), 830–839. https://doi.org/10.1016/S2215-0366(19)30286-X

Harbord, R. M., Egger, M., & Sterne, J. A. (2006). A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Statistics in medicine*, *25*(20), 3443–3457. https://doi.org/10.1002/sim.2380

Harrington, C. A., & English, C. (2010). Tolerability of paliperidone: a meta-analysis of randomized, controlled trials. *International clinical psychopharmacology*, *25*(6), 334–341. https://doi.org/10.1097/YIC.0b013e32833db3d8

Hay, A., Byers, A., Sereno, M., Basra, M. K., & Dutta, S. (2015). Asenapine versus placebo for schizophrenia. *The Cochrane database of systematic reviews*, *2015*(11), CD011458. https://doi.org/10.1002/14651858.CD011458.pub2

He, Y., Li, X., Gasevic, D., Brunt, E., McLachlan, F., Millenson, M., Timofeeva, M., Ioannidis, J., Campbell, H., & Theodoratou, E. (2018). Statins and Multiple Noncardiovascular Outcomes: Umbrella Review of Meta-analyses of Observational Studies and Randomized Controlled Trials. *Annals of internal medicine*, *169*(8), 543–553. https://doi.org/10.7326/M18-0808

Hernán M. A. (2010). The hazards of hazard ratios. *Epidemiology (Cambridge, Mass.)*, *21*(1), 13–15. https://doi.org/10.1097/EDE.0b013e3181c1ea43

Higashi, K., Medic, G., Littlewood, K. J., Diez, T., Granström, O., & De Hert, M. (2013). Medication adherence in schizophrenia: factors influencing adherence and consequences of nonadherence, a systematic literature review. *Therapeutic advances in psychopharmacology*, *3*(4), 200–218. https://doi.org/10.1177/2045125312474019

Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2019). *Cochrane handbook for systematic reviews of interventions*. John Wiley & Sons.

Higgins, J. P., Thompson, S. G., & Spiegelhalter, D. J. (2009). A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society. Series A, (Statistics in Society)*, *172*(1), 137–159. https://doi.org/10.1111/j.1467-985X.2008.00552.x

Hirsch, L., Patten, S. B., Bresee, L., Jette, N., & Pringsheim, T. (2018). Second-generation antipsychotics and metabolic side-effects: Canadian population-based study. *BJPsych open*, *4*(4), 256–261. https://doi.org/10.1192/bjo.2018.33

Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I2 index?. *Psychological methods*, *11*(2), 193–206. https://doi.org/10.1037/1082-989X.11.2.193

Huhn, M., Nikolakopoulou, A., Schneider-Thoma, J., Krause, M., Samara, M., Peter, N., Arndt, T., Bäckers, L., Rothe, P., Cipriani, A., Davis, J., Salanti, G., & Leucht, S. (2019). Comparative efficacy and tolerability of 32 oral antipsychotics for the acute treatment of adults with multi-episode schizophrenia: a systematic review and network meta-analysis. *Lancet (London, England)*, *394*(10202), 939–951. https://doi.org/10.1016/S0140-6736(19)31135-3

IntHout, J., Ioannidis, J. P., Rovers, M. M., & Goeman, J. J. (2016). Plea for routinely presenting prediction intervals in meta-analysis. *BMJ open*, *6*(7), e010247. https://doi.org/10.1136/bmjopen-2015-010247

Ioannidis, J. P., & Trikalinos, T. A. (2007). An exploratory test for an excess of significant findings. *Clinical trials (London, England)*, *4*(3), 245–253. https://doi.org/10.1177/1740774507079441

Ioannidis, J. P., Patsopoulos, N. A., & Evangelou, E. (2007). Uncertainty in heterogeneity estimates in meta-analyses. *BMJ (Clinical research ed.)*, *335*(7626), 914–916. https://doi.org/10.1136/bmj.39343.408449.80

John, A., McGregor, J., Jones, I., Lee, S. C., Walters, J., Owen, M. J., O'Donovan, M., DelPozo-Banos, M., Berridge, D., & Lloyd, K. (2018). Premature mortality among people with severe

mental illness - New evidence from linked primary care data. *Schizophrenia research*, *199*, 154–162. https://doi.org/10.1016/j.schres.2018.04.009

Jönsson, A. K., Schill, J., Olsson, H., Spigset, O., & Hägg, S. (2018). Venous Thromboembolism During Treatment with Antipsychotics: A Review of Current Evidence. *CNS drugs*, *32*(1), 47–64. https://doi.org/10.1007/s40263-018-0495-7

Kapur, S., & Seeman, P. (2001). Does fast dissociation from the dopamine d(2) receptor explain the action of atypical antipsychotics?: A new hypothesis. *The American journal of psychiatry*, *158*(3), 360–369. https://doi.org/10.1176/appi.ajp.158.3.360

Kate, N. S., Pawar, S. S., Parkar, S. R., & Sawant, N. S. (2015). Adverse drug reactions due to antipsychotics and sedative-hypnotics in the elderly. *Journal of Geriatric Mental Health*, *2*(1), 16.

Kavvoura, F. K., McQueen, M. B., Khoury, M. J., Tanzi, R. E., Bertram, L., & Ioannidis, J. P. (2008). Evaluation of the potential excess of statistically significant findings in published genetic association studies: application to Alzheimer's disease. *American journal of epidemiology*, *168*(8), 855–865. https://doi.org/10.1093/aje/kwn206

Keeney, B. J., Fulton-Kehoe, D., Wickizer, T. M., Turner, J. A., Chan, K. C., & Franklin, G. M. (2013). Clinically significant weight gain 1 year after occupational back injury. *Journal of occupational and environmental medicine*, *55*(3), 318–324. https://doi.org/10.1097/JOM.0b013e31827943c6

Kim, S. F., Huang, A. S., Snowman, A. M., Teuscher, C., & Snyder, S. H. (2007). From the Cover: Antipsychotic drug-induced weight gain mediated by histamine H1 receptor-linked activation of hypothalamic AMP-kinase. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(9), 3456–3459. https://doi.org/10.1073/pnas.0611417104

Kimachi, M., Onishi, A., Tajika, A., Kimachi, K., & Furukawa, T. A. (2021). Systematic differences in effect estimates between observational studies and randomized control trials in meta-analyses in nephrology. *Scientific reports*, *11*(1), 6088. https://doi.org/10.1038/s41598-021-85519-5

Klok, M., van Eijndhoven, P. F., Argyelan, M., Schene, A. H., & Tendolkar, I. (2019). Structural brain characteristics in treatment-resistant depression: review of magnetic resonance imaging studies. *BJPsych open*, *5*(5), e76. https://doi.org/10.1192/bjo.2019.58

Knol, W., van Marum, R. J., Jansen, P. A., Souverein, P. C., Schobben, A. F., & Egberts, A. C. (2008). Antipsychotic drug use and risk of pneumonia in elderly people. *Journal of the American Geriatrics Society*, *56*(4), 661–666. https://doi.org/10.1111/j.1532-5415.2007.01625.x

Kolotkin, R. L., Norquist, J. M., Crosby, R. D., Suryawanshi, S., Teixeira, P. J., Heymsfield, S. B., Erondu, N., & Nguyen, A. M. (2009). One-year health-related quality of life outcomes in weight loss trial participants: comparison of three measures. *Health and quality of life outcomes*, *7*, 53. https://doi.org/10.1186/1477-7525-7-53

Kravitz, R. L., Duan, N., & Braslow, J. (2004). Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. *The Milbank quarterly*, *82*(4), 661–687. https://doi.org/10.1111/j.0887-378X.2004.00327.x

Kusumi, I., Boku, S., & Takahashi, Y. (2015). Psychopharmacology of atypical antipsychotic drugs: From the receptor binding profile to neuroprotection and neurogenesis. *Psychiatry and clinical neurosciences*, *69*(5), 243–258. https://doi.org/10.1111/pcn.12242

Lan, C. C., Liu, C. C., Lin, C. H., Lan, T. Y., McInnis, M. G., Chan, C. H., & Lan, T. H. (2015). A reduced risk of stroke with lithium exposure in bipolar disorder: a population-based retrospective cohort study. *Bipolar disorders*, *17*(7), 705–714. https://doi.org/10.1111/bdi.12336

Lao, K. S., He, Y., Wong, I. C., Besag, F. M., & Chan, E. W. (2016). Tolerability and Safety Profile of Cariprazine in Treating Psychotic Disorders, Bipolar Disorder and Major Depressive

Disorder: A Systematic Review with Meta-Analysis of Randomized Controlled Trials. *CNS drugs*, *30*(11), 1043–1054. https://doi.org/10.1007/s40263-016-0382-z

Leppien, E., Bystrak, T., & Doughty, B. (2019). Antipsychotic agents. *Side Effects of Drugs Annual*, *41*, 51-63.

Leucht, S., Arbter, D., Engel, R. R., Kissling, W., & Davis, J. M. (2009). How effective are second-generation antipsychotic drugs? A meta-analysis of placebo-controlled trials. *Molecular psychiatry*, *14*(4), 429–447. https://doi.org/10.1038/sj.mp.4002136

Leucht, S., Crippa, A., Siafis, S., Patel, M. X., Orsini, N., & Davis, J. M. (2020). Dose-Response Meta-Analysis of Antipsychotic Drugs for Acute Schizophrenia. *The American journal of psychiatry*, *177*(4), 342–353. https://doi.org/10.1176/appi.ajp.2019.19010034

Li, P., Snyder, G. L., & Vanover, K. E. (2016). Dopamine Targeting Drugs for the Treatment of Schizophrenia: Past, Present and Future. *Current topics in medicinal chemistry*, *16*(29), 3385–3403. https://doi.org/10.2174/1568026616666160608084834

Lin, L., & Chu, H. (2018). Quantifying publication bias in meta-analysis. *Biometrics*, *74*(3), 785–794. https://doi.org/10.1111/biom.12817

Lind, C. C., Carchedi, C. L., Staudenmeier, L. J., & Diebold, L. P. (2005). Atypical presentations of atypical antipsychotics. *Psychiatry (Edgmont (Pa. : Township))*, *2*(6), 32–39.

Liperoti, R., Onder, G., Landi, F., Lapane, K. L., Mor, V., Bernabei, R., & Gambassi, G. (2009). All-cause mortality associated with atypical and conventional antipsychotics among nursing home residents with dementia: a retrospective cohort study. *The Journal of clinical psychiatry*, *70*(10), 1340–1347. https://doi.org/10.4088/JCP.08m04597yel

Liu, Y., Xu, J., Fang, K., Xu, Y., Gao, J., Zhou, C., Tang, X., Fang, X., Chen, J., Xie, C., Zhang, F., Zhang, X., & Wang, C. (2021). Current antipsychotic agent use and risk of venous thromboembolism and pulmonary embolism: a systematic review and meta-analysis of observational studies. *Therapeutic advances in psychopharmacology*, *11*, 2045125320982720. https://doi.org/10.1177/2045125320982720

Loke, Y. K., Price, D., Herxheimer, A., & Cochrane Adverse Effects Methods Group (2007). Systematic reviews of adverse effects: framework for a structured approach. *BMC medical research methodology*, *7*, 32. https://doi.org/10.1186/1471-2288-7-32

Lorenz, R. C., Matthias, K., Pieper, D., Wegewitz, U., Morche, J., Nocon, M., Rissling, O., Schirm, J., & Jacobs, A. (2019). A psychometric study found AMSTAR 2 to be a valid and moderately reliable appraisal tool. *Journal of clinical epidemiology*, *114*, 133–140. https://doi.org/10.1016/j.jclinepi.2019.05.028

Luan, S., Wan, H., Wang, S., Li, H., & Zhang, B. (2017). Efficacy and safety of olanzapine/fluoxetine combination in the treatment of treatment-resistant depression: a meta-analysis of randomized controlled trials. *Neuropsychiatric disease and treatment*, *13*, 609–620. https://doi.org/10.2147/NDT.S127453

Luan, S., Wan, H., Zhang, L., & Zhao, H. (2018). Efficacy, acceptability, and safety of adjunctive aripiprazole in treatment-resistant depression: a meta-analysis of randomized controlled trials. *Neuropsychiatric disease and treatment*, *14*, 467–477. https://doi.org/10.2147/NDT.S156619

Ma, H., Huang, Y., Cong, Z., Wang, Y., Jiang, W., Gao, S., & Zhu, G. (2014). The efficacy and safety of atypical antipsychotics for the treatment of dementia: a meta-analysis of randomized placebo-controlled trials. *Journal of Alzheimer's disease : JAD*, *42*(3), 915–937. https://doi.org/10.3233/JAD-140579

Maglione, M., Maher, A. R., Hu, J., Wang, Z., Shanman, R., Shekelle, P. G., Roth, B., Hilton, L., Suttorp, M. J., Ewing, B. A., Motala, A., & Perry, T. (2011). *Off-Label Use of Atypical Antipsychotics: An Update*. Agency for Healthcare Research and Quality (US).

Maher, A. R., Maglione, M., Bagley, S., Suttorp, M., Hu, J. H., Ewing, B., Wang, Z., Timmer, M., Sultzer, D., & Shekelle, P. G. (2011). Efficacy and comparative effectiveness of atypical

antipsychotic medications for off-label uses in adults: a systematic review and meta-analysis. *JAMA*, *306*(12), 1359–1369. https://doi.org/10.1001/jama.2011.1360

Mangoni, A. A., & Jackson, S. H. (2004). Age-related changes in pharmacokinetics and pharmacodynamics: basic principles and practical applications. *British journal of clinical pharmacology*, *57*(1), 6–14. https://doi.org/10.1046/j.1365-2125.2003.02007.x

Marston, L., Nazareth, I., Petersen, I., Walters, K., & Osborn, D. P. (2014). Prescribing of antipsychotics in UK primary care: a cohort study. *BMJ open*, *4*(12), e006135. https://doi.org/10.1136/bmjopen-2014-006135

Martino, D., Karnik, V., Osland, S., Barnes, T., & Pringsheim, T. M. (2018). Movement Disorders Associated With Antipsychotic Medication in People With Schizophrenia: An Overview of Cochrane Reviews and Meta-Analysis. *Canadian journal of psychiatry. Revue canadienne de psychiatrie*, *63*(11), 706743718777392. Advance online publication. https://doi.org/10.1177/0706743718777392

Medicines and Healthcare products Regulatory Agency. (2005, August 25). *Antipsychotics medicines*. https://www.gov.uk/government/publications/antipsychotic-medicines-licensed-products-uses-and-side-effects/antipsychotic-medicines.

Medicines and Healthcare products Regulatory Agency. (n.d.). *Guidance on adverse drug reactions*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/949130/Guidance_on_adverse_drug_reactions.pdf.

Meltzer, H. Y., Matsubara, S., & Lee, J. C. (1989). Classification of typical and atypical antipsychotic drugs on the basis of dopamine D-1, D-2 and serotonin2 pKi values. *The Journal of pharmacology and experimental therapeutics*, *251*(1), 238–246.

Mihanović, M., Restek-Petrović, B., Bodor, D., Molnar, S., Oresković, A., & Presecki, P. (2010). Suicidality and side effects of antidepressants and antipsychotics. *Psychiatria Danubina*, *22*(1), 79–84.

Miyamoto, S., Miyake, N., Jarskog, L. F., Fleischhacker, W. W., & Lieberman, J. A. (2012). Pharmacological treatment of schizophrenia: a critical review of the pharmacology and clinical effects of current and future therapeutic agents. *Molecular psychiatry*, *17*(12), 1206–1227. https://doi.org/10.1038/mp.2012.47

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Mortensen, S. J., Mohamadi, A., Wright, C. L., Chan, J. J., Weaver, M. J., von Keudell, A., & Nazarian, A. (2020). Medications as a Risk Factor for Fragility Hip Fractures: A Systematic Review and Meta-analysis. *Calcified tissue international*, *107*(1), 1–9. https://doi.org/10.1007/s00223-020-00688-1

Muench, J., & Hamer, A. M. (2010). Adverse effects of antipsychotic medications. *American family physician*, *81*(5), 617–622.

Nasrallah H. A. (2007). The roles of efficacy, safety, and tolerability in antipsychotic effectiveness: practical implications of the CATIE schizophrenia trial. *The Journal of clinical psychiatry*, *68 Suppl 1*, 5–11.

Nasrallah H. A. (2008). Atypical antipsychotic-induced metabolic side effects: insights from receptor-binding profiles. *Molecular psychiatry*, *13*(1), 27–35. https://doi.org/10.1038/sj.mp.4002066

Nebeker, J. R., Barach, P., & Samore, M. H. (2004). Clarifying adverse drug events: a clinician's guide to terminology, documentation, and reporting. *Annals of internal medicine*, *140*(10), 795–801. https://doi.org/10.7326/0003-4819-140-10-200405180-00009

Nelson, J. C., & Papakostas, G. I. (2009). Atypical antipsychotic augmentation in major depressive disorder: a meta-analysis of placebo-controlled randomized trials. *The American journal of psychiatry*, *166*(9), 980–991. https://doi.org/10.1176/appi.ajp.2009.09030312

NHS. (2020, December 9). *Medicines Used in Mental Health – England – 2015/16 to 2019/20*. https://nhsbsa-opendata.s3-eu-west-2.amazonaws.com/mh-annual-narrative-final.html.

NICE. (2009) *Clinical guideline 90: Depression in adults. The treatment and management of depression in adults*. National Institute for Health and Care Excellence (UK).

NIH National Cancer Institute. (2017, November 27). *Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0*. https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/ctcae_v5_quick_reference_5x7.pdf.

Noble, S. (2022) Symptomatic and Supportive Care. In S. M. Janes (2nd Ed.), *Encyclopedia of Respiratory Medicine* (pp. 921-930). Academic Press.

Nosè, M., Recla, E., Trifirò, G., & Barbui, C. (2015). Antipsychotic drug exposure and risk of pneumonia: a systematic review and meta-analysis of observational studies. *Pharmacoepidemiology and drug safety*, *24*(8), 812–820. https://doi.org/10.1002/pds.3804

Nussbaum, A. M., & Stroup, T. S. (2012). Paliperidone palmitate for schizophrenia. *Schizophrenia bulletin*, *38*(6), 1124–1127. https://doi.org/10.1093/schbul/sbs099

Nussbaum, A., & Stroup, T. S. (2008). Paliperidone for schizophrenia. *The Cochrane database of systematic reviews*, (2), CD006369. https://doi.org/10.1002/14651858.CD006369.pub2

Papageorgiou, S. N., & Biondi-Zoccai, G. (2016). Designing the review. In *Umbrella Reviews* (pp. 57-80). Springer.

Papatheodorou, S. (2019). Umbrella reviews: what they are and why we need them. *European journal of epidemiology*, *34*(6), 543–546. https://doi.org/10.1007/s10654-019-00505-6

Papola, D., Ostuzzi, G., Gastaldon, C., Morgano, G. P., Dragioti, E., Carvalho, A. F., Fusar-Poli, P., Correll, C. U., Solmi, M., & Barbui, C. (2019). Antipsychotic use and risk of life-threatening medical events: umbrella review of observational studies. *Acta psychiatrica Scandinavica*, *140*(3), 227–243. https://doi.org/10.1111/acps.13066

Papola, D., Ostuzzi, G., Thabane, L., Guyatt, G., & Barbui, C. (2018). Antipsychotic drug exposure and risk of fracture: a systematic review and meta-analysis of observational studies. *International clinical psychopharmacology*, *33*(4), 181–196. https://doi.org/10.1097/YIC.0000000000000221

Pariente, A., Sanctussy, D. J., Miremont-Salamé, G., Moore, N., Haramburu, F., Fourrier-Réglat, A., & l'Association Française des Centres Régionaux de Pharmacovigilance (CRPV) (2010). Factors associated with serious adverse reactions to cholinesterase inhibitors: a study of spontaneous reporting. *CNS drugs*, *24*(1), 55–63. https://doi.org/10.2165/11530300-000000000-00000

Patil, P., & Schwartz, T. L. (2018). Fine tuning the use of second generation antipsychotics. *Journal of Mental Health & Clinical Psychology*, *2*(5), 22–39. https://doi.org/10.29245/2578-2959/2018/5.1138

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2006). Comparison of two methods to detect publication bias in meta-analysis. *JAMA*, *295*(6), 676–680. https://doi.org/10.1001/jama.295.6.676

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2007). Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity. *Statistics in medicine*, *26*(25), 4544–4562. https://doi.org/10.1002/sim.2889

Phillips, R., Hazell, L., Sauzet, O., & Cornelius, V. (2019). Analysis and reporting of adverse events in randomised controlled trials: a review. *BMJ open*, *9*(2), e024537. https://doi.org/10.1136/bmjopen-2018-024537

Pillay, J., Boylan, K., Newton, A., Hartling, L., Vandermeer, B., Nuspl, M., MacGregor, T., Featherstone, R., & Carrey, N. (2018). Harms of Antipsychotics in Children and Young

Adults: A Systematic Review Update. *Canadian journal of psychiatry. Revue canadienne de psychiatrie*, *63*(10), 661–678. https://doi.org/10.1177/0706743718779950

Polcwiartek, C., Sneider, B., Graff, C., Taylor, D., Meyer, J., Kanters, J. K., & Nielsen, J. (2015). The cardiac safety of aripiprazole treatment in patients at high risk for torsade: a systematic review with a meta-analytic approach. *Psychopharmacology*, *232*(18), 3297–3308. https://doi.org/10.1007/s00213-015-4024-9

Pomares, F. B., Boucetta, S., Lachapelle, F., Steffener, J., Montplaisir, J., Cha, J., Kim, H., & Dang-Vu, T. T. (2019). Beyond sleepy: structural and functional changes of the default-mode network in idiopathic hypersomnia. *Sleep*, *42*(11), zsz156. https://doi.org/10.1093/sleep/zsz156

Pope, A., Adams, C., Paton, C., Weaver, T., & Barnes, T. R. (2010). Assessment of adverse effects in clinical studies of antipsychotic medication: survey of methods used. *The British journal of psychiatry : the journal of mental science*, *197*(1), 67–72. https://doi.org/10.1192/bjp.bp.109.070961

Pretorius, R. W., Gataric, G., Swedlund, S. K., & Miller, J. R. (2013). Reducing the risk of adverse drug events in older adults. *American family physician*, *87*(5), 331–336.

Ranstam J. (2012). Why the P-value culture is bad and confidence intervals a better alternative. *Osteoarthritis and cartilage*, *20*(8), 805–808. https://doi.org/10.1016/j.joca.2012.04.001

Richardson, M., Garner, P., & Donegan, S. (2019). Interpretation of subgroup analyses in systematic reviews: a tutorial. *Clinical Epidemiology and Global Health*, *7*(2), 192-198.

Riley, R. D., Higgins, J. P., & Deeks, J. J. (2011). Interpretation of random effects meta-analyses. *BMJ (Clinical research ed.)*, *342*, d549. https://doi.org/10.1136/bmj.d549

Ringen, P. A., Engh, J. A., Birkenaes, A. B., Dieset, I., & Andreassen, O. A. (2014). Increased mortality in schizophrenia due to cardiovascular disease - a non-systematic review of epidemiology, possible causes, and interventions. *Frontiers in psychiatry*, *5*, 137. https://doi.org/10.3389/fpsyt.2014.00137

Rojo, L. E., Gaspar, P. A., Silva, H., Risco, L., Arena, P., Cubillos-Robles, K., & Jara, B. (2015). Metabolic syndrome and obesity among users of second generation antipsychotics: A global challenge for modern psychopharmacology. *Pharmacological research*, *101*, 74–85. https://doi.org/10.1016/j.phrs.2015.07.022

Romeo, B., Blecha, L., Locatelli, K., Benyamina, A., & Martelli, C. (2018). Meta-analysis and review of dopamine agonists in acute episodes of mood disorder: Efficacy and safety. *Journal of Psychopharmacology*, *32*(4), 385–396. http://dx.doi.org/10.1177/0269881118760661

Rotella, F., Cassioli, E., Calderani, E., Lazzeretti, L., Ragghianti, B., Ricca, V., & Mannucci, E. (2020). Long-term metabolic and cardiovascular effects of antipsychotic drugs. A meta-analysis of randomized controlled trials. *European Neuropsychopharmacology*, *32*, 56–65. http://dx.doi.org/10.1016/j.euroneuro.2019.12.118

Roth, B. L., Hanizavareh, S. M., & Blum, A. E. (2004). Serotonin receptors represent highly favorable molecular targets for cognitive enhancement in schizophrenia and other disorders. *Psychopharmacology*, *174*(1), 17–24. https://doi.org/10.1007/s00213-003-1683-8

Rücker, G., Schwarzer, G., & Carpenter, J. (2008). Arcsine test for publication bias in meta-analyses with binary outcomes. *Statistics in medicine*, *27*(5), 746–763. https://doi.org/10.1002/sim.2971

Rush, A. J., Trivedi, M. H., Wisniewski, S. R., Nierenberg, A. A., Stewart, J. W., Warden, D., Niederehe, G., Thase, M. E., Lavori, P. W., Lebowitz, B. D., McGrath, P. J., Rosenbaum, J. F., Sackeim, H. A., Kupfer, D. J., Luther, J., & Fava, M. (2006). Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D

report. *The American journal of psychiatry*, *163*(11), 1905–1917. https://doi.org/10.1176/ajp.2006.163.11.1905

Scherk, H., Pajonk, F. G., & Leucht, S. (2007). Second-generation antipsychotic agents in the treatment of acute mania: a systematic review and meta-analysis of randomized controlled trials. *Archives of general psychiatry*, *64*(4), 442–455. https://doi.org/10.1001/archpsyc.64.4.442

Schlesinger, S., Schwingshackl, L., Neuenschwander, M., & Barbaresko, J. (2019). A critical reflection on the grading of the certainty of evidence in umbrella reviews. *European journal of epidemiology*, *34*(9), 889–890. https://doi.org/10.1007/s10654-019-00531-4

Schneider-Thoma, J., Efthimiou, O., Bighelli, I., Dörries, C., Huhn, M., Krause, M., Reichelt, L., Röder, H., Furukawa, T. A., Davis, J. M., & Leucht, S. (2019). Second-generation antipsychotic drugs and short-term somatic serious adverse events: a systematic review and meta-analysis. *The lancet. Psychiatry*, *6*(9), 753–765. https://doi.org/10.1016/S2215-0366(19)30223-8

Schneider-Thoma, J., Efthimiou, O., Huhn, M., Krause, M., Reichelt, L., Röder, H., Davis, J. M., Salanti, G., & Leucht, S. (2018). Second-generation antipsychotic drugs and short-term mortality: a systematic review and meta-analysis of placebo-controlled randomised controlled trials. *The lancet. Psychiatry*, *5*(8), 653–663. https://doi.org/10.1016/S2215-0366(18)30177-9

Seppala, L. J., Wermelink, A., de Vries, M., Ploegmakers, K. J., van de Glind, E., Daams, J. G., van der Velde, N., & EUGMS task and Finish group on fall-risk-increasing drugs (2018). Fall-Risk-Increasing Drugs: A Systematic Review and Meta-Analysis: II. Psychotropics. *Journal of the American Medical Directors Association*, *19*(4), 371.e11–371.e17. https://doi.org/10.1016/j.jamda.2017.12.098

Shahid, M., Walker, G. B., Zorn, S. H., & Wong, E. H. (2009). Asenapine: a novel psychopharmacologic agent with a unique human receptor signature. *Journal of psychopharmacology (Oxford, England)*, *23*(1), 65–73. https://doi.org/10.1177/0269881107082944

Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., Moher, D., Tugwell, P., Welch, V., Kristjansson, E., & Henry, D. A. (2017). AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ (Clinical research ed.)*, *358*, j4008. https://doi.org/10.1136/bmj.j4008

Siskind, D., McCartney, L., Goldschlager, R., & Kisely, S. (2016). Clozapine v. first- and second-generation antipsychotics in treatment-refractory schizophrenia: systematic review and meta-analysis. *The British journal of psychiatry : the journal of mental science*, *209*(5), 385–392. https://doi.org/10.1192/bjp.bp.115.177261

Solmi, M., Correll, C. U., Carvalho, A. F., & Ioannidis, J. (2018). The role of meta-analyses and umbrella reviews in assessing the harms of psychotropic medications: beyond qualitative synthesis. *Epidemiology and psychiatric sciences*, *27*(6), 537–542. https://doi.org/10.1017/S204579601800032X

Solmi, M., Radua, J., Stubbs, B., Ricca, V., Moretti, D., Busatta, D., Carvalho, A. F., Dragioti, E., Favaro, A., Monteleone, A. M., Shin, J. I., Fusar-Poli, P., & Castellini, G. (2021). Risk factors for eating disorders: an umbrella review of published meta-analyses. *Revista brasileira de psiquiatria (Sao Paulo, Brazil : 1999)*, *43*(3), 314–323. https://doi.org/10.1590/1516-4446-2020-1099

Spielmans, G. I., Berman, M. I., Linardatos, E., Rosenlicht, N. Z., Perry, A., & Tsai, A. C. (2016). Adjunctive Atypical Antipsychotic Treatment for Major Depressive Disorder: A Meta-Analysis of Depression, Quality of Life, and Safety Outcomes. *Focus (American Psychiatric Publishing)*, *14*(2), 244–265. https://doi.org/10.1176/appi.focus.140202

Spineli, L. M., & Pandis, N. (2020). Statistical heterogeneity: Notion and estimation in meta-analysis. *American journal of orthodontics and dentofacial orthopedics : official publication of the American Association of Orthodontists, its constituent societies, and the American Board of Orthodontics*, *157*(6), 856–859.e2. https://doi.org/10.1016/j.ajodo.2020.03.009

Sterne, J. A., & Egger, M. (2001). Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *Journal of clinical epidemiology*, *54*(10), 1046–1055. https://doi.org/10.1016/s0895-4356(01)00377-8

Stevens, J., Truesdale, K. P., McClain, J. E., & Cai, J. (2006). The definition of weight maintenance. *International journal of obesity (2005)*, *30*(3), 391–399. https://doi.org/10.1038/sj.ijo.0803175

Straus, S. M., Bleumink, G. S., Dieleman, J. P., van der Lei, J., 't Jong, G. W., Kingma, J. H., Sturkenboom, M. C., & Stricker, B. H. (2004). Antipsychotics and the risk of sudden cardiac death. *Archives of internal medicine*, *164*(12), 1293–1297. https://doi.org/10.1001/archinte.164.12.1293

Suzuki T. (2019). Antipsychotic serious adverse events in context. *The lancet. Psychiatry*, *6*(9), 717–718. https://doi.org/10.1016/S2215-0366(19)30274-3

Szucs, D., & Ioannidis, J. P. (2017). Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature. *PLoS biology*, *15*(3), e2000797. https://doi.org/10.1371/journal.pbio.2000797

Szumilas M. (2010). Explaining odds ratios. *Journal of the Canadian Academy of Child and Adolescent Psychiatry = Journal de l'Academie canadienne de psychiatrie de l'enfant et de l'adolescent*, *19*(3), 227–229.

Taylor, R. W., Marwood, L., Oprea, E., DeAngel, V., Mather, S., Valentini, B., Zahn, R., Young, A. H., & Cleare, A. J. (2020). Pharmacological Augmentation in Unipolar Depression: A Guide to the Guidelines. *The international journal of neuropsychopharmacology*, *23*(9), 587–625. https://doi.org/10.1093/ijnp/pyaa033

Teferra, S., Hanlon, C., Beyero, T., Jacobsson, L., & Shibre, T. (2013). Perspectives on reasons for non-adherence to medication in persons with schizophrenia in Ethiopia: a qualitative study of patients, caregivers and health workers. *BMC psychiatry*, *13*, 168. https://doi.org/10.1186/1471-244X-13-168

Tenback, D. E., van Harten, P. N., & van Os, J. (2009). Non-therapeutic risk factors for onset of tardive dyskinesia in schizophrenia: a meta-analysis. *Movement disorders : official journal of the Movement Disorder Society*, *24*(16), 2309–2315. https://doi.org/10.1002/mds.22707

Tenny, S., & Hoffman, M. R. (2021). Odds Ratio. In *StatPearls*. StatPearls Publishing.

Thompson, S. G., & Sharp, S. J. (1999). Explaining heterogeneity in meta-analysis: a comparison of methods. *Statistics in medicine*, *18*(20), 2693–2708. https://doi.org/10.1002/(sici)1097-0258(19991030)18:20<2693::aid-sim235>3.0.co;2-v

Trifirò, G., Gambassi, G., Sen, E. F., Caputi, A. P., Bagnardi, V., Brea, J., & Sturkenboom, M. C. (2010). Association of community-acquired pneumonia with antipsychotic drug use in elderly patients: a nested case-control study. *Annals of internal medicine*, *152*(7), 418–W140. https://doi.org/10.7326/0003-4819-152-7-201004060-00006

Trivedi, M. H., Fava, M., Wisniewski, S. R., Thase, M. E., Quitkin, F., Warden, D., Ritz, L., Nierenberg, A. A., Lebowitz, B. D., Biggs, M. M., Luther, J. F., Shores-Wilson, K., Rush, A. J., & STAR*D Study Team (2006). Medication augmentation after the failure of SSRIs for depression. *The New England journal of medicine*, *354*(12), 1243–1252. https://doi.org/10.1056/NEJMoa052964

Tschoner, A., Engl, J., Laimer, M., Kaser, S., Rettenbacher, M., Fleischhacker, W. W., Patsch, J. R., & Ebenbichler, C. F. (2007). Metabolic side effects of antipsychotic medication. *International journal of clinical practice*, *61*(8), 1356–1370. https://doi.org/10.1111/j.1742-1241.2007.01416.x

Tschoner, A., Engl, J., Rettenbacher, M., Edlinger, M., Kaser, S., Tatarczyk, T., Effenberger, M., Patsch, J. R., Fleischhacker, W. W., & Ebenbichler, C. F. (2009). Effects of six second generation antipsychotics on body weight and metabolism - risk assessment and results from a prospective study. *Pharmacopsychiatry*, *42*(1), 29–34. https://doi.org/10.1055/s-0028-1100425

Tsilidis, K. K., Panagiotou, O. A., Sena, E. S., Aretouli, E., Evangelou, E., Howells, D. W., Al-Shahi Salman, R., Macleod, M. R., & Ioannidis, J. P. (2013). Evaluation of excess significance bias in animal studies of neurological diseases. *PLoS biology*, *11*(7), e1001609. https://doi.org/10.1371/journal.pbio.1001609

Uchida, H., Mamo, D. C., Mulsant, B. H., Pollock, B. G., & Kapur, S. (2009). Increased antipsychotic sensitivity in elderly patients: evidence and mechanisms. *The Journal of clinical psychiatry*, *70*(3), 397–405. https://doi.org/10.4088/jcp.08r04171

Vancampfort, D., Firth, J., Schuch, F., Rosenbaum, S., De Hert, M., Mugisha, J., Probst, M., & Stubbs, B. (2016). Physical activity and sedentary behavior in people with bipolar disorder: A systematic review and meta-analysis. *Journal of affective disorders*, *201*, 145–152. https://doi.org/10.1016/j.jad.2016.05.020

Vancampfort, D., Stubbs, B., Mitchell, A. J., De Hert, M., Wampers, M., Ward, P. B., Rosenbaum, S., & Correll, C. U. (2015). Risk of metabolic syndrome and its components in people with schizophrenia and related psychotic disorders, bipolar disorder and major depressive disorder: a systematic review and meta-analysis. *World psychiatry : official journal of the World Psychiatric Association (WPA)*, *14*(3), 339–347. https://doi.org/10.1002/wps.20252

Vermeulen, J., van Rooijen, G., Doedens, P., Numminen, E., van Tricht, M., & de Haan, L. (2017). Antipsychotic medication and long-term mortality risk in patients with schizophrenia; a systematic review and meta-analysis. *Psychological medicine*, *47*(13), 2217–2228. https://doi.org/10.1017/S0033291717000873

von Hippel P. T. (2015). The heterogeneity statistic I(2) can be biased in small meta-analyses. *BMC medical research methodology*, *15*, 35. https://doi.org/10.1186/s12874-015-0024-z

Wadden, T. A., Neiberg, R. H., Wing, R. R., Clark, J. M., Delahanty, L. M., Hill, J. O., Krakoff, J., Otto, A., Ryan, D. H., Vitolins, M. Z., & Look AHEAD Research Group (2011). Four-year weight losses in the Look AHEAD study: factors associated with long-term success. *Obesity (Silver Spring, Md.)*, *19*(10), 1987–1998. https://doi.org/10.1038/oby.2011.230

Wayant, C., Scott, J., & Vassar, M. (2018). Evaluation of Lowering the P Value Threshold for Statistical Significance From .05 to .005 in Previously Published Randomized Clinical Trials in Major Medical Journals. *JAMA*, *320*(17), 1813–1815. https://doi.org/10.1001/jama.2018.12288

Weiden P. J. (2007). EPS profiles: the atypical antipsychotics are not all the same. *Journal of psychiatric practice*, *13*(1), 13–24. https://doi.org/10.1097/00131746-200701000-00003

Wong, D. F., Pearlson, G. D., Tune, L. E., Young, L. T., Meltzer, C. C., Dannals, R. F., Ravert, H. T., Reith, J., Kuhar, M. J., & Gjedde, A. (1997). Quantification of neuroreceptors in the living human brain: IV. Effect of aging and elevations of D2-like receptors in schizophrenia and bipolar illness. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*, *17*(3), 331–342. https://doi.org/10.1097/00004647-199703000-00010

World Health Organization. (2002). *The importance of pharmacovigilance*. https://apps.who.int/iris/bitstream/handle/10665/42493/a75646.pdf.

Wu, C. S., Tsai, Y. T., & Tsai, H. J. (2015). Antipsychotic drugs and the risk of ventricular arrhythmia and/or sudden cardiac death: a nation-wide case-crossover study. *Journal of the American Heart Association*, *4*(2), e001568. https://doi.org/10.1161/JAHA.114.001568

Yang, C., Hao, Z., Tian, J., Zhang, W., Li, W., Zhang, L. L., & Song, F. (2018). Does antipsychotic drug use increase the risk of long term mortality? A systematic review and meta-analysis of

observational studies. *Oncotarget*, *9*(19), 15101–15110. https://doi.org/10.18632/oncotarget.24120

Yoshida, K., & Takeuchi, H. (2021). Dose-dependent effects of antipsychotics on efficacy and adverse effects in schizophrenia. *Behavioural brain research*, *402*, 113098. https://doi.org/10.1016/j.bbr.2020.113098

Zähringer, J., Schwingshackl, L., Movsisyan, A., Stratil, J. M., Capacci, S., Steinacker, J. M., Forberger, S., Ahrens, W., Küllenberg de Gaudry, D., Schünemann, H. J., Meerpohl, J. J., & PEN consortium (2020). Use of the GRADE approach in health policymaking and evaluation: a scoping review of nutrition and physical activity policies. *Implementation science : IS*, *15*(1), 37. https://doi.org/10.1186/s13012-020-00984-2

Zivkovic, S., Koh, C. H., Kaza, N., & Jackson, C. A. (2019). Antipsychotic drug use and risk of stroke and myocardial infarction: a systematic review and meta-analysis. *BMC psychiatry*, *19*(1), 189. https://doi.org/10.1186/s12888-019-2177-5

Zorzela, L., Loke, Y. K., Ioannidis, J. P., Golder, S., Santaguida, P., Altman, D. G., Moher, D., Vohra, S., & PRISMAHarms Group (2016). PRISMA harms checklist: improving harms reporting in systematic reviews. *BMJ (Clinical research ed.)*, *352*, i157. https://doi.org/10.1136/bmj.i157

Appendix A: PRISMA Harms Checklist

Table 2 | PRISMA harms checklist

Section/topic (page no)	Item	PRISMA checklist item	PRISMA harms (minimum)	Recommendations for reporting harms in systematic reviews (desirable)	Check if done
Title					
Title (3)	1	Identify the report as a systematic review, meta-analysis, or both.	Specifically mention "harms" or other related terms, or the harm of interest in the review.	—	✓
Abstract					
Structured summary (4)	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	—	Abstracts should report any analysis of harms undertaken in the review, if harms are a primary or secondary outcome.	✓
Introduction					
Rationale (5)	3	Describe the rationale for the review in the context of what is already known.	—	It should clearly describe in introduction or in methods section which events are considered harms and provide a clear rationale for the specific harm(s), condition(s), and patient group(s) included in the review.	✓
Objectives (5)	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	—	PICOS format should be specified, although in systematic reviews of harms the selection criteria for P, C, and O may be very broad (same intervention may have been used for heterogeneous indications in a diverse range of patients)	✓
Methods					
Protocol and registration (6)	5	Indicate if a review protocol exists, if and where it can be accessed (eg, web address), and, if available, provide registration information including registration number.	—	No specific additional information is required for systematic reviews of harms.	✓
Eligibility criteria (6)	6	Specify study characteristics (eg, PICOS, length of follow-up) and report characteristics (eg, years considered, language, publication status) used as criteria for eligibility, giving rationale.	—	Report how handled relevant studies (based on population and intervention) when the outcomes of interest were not reported.Report choices for specific study designs and length of follow-up.	✓
Information sources (7)	7	Describe all information sources (eg, databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	—	Report if only searched for published data, or also sought data from unpublished sources, from authors, drug manufacturers and regulatory agencies. If includes unpublished data, provide the source and the process of obtaining it.	✓
Search (7)	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	—	If additional searches were used specifically to identify adverse events, authors should present the full search process so it can be replicated.	✓
Study selection (8)	9	State the process for selecting studies (ie, screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	—	If only included studies reporting on adverse events of interest, defined if screening was based on adverse event reporting in title/abstract or full text. If no harms reported in the text, report if any attempt was made to retrieve relevant data from authors.	✓
Data collection process (9)	10	Describe method of data extraction from reports (eg, piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	—	No specific additional information is required for systematic reviews of harms.	✓
Data items (9)	11	List and define all variables for which data were sought (eg, PICOS, funding sources) and any assumptions and simplifications made.	—	Report the definition of the harm and seriousness used by each included study (if applicable). Report if multiple events occurred in the same individuals, if this information is available. Consider if the harm may be related to factors associated with participants (eg, age, sex, use of medications) or provider (eg, years of practice, level of training). Specify if information was extracted and how it was used in subsequent results. Specify if extracted details regarding the specific methods used to capture harms (active/passive and timing of adverse event).	✓
Risk of bias in individual studies (10)	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	—	The risk of bias assessment should be considered separately for outcomes of benefit and harms.	✓
Summary measures (11)	13	State the principal summary measures (eg, risk ratio, difference in means).	—	No specific additional information is required for systematic reviews of harms.	✓

Table 2 | PRISMA harms checklist

Section/topic (page no)	Item	PRISMA checklist item	PRISMA harms (minimum)	Recommendations for reporting harms in systematic reviews (desirable)	Check if done
Synthesis of results (11)	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (eg, I^2) for each meta-analysis.	Specify how zero events were handled, if relevant.		✓
Risk of bias across studies (11)	15	Specify any assessment of risk of bias that may affect the cumulative evidence (eg, publication bias, selective reporting within studies).	—	Present the extent of missing information (studies without harms outcomes), any factors that may account for their absence, and whether these reasons may be related to the results.	✓
Additional analyses (12)	16	Describe methods of additional analyses (eg, sensitivity or subgroup analyses, meta-regression), if done, indicating which were prespecified.	—	Sensitivity analyses may be affected by different definitions, grading, and attribution of adverse events, as adverse events are typically infrequent or reported using heterogeneous classifications. Report the number of participants and studies included in each subgroup.	✓
Results					
Study selection (13)	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	—	If a review addresses both efficacy and harms, display a flow diagram specific for each (efficacy and harm).	✓
Study characteristics (14)	18	For each study, present characteristics for which data were extracted (eg, study size, PICOS, follow-up period) and provide the citations.	Define each harm addressed, how it was ascertained (eg, patient report, active search), and over what time period.	Add additional characteristics to: "P" (population) patient risk factors that were considered as possibly affecting the risk of the harm outcome. "I" (intervention) professional expertise/skills if relevant (for example if the intervention is a procedure). "T" (time) timing of all harms assessments and the length of follow-up.	✓
Risk of bias within studies (15)	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	—	Consider the possible sources of biases that could affect the specific harm under consideration within the review. Sample selection, dropouts and measurement of adverse events should be evaluated separately from the outcomes of benefit as described in item 12, above.	✓
Results of individual studies (16)	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	—	Report the actual numbers of adverse events in each study, separately for each intervention.	✓
Synthesis of results (17)	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	Describe any assessment of possible causality.	If included data from unpublished sources, report clearly the data source and the impact of these studies to the final systematic review.	✓
Risk of bias across studies (18)	22	Present results of any assessment of risk of bias across studies (see item 15).	—	No specific additional information is required for systematic reviews of harms. See item 15 above.	✓
Additional analysis (18)	23	Give results of additional analyses, if done (eg, sensitivity or subgroup analyses, meta-regression (see item 16)).	—	No specific additional information is required for systematic reviews of harms.	✓
Discussion					
Summary of evidence (18)	24	Summarise the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (eg, healthcare providers, users, and policy makers).	—	No specific additional information is required for systematic reviews of harms.	✓
Limitations (18)	25	Discuss limitations at study and outcome level (eg, risk of bias), and at review level (eg, incomplete retrieval of identified research, reporting bias).	—	Recognise possible limitations of meta-analysis for rare adverse events (ie, quality and quantity of data), issues noted previously related to collection and reporting.	✓
Conclusions (19)	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	—	State conclusions in coherence with the review findings. When adverse events were not identified we caution against the conclusion that the intervention is "safe," when, in reality, its safety remains unknown.	✓
Funding					
Funding (19)	27	Describe sources of funding for the systematic review and other support (eg, supply of data); role of funders for the systematic review.	—	No specific additional information is required for systematic reviews of harms.	✓

Appendix B: Search Strategy for PubMed

(antipsychotic agents/ae OR antipsychotic agents/to OR antipsychotic agents/po OR chlorpromazine[tiab] OR fluphenazine[tiab] OR haloperidol[tiab] OR pimozide[tiab] OR prochlorperazine[tiab] OR trifluoperazine[tiab] OR benperidol[tiab] OR flupentixol[tiab] OR levomepromazine[tiab] OR pericyazine[tiab] OR promazine[tiab] OR sulpiride[tiab] OR zuclopenthixol[tiab] OR acepromazine[tiab] OR droperidol[tiab] OR loxapine[tiab] OR trifluprомazine[tiab] OR perphenazine[tiab] OR thioridazine[tiab] OR molindone[tiab] OR thiothixene[tiab] OR aripiprazole[tiab] OR clozapine[tiab] OR olanzapine[tiab] OR quetiapine[tiab] OR risperidone[tiab] OR amisulpride[tiab] OR paliperidone[tiab] OR ziprasidone[tiab] OR asenapine[tiab] OR cariprazine[tiab] OR lurasidone[tiab]) AND (((((systematic review[tiab] OR systematic literature review[tiab] OR systematic scoping review[tiab] OR systematic narrative review[tiab] OR systematic qualitative review[tiab] OR systematic evidence review[tiab] OR systematic quantitative review[tiab] OR systematic meta-review[tiab] OR systematic critical review[tiab] OR systematic mixed studies review[tiab] OR systematic mapping review[tiab] OR systematic cochrane review[tiab] OR systematic search and review[tiab] OR systematic integrative review[tiab]) NOT comment[pt] NOT (protocol[tiab] OR protocols[tiab])) NOT MEDLINE [subset]) OR (Cochrane Database Syst Rev[ta] AND review[pt]) OR systematic review[pt]) OR (meta-analysis[pt] OR meta-analysis[tiab] OR metaanaly*[tiab] OR meta-analy*[tiab]))

Appendix C: Screening Template

Template for screening titles and abstracts

Review Title: Adverse Health Outcomes Associated with Antipsychotic Use: An Umbrella Review	
Purpose: To synthesise the evidence for the reported adverse drug events associated with antipsychotic use	
Inclusion criteria ('yes' indicates inclusion)	
Study design	
Is the study an observational or experimental study with a defined control group (include RCTs, cohort studies, case-control studies and nested case-control studies)?	Yes/No/Unclear
Population	
Is the study population community-living?	Yes/No/Unclear
Exposure	
Were the study participants exposed to any of the antipsychotics listed in the BNF, USP-NF and WHO Model List of Essential Medicines (2019)[1]?	Yes/No/Unclear
Measures	
Are there measurable quantitative data? AND Are there data on at least one antipsychotics-associated adverse event?	Yes/No/Unclear Yes/No/Unclear
Exclusion criteria ('yes' indicates exclusion)	
Was the paper written in a language other than English?	Yes/No/Unclear
Was study participants' exposure to antipsychotics in-utero?	Yes/No/Unclear
Did the study only investigate antipsychotic overdose as an exposure or outcome?	Yes/No/Unclear

Notes for screening

- This template is to be used in the first stage of screening for study titles and abstracts.
- To be included, studies must fulfil all the inclusion criteria and not have any of the exclusion criteria.
- For those marked 'unclear', full text will be retrieved to determine inclusion.

Endnote ID Author Last name + year	Include (YES/NO/UNCLEAR)	Reason for exclusion* (A-J)

Add rows as appropriate

***Reason for exclusion:**

A= Not a systematic review or meta-analysis
B = In-utero exposure to antipsychotics
C = Does not include specified control/comparison group
D = No exposure to any of the antipsychotics listed in the BNF, USP-NF and WHO Model List of Essential Medicines (2019)
E = Primarily investigated antipsychotic overdose
F = Did not specify specific adverse outcome(s) as study variables
G = No quantitative data reported on any harm associated with antipsychotic use
H = Participants are non-human
I = Paper not in English

Appendix D: AMSTAR 2 Critical Appraisal Instrument

Table S1. AMSTAR 2 checklist, with critical items bolded. Adapted from Shea et al. (2017)

Domain questions
1. Did the research questions and inclusion criteria for the review include the components of PICO?
2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?
3. Did the review authors explain their selection of the study designs for inclusion in the review?
4. Did the review authors use a comprehensive literature search strategy?
5. Did the review authors perform study selection in duplicate?
6. Did the review authors perform data extraction in duplicate?
7. Did the review authors provide a list of excluded studies and justify the exclusions?
8. Did the review authors describe the included studies in adequate detail?
9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?
10. Did the review authors report on the sources of funding for the studies included in the review?
11. If meta-analysis was performed, did the review authors use appropriate methods for statistical combination of results?
12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?
13. Did the review authors account for RoB in primary studies when interpreting/discussing the results of the review?
14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?
15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?
16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?

Table S2. Overall confidence ratings in the AMSTAR 2 instrument.

Overall confidence	**Description and indication**
High	With no or one non-critical weakness. Provides an accurate and comprehensive summary of the results of the available studies that address the question of interest
Moderate	With more than one non-critical weakness – has more than one weakness but no critical flaws. May provide an accurate summary of the results of the available studies that were included in the review.

Low	With one critical flaw with or without non-critical weaknesses. May not provide an accurate and comprehensive summary of the available studies that address the question of interest.
Critically low	With more than one critical flaw with or without non-critical weaknesses. Should not be relied on to provide an accurate and comprehensive summary of the available studies.

Appendix E: Formulae Used in Meta-Analysis

1. Conversion of 2x2 contingency data to risk ratio

$$RR = \frac{a/(a+b)}{c/(c+d)}$$

where:
a = number of antipsychotic-exposed participants who developed adverse outcomes
b = number of antipsychotic-exposed participants who did not develop adverse outcomes
c = number of antipsychotic-unexposed participants who developed adverse outcomes
d = number of antipsychotic-unexposed participants who did not develop adverse outcomes

2. Conversion of attributable risk to risk ratio

$$RR = \frac{100}{100 - AR}$$

3. Calculation of *p* value for effect sizes

$$p = e^{-0.717 \times z - 0.416 \times z^2}$$

z is the test statistic and is given by:

$$z = \frac{\log(ES)}{SE}$$

where *ES* is the effect size and could refer to either RR or OR. *SE* is the standard error of the log(ES) and could be calculated with:

$$SE = \frac{UCI - LCI}{3.92}$$

where UCI and LCI are the upper and lower limits of the log(effect size)'s 95% confidence intervals respectively

4. Calculation of ratio of odds ratios:

$$\frac{OR_{RCT}}{OR_{observational\ study}}$$

5. Conversion of Cochran's Q to I^2

$$I^2 = \left(\frac{Q - df}{Q}\right) \times 100\%$$

where *df* = *k*-1 and *k* = the number of primary studies

6. Calculation of 95% prediction interval

$$\log(ES) \pm t_{1-0.05/2,k-1} \times \sqrt{\tau^2 + SE^2}$$

where *ES* is effect size and could be RR or OR, $t_{1-0.05/2,k-1}$ is the 97.5th percentile of *t* distribution at significance level 0.05 and *k*-2 degrees of freedom when *k* primary studies are included in the meta-analysis, and SE is given by (Inthout et al., 2016):

$$SE = \frac{UCI - LCI}{3.92}$$

where UCI and LCI are the natural log of upper and lower limits of the log(effect size)'s 95% confidence intervals respectively

Appendix F: Outcomes Excluded from Analyses with Reasons

Adverse outcome	Reason for exclusion
Abdominal discomfort	With less than 3 primary studies
Abdominal pain	Not a diagnosis - symptomatic
Abnormal thoughts	With less than 3 primary studies
Abnormal/ blurred vision	Non-specific diagnosis
Absolute BMI change	Non-clinically significant
Absolute death rate from myocarditis	Unusable data
Absolute weight change	Non-clinically significant
Absolute weight change (kg)	Non-clinically significant
Acute coronary syndrome	With less than 3 primary studies
Agitation/ aggression	Non-specific diagnosis
Agranulocytosis	Unusable data
Akathisia-related events	Non-specific diagnosis
Anticholinergic events	With less than 3 primary studies
Anxiety/ depression	Non-specific diagnosis
Asthenia/ lassitude/ increased fatigability	Non-specific diagnosis
Average change in AIMS from baseline	Non-clinically significant
Average change in alanine aminotransferase (U/L)	Non-clinically significant
Average change in alkaline phosphatase (U/L)	Non-clinically significant
Average change in aspartate amino-transferase (U/L)	Non-clinically significant
Average change in BARS from baseline	Non-clinically significant
Average change in bilirubin (mg/dL)	Non-clinically significant
Average change in cholesterol serum level	Non-clinically significant
Average change in creatine kinase level (mg/dL)	Non-clinically significant
Average change in diastolic blood pressure (mmHg)	Non-clinically significant
Average change in fasting glucose	With less than 3 primary studies
Average change in glucose serum level (mmol/L)	Non-clinically significant
Average change in high-density lipoprotein serum level (mg/dL)	With less than 3 primary studies
Average change in insulin serum level (mmol/L)	Non-clinically significant
Average change in low-density lipoprotein serum level (mg/dL)	With less than 3 primary studies
Average change in movement disorder score	Non-clinically significant
Average change in prolacin levels (miu/L)	Non-clinically significant

Average change in prolactin serum level (female) (ng/ml)	Non-clinically significant
Average change in prolactin serum level (male) (ng/ml)	Non-clinically significant
Average change in QTbc	Non-clinically significant
Average change in QTc LD (msec)	Non-clinically significant
Average change in SAS from baseline	Non-clinically significant
Average change in SAS score	Non-clinically significant
Average change in serum level (mmol/L)	Non-clinically significant
Average change in systolic blood pressure (mmHg)	Non-clinically significant
Average change in triglyceride serum level (mg/dL)	Non-clinically significant
Average change in triglyceride serum level (mmol/L)	Non-clinically significant
Average change in waist circumference	Non-clinically significant
Average change in weight (kg)	Non-clinically significant
Average change in weight at last visit	Non-clinically significant
Average total change in cholesterol (mg/dL)	Non-clinically significant
Average total change in prolactin (mg/dL)	Non-clinically significant
Average total change in sitting diastolic blood pressure	Non-clinically significant
Average total change in sitting pulse rate (BPM)	Non-clinically significant
Average total change in sitting systolic blood pressure	Non-clinically significant
Average total change in standing diastolic blood pressure	Non-clinically significant
Average total change in standing pulse rate (BPM)	Non-clinically significant
Average total change in standing systolic blood pressure	Non-clinically significant
Average total sitting pulse score	Non-clinically significant
Average total standing pulse score	Non-clinically significant
Average total weight change	Non-clinically significant
Average weight change	Non-clinically significant
Average weight gain	Non-clinically significant
Average weight gain (kg)	Non-clinically significant
Blood glucose abnormalities	Non-clinically significant
BMI change (kg/m^2)	Non-clinically significant
BMI change from baseline (kg/m^2)	With less than 3 primary studies
Body pain	With less than 3 primary studies
Cardiomyopathy	Unusable data
Cardiovascular - other	Non-specific diagnosis
Cardiovascular event	Non-specific diagnosis

Case fatality rate from cardiomyopathy	Unusable data
Case fatality rate from myocarditis	Unusable data
Cerebrovascular events	Non-specific diagnosis
Change in akathsia scores	Non-clinically significant
Change in alanine transaminase	Non-clinically significant
Change in aspartate transaminase	Non-clinically significant
Change in basophils count	Non-clinically significant
Change in diastolic blood pressure from baseline (mmHg)	Non-clinically significant
Change in fasting glucose (mmol/L)	Non-clinically significant
Change in fasting glycerides	Non-clinically significant
Change in high-density lipoprotein cholesterol	Non-clinically significant
Change in low-density lipoprotein cholesterol	Non-clinically significant
Change in total cholesterol from baseline (mg/dL)	Non-clinically significant
Change in total cholesterol level	Non-clinically significant
Change in y-glutamyl transferase	Non-clinically significant
Changes in blood calcium	Non-clinically significant
Changes in fasting glucose	Non-clinically significant
Changes in leukocytes count	Non-clinically significant
Changes in monocytes count	Non-clinically significant
Changes in segmented neutrophils count	Non-clinically significant
Changes in triglycerides count	Non-clinically significant
Cholesterol variation from baseline	Non-clinically significant
Clinically significant alanine aminotransferase (ALT) increase	With less than 3 primary studies
Clinically significant alanine aminotransferase levels	With less than 3 primary studies
Clinically significant fasting triglyceride levels	With less than 3 primary studies
Clinically significant weight loss	With less than 3 primary studies
Clinically significant y-glutamyl transpeptidase levels	With less than 3 primary studies
Concentration difficulties	With less than 3 primary studies
Constipation	Not a diagnosis - symptomatic
Coughing	With less than 3 primary studies
Death from agranulocytosis	Unusable data
Death from neutropenia	Unusable data
Death in patients with clozapine-associated agranulocytosis	Unusable data
Decreased appetite	With less than 3 primary studies
Decreased duration of sleep	With less than 3 primary studies
Decreased high-density lipoprotein at endpoint	With less than 3 primary studies
Decreased libido	Not a diagnosis - symptomatic
Deliberate self-injury or violent behaviour	With less than 3 primary studies

Diarrhoea	Not a diagnosis - symptomatic
Direct bilirubin change	Non-clinically significant
Dizziness	Not a diagnosis - symptomatic
Drowsiness	Not a diagnosis - symptomatic
Dysathria	With less than 3 primary studies
Dyskinesia	With less than 3 primary studies
Dyskinetic movement total score	Non-clinically significant
Dyspepsia	Not a diagnosis - symptomatic
Dystonic events	Non-specific diagnosis
EPS-related events	Non-specific diagnosis
Failing memory	With less than 3 primary studies
Female breast pain	Not a diagnosis - symptomatic
Galactorrhoea	Not a diagnosis - symptomatic
Gynaecomastia	Not a diagnosis - symptomatic
Headache	Not a diagnosis - symptomatic
High triglycerides any time during treatment	Non-clinically significant
Hyperkinesia	With less than 3 primary studies
Hyperkinesia/ akathisia	Non-specific diagnosis
Hypersalivation	Not a diagnosis - symptomatic
Increase of ≥15 beats/ min	With less than 3 primary studies
Increased appetite	Not a diagnosis - symptomatic
Increased appetite/ weight gain	Non-specific diagnosis
Increased dream activity	With less than 3 primary studies
Increased duration of sleep	With less than 3 primary studies
Increased fasting blood glucose	With less than 3 primary studies
Increased fasting triglycerides at endpoint	With less than 3 primary studies
Increased lipids	Non-clinically significant
Increased low-density lipoprotein at endpoint	With less than 3 primary studies
Increased prolactin levels	With less than 3 primary studies
Increased tendency to sweating	With less than 3 primary studies
Injection site pain	Not a diagnosis - symptomatic
Irritability	With less than 3 primary studies
Joint disorder	With less than 3 primary studies
Liver function test abnormalities	With less than 3 primary studies
Metabolic lab abnormalities	Non-clinically significant
Muscle rigidity	With less than 3 primary studies
Myocarditis	Unusable data
Nasopharyngitis	With less than 3 primary studies
Nausea	Not a diagnosis - symptomatic
Nausea/ vomiting	With less than 3 primary studies
Nervousness	With less than 3 primary studies
Neurological events	Non-specific diagnosis

Neutropenia at a threshold of <1000	Unusable data
Neutropenia at a threshold of <1500	Unusable data
Neutropenia case fatalities	Unusable data
Oral hypoesthesia	With less than 3 primary studies
Pain	With less than 3 primary studies
Pain in extremity	Not a diagnosis - symptomatic
Palpitation/ tachycardia	With less than 3 primary studies
Parkinsonism/ hypertonia	Duplicated outcomes
Participants receiving medications of extrapyramidal Aes	With less than 3 primary studies
Polyuria/ Polydipsia	With less than 3 primary studies
Potential prolactin-associated events	With less than 3 primary studies
Prolacin increase-associated events	With less than 3 primary studies
Prolactin increase	Non-clinically significant
Prolactin increase (ng/mL)	Non-clinically significant
Prolactin-associated events	Non-specific diagnosis
Pruritis	Not a diagnosis - symptomatic
Pyrexia	Not a diagnosis - symptomatic
Restlessness	Not a diagnosis - symptomatic
Rhinitis	With less than 3 primary studies
Salivation	Not a diagnosis - symptomatic
Sedation-associated events	Non-specific diagnosis
Seizure	With less than 3 primary studies
Shift from <3 to ≥3 metabolic risk factors	Non-clinically significant
Sinus tachycardia	With less than 3 primary studies
Sleepiness	Not a diagnosis - symptomatic
Sleepiness/ sedation	With less than 3 primary studies
Somnolence/ sedation	Duplicated outcomes
Suicidal or homicidal ideation	Non-specific outcome
Tachycardia	Not a diagnosis - symptomatic
Tension	With less than 3 primary studies
Toothache	Not a diagnosis - symptomatic
Total change in bilirubin	Non-clinically significant
Total change in cholesterol	Non-clinically significant
Total movement disorder events	Non-clinically significant
Toxicity	With less than 3 primary studies
Upper respiratory infection	Not a diagnosis - symptomatic
Urinary symptoms	Non-specific diagnosis
Use of anti-EPS medication	Not a diagnosis
Use of anticholingeric medication	Not a diagnosis
Use of antiparkinson medication	Not a diagnosis
Use of anxiolytic medication	Not a diagnosis

Use of benzodiazepines	Not a diagnosis
Use of beta-blockers medication for akathisia	Not a diagnosis
Venous thromboembolism and pulmonary embolism	Non-specific diagnosis
Viral gastroenteritis	With less than 3 primary studies
Vomiting	Not a diagnosis - symptomatic
Weight change	Non-clinically significant
Weight change (kg)	Non-clinically significant
Weight change (lb)	Non-clinically significant
Weight gain	Non-clinically significant
Weight gain (kg)	Non-clinically significant
Weight gain (lb)	Non-clinically significant
Weight loss	Non-clinically significant
Worsening psychotic symptoms	With less than 3 primary studies
Xerostomia	Not a diagnosis - symptomatic

Appendix G: Potentially Relevant But Excluded Studies

Author (Year)	Reason for exclusion
Arbaizar et al. (2009)	Overlapping meta-analysis reporting the same adverse outcome
Ardozzone et al. (2010)	Overlapping meta-analysis reporting the same adverse outcome
Barbui et al. (2014)	Overlapping meta-analysis reporting the same adverse outcome
Barton et al. (2020)	Overlapping meta-analysis reporting the same adverse outcome
Chang et al. (2017)	Overlapping meta-analysis reporting the same adverse outcome
Chung et al. (2010)	Overlapping meta-analysis reporting the same adverse outcome
Citrome (2019)	Overlapping meta-analysis reporting the same adverse outcome
Cipriani et al. (2009)	Overlapping meta-analysis reporting the same adverse outcome
Cooper et al. (2019)	Overlapping meta-analysis reporting the same adverse outcome
Davidson et al. (2000)	Overlapping meta-analysis reporting the same adverse outcome
De Hert et al. (2012)	Overlapping meta-analysis reporting the same adverse outcome
Depping et al. (2010)	Overlapping meta-analysis reporting the same adverse outcome
Domecq et al. (2015)	Overlapping meta-analysis reporting the same adverse outcome
Fang et al. (2016)	Overlapping meta-analysis reporting the same adverse outcome
Fang et al. (2017)	Overlapping meta-analysis reporting the same adverse outcome
Flank et al. (2014)	Overlapping meta-analysis reporting the same adverse outcome
Flank et al. (2014)	Overlapping meta-analysis reporting the same adverse outcome
Fung et al. (2016)	Overlapping meta-analysis reporting the same adverse outcome
Fusar-Poli et al. (2013)	Overlapping meta-analysis reporting the same adverse outcome
Gao et al. (2008)	Overlapping meta-analysis reporting the same adverse outcome
Gao et al. (2008)b	Overlapping meta-analysis reporting the same adverse outcome
Harrington et al. (2010)	Overlapping meta-analysis reporting the same adverse outcome

Hay et al. (2015)	Overlapping meta-analysis reporting the same adverse outcome
Hirsch & Pringsheim (2016)	Overlapping meta-analysis reporting the same adverse outcome
Hollis et al. (2016)	Overlapping meta-analysis reporting the same adverse outcome
Hsu et al. (2017)	Overlapping meta-analysis reporting the same adverse outcome
Huang et al. (2017)	Overlapping meta-analysis reporting the same adverse outcome
Kishi et al. (2013)a	Overlapping meta-analysis reporting the same adverse outcome
Kishi et al. (2013)b	Overlapping meta-analysis reporting the same adverse outcome
Kishi et al. (2016)a	Overlapping meta-analysis reporting the same adverse outcome
Kishi et al. (2016)b	Overlapping meta-analysis reporting the same adverse outcome
Komossa et al. (2010)a	Overlapping meta-analysis reporting the same adverse outcome
Komossa et al. (2010)b	Overlapping meta-analysis reporting the same adverse outcome
Kuharic et al. (2019)	Overlapping meta-analysis reporting the same adverse outcome
Leucht et al. (2009)	Overlapping meta-analysis reporting the same adverse outcome
Li et al. (2017)	Overlapping meta-analysis reporting the same adverse outcome
Lieb et al. (2018)	Overlapping meta-analysis reporting the same adverse outcome
Ma et al. (2014)	Overlapping meta-analysis reporting the same adverse outcome
Maayan et al. (2015)	Overlapping meta-analysis reporting the same adverse outcome
Majer et al. (2015)	Overlapping meta-analysis reporting the same adverse outcome
Marshall & Rathbone (2011)	Overlapping meta-analysis reporting the same adverse outcome
McQuire et al. (2015)	Overlapping meta-analysis reporting the same adverse outcome
Nussbaum & Stroup (2013)	Overlapping meta-analysis reporting the same adverse outcome
Odera et al. (2012)	Overlapping meta-analysis reporting the same adverse outcome
Orr et al. (2017)	Overlapping meta-analysis reporting the same adverse outcome
Oya et al. (2015)	Overlapping meta-analysis reporting the same adverse outcome

Prajapati et al. (2018)	Overlapping meta-analysis reporting the same adverse outcome
Pringsheim et al. (2011)	Overlapping meta-analysis reporting the same adverse outcome
Rodrigues et al. (2020)	Overlapping meta-analysis reporting the same adverse outcome
Ruxton et al. (2015)	Overlapping meta-analysis reporting the same adverse outcome
Sampson et al. (2016)	Overlapping meta-analysis reporting the same adverse outcome
Silva et al. (2013)	Overlapping meta-analysis reporting the same adverse outcome
Spielmans et al. (2013)	Overlapping meta-analysis reporting the same adverse outcome
Srisurapanont et al. (2004)	Overlapping meta-analysis reporting the same adverse outcome
Stafford et al. (2015)	Overlapping meta-analysis reporting the same adverse outcome
Suttajit et al. (2014)	Overlapping meta-analysis reporting the same adverse outcome
Tek et al. (2016)	Overlapping meta-analysis reporting the same adverse outcome
Vancampfort et al. (2013)	Overlapping meta-analysis reporting the same adverse outcome
Vita et al. (2013)	Overlapping meta-analysis reporting the same adverse outcome
Wang et al. (2016)	Overlapping meta-analysis reporting the same adverse outcome
Willems et al. (2014)	Overlapping meta-analysis reporting the same adverse outcome
Yu et al. (2016)	Overlapping meta-analysis reporting the same adverse outcome
Zhai et al. (2016)	Overlapping meta-analysis reporting the same adverse outcome
Zhao et al. (2018)	Overlapping meta-analysis reporting the same adverse outcome
Alfageh et al. (2019)	Reported adverse outcome excluded during AE classification
Almandil et al. (2013)	Reported adverse outcome excluded during AE classification
Douglas-Hall et al. (2011)	Reported adverse outcome excluded during AE classification
Jenser et al. (2007)	Reported adverse outcome excluded during AE classification
Lalonde et al. (2011)	Reported adverse outcome excluded during AE classification
Loy et al. (2017)	Reported adverse outcome excluded during AE classification

Myles et al. (2018)	Reported adverse outcome excluded during AE classification
Nashed et al. (2011)	Reported adverse outcome excluded during AE classification
Pae et al. (2008)	Reported adverse outcome excluded during AE classification
Tan et al. (2018)	Reported adverse outcome excluded during AE classification
Siskind et al. (2020)	Data could not be quantitatively analysed

Appendix H: Individual Effect Sizes Reported in Meta-Analyses

Figure S1. Myocardial infarction. All associations extracted from Rotella et al. (2018)

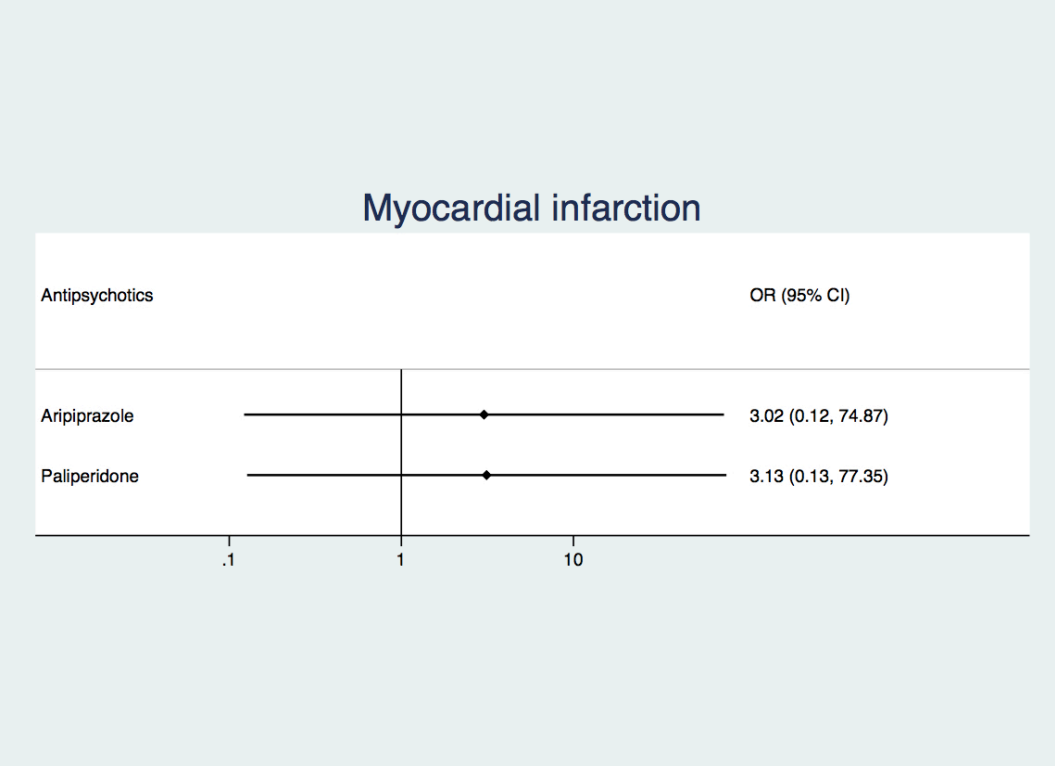


Figure S2. Endocrine abnormalities, except diabetes. All associations extracted from Maglione et al. (2011)

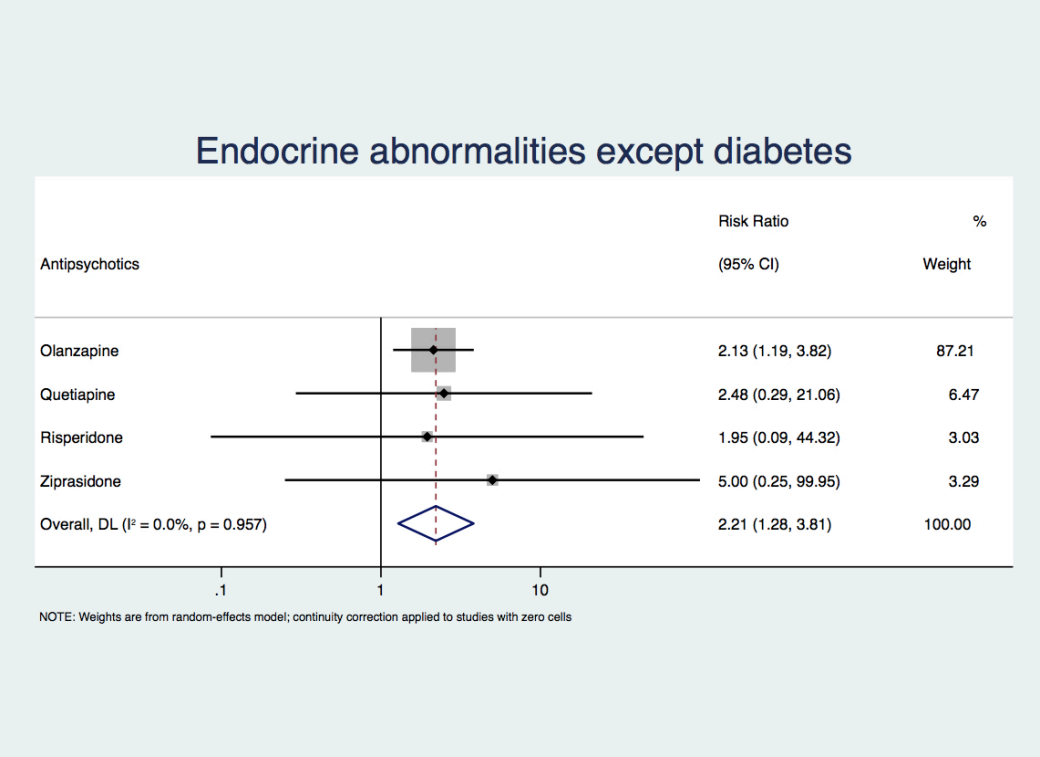


Figure S3. Diabetes mellitus. All associations extracted from Maglione et al. (2011)

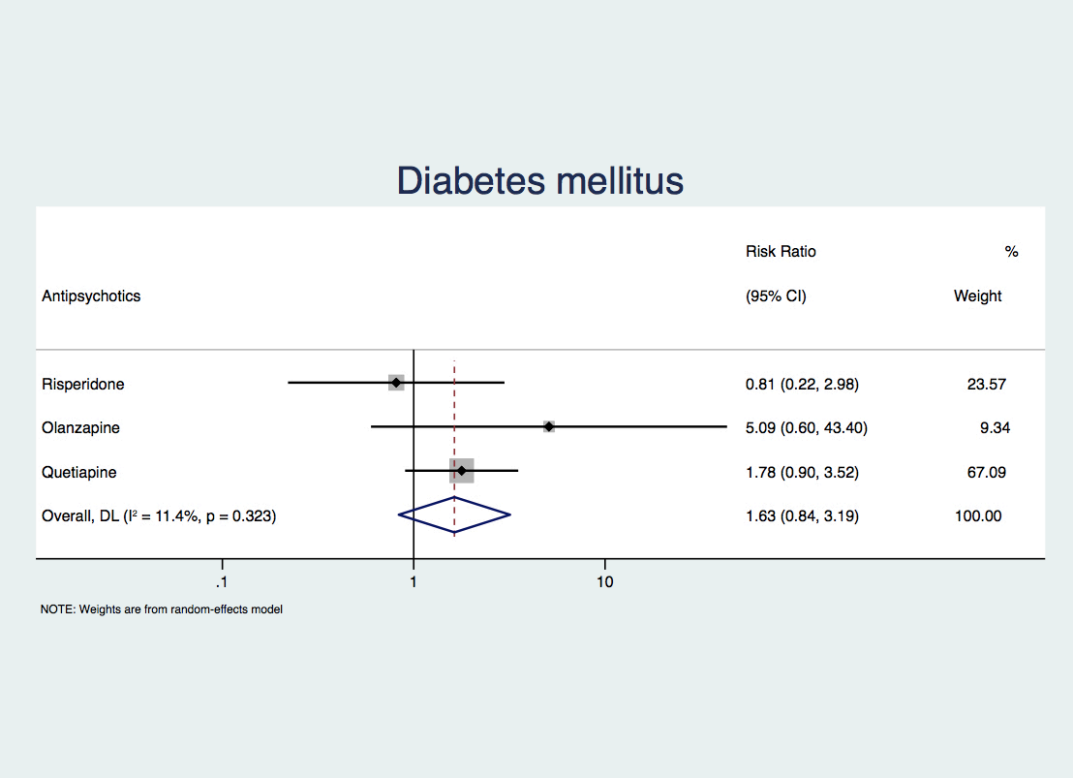


Figure S4. Fatigue. All associations extracted from Maher et al. (2011)

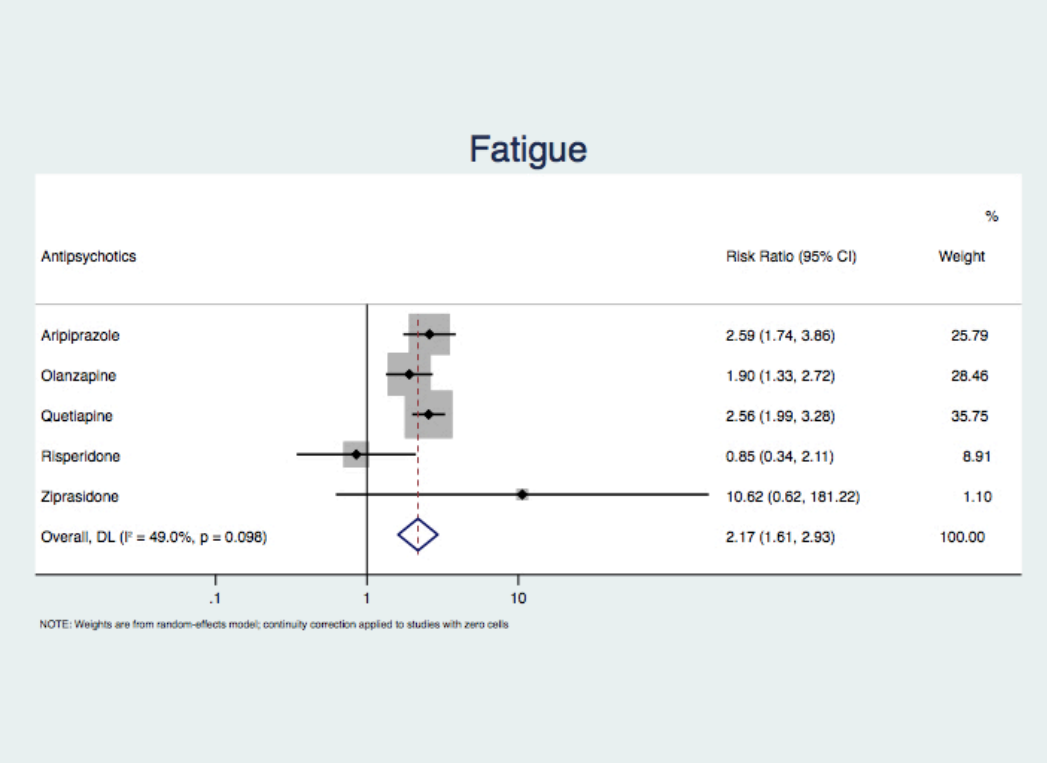


Figure S5. Injury. All associations extracted from Ma et al. (2014)

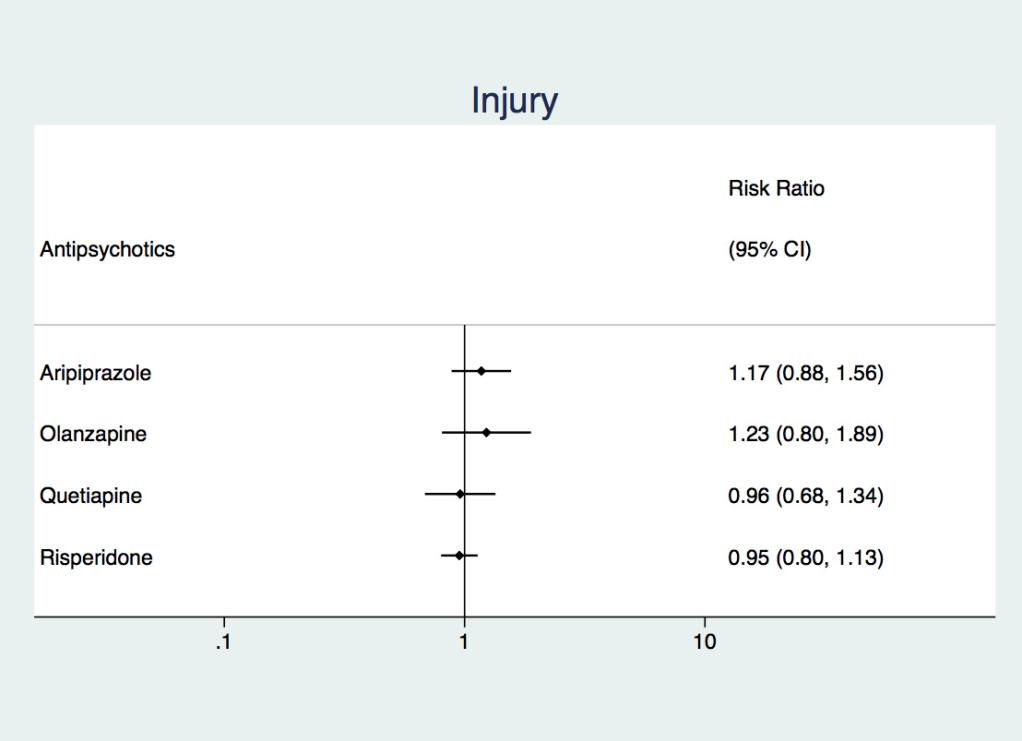


Figure S6. Clinically significant weight gain. All associations extracted from Bai et al. (2020)

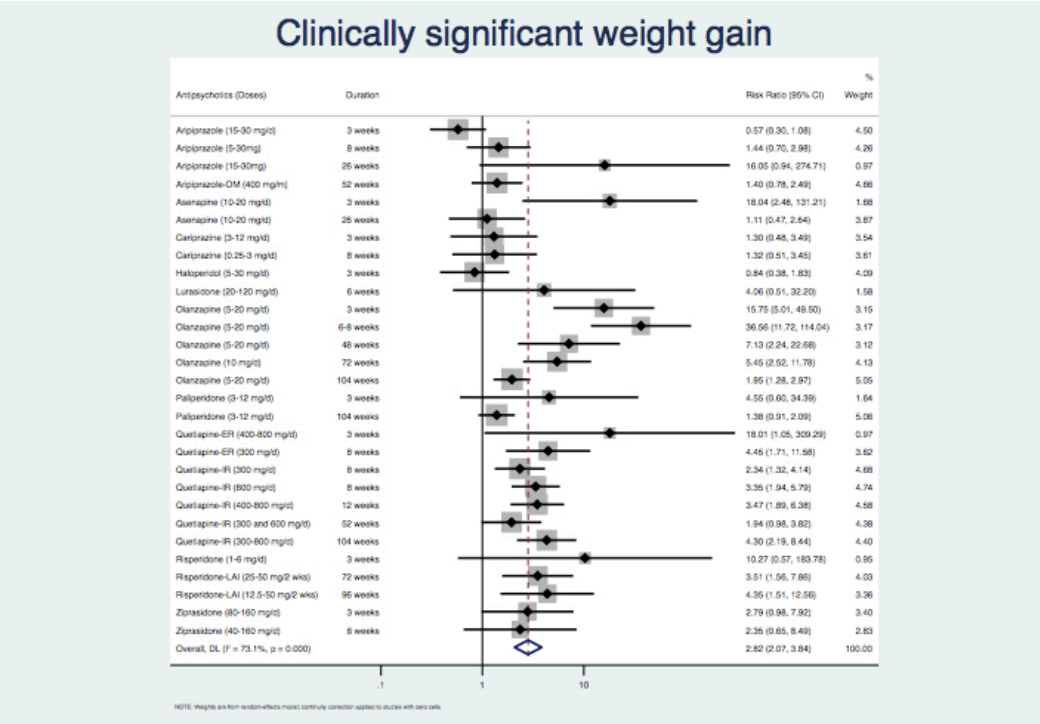


Figure S7. All associations extracted from Vancampfort et al. (2015)

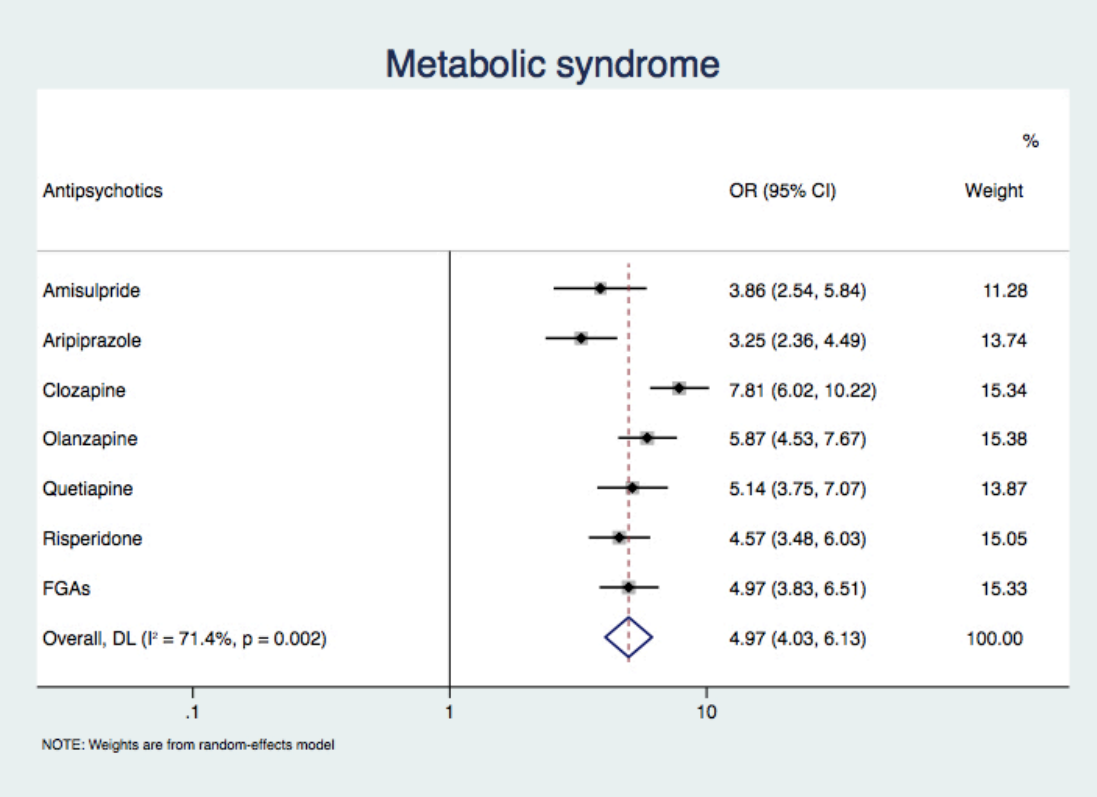


Figure S8. Insomnia. All associations extracted from Romeo et al. (2018)

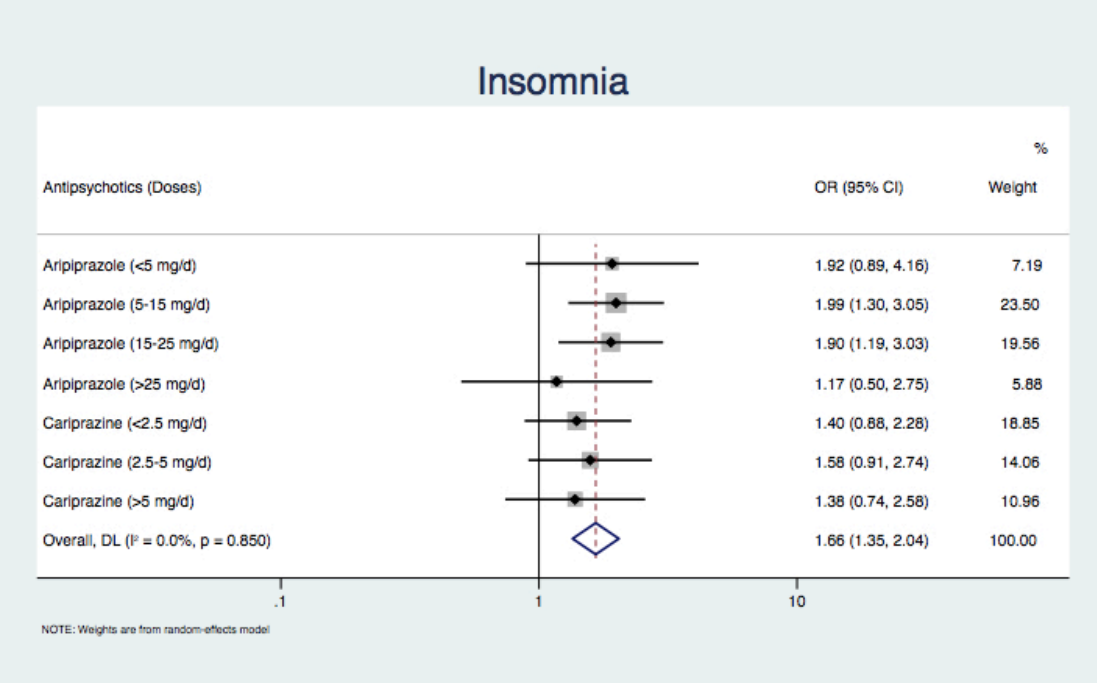


Figure S9. Sedation. All associations extracted from Maher et al. (2011)

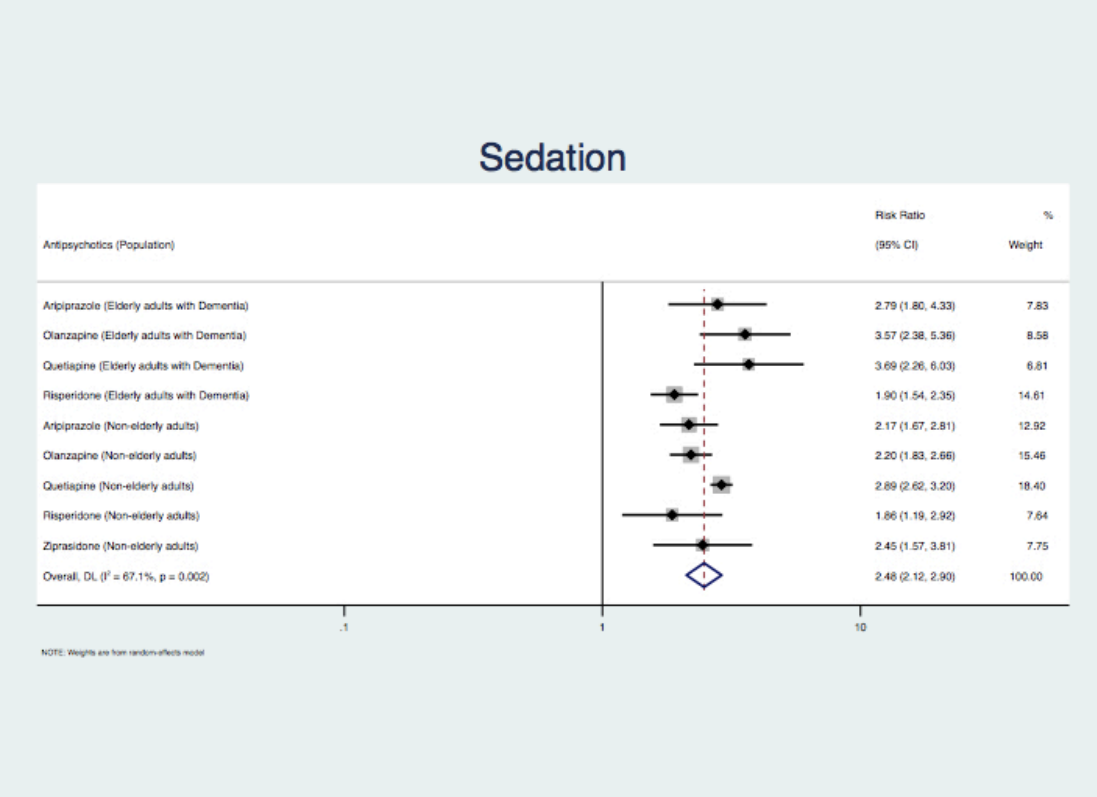
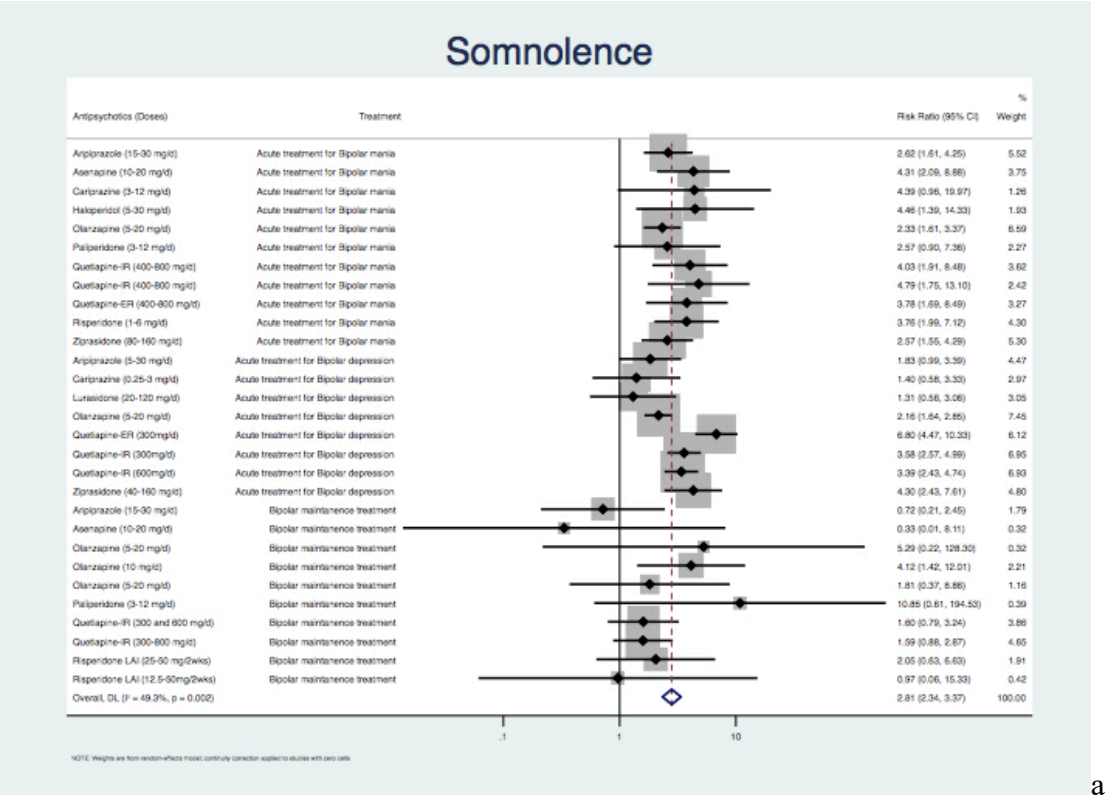


Figure S10. Somnolence. All associations extracted from Bai et al. (2020)



a

Figure S11. Extrapyramidal disorder. All associations extracted from Demyttenaere et al. (2019).

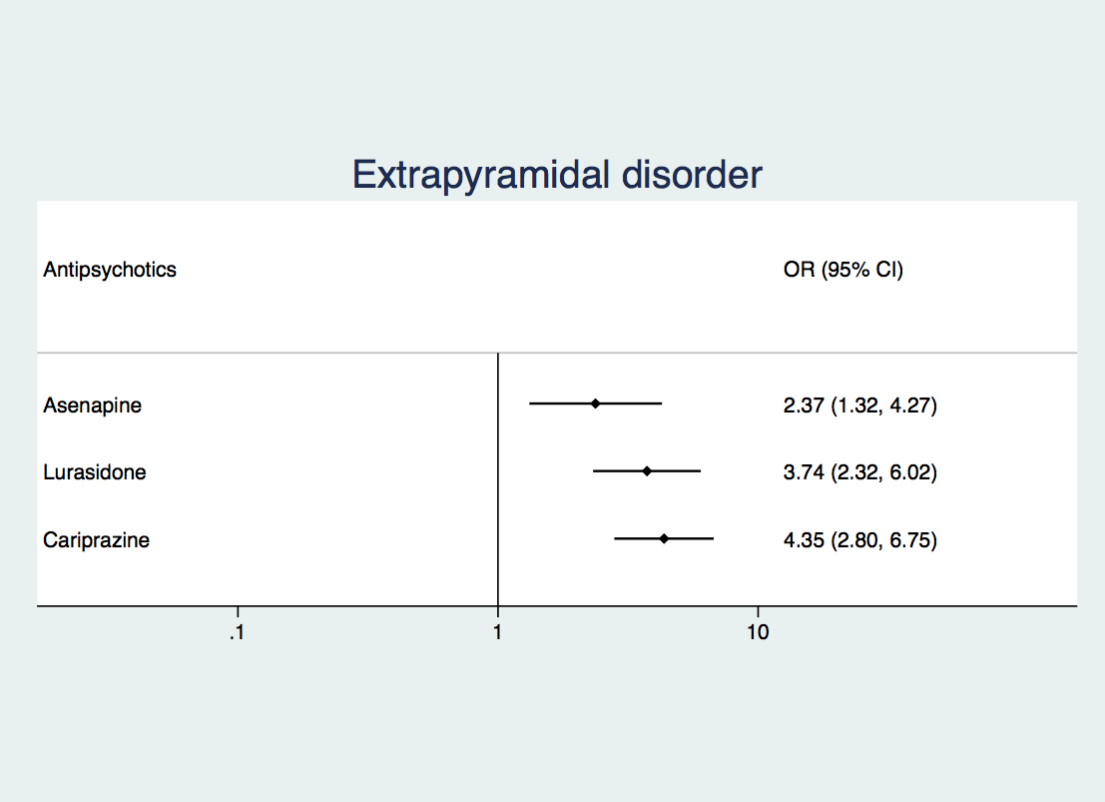


Figure S12. Gait abnormality. All associations extracted from Ma et al. (2014)

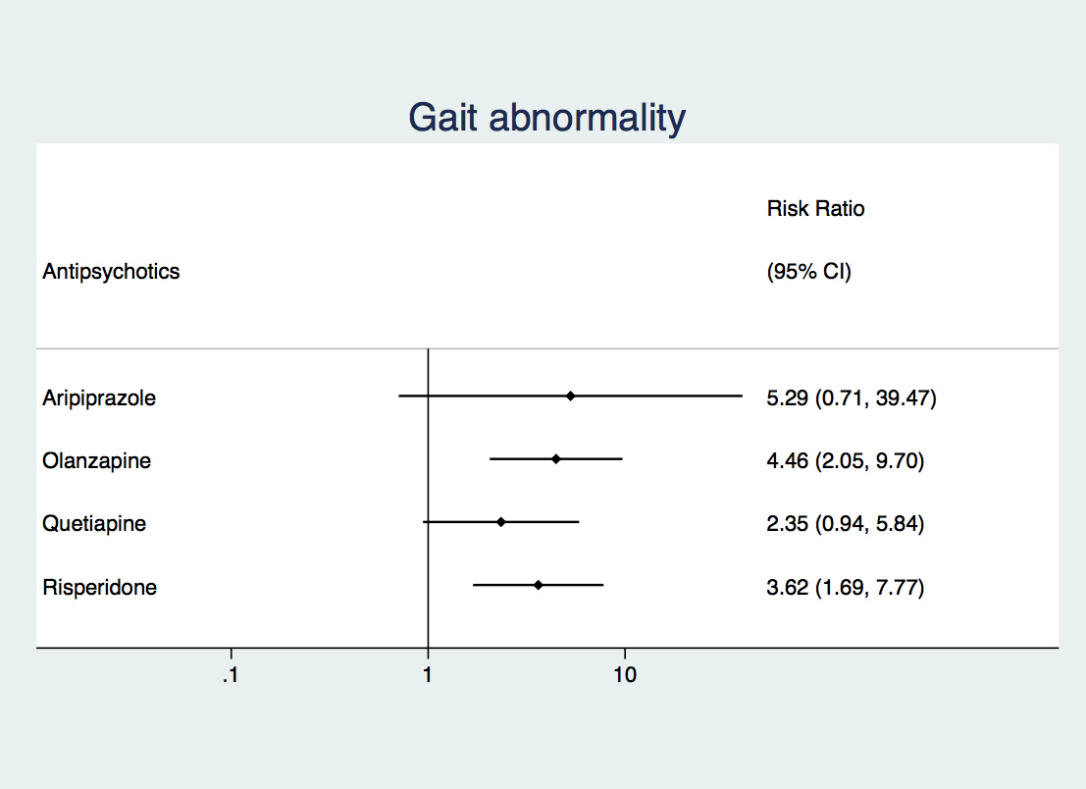


Figure S13. Stroke. All associations extracted from Maher et al. (2011)

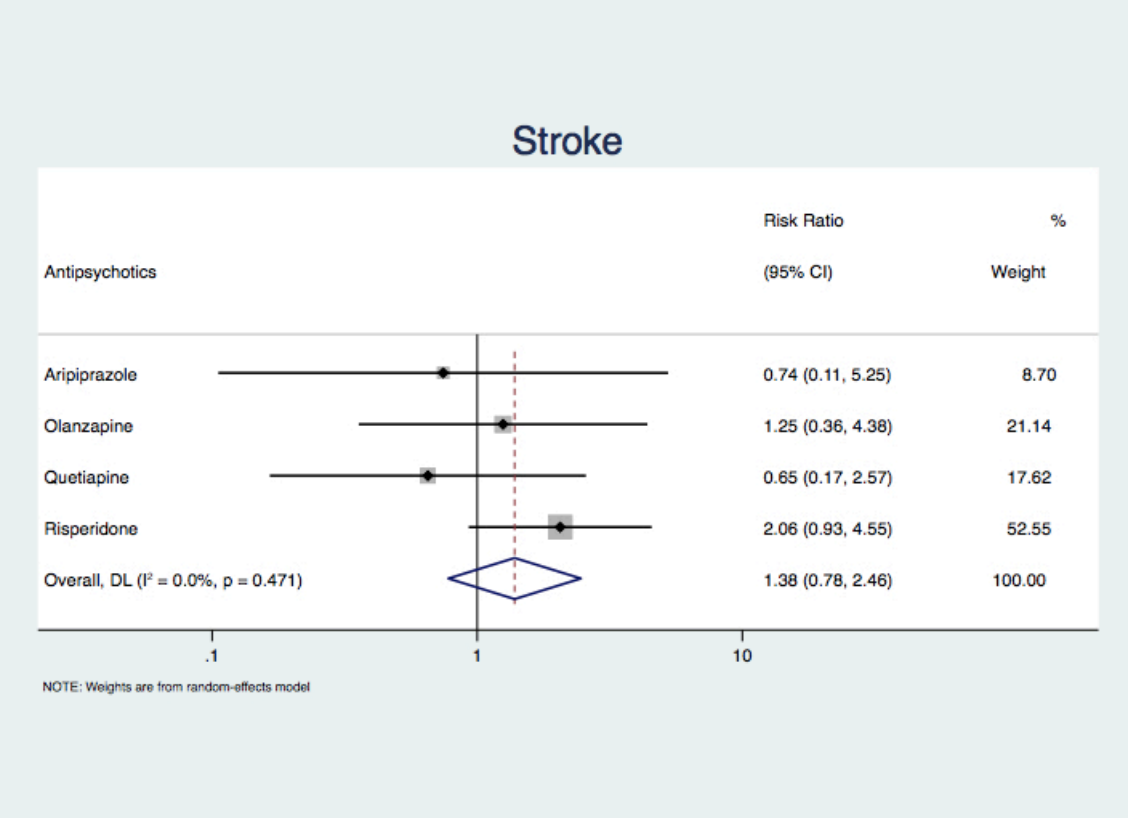


Figure S14. Agitation. All associations extracted from Romeo et al. (2018)

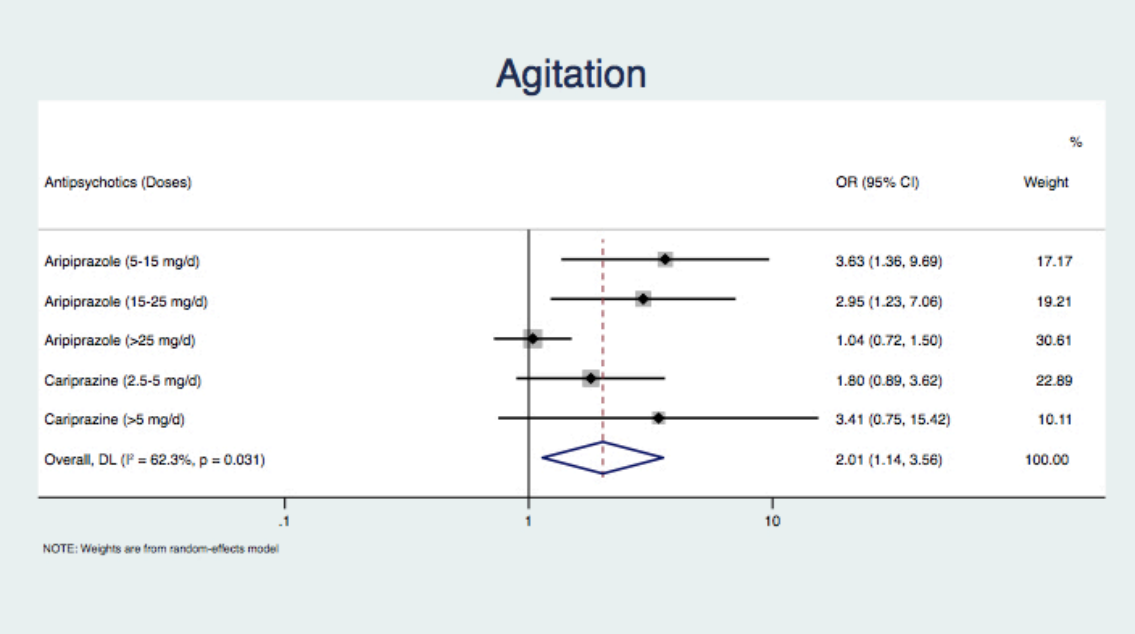


Figure S15. Depression. All associations extracted from Derry and Moore (2007)

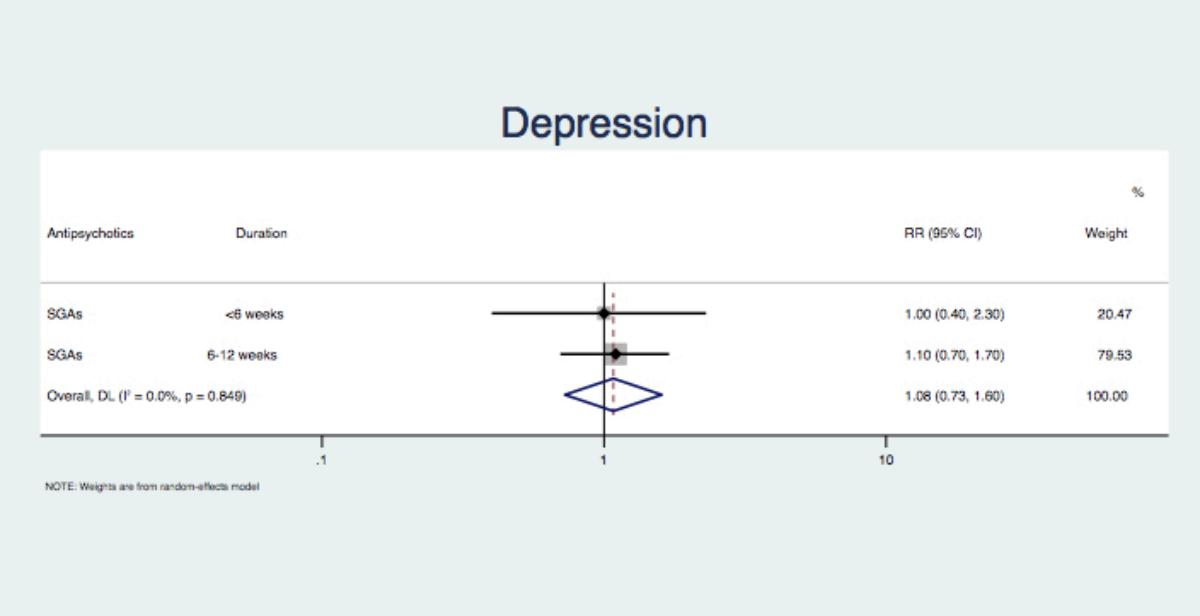


Figure S16. Recurrence of psychotic symptoms. All associations extracted from Defruyt et al. (2011)

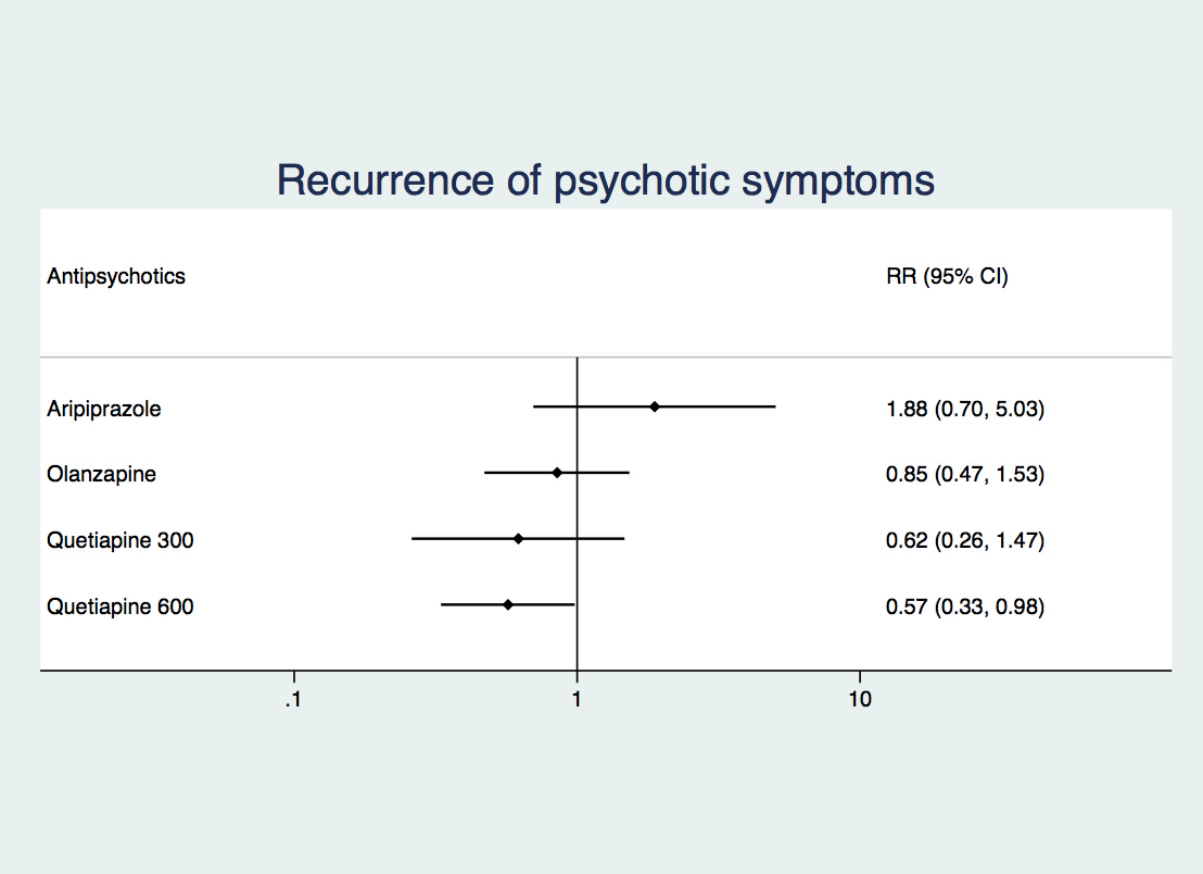


Figure S17. All-cause death. All associations extracted from Schneider-Thoma et al. (2018).

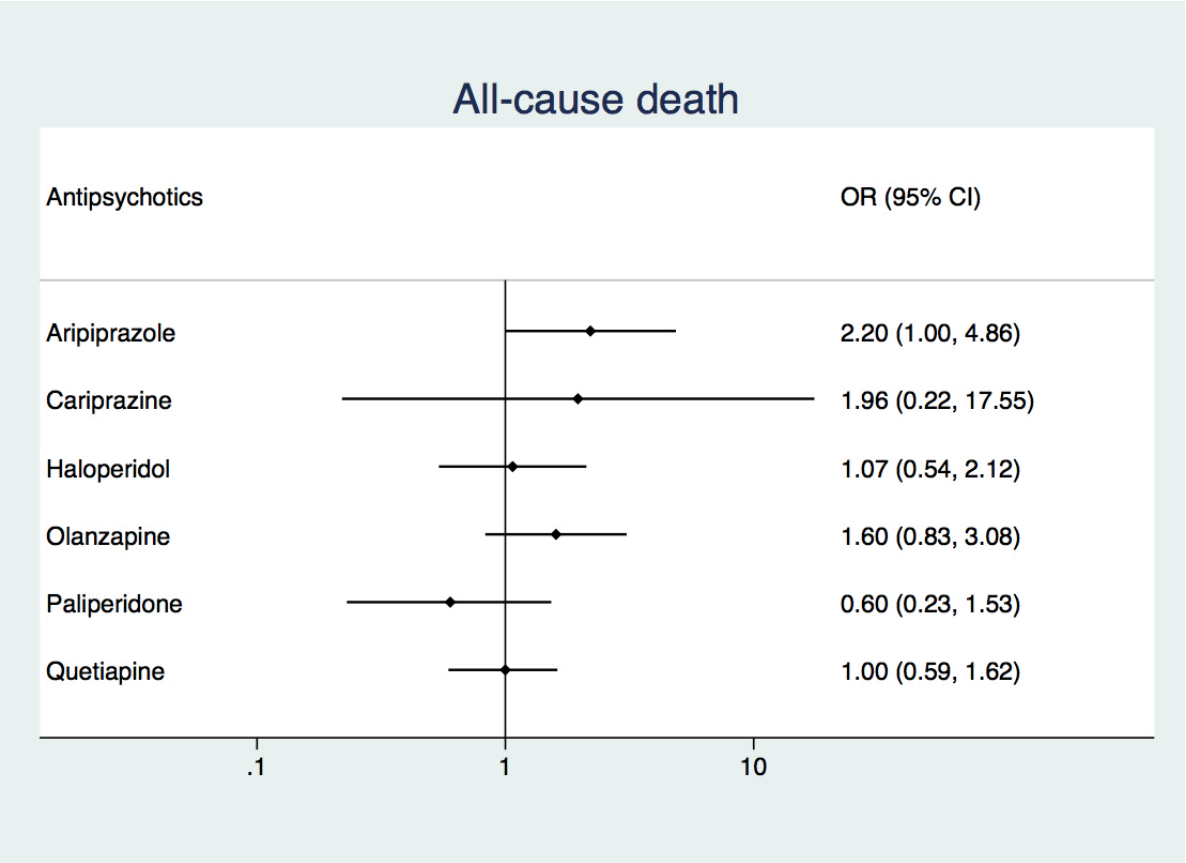
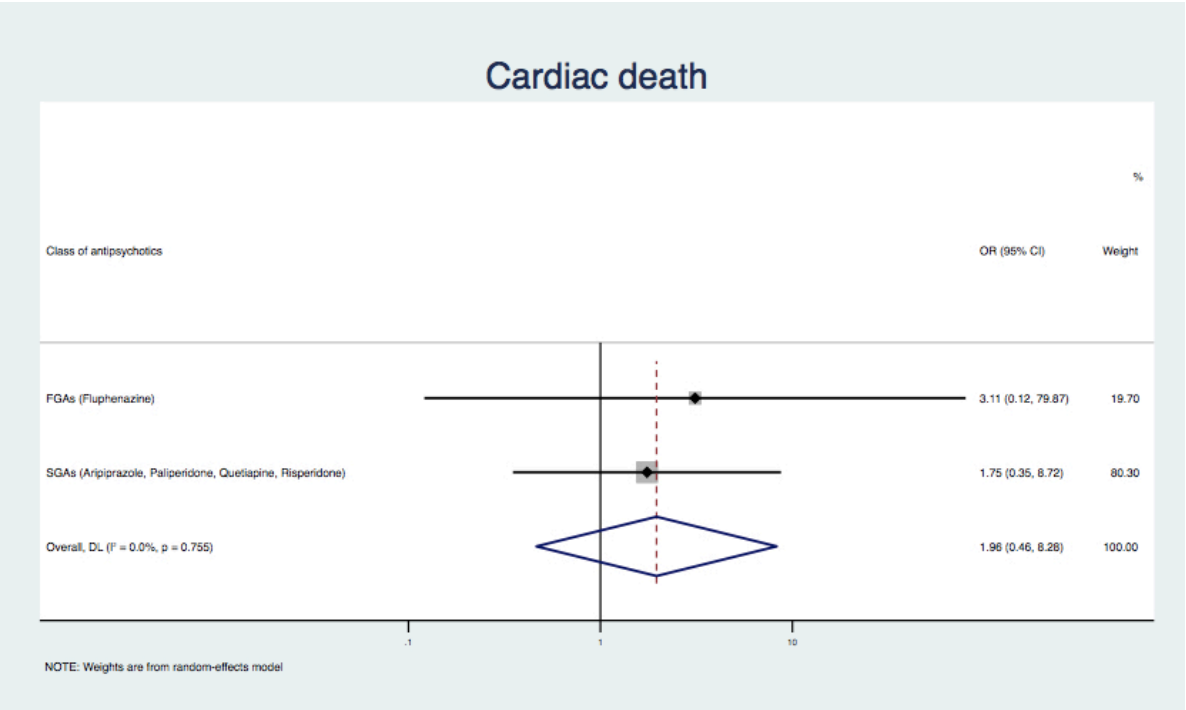
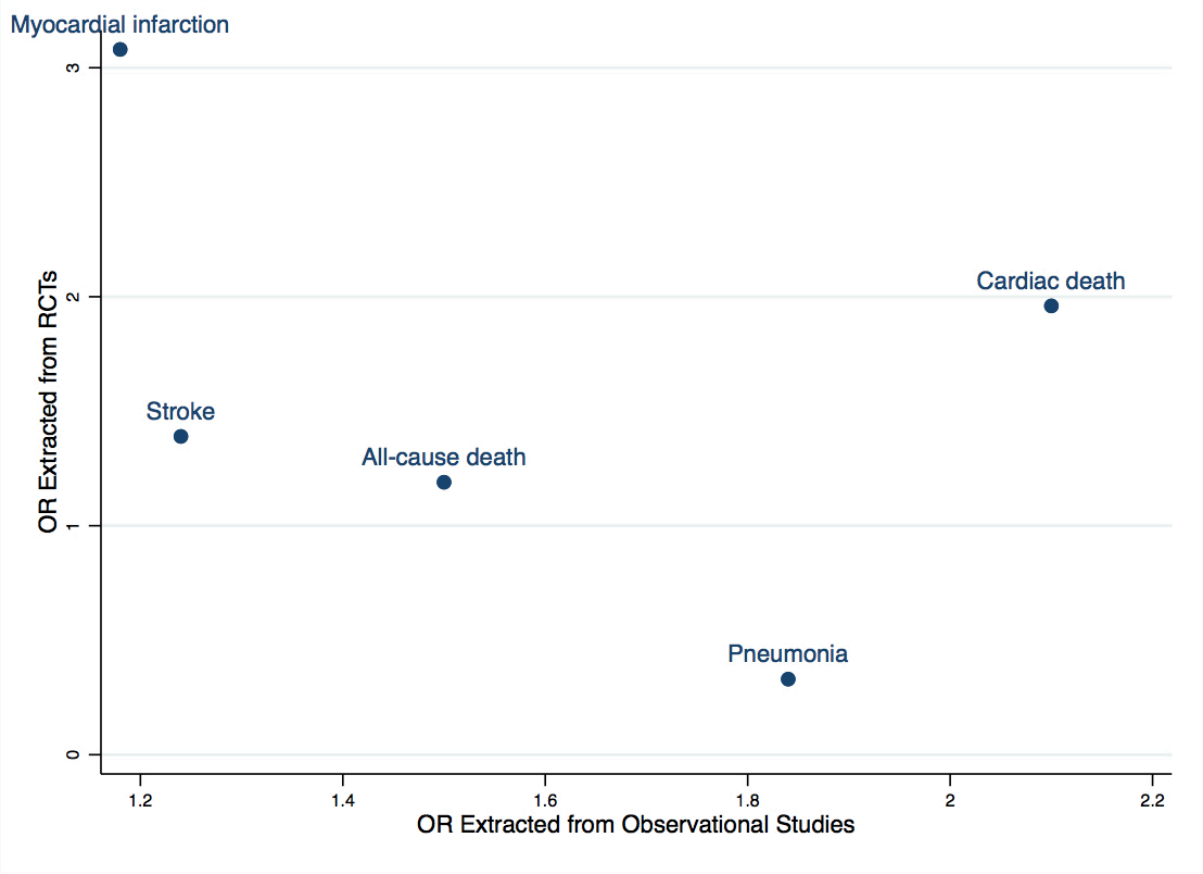


Figure S18. Cardiac death. All associations extracted from Rotella et al. (2020).



Appendix I: Distribution of RORs



Appendix J: Strength Assessment of Risk Ratios and Odds Ratios

Table S1. Strength assessment of RRs, adapted from Monson (1990) and Craun and Calderon (2006)

Increased risk	Decreased risk	Strength of association
1.0 - 1.2	0.9 - 1.0	None
1.2 - 1.5	0.7 - 0.9	Weak
> 1.5	< 0.9	Moderate to Strong

Table S2. Strength assessment of ORs, adapted from Rosenthal (1996)

Increased risk	Decreased risk	Strength of association
1.0 - 1.5	0.67 - 1.0	Weak
1.6 - 2.5	0.4 - 0.66	Moderate
2.6 - 9.9	0.10 - 0.39	Strong
>10.0	<0.10	Very strong

Supplementary References

Alfageh, B. H., Wang, Z., Mongkhon, P., Besag, F., Alhawassi, T. M., Brauer, R., & Wong, I. (2019). Safety and Tolerability of Antipsychotic Medication in Individuals with Autism Spectrum Disorder: A Systematic Review and Meta-Analysis. *Paediatric drugs*, *21*(3), 153–167. https://doi.org/10.1007/s40272-019-00333-x

Almandil, N. B., Liu, Y., Murray, M. L., Besag, F. M., Aitchison, K. J., & Wong, I. C. (2013). Weight gain and other metabolic adverse effects associated with atypical antipsychotic treatment of children and adolescents: a systematic review and meta-analysis. *Paediatric drugs*, *15*(2), 139–150. https://doi.org/10.1007/s40272-013-0016-6

Anonymous. (2000). Novel antipsychotics in the treatment of psychosis and aggression associated with dementia: A meta-analysis of randomized controlled clinical trials: Discussion 5. *International Psychogeriatrics*, *12*(SUPPL. 1), 279-280.

Arbaizar, B., Dierssen-Sotos, T., Gómez-Acebo, I., & Llorca, J. (2009). Aripiprazole in major depression and mania: meta-analyses of randomized placebo-controlled trials. *General hospital psychiatry*, *31*(5), 478–483. https://doi.org/10.1016/j.genhosppsych.2009.05.005

Ardizzone, I., Nardecchia, F., Marconi, A., Carratelli, T. I., & Ferrara, M. (2010). Antipsychotic medication in adolescents suffering from schizophrenia: a meta-analysis of randomized controlled trials. *Psychopharmacology bulletin*, *43*(2), 45–66.

Barbui, C., Conti, V., & Cipriani, A. (2014). Antipsychotic drug exposure and risk of venous thromboembolism: a systematic review and meta-analysis of observational studies. *Drug safety*, *37*(2), 79–90. https://doi.org/10.1007/s40264-013-0127-6

Barton, B. B., Segger, F., Fischer, K., Obermeier, M., & Musil, R. (2020). Update on weight-gain caused by antipsychotics: a systematic review and meta-analysis. *Expert opinion on drug safety*, *19*(3), 295–314. https://doi.org/10.1080/14740338.2020.1713091

Bosnjak Kuharic, D., Kekin, I., Hew, J., Rojnic Kuzman, M., & Puljak, L. (2019). Interventions for prodromal stage of psychosis. *The Cochrane database of systematic reviews*, *2019*(11), CD012236. https://doi.org/10.1002/14651858.CD012236.pub2

Chang, H. Y., Tseng, P. T., Stubbs, B., Chu, C. S., Li, D. J., Fornaro, M., Carvalho, A., Solmi, M., Veronese, N., Vieta, E., Chen, T. Y., Chen, Y. W., Lin, P. Y., & Chow, P. C. (2017). The efficacy and tolerability of paliperidone in mania of bipolar disorder: A preliminary meta-analysis. *Experimental and clinical psychopharmacology*, *25*(5), 422–433. https://doi.org/10.1037/pha0000148

Chung, A. K., & Chua, S. E. (2011). Effects on prolongation of Bazett's corrected QT interval of seven second-generation antipsychotics in the treatment of schizophrenia: a meta-analysis. *Journal of psychopharmacology (Oxford, England)*, *25*(5), 646–666. https://doi.org/10.1177/0269881110376685

Cipriani, A., Rendell, J. M., & Geddes, J. (2009). Olanzapine in long-term treatment for bipolar disorder. *The Cochrane database of systematic reviews*, (1), CD004367. https://doi.org/10.1002/14651858.CD004367.pub2

Citrome L. (2019). Cariprazine for bipolar depression: What is the number needed to treat, number needed to harm and likelihood to be helped or harmed?. *International journal of clinical practice*, *73*(10), e13397. https://doi.org/10.1111/ijcp.13397

Craun, G. F., & Calderon, R. L. (2005). How to interpret epidemiological associations. *Nutrients in Drinking Water*, *108*, 108-115.

Cooper, H., Mishriky, R., & Reyad, A. A. (2020). Efficacy and Safety of Cariprazine in Acute Management of Psychiatric Disorders: a Meta-Analysis of Randomized Controlled Trials. *Psychiatria Danubina*, *32*(1), 36–45. https://doi.org/10.24869/psyd.2020.36

De Hert, M., Yu, W., Detraux, J., Sweers, K., van Winkel, R., & Correll, C. U. (2012). Body weight and metabolic adverse effects of asenapine, iloperidone, lurasidone and paliperidone in the treatment of schizophrenia and bipolar disorder: a systematic review and exploratory meta-analysis. *CNS drugs*, *26*(9), 733–759. https://doi.org/10.2165/11634500-000000000-00000

Demyttenaere, K., Detraux, J., Racagni, G., & Vansteelandt, K. (2019). Medication-Induced Akathisia with Newly Approved Antipsychotics in Patients with a Severe Mental Illness: A Systematic Review and Meta-Analysis. *CNS drugs*, *33*(6), 549–566. https://doi.org/10.1007/s40263-019-00625-3

Depping, A. M., Komossa, K., Kissling, W., & Leucht, S. (2010). Second-generation antipsychotics for anxiety disorders. *The Cochrane database of systematic reviews*, (12), CD008120. https://doi.org/10.1002/14651858.CD008120.pub2

Derry, S., & Moore, R. A. (2007). Atypical antipsychotics in bipolar disorder: systematic review of randomised trials. *BMC psychiatry*, *7*, 40. https://doi.org/10.1186/1471-244X-7-40

Domecq, J. P., Prutsky, G., Leppin, A., Sonbol, M. B., Altayar, O., Undavalli, C., Wang, Z., Elraiyah, T., Brito, J. P., Mauck, K. F., Lababidi, M. H., Prokop, L. J., Asi, N., Wei, J., Fidahussein, S., Montori, V. M., & Murad, M. H. (2015). Clinical review: Drugs commonly associated with weight change: a systematic review and meta-analysis. *The Journal of clinical endocrinology and metabolism*, *100*(2), 363–370. https://doi.org/10.1210/jc.2014-3421

Douglas-Hall, P., Curran, S., Bird, V., & Taylor, D. (2011). Aripiprazole: a review of its use in the treatment of irritability associated with autistic disorder patients aged 6-17. *Journal of central nervous system disease*, *3*, 143–153. https://doi.org/10.4137/JCNSD.S4140

Fang, F., Sun, H., Wang, Z., Ren, M., Calabrese, J. R., & Gao, K. (2016). Antipsychotic Drug-Induced Somnolence: Incidence, Mechanisms, and Management. *CNS drugs*, *30*(9), 845–867. https://doi.org/10.1007/s40263-016-0352-5

Fang, F., Wang, Z., Wu, R., Calabrese, J. R., & Gao, K. (2017). Is there a 'weight neutral' second-generation antipsychotic for bipolar disorder?. *Expert review of neurotherapeutics*, *17*(4), 407–418. https://doi.org/10.1080/14737175.2016.1276284

Flank, J., Sung, L., Dvorak, C. C., Spettigue, W., & Dupuis, L. L. (2014). The safety of olanzapine in young children: a systematic review and meta-analysis. *Drug safety*, *37*(10), 791–804. https://doi.org/10.1007/s40264-014-0219-y

Fung, L. K., Mahajan, R., Nozzolillo, A., Bernal, P., Krasner, A., Jo, B., Coury, D., Whitaker, A., Veenstra-Vanderweele, J., & Hardan, A. Y. (2016). Pharmacologic Treatment of Severe Irritability and Problem Behaviors in Autism: A Systematic Review and Meta-analysis. *Pediatrics*, *137 Suppl 2*, S124–S135. https://doi.org/10.1542/peds.2015-2851K

Fusar-Poli, P., Kempton, M. J., & Rosenheck, R. A. (2013). Efficacy and safety of second-generation long-acting injections in schizophrenia: a meta-analysis of randomized-controlled trials. *International clinical psychopharmacology*, *28*(2), 57–66. https://doi.org/10.1097/YIC.0b013e32835b091f

Gao, K., Ganocy, S. J., Gajwani, P., Muzina, D. J., Kemp, D. E., & Calabrese, J. R. (2008). A review of sensitivity and tolerability of antipsychotics in patients with bipolar disorder or schizophrenia: focus on somnolence. *The Journal of clinical psychiatry*, *69*(2), 302–309. https://doi.org/10.4088/jcp.v69n0217

Gao, K., Kemp, D. E., Ganocy, S. J., Gajwani, P., Xia, G., & Calabrese, J. R. (2008). Antipsychotic-induced extrapyramidal side effects in bipolar disorder and schizophrenia: a systematic review. *Journal of clinical psychopharmacology*, *28*(2), 203–209. https://doi.org/10.1097/JCP.0b013e318166c4d5

Han, C., Pae, C. U., Wang, S. M., Lee, S. J., Patkar, A. A., Masand, P. S., & Serretti, A. (2014). The potential role of atypical antipsychotics for the treatment of posttraumatic stress

disorder. *Journal of psychiatric research*, *56*, 72–81. https://doi.org/10.1016/j.jpsychires.2014.05.003

Hirsch, L. E., & Pringsheim, T. (2016). Aripiprazole for autism spectrum disorders (ASD). *The Cochrane database of systematic reviews*, *2016*(6), CD009043. https://doi.org/10.1002/14651858.CD009043.pub3

Hollis, C., Pennant, M., Cuenca, J., Glazebrook, C., Kendall, T., Whittington, C., Stockton, S., Larsson, L., Bunton, P., Dobson, S., Groom, M., Hedderly, T., Heyman, I., Jackson, G. M., Jackson, S., Murphy, T., Rickards, H., Robertson, M., & Stern, J. (2016). Clinical effectiveness and patient perspectives of different treatment strategies for tics in children and adolescents with Tourette syndrome: a systematic review and qualitative analysis. *Health technology assessment (Winchester, England)*, *20*(4), 1–viii. https://doi.org/10.3310/hta20040

Hsu, W. T., Esmaily-Fard, A., Lai, C. C., Zala, D., Lee, S. H., Chang, S. S., & Lee, C. C. (2017). Antipsychotics and the Risk of Cerebrovascular Accident: A Systematic Review and Meta-Analysis of Observational Studies. *Journal of the American Medical Directors Association*, *18*(8), 692–699. https://doi.org/10.1016/j.jamda.2017.02.020

Huang, K. L., Fang, C. J., Hsu, C. C., Wu, S. I., Juang, J. J., & Stewart, R. (2017). Myocardial infarction risk and antipsychotics use revisited: a meta-analysis of 10 observational studies. *Journal of psychopharmacology (Oxford, England)*, *31*(12), 1544–1555. https://doi.org/10.1177/0269881117714047

Jesner, O. S., Aref-Adib, M., & Coren, E. (2007). Risperidone for autism spectrum disorder. *The Cochrane database of systematic reviews*, (1), CD005040. https://doi.org/10.1002/14651858.CD005040.pub2

Kishi, T., Matsuda, Y., Iwata, N., & Correll, C. U. (2013). Antipsychotics for cocaine or psychostimulant dependence: systematic review and meta-analysis of randomized, placebo-controlled trials. *The Journal of clinical psychiatry*, *74*(12), e1169–e1180. https://doi.org/10.4088/JCP.13r08525

Kishi, T., Matsunaga, S., & Iwata, N. (2016). Mortality Risk Associated With Long-acting Injectable Antipsychotics: A Systematic Review and Meta-analyses of Randomized Controlled Trials. *Schizophrenia bulletin*, *42*(6), 1438–1445. https://doi.org/10.1093/schbul/sbw043

Kishi, T., Oya, K., & Iwata, N. (2016). Long-Acting Injectable Antipsychotics for Prevention of Relapse in Bipolar Disorder: A Systematic Review and Meta-Analyses of Randomized Controlled Trials. *The international journal of neuropsychopharmacology*, *19*(9), pyw038. https://doi.org/10.1093/ijnp/pyw038

Kishi, T., Sevy, S., Chekuri, R., & Correll, C. U. (2013). Antipsychotics for primary alcohol dependence: a systematic review and meta-analysis of placebo-controlled trials. *The Journal of clinical psychiatry*, *74*(7), e642–e654. https://doi.org/10.4088/JCP.12r08178

Komossa, K., Depping, A. M., Gaudchau, A., Kissling, W., & Leucht, S. (2010). Second-generation antipsychotics for major depressive disorder and dysthymia. *The Cochrane database of systematic reviews*, (12), CD008121. https://doi.org/10.1002/14651858.CD008121.pub2

Komossa, K., Depping, A. M., Meyer, M., Kissling, W., & Leucht, S. (2010). Second-generation antipsychotics for obsessive compulsive disorder. *The Cochrane database of systematic reviews*, (12), CD008141. https://doi.org/10.1002/14651858.CD008141.pub2

Krause, M., Huhn, M., Schneider-Thoma, J., Rothe, P., Smith, R. C., & Leucht, S. (2018). Antipsychotic drugs for elderly patients with schizophrenia: A systematic review and meta-analysis. *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology*, *28*(12), 1360–1370. https://doi.org/10.1016/j.euroneuro.2018.09.007

Kucukgoncu, S., Guloksuz, S., Celik, K., Bahtiyar, M. O., Luykx, J. J., Rutten, B., & Tek, C. (2020). Antipsychotic Exposure in Pregnancy and the Risk of Gestational Diabetes: A Systematic

Review and Meta-analysis. *Schizophrenia bulletin*, *46*(2), 311–318. https://doi.org/10.1093/schbul/sbz058

LaLonde, C. D., & Van Lieshout, R. J. (2011). Treating generalized anxiety disorder with second generation antipsychotics: a systematic review and meta-analysis. *Journal of clinical psychopharmacology*, *31*(3), 326–333. https://doi.org/10.1097/JCP.0b013e31821b2b3f

Leucht, S., Arbter, D., Engel, R. R., Kissling, W., & Davis, J. M. (2009). How effective are second-generation antipsychotic drugs? A meta-analysis of placebo-controlled trials. *Molecular psychiatry*, *14*(4), 429–447. https://doi.org/10.1038/sj.mp.4002136

Leucht, S., Corves, C., Arbter, D., Engel, R. R., Li, C., & Davis, J. M. (2009). Second-generation versus first-generation antipsychotic drugs for schizophrenia: a meta-analysis. *Lancet (London, England)*, *373*(9657), 31–41. https://doi.org/10.1016/S0140-6736(08)61764-X

Leucht, S., Pitschel-Walz, G., Abraham, D., & Kissling, W. (1999). Efficacy and extrapyramidal side-effects of the new antipsychotics olanzapine, quetiapine, risperidone, and sertindole compared to conventional antipsychotics and placebo. A meta-analysis of randomized controlled trials. *Schizophrenia research*, *35*(1), 51–68. https://doi.org/10.1016/s0920-9964(98)00105-4

Li, D. J., Tseng, P. T., Stubbs, B., Chu, C. S., Chang, H. Y., Vieta, E., Fornaro, M., Carvalho, A. F., Solmi, M., Veronese, N., Chen, T. Y., Chen, Y. W., Lin, P. Y., & Chow, P. C. (2017). Efficacy, safety and tolerability of aripiprazole in bipolar disorder: An updated systematic review and meta-analysis of randomized controlled trials. *Progress in neuro-psychopharmacology & biological psychiatry*, *79*(Pt B), 289–301. https://doi.org/10.1016/j.pnpbp.2017.06.023

Li, X. H., Zhong, X. M., Lu, L., Zheng, W., Wang, S. B., Rao, W. W., Wang, S., Ng, C. H., Ungvari, G. S., Wang, G., & Xiang, Y. T. (2020). The prevalence of agranulocytosis and related death in clozapine-treated patients: a comprehensive meta-analysis of observational studies. *Psychological medicine*, *50*(4), 583–594. https://doi.org/10.1017/S0033291719000369

Lieb, K., Völlm, B., Rücker, G., Timmer, A., & Stoffers, J. M. (2010). Pharmacotherapy for borderline personality disorder: Cochrane systematic review of randomised trials. *The British journal of psychiatry : the journal of mental science*, *196*(1), 4–12. https://doi.org/10.1192/bjp.bp.108.062984

Loy, J. H., Merry, S. N., Hetrick, S. E., & Stasiak, K. (2017). Atypical antipsychotics for disruptive behaviour disorders in children and youths. *The Cochrane database of systematic reviews*, *8*(8), CD008559. https://doi.org/10.1002/14651858.CD008559.pub3

Ma, H., Huang, Y., Cong, Z., Wang, Y., Jiang, W., Gao, S., & Zhu, G. (2014). The efficacy and safety of atypical antipsychotics for the treatment of dementia: a meta-analysis of randomized placebo-controlled trials. *Journal of Alzheimer's disease : JAD*, *42*(3), 915–937. https://doi.org/10.3233/JAD-140579

Maayan, N., Quraishi, S. N., David, A., Jayaswal, A., Eisenbruch, M., Rathbone, J., Asher, R., & Adams, C. E. (2015). Fluphenazine decanoate (depot) and enanthate for schizophrenia. *The Cochrane database of systematic reviews*, (2), CD000307. https://doi.org/10.1002/14651858.CD000307.pub2

Majer, I. M., Gaughran, F., Sapin, C., Beillat, M., & Treur, M. (2015). Efficacy, tolerability, and safety of aripiprazole once-monthly versus other long-acting injectable antipsychotic therapies in the maintenance treatment of schizophrenia: a mixed treatment comparison of double-blind randomized clinical trials. *Journal of market access & health policy*, *3*, 10.3402/jmahp.v3.27208. https://doi.org/10.3402/jmahp.v3.27208

Marshall, M., & Rathbone, J. (2011). Early intervention for psychosis. *The Cochrane database of systematic reviews*, (6), CD004718. https://doi.org/10.1002/14651858.CD004718.pub3

McQuire, C., Hassiotis, A., Harrison, B., & Pilling, S. (2015). Pharmacological interventions for challenging behaviour in children with intellectual disabilities: a systematic review and meta-analysis. *BMC psychiatry*, *15*, 303. https://doi.org/10.1186/s12888-015-0688-2

Monson, R. R. (1990). *Occupational epidemiology*. CRC press.

Myles, N., Myles, H., Xia, S., Large, M., Kisely, S., Galletly, C., Bird, R., & Siskind, D. (2018). Meta-analysis examining the epidemiology of clozapine-associated neutropenia. *Acta psychiatrica Scandinavica*, *138*(2), 101–109. https://doi.org/10.1111/acps.12898

Nashed, M. G., Restivo, M. R., & Taylor, V. H. (2011). Olanzapine-induced weight gain in patients with bipolar I disorder: a meta-analysis. *The primary care companion for CNS disorders*, *13*(6), PCC.11r01174. https://doi.org/10.4088/PCC.11r01174

Nussbaum, A. M., & Stroup, T. S. (2013). Drug information update: Paliperidone palmitate for schizophrenia. *Psychiatrist*, *37*(5), 164-166.

Oderda, L. H., Young, J. R., Asche, C. V., & Pepper, G. A. (2012). Psychotropic-related hip fractures: meta-analysis of first-generation and second-generation antidepressant and antipsychotic drugs. *The Annals of pharmacotherapy*, *46*(7-8), 917–928. https://doi.org/10.1345/aph.1Q589

Orr, C., Deshpande, S., Sawh, S., Jones, P. M., & Vasudev, K. (2017). Asenapine for the Treatment of Psychotic Disorders. *Canadian journal of psychiatry. Revue canadienne de psychiatrie*, *62*(2), 123–137. https://doi.org/10.1177/0706743716661324

Oya, K., Kishi, T., & Iwata, N. (2015). Efficacy and tolerability of aripiprazole once monthly for schizophrenia: a systematic review and meta-analysis of randomized controlled trials. *Neuropsychiatric disease and treatment*, *11*, 2299–2307. https://doi.org/10.2147/NDT.S91397

Pae, C. U., Lim, H. K., Peindl, K., Ajwani, N., Serretti, A., Patkar, A. A., & Lee, C. (2008). The atypical antipsychotics olanzapine and risperidone in the treatment of posttraumatic stress disorder: a meta-analysis of randomized, double-blind, placebo-controlled clinical trials. *International clinical psychopharmacology*, *23*(1), 1–8. https://doi.org/10.1097/YIC.0b013e32825ea324

Prajapati, A. R., Wilson, J., Song, F., & Maidment, I. (2018). Second-generation antipsychotic long-acting injections in bipolar disorder: Systematic review and meta-analysis. *Bipolar disorders*, *20*(8), 687–696. https://doi.org/10.1111/bdi.12707

Pringsheim, T., Lam, D., Ching, H., & Patten, S. (2011). Metabolic and neurological complications of second-generation antipsychotic use in children: a systematic review and meta-analysis of randomized controlled trials. *Drug safety*, *34*(8), 651–668. https://doi.org/10.2165/11592020-000000000-00000

Rodrigues, R., Lai, M. C., Beswick, A., Gorman, D. A., Anagnostou, E., Szatmari, P., Anderson, K. K., & Ameis, S. H. (2021). Practitioner Review: Pharmacological treatment of attention-deficit/hyperactivity disorder symptoms in children and youth with autism spectrum disorder: a systematic review and meta-analysis. *Journal of child psychology and psychiatry, and allied disciplines*, *62*(6), 680–700. https://doi.org/10.1111/jcpp.13305

Rosenthal, J. A. (1996). Qualitative descriptors of strength of association and effect size. *Journal of Social Service Research*, *21*(4), 37-59.

Ruxton, K., Woodman, R. J., & Mangoni, A. A. (2015). Drugs with anticholinergic effects and cognitive impairment, falls and all-cause mortality in older adults: A systematic review and meta-analysis. *British journal of clinical pharmacology*, *80*(2), 209–220. https://doi.org/10.1111/bcp.12617

Sampson, S., Hosalli, P., Furtado, V. A., & Davis, J. M. (2016). Risperidone (depot) for schizophrenia. *The Cochrane database of systematic reviews*, *4*, CD004161. https://doi.org/10.1002/14651858.CD004161.pub2

Silva, M. T., Zimmermann, I. R., Galvao, T. F., & Pereira, M. G. (2013). Olanzapine plus fluoxetine for bipolar disorder: a systematic review and meta-analysis. *Journal of affective disorders*, *146*(3), 310–318. https://doi.org/10.1016/j.jad.2012.11.001

Siskind, D., Sidhu, A., Cross, J., Chua, Y. T., Myles, N., Cohen, D., & Kisely, S. (2020). Systematic review and meta-analysis of rates of clozapine-associated myocarditis and cardiomyopathy. *The Australian and New Zealand journal of psychiatry*, *54*(5), 467–481. https://doi.org/10.1177/0004867419898760

Spielmans, G. I., Berman, M. I., Linardatos, E., Rosenlicht, N. Z., Perry, A., & Tsai, A. C. (2013). Adjunctive atypical antipsychotic treatment for major depressive disorder: a meta-analysis of depression, quality of life, and safety outcomes. *PLoS medicine*, *10*(3), e1001403. https://doi.org/10.1371/journal.pmed.1001403

Srisurapanont, M., Maneeton, B., & Maneeton, N. (2004). Quetiapine for schizophrenia. *The Cochrane database of systematic reviews*, *2004*(2), CD000967. https://doi.org/10.1002/14651858.CD000967.pub2

Stafford, M. R., Mayo-Wilson, E., Loucas, C. E., James, A., Hollis, C., Birchwood, M., & Kendall, T. (2015). Efficacy and safety of pharmacological and psychological interventions for the treatment of psychosis and schizophrenia in children, adolescents and young adults: a systematic review and meta-analysis. *PloS one*, *10*(2), e0117166. https://doi.org/10.1371/journal.pone.0117166

Suttajit, S., Srisurapanont, M., Maneeton, N., & Maneeton, B. (2014). Quetiapine for acute bipolar depression: a systematic review and meta-analysis. *Drug design, development and therapy*, *8*, 827–838. https://doi.org/10.2147/DDDT.S63779

Tan, E., Lexomboon, D., Sandborgh-Englund, G., Haasum, Y., & Johnell, K. (2018). Medications That Cause Dry Mouth As an Adverse Effect in Older People: A Systematic Review and Metaanalysis. *Journal of the American Geriatrics Society*, *66*(1), 76–84. https://doi.org/10.1111/jgs.15151

Tek, C., Kucukgoncu, S., Guloksuz, S., Woods, S. W., Srihari, V. H., & Annamalai, A. (2016). Antipsychotic-induced weight gain in first-episode psychosis patients: a meta-analysis of differential effects of antipsychotic medications. *Early intervention in psychiatry*, *10*(3), 193–202. https://doi.org/10.1111/eip.12251

Vancampfort, D., Vansteelandt, K., Correll, C. U., Mitchell, A. J., De Herdt, A., Sienaert, P., Probst, M., & De Hert, M. (2013). Metabolic syndrome and metabolic abnormalities in bipolar disorder: a meta-analysis of prevalence rates and moderators. *The American journal of psychiatry*, *170*(3), 265–274. https://doi.org/10.1176/appi.ajp.2012.12050620

Vita, A., De Peri, L., Siracusano, A., Sacchetti, E., & ATOM group (2013). Efficacy and tolerability of asenapine for acute mania in bipolar I disorder: meta-analyses of randomized-controlled trials. *International clinical psychopharmacology*, *28*(5), 219–227. https://doi.org/10.1097/YIC.0b013e32836290d2

Wang, H., Xiao, L., Wang, H. L., & Wang, G. H. (2020). Efficacy and safety of lurasidone versus placebo as adjunctive to mood stabilizers in bipolar I depression: A meta-analysis. *Journal of affective disorders*, *264*, 227–233. https://doi.org/10.1016/j.jad.2019.11.031

Willems, A. E., Tenback, D. E., Ingenhoven, T. J., & van Harten, P. N. (2014). Acute movement disorders associated with the use of second-generation antipsychotics in borderline personality disorder: a meta-analysis. *Journal of clinical psychopharmacology*, *34*(4), 520–522. https://doi.org/10.1097/JCP.0000000000000152

Yu, Z. H., Jiang, H. Y., Shao, L., Zhou, Y. Y., Shi, H. Y., & Ruan, B. (2016). Use of antipsychotics and risk of myocardial infarction: a systematic review and meta-analysis. *British journal of clinical pharmacology*, *82*(3), 624–632. https://doi.org/10.1111/bcp.12985

Zhao, M. J., Qin, B., Wang, J. B., Zhang, Y. P., Zhao, J. T., Mao, Y. G., Zhang, X. Y., & Zhang, R. L. (2018). Efficacy and Acceptability of Cariprazine in Acute Exacerbation of Schizophrenia:

Meta-Analysis of Randomized Placebo-Controlled Trials. *Journal of clinical psychopharmacology*, *38*(1), 55–59. https://doi.org/10.1097/JCP.0000000000000834