Applications of Meta-Analytic Approaches to Inform Estimates for the Global Burden of Mental Disorders

Modhurima Moitra

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2021

Reading Committee
Pamela Y. Collins, Chair
Alize Ferrari
Damian Santomauro
Theo Vos

Program Authorized to Offer Degree:
Global Health

University of Washington

Abstract

Applications of Meta-Analytic Approaches to Inform Estimates for the Global Burden of Mental Disorders

Modhurima Moitra

Chair of the Supervisory Committee:
Pamela Y. Collins, MD, MPH
Department of Global Health
Department of Psychiatry and Behavioral Sciences

Mental disorders are some of the most important contributors to disability and indirectly to mortality worldwide. Despite the existence of effective interventions, treatment coverage for common and highly prevalent mental disorders such as major depressive disorder (MDD) is remarkably low. Given the important role mental disorders play in contributing to health loss, it is important to develop improved population-level estimates of the burden of mental disorders that may better inform resource allocation and prioritize prevention and treatment for mental health.

The process of developing population-level estimates such as those from the Global Burden of Disease (GBD) study typically involves thorough reviews of the available evidence and application of appropriate methods to arrive at usable estimates of the global burden of mental disorders. However, certain challenges exist in this process. Available data may be derived from a variety of study settings or constrained to select locations that may contribute to bias in existing population-level estimates. Additionally, important components of MDD care such as treatment coverage and efficacy that will likely alter the known burden of MDD have not been systematically quantified in the recent literature.

To this end, this work utilizes a meta-analytic measurement framework to address these limitations in three areas that are important for developing better estimates but remain relatively unexplored to date. First, a comparative assessment of the risk of suicide associated with mental disorders is conducted via an updated systematic review and updated meta-regression methods. Second, the efficacy of available interventions for MDD are explored in a novel network meta-analytic approach to better account for between-study heterogeneity. Third, an updated systematic review and Bayesian meta-regression analysis of treatment coverage data between 2000 and 2019 was conducted to obtain treatment coverage estimates for six modalities of MDD care across income levels and geography.

The work reported here collectively provide more accurate estimates that may inform the larger process of developing burden measures for mental disorders. There are several important implications of this work. Updated estimates of the risk of suicide associated with mental disorders may better inform population attributable fractions for suicide and subsequently fatal burden associated with mental disorders. Incorporating estimates of treatment efficacy and coverage for MDD may potentially inform us about the burden of MDD that can be reduced or altered in the presence of adequate and effective treatment. Overall, this work represents emerging areas in global mental health that can inform evidence-based benchmarking and priority setting to reduce the global burden of mental disorders.

Applications of Meta-Analytic Approaches to Inform Estimates for the Global Burden of Mental Disorders

Modhurima Moitra, MPH
moitra@uw.edu

Advisor:
Pamela Y. Collins, MD, MPH
Professor, Global Health, University of Washington
Professor, Psychiatry and Behavioral Sciences, University of Washington

Reading Committee:
Alize Ferrari, PhD
Affiliate Assistant Professor, Health Metrics Sciences, University of Washington

Damian Santomauro, PhD
Affiliate Assistant Professor, Health Metrics Sciences, University of Washington

Theo Vos, MD, PhD
Professor, Health Metrics Sciences, University of Washington

Graduate Student Representative:
Ali Rowhani-Rahbar, MD, PhD
Professor, Epidemiology, University of Washington

List of Contents:

0. Introduction.

0.1 Overview.

Mental disorders (MDs) are among the leading causes of years lived with disability (YLDs) globally. In addition to disability, they are also known to precipitate fatal outcomes such as suicide [1-3]. Large gaps exist in the treatment of MDs, particularly in low and middle-income settings where up to 90% of those with MDs do not receive appropriate treatment or care [4]. The Global Burden of Disease (GBD) study currently provides yearly estimates of the disease burden imposed by twelve mental disorders [5]. These GBD studies have played an instrumental role in recognizing MDs to be a major contributor to disease burden in comparison to more salient medical conditions. Given the major role that MDs play in contributing to overall health loss and indirectly to mortality, it is important to obtain accurate and precise estimates of disease burden imposed by MDs in order to effectively inform priority-setting efforts for mental health at national and global levels including mental health financing, and the development and delivery of services and interventions. This involves a thorough review of existing data and the application of appropriate statistical techniques to generate usable measures that directly quantify dimensions of disease burden (mortality, disability, and quality of health access to name a few). However, this process of quantifying MD burden estimates is difficult for several reasons. Firstly, there are sparse data especially from low and middle-income countries (LMICs) where MD burden is understudied and likely goes unrecognized and untreated . Secondly, available data are subject to methodological constraints that influence how cases are identified, how comparisons can be made between studies, and subsequently how outcomes of MDs, such as suicide risk, is measured. Thirdly, little is known about the severity distribution of MDs and how this varies by country and health care access. Severity distributions quantify the range of functional health loss in a population due to health conditions. Efforts to quantify the efficacy of available MD interventions, treatment rates and how these alter severity distributions have not been incorporated within a burden of disease framework.

0.2 Specific Aims.

We seek to address these limitations in the form of three research aims. The first aim addresses the problems of data sparsity and methodological constraints via a systematic review and analysis of all available and usable evidence on MDs as risk factors for suicide. This work will improve upon existing modeling techniques to better handle sparse and complex data environments and incorporate important methodological information in order to generate robust estimates of suicide risk. These estimates will also be useful in calculating more precise population attributable fractions (PAFs) for suicide within the GBD framework. The second and third aims focus on major depressive disorder (MDD) as case studies as MDD is the most prevalent mood disorder and the greatest cause of MD-related burden. However, methods proposed in these two aims may be applied more generally to other MDs as well. The second aim seeks to generate evidence on treatment efficacy for MDD using both direct and indirect comparisons. The third aim seeks to review the evidence on treatment rates for MDD and analytically quantify estimated treatment rates and corresponding gaps in treatment coverage by geography and resource-setting. The results generated via the second and third aims will be useful inputs to the re-estimation process to obtain more informed severity distributions for MDD within the GBD framework.

Aim I: Apply multilevel meta-regression techniques to obtain updated estimates of suicide risk.

a. Conduct an updated systematic review on MDs as risk factors for suicide.
b. Apply multilevel meta-regression techniques to explore characteristics that contribute to variation in suicide risk due to MDs.

Aim II: Assess the efficacy of treatments for MDD using an application of network meta-analytic techniques.

a. Compile updated datasets from existing systematic reviews conducted by leading collaborators.
b. Apply network meta-analyses to quantify treatment efficacy for multiple treatment comparisons by treatment class.

Aim III: Review the evidence on treatment rates for MDD from an updated systematic review:

a. Update a systematic review to estimate treatment rates of MDD by type of service used.
b. Analytically assess sources of heterogeneity such as country income level, sample characteristics, treatment type, disorder severity, and adequacy of treatment.
c. Estimate treatment rates and corresponding gaps in treatment coverage by geography and resource-setting.

0.3 Background and Scientific Context:

In the first research aim, we estimate the risk of suicide attributable to MDs using improved meta-analytic approaches to account for sources of measurement error. This area of work is of particular interest because existing reviews of the evidence do not fully take measurement error into account and therefore limit our ability to accurately assess suicide risk. This is an important limitation especially because suicide is well-established in the literature to be a preventable mental health outcome. Therefore, an accurate assessment of the attributable risk would make a more compelling case for MDs to be considered as a risk factor for suicide mortality in population-based estimation efforts. The Global Burden of Disease Study 2019 (GBD 2019) estimated suicide to be the 12th leading cause of years of life lost (YLLs) in 2019, up from the 13th in 2010. Suicide is a major public health concern globally; over 700,000 deaths occur annually due to suicide worldwide. Suicide affects populations from all age groups. Of all suicide deaths, 58% occur between the ages of 15–49 years which indicates the large magnitude of potentially productive years of life lost [6]. However, older age is also typically associated with major life events and complex transitions, which may be precipitants for suicide. According to GBD 2019 results, death rates from suicide are much higher in older age groups than in younger ones (24.53 per 100,000 for ages greater than 70 and 14.25 per 100,000 between 50-69 years vs 11.19 per 100,000 between 15-49 years) [5, 6]. Existing literature reviews indicate that suicide risk is elevated with diagnoses of mood disorders, schizophrenia, and alcohol use disorders [7-10].However, the GBD captures suicide related deaths under ‘Injuries’ in the GBD cause hierarchy as the direct (rather than underlying) cause of death, while MDs are grouped separately within ‘Non-Communicable’ diseases. To estimate the proportion of suicide related deaths due to mental and substance use disorders (MSDs), previous work by Ferrari and collaborators applied comparative risk assessment (CRA) methodology to GBD 2010 data. Their review showed that MSDs moved from the 5th to the 3rd leading contributor to disease burden once suicide YLLs were reassigned to MSDs [2].

Given these findings, it is important to periodically re-assess the effect of MDs on suicide risk by updating the evidence base for the relationship between MDs and suicide. The work by Ferrari and Collaborators was conducted using GBD 2010 data and systematic literature reviews conducted in 2014. However, analyses from this previous work mostly comprised pooled RRs obtained using a simple meta-analytic approach. It is important to understand the impact of differences in study methods and subgroups on reported magnitude of suicide risk which, to date, remain largely unexplored. Existing reviews typically define broad inclusion criteria including non-representative samples or high-risk subgroups. These may introduce measurement error and subsequently result in potentially inflated risk estimates. The analytic approaches in these reviews also typically involve pooled analyses that do not fully adjust for sources of measurement error (e.g. study design, sampling strategy, case definition, etc.) or assess the impact of other ecological variables (e.g. location or individual characteristics).

This current systematic literature review aims to account for these limitations by providing updated suicide risk estimates adjusted in a meta-regression for key demographic and methodological covariates that may contribute to variation in study-reported estimates. We will use the earlier review by Ferrari et al. (2014) as the basis and template for this work. New data will be added, and the search strings for relevant papers will be expanded to incorporate self-harm and self-injury (deliberate suicidal behaviors with an intent to die) to maximize the scope of potentially eligible studies. A multilevel meta-regression approach will be used to account for data that are inherently nested or clustered by study. Relative risk estimates for suicide adjusted for select covariates will be predicted by disorder and sex.

In addition to an analysis of risk factors covered in Aim I thus far, efforts to improve MD burden estimates also need to account for treatment efficacy and geographic variation in treatment use. These are important to consider as they may influence the severity distribution or the range of functional health loss due to MDs that is experienced across populations. Re-estimating the severity distribution of major depressive disorder (MDD) is of particular interest as MDD is one of the most prevalent MDs globally and a major risk factor for suicide. It is also one of the common MDs for which effective prevention and treatment strategies are known to exist. Current methods for estimating the severity distribution of MDD involve using data collected from the Short-Form 12 instrument, from a single mental health survey conducted in Australia. The Short-Form 12 is a self-report outcome measure that assesses the impact of physical and mental health on quality of life [11]. However, data from a single high-income country may not be entirely representative of severity of symptoms other countries. A potential solution is to incorporate in this analysis of severity, more globally representative data on treatment coverage by different forms of treatment. In addition to treatment coverage, it is reasonable to expect that the efficacy of available treatment options for MDD may vary by location and time. Therefore, it is important to incorporate MDD treatment efficacy as another input in the MDD severity estimation process. The second and third research aims sequentially explore these two components of the MDD severity estimation process in the form of two research questions. In the second aim, we will apply a Bayesian network meta-analytic approach to review the efficacy of available treatments for MDD. Several pharmacological, physical, and psychosocial interventions can effectively treat MDD [12]. However, there is mixed evidence for treatment efficacy from existing reviews

due to the limitation of direct comparisons. Existing efficacy reviews typically summarize treatment effects from published psychiatric trials of specific treatment comparisons to report a pooled treatment effect [13-17]. However, individual pairwise direct comparisons commonly reported in trials may not provide comparative efficacy estimates for all known treatment types. In order to address this limitation, a network meta-analysis (NMA) approach will be used to generate indirect evidence for treatment efficacy not otherwise directly estimated. Subsequently, a global effect size for the entire treatment 'network' can be generated using both direct and indirect evidence. In this analysis, we will apply a novel Bayesian network meta-analysis approach to produce pooled treatment efficacy estimates for MDD for major classes of pharmacological and psychological interventions.

In the third research aim, we a) review the evidence on variation in MDD treatment rates via a systematic review and b) analytically assess sources of heterogeneity. This work is of particular interest because treatment coverage for MDD is relatively low. Treatment coverage for MDD refers to the proportion of those with MDD who received treatment for their disorder. An estimated 7% to 28% of those with depression actually receive appropriate care and treatment [18]. As the quality and availability of healthcare will vary by geography and resource availability, it is plausible to assume that this will then affect the severity of symptoms (and therefore disability) experienced by those with MDD in the population. Therefore, it is also important to synthesize and consider the available evidence on treatment rates when estimating the severity and burden due to MDD. An updated systematic review of the literature from 2000 – present was conducted covering 83 countries. The second part of this aim analytically explored sources of heterogeneity that exist in the treatment rates dataset. These included variation by country income level, disorder type, disorder severity, treatment type, definition of adequate treatment, urbanicity, sample response rate, disorder recall thresholds (12 months or less), and others. Pooled estimates of treatment rates adjusted for important sources of heterogeneity were reported by location and treatment type. <>.

Scientific Context: Given the large population health impact of MDs, estimation of MD burden is particularly important. In this thesis, key parameters within the burden estimation process for MDs will be reviewed. These are premature suicide mortality due to MDs, variation in treatment efficacy and treatment rates. We examine suicide because although MDs are well-established risk factors for suicide, there are methodological limitations that may result in potentially biased risk estimates. Therefore, there is a need to improve statistical techniques to obtain more precise suicide risk estimates. Next, we choose to review treatment efficacy and rates. MDD is one of the leading causes of nonfatal burden worldwide. It is also associated with a high magnitude of suicide risk and despite being one of the most prevalent mental disorders, it is largely left untreated in many parts of the world. Therefore, it is important to quantify estimated gaps in treatment coverage by resource setting and geography. Quantifying these parameters has several challenges discussed above. The research aims outlined above seek to apply improved meta-analytic methods to better inform disease burden estimates for MDs by examining these three areas of work. This section provides general background to the three research aims along with the gaps in the literature they address.

0.4 Significance:

The primary goal of this work is to improve upon existing methods applied to MD burden estimation. The three proposed aims described above collectively serve to address

specific methodological limitations and subsequently arrive at more precise estimates of mental disorder burden. Comparative epidemiological assessments such as estimating the risk of suicide associated with MDs provides risk estimates that are useful in understanding the fatal burden attributed to MDs. The second and third aims provide important inputs on the efficacy of available treatment options and global treatment coverage for MDD. A thorough understanding of the efficacy of treatment options and the evidence on treatment coverage for MDD can in turn inform us about the burden of MDD that may be influenced (or reduced or altered) by the availability and quality of effective care. . Overall, this work aligns with population health priorities defined by major institutions. The United Nations 2030 Agenda for Sustainable Development calls for the reduction by one-third of premature mortality from non-communicable diseases and the promotion of mental health and well-being [19]. The WHO Special Initiative on Mental Health aims to ramp up services particularly in low and middle-income countries (LMICs) [20].The Grand Challenges in Global Mental Health initiative laid out by Collins et al. includes identification of risk and protective factors (Goal A), improve treatments and expand access to care (Goal C) and raising awareness of the global burden (Goal D) [21].The research aims proposed here collectively serve to inform these efforts as essential inputs to our understanding of the burden of MDs and subsequently contribute to the larger goal of evidence-based priority setting for health interventions and policy-making in global mental health.

The research aims proposed here seek to apply methodological improvements to specific areas to improve the MSD burden estimation process. The novel contributions for each aim are as follows:

Aim I: The first aim improves upon existing meta-analytic approaches to estimate suicide risk by using a multi-level meta-regression model to better account for sources of measurement error that have not been fully explored in earlier reviews. This approach is best suited for the data because they are inherently nested or clustered within larger groups. Most studies in our dataset report multiple effect sizes by disorder, age and/or sex and therefore are not independent observations. This multi-level approach will allow us to account for any within and between study dependencies in the data by: (1) better utilizing (in its nested approach) all the data available across studies and (2) more robustly explaining the heterogeneity in the data reported between studies to separate measurement error from variability that is ‘real’ and due to ecological variables. As an additional methodological improvement, we also use the Bayesian Meta-regression approach to obtain pooled relative risk estimates. These inputs generated here will inform population attributable fractions (PAFs) for suicide and subsequently help estimate the proportion of fatal burden attributed to MDs. .

Aim II: In our network meta-analysis for depression interventions, we make use of a novel Bayesian meta-regression application that incorporates between-study heterogeneity in the uncertainty intervals for treatment effects. This particular method has not been used to model depression treatment effect sizes before and therefore will allow us to obtain estimates with higher accuracy and a more informative margin of error in the available data. The inputs generated here may also contribute to re-estimating a severity distribution for MDD that incorporates comprehensive and geographically representative data on treatment efficacy. .

Aim III: We review the evidence on MDD treatment rates from an exhaustive systematic literature review, spanning two decades and eighty-three countries. We quantify important sources of heterogeneity arising from variation in disorder and treatment characteristics and report pooled adjusted treatment rates by location. This work will inform efforts to a) compile evidence on coverage and efficacy of interventions and b) disseminate evidence-based information to decision-makers on the impact of depression interventions. This work may also contribute to modeling efforts for future scenarios and potential trajectories of MDD burden given this current evidence on global variation in availability, treatment modalities, and quality of MDD care.

0.5 References:

1. Vigo D, Thornicroft G, Atun R. Estimating the true global burden of mental illness. The Lancet Psychiatry. 2016;3(2):171-8.
2. Ferrari AJ, Norman RE, Freedman G, Baxter AJ, Pirkis JE, Harris MG, et al. The burden attributable to mental and substance use disorders as risk factors for suicide: findings from the Global Burden of Disease Study 2010. PloS one. 2014;9(4):e91936.
3. Harris EC, Barraclough B. Suicide as an outcome for mental disorders. A meta-analysis. British journal of psychiatry. 1997;170(3):205-28
4. Patel V. Why mental health matters to global health. Transcultural psychiatry. 2014 Dec;51(6):777-89.
5. (IHME). GBD Compare [Available from: https://vizhub.healthdata.org/gbd-compare/.
6. Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, Abbasi-Kangevari M, Abbastabar H, Abd-Allah F, Abdelalim A, Abdollahi M. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet. 2020 Oct 17;396(10258):1204-22.
7. Inskip H, Harris C, Barraclough B. Lifetime risk of suicide for affective disorder, alcoholism and schizophrenia. The British Journal of Psychiatry. 1998;172(1):35-7.
8. Harris EC, Barraclough B. Suicide as an outcome for mental disorders. A meta-analysis. British journal of psychiatry. 1997;170(3):205-28.
9. Cavanagh JT, Carson AJ, Sharpe M, Lawrie SM. Psychological autopsy studies of suicide: a systematic review. Psychological medicine. 2003;33(3):395-405.
10. Harris C, Barraclough B. Excess mortality of mental disorder. The British journal of psychiatry. 1998;173(1):11-53.
11. Ware Jr JE, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Medical care. 1996 Mar 1:220-33.
12. Patel V, Chisholm D, Parikh R, Charlson FJ, Degenhardt L, Dua T, Ferrari AJ, Hyman S, Laxminarayan R, Levin C, Lund C. Global priorities for addressing the burden of mental, neurological, and substance use disorders.
13. Furukawa TA, Streiner D, Young LT, Kinoshita Y. Antidepressants plus benzodiazepines for major depression. Cochrane Database of Systematic Reviews. 2001(3).
14. Hetrick SE, McKenzie JE, Cox GR, Simmons MB, Merry SN. Newer generation antidepressants for depressive disorders in children and adolescents. Cochrane Database of Systematic Reviews. 2012(11).
15. Komossa K, Depping AM, Gaudchau A, Kissling W, Leucht S. Second-generation antipsychotics for major depressive disorder and dysthymia. Cochrane Database of Systematic Reviews. 2010(12).
16. Churchill R, Moore TH, Furukawa TA, Caldwell DM, Davies P, Jones H, Shinohara K, Imai H, Lewis G, Hunot V. 'Third wave'cognitive and behavioural therapies versus treatment as usual for depression. Cochrane Database of Systematic Reviews. 2013(10).

17. Peng L, Min S, Wei K, Ziemann-Gimmel P. Different regimens of intravenous sedatives or hypnotics for electroconvulsive therapy (ECT) in adult patients with depression. Cochrane database of systematic reviews. 2014(4).
18. Chisholm D, Sweeny K, Sheehan P, Rasmussen B, Smit F, Cuijpers P, Saxena S. Scaling-up treatment of depression and anxiety: a global return on investment analysis. The Lancet Psychiatry. 2016 May 1;3(5):415-24.
19. Persaud N, Dagher R. The United Nations: 2030 Sustainable Development Goals Agenda. InThe Role of Monitoring and Evaluation in the UN 2030 SDGs Agenda 2021 (pp. 1-41). Palgrave Macmillan, Cham.
20. World Health Organization. The WHO special initiative for mental health (2019-2023). Universal Health Coverage for Mental Health. 2019.
21. Collins PY, Patel V, Joestl SS, March D, Insel TR, Daar AS, Bordin IA, Costello EJ, Durkin M, Fairburn C, Glass RI. Grand challenges in global mental health. Nature. 2011 Jul;475(7354):27-30.

1. Chapter 1: Estimating the risk of suicide attributable to mental disorders: A systematic review and meta-regression analysis.

[Note: The contents of this chapter are published in the following paper: Moitra M, Santomauro D, Degenhardt L, Collins PY, Whiteford H, Vos T, Ferrari A. Estimating the risk of suicide associated with mental disorders: A systematic review and meta-regression analysis. Journal of psychiatric research. 2021 Mar 2.]

1.1 **Introduction:**

The Global Burden of Disease Study (GBD) estimated suicide to be the 15th leading cause of years of life lost (YLLs) in 2019, up from the 18th in 2010 (Vos et al. 2020). Suicide is a major public health concern globally with nearly 800,000 deaths due to suicide worldwide (Ferrari et al. 2014, Vigo et al. 2016). It also affects people across the life course – children, adolescents, as well as older adults. Of all suicide deaths, 58% occur between the ages of 15-49 years which indicates the large magnitude of potentially productive years of life lost (Vos et al. 2020). In older age, major life events and complex transitions may be precipitants for suicide. According to GBD 2019 results, death rates from suicide are much higher in older age groups than in younger ones (24·53 per 100,000 for ages greater than 70 and 14·25 per 100,000 between 50-69 years vs 11·19 per 100,000 between 15-49 years). There are also differences in sex-specific suicide rates, with the suicide death rate among males being consistently higher than that of females over time (13·5 per 100,000 in males vs 6·1 in females in 2019) (Vos et al. 2020). Globally, India and China continue to report the highest number of deaths by suicide along with distinctly dominant suicide methods such as poisoning, hanging, and self-immolation (Ferrari et al. 2014, James et al. 2017, Wu et al. 2012, Adjacic-Gross et al. 2008). This suggests a continued need to assess underlying risk factors and potential target areas for suicide prevention interventions in low and middle-income countries.

The risk of suicide attributable to mental disorders (MDs) is well established in the literature (Ferrari et al. 2014). Existing reviews indicate that suicide risk is particularly high at the time of diagnosis for affective disorders, and schizophrenia (Inskip et al. 1998, Harris et al. 1997, Cavanagh et al. 2003, Harris et al. 1998, Li et al. 2011, Walker et al. 2015). Up to 80% of suicide deaths have been attributed to a mental or substance use disorder in high income countries compared to approximately 70% in low income countries (Ferrari et al. 2014). However, the World Health Organization (WHO) classification followed by the GBD study captures suicide related deaths under 'Injuries' as the direct (rather than underlying) cause of death, while MDs are grouped separately within 'Non-Communicable diseases (Ferrari et al. 2014).To estimate the proportion of suicide related deaths due to mental and substance use disorders (MSDs), previous work by Ferrari and collaborators applied comparative risk assessment (CRA) methodology to GBD 2010 data. They obtained RR estimates from a review of the literature which were then pooled and combined with GBD 2010 prevalence estimates to generate population-attributable fractions (PAFs). In their review of this literature, Ferrari and collaborators indicated that 62·2% of the burden estimated for suicide in an earlier iteration of the GBD study could be attributed to MSDs. MSDs moved from the 5th to the 3rd leading class of disease burden once suicide YLLs were reassigned to MSDs (Ferrari et al. 2014).

Given these findings, it is important to periodically re-assess the effect of MDs on premature mortality risk by updating the evidence base for the relationship between MDs and suicide. The work by Ferrari and Collaborators was conducted using GBD 2010 data and systematic literature reviews conducted in 2014. Analyses from this previous work largely comprised pooled RRs obtained using a simple meta-analytic approach. This approach did not account for sources of variation that may contribute to differences in reported suicide risk. These are important to consider in order to understand the impact of differences in study methods and subgroups on reported magnitude of suicide risk which, to date, remain largely unexplored. Existing reviews on this topic typically define broad inclusion criteria including vulnerable or high-risk subgroups (Rotenstein et al.

2016, da Silva et al. 2015). These may introduce measurement error and subsequently result in potentially inflated risk estimates. The analytic approach in these reviews also typically involve pooled analyses that do not fully adjust for sources of measurement error (e.g. study design, sampling strategy, etc.) or assess the impact of other ecological variables (e.g. region or individual characteristics). This current systematic literature review aims to account for these limitations by providing updated suicide risk estimates adjusted in a meta-regression for key demographic and methodological covariates that may contribute to variation in study-reported estimates. We use the earlier review by Ferrari et al. (2014) as the basis and template for this work. We expand the search strings to incorporate self-harm and self-injury to maximize the scope of potentially eligible studies. Relevant data from 2010 – 2019 updates this systematic review, this the most comprehensive review of this risk-outcome pair to the best of our knowledge. This review will also contribute to measurement of mortality that can be indirectly attributed to mental disorders.

1.2 Methods:

Case definition:
The MDs included in this paper were those identified previously as risk factors for suicide.[2] They were defined according to the Diagnostic and Statistical Manual of Mental Disorders (DSM) or the International Classification of Diseases (ICD) diagnostic criteria (WHO 1992, American Psychiatric Association 2000). The mental disorders included were major depressive disorder (MDD), dysthymia, bipolar disorder, anxiety disorders, and schizophrenia. Suicide was defined as cases meeting criteria for ICD cause of death codes for intentional self-inflicted poisoning or injury.

Systematic Review and Assessment of the Literature:
We expanded upon the 2014 systematic review by Ferrari and colleagues which covered the literature between 1966 and 2010 (See appendix). The search protocol adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines and are described in the appendix (Moher et al. 2009). Three major scholarly databases, namely Medline, Embase, and PsychINFO were searched between January 2010 and June 2019 Studies were eligible for inclusion based on the following criteria:

(1) Considered MDs (as defined by DSM or ICD) as a risk factor associated with death by suicide

(2) Reported a relative risk (RR) with 95% uncertainty, or provided sufficient information for these to be calculated

(3) Were individual-level case-control or cohort studies defining a clear temporal association between the exposure (presence of mental disorder(s)) and outcome (suicide). Prospective cohort studies were considered the gold standard for inclusion. However, retrospective cohort studies and case-control studies were also included and adjusted for in the main analysis.

(4) Study sample was representative of the general population of a given location (e.g.: community-based studies). Therefore, studies based on treated samples, linked

psychiatric treatment registries, health insurance or claims data, or vulnerable populations (prisoners, pregnant women, veterans, homeless, etc.) were excluded.

In addition to the electronic database search, experts in the field were consulted and recently published reviews on this topic were reviewed for any additional data sources. All data from earlier reviews were first re-assessed for eligibility and accepted only if they met criteria for this review. Relevant information on effect sizes, study methodology and sample characteristics were extracted and compiled for each study. References were managed and duplicates were removed via EndNote citation management software. MM conducted the data searches and data extraction. AF and DS re-assessed study eligibility as needed and reconciled any questions around study inclusion.

Statistical Analysis:
A multi-level meta-regression approach was used to obtain pooled relative risk estimates (RRs) and 95% uncertainty intervals for each disorder. We chose a multi-level meta-regression approach as it is best suited for data that are inherently nested or clustered within larger groups, as was the case here. Most studies in our dataset reported multiple effect sizes by disorder, age and/or sex and therefore were not independent observations. The multi-level approach with random effect terms for studies and disorder groups allowed us to account for any within and between study dependencies in the data. An unstructured variance-covariance matrix was used for random effects variances. The details and applications of this method are described in detail elsewhere (Cheung et al. 2014, Assink et al. 2016). Exploratory forest plots for disorder-specific RRs and variance-covariance matrices of models with different correlation thresholds are reported in the appendix.

We made use of a series of covariates to quantify and adjust for (where appropriate) sources of heterogeneity in the data. A detailed explanation of covariates is provided in Table 1. The inclusion of covariates in the model occurred sequentially, starting with (1) demographic variables known to be primary sources of variation (age, sex, disorder type, location); (2) methodological variables (follow-up time, sampling type, response rate, study design, estimate adjustment); and (3) the Sociodemographic Index (SDI) and the Health Access Quality Index (HAQI). Changes to effect-sizes due to the inclusion of new variables were reviewed at each step.

Due to considerable variation in the data collection and analytical methodology used between studies, several analytical choices were made in order to maximize data use. If studies reported odds ratios for suicide, these were converted to RRs using methods described in the appendix. Data from categories such as ‘mental disorders’, ‘mood and psychotic disorders’ or ‘mood and alcohol use disorders’ were not used as these very broad categories were less informative than more specific disorder categories. We considered the use of a location-level covariate using either GBD super-region classifications (Vos et al. 2020) or World Bank income level groups (Fantom et al. 2016). However, only 6 out of 20 studies were from low or middle-income settings, which may have limited statistical power to detect a significant effect in our analysis. For this reason, a location covariate was not used in the final model. Covariates in our final model included age, sex, disorder type, estimate adjustment, use of psychological autopsy method, and study design.

In order to test if specific estimates had greater influence on pooled results compared to others, exploratory Baujat plots were generated (Baujat et al. 2002). These plots display

the squared residuals on the X-axis and influence on the model on the Y-axis. Any estimates that had high residuals or high influence were determined to be outliers. Studies with outliers were re-assessed and those determined to have limitations in quality and/or measurement methods were excluded. Sensitivity analyses were conducted by assessing if overall results changed significantly upon removal of outliers as observed from the Baujat plots (Appendix). The distribution of variance across the levels included in our model was calculated using methods described in Cheung et al., 2014. This provided the proportion of the overall variance attributed to each level. All meta-regression analyses were conducted using the metafor package in R v3·5·2 (R Foundation for Statistical Computing 2019, Viechtbauer 2010).

1.3 Results:

A total of 20 studies and 69 individual RRs were included in the final analysis (See Figure 1). These studies covered 10 countries and 4 GBD regions. Anxiety disorders and MDD were the most commonly assessed MDs. (See Table 2 for a summary of selected studies). There was wide variation in reported sample sizes (84 to 68378) and follow-up time (1 year to 64 years)

Multi-level Meta-Regression:

Table 3 presents the meta-regression model coefficients and Table 4 presents the predicted RRs adjusted for select covariates. Disorder type, sex, age, study design, estimate adjustment, and use of psychological autopsy (PA) were tested in the model for suicide risk (Table 3). The amount of the overall variance attributed to random effects was 73.14%. We found no significant associations between suicide risk and response rate, follow-up time, sampling type, high vs low/middle-income locations, SDI, or HAQI. These variables were therefore not included in the final model. Our sensitivity analyses also showed no major change in significance and direction of estimates upon removal of outliers (4 observations – see appendix).

Age was associated with an increase in suicide risk (Table 3). Among methodological covariates, the use of retrospective study design was associated with lower suicide risk compared to prospective study design. There was no significant effect of estimate adjustment or use of the psychological autopsy method.

Suicide risk was significantly associated with the presence of MDD, anxiety disorders, bipolar disorder, dysthymia, and schizophrenia (Table 3). Predicted RRs were the highest for MDD, bipolar disorder, and schizophrenia (Table 4). RRs for females were consistently lower than RRs for males across all disorders (although this association was not significant).

1.4 Discussion:

Our results confirm that MDs are important risk factors for suicide. This association remained after adjusting for methodological factors. The earlier Ferrari 2014 review provided pooled estimates of suicide risk for each disorder obtained via a random effects meta-analytic approach. This updated review improves upon this approach by using a multi-level meta-regression model with predictive covariates that can (1) better utilize (in

its nested approach) all the data available across studies and (2) more robustly explain the heterogeneity in the data reported between studies to separate measurement error from variability that is 'real' and due to ecological variables.

Our findings showed that MDD, dysthymia, anxiety disorders, bipolar disorder, and schizophrenia were significantly associated with suicide risk. The selection of disorders attaining statistical significance, are consistent with meta-analytic results from other studies exploring suicide mortality due to MDs (Chesney et al. 2014, Hawton et al. 2013, Darvishi et al. 2015, Baxter et al. 2013). An earlier review on anxiety as a risk factor for suicide did not find anxiety disorders to be an important predictor of death by suicide (Bentley et al. 2016). This may have been due to the inclusion of studies that used symptom scales to assess anxiety disorders – which are best used for screening or symptomatic evaluation rather than provide a clinical diagnosis. Secondly, the inclusion of any individual anxiety diagnosis may have attenuated the pooled risk of suicide. Reviews may also focus only on specific disorders (da Silva et al. 2015). Although these provide an important profile of disorder-specific suicide risk, they exclude psychiatric comorbidity - an important contributor to and perhaps catalyst for suicide risk. Reviews may also draw from national treatment registers for data from samples with mental disorders (Too et al. 2019). National treatment registers are a useful source of data on mental disorders often drawn from large samples and allow for a more efficient study of rare outcomes such as death by suicide. However, it is important to consider that this mode of data collection may also be inherently biased as it does not capture mental disorder cases among those who may not have made contacts with any health service. It is likely that register-based diagnoses may reflect a higher severity of mental disorders than what may be present in the general population. Therefore, the use of registry-based data is expected to be associated with an increase in or overestimation of pooled estimates of suicide risk. Some reviews also examine suicidal behavior (such as suicide attempts and/or self-injury with an intent to die). A recent review by Gili and colleagues examine the evidence for mental and substance use disorders as risk factors for suicidal behavior (suicidal acts and attempts) among young people (Gili et al. 2019). Although this review provides important insights into the risk for suicide attempts that frequently precede suicide deaths in younger populations, the analyses use ORs instead of RRs and use broad groups of disorders which may limit interpretation for disorder-specific risks especially for depressive disorders which are known to account for a large proportion of suicide risk. It is important to highlight that the disorder-specific pooled RRs obtained here are considerably lower than those found in the earlier review by Ferrari 2014. The use of demographic, methodological, and location-level covariates may have contributed to this difference. We adjusted for study-level random effects using a hierarchical approach that is also likely to yield lower adjusted RRs for suicide compared to a standard meta-analytic approach that has typically been used to generate pooled effect sizes in the past. We also used stricter inclusion criteria and data prepping procedures than what has been used in the existing literature. Furthermore, meta-analytical studies will often use odds ratios as an approximation for RRs if the prevalence of suicide is known to be very low. This may well overestimate RRs and lead to higher effect sizes we usually see in other studies (Walker et al. 2015). This study attempts to take suicide risk measurement a step further by including covariates that account for measurement error. Although uncertainty in the model remains, we believe that the estimates are more robust as a result of this.

In our model, older age was associated with higher RRs. This is consistent with findings in the literature where older age is likely to be associated with an increased risk of

suicide (Whiteford et al. 2015). Among methodological variables, the use of a retrospective study design was associated with lowered suicide risk. It is possible that a prospective study design is better at establishing a temporal association between mental disorders and suicide which may not be possible in a retrospective design. Additionally, studies using a retrospective design may not be able to capture the impact of confounders on mental disorders and suicide as accurately as those using a prospective design.

We found no significant associations between the use of psychological autopsy and suicide risk. This may be due to the low number of non-psych autopsy studies that were included in the analyses. Overestimation or underestimation of suicide risk in psych-autopsy studies may be due to biases in how how psychopathology is measured, specifically the choice of informants, choice of comparison group, and lowered event recall - especially if the assessment is conducted much later than the date of suicide. Although our study did not empirically assess these components beyond a dichotomous measure of psychological autopsy use, it is reasonable to assume that they may contribute to the heterogeneity in reported suicide risk beyond what we found in our analyses.

Adjusted RRs were not significantly different from unadjusted RRs in our study. However, it was important to include this covariate in our model since RRs that controlled for known potential confounders were likely be more accurate in estimating true suicide risk compared to crude RRs. A possible reason for not detecting a significant effect may be the low number of studies adjusted RRs. There is considerable variation in the number and type of confounders controlled for between studies. Some studies may control for a single confounder (such as age, or educational attainment). Other studies may control for a larger number and range of confounders such as psychiatric comorbidity, family history of MDs, socioeconomic status, physical illnesses, and others. Although our current analyses did not empirically assess the impact of this variation on suicide risk, it is plausible that studies that control for as many potential confounders as possible likely report estimates of a higher quality and reliability compared to studies that do not.

Our earlier modeling iterations did not find study follow-up period to be a statistically significant predictor of suicide risk. Therefore, it was not included in our final model. However, it is important to acknowledge that RRs from studies with very long follow-up periods between an MD diagnosis and death by suicide may be susceptible to bias as MDs can manifest as recurrent episodes with varying levels of functioning and periods of remission over time. Therefore, it may be challenging to determine the extent to which the risk of suicide may be attributed entirely to MDs in such studies. It is plausible that risk factors during a long follow-up period may be different than those during a short follow-up period (Suokas et al. 2001). Future work in this area could potentially incorporate data on disorder remission periods and other competing proximal and distal risk factors that may be important risk factors for suicide (such as prior suicide attempts, treatment history, etc.)

We were unable to empirically assess location-level differences in RRs due to the lack of sufficient studies from LMICs. Out of 20 studies, 14 were from high-income locations - mostly Western Europe and North America. Given the lack of income level-specific estimates across disorders, the uncertainty across the results produced is greater and interpretability of the resulting sex- and region-specific RR are lowered. The sparsity of

data from LMICs may be due to variation and constraints in country-level reporting of suicide deaths. Suicide continues to be a criminal offence in many countries in Africa, Asia, and South America, which potentially leads to systematic underreporting of suicide cases (Khan 2005, Fleischmann and De Leo 2014). The global data coverage for MDs is also low for LMICs (Baxter et al. 2013). Better quality data from LMICs along with more sex-specific data on MDs and suicide may allow us to make more substantive conclusions about regional and sex-specific patterns in RRs for suicide.

This systematic review provides a useful insight into the recent available evidence on the risk of suicide among those with versus without MDs and the heterogeneity that may exist in the data based on study-specific attributes. Our findings of MDs being significant predictors of suicide risk are largely consistent with findings from other individual-level and meta-analytic studies (Walker et al. 2015, Rotenstein et al. 2016, da Silva et al. 2015). Based on the GBD study, MDs are the some of the leading causes of non-fatal disability (Whiteford et al. 2013, Whiteford et al. 2105). However, very few MDs are attributed as direct causes of suicide despite overwhelming evidence in the literature. Findings from this study may better inform the burden from premature mortality directly attributed to MDs in GBD.

These findings also support the need for adequate mental healthcare as an important suicide prevention measure. Research using World Mental Health surveys shows that a large treatment gap exists for those exhibiting suicidal ideation and behaviors. Only 39% of those with suicidality receive any form of treatment worldwide. This proportion is even lower in low-income countries where only 17% of those with suicidality receive any form of treatment (Bruffaerts et al. 2011). Although there exists evidence in support of select interventions for suicide prevention, improving prevention efforts at earlier stages of care may increase the likelihood of fulfilling this unmet need for care (Mann et al. 2005). These findings also emphasize the continued need for primary interventions for suicide prevention that focus on adequate screening for and detection of MDs in different contexts such as community, primary care, and school-based settings (Dowdy et al. 2015, Kline et al. 2014, Siu et al. 2016, Ali et al. 20166, Guntuku et al. 2017, Bruffaerts et al. 2011). However, precision in MD detection is dependent on the screening tool and classification thresholds used - particularly in varying cultural contexts. Arango and colleagues define secondary and tertiary prevention efforts in addition to primary prevention efforts in their review (Arango et al. 2016). Secondary prevention efforts that focus on reducing the incidence of MDs and tertiary prevention efforts that treat those with established MDs are important as they address the role of MDs in suicide prevention (Arango et al. 2016). A systematic review on suicide prevention strategies found that in addition to direct suicide prevention strategies such as restricting access to lethal means, pharmacological and psychological treatment of MDs are also important in suicide prevention (Zalsman et al. 2016).

Several limitations constrained our analyses. First, the geographic representativeness of our analysis is limited. Although our inclusion and exclusion criteria ensured the selection of high quality studies, these were mostly from high-income countries in North America and Western Europe. Therefore, studies from low and middle-income countries are under-represented in this analysis indicating the need for more published data from these regions. Second, we were unable to provide estimates for additional psychiatric disorders such as eating disorders due to the lack of studies that fulfilled our selection criteria. Third, we acknowledge that the uncertainty intervals for predicted RRs were at times wide and overlapping between disorders due to lack of data for some disorders

and by income-level. Fourth, analyzing data for all available MDs in a single meta-regression model may either attenuate or amplify the magnitude of disorder-specific RRs. Fifth, the standard meta-regression framework allows for uncertainty intervals of predicted RRs to capture error attributed to fixed effects only and not between-study heterogeneity. Lastly, although we tried to quantify as many methodological sources of variation as possible, we were unable to account for all sources of measurement error. We excluded methodological covariates such as sampling type, follow-up duration, and response rate as there was insufficient power to detect an effect.

1.5 Conclusion:

MDs are associated with an increased risk for suicide. This synthesized analysis shows that the magnitude of study-reported risk may vary depending on study methodology and disorders examined. Therefore, study quality and choice of methods are important to consider in public reports on aggregated evidence. More research is needed to better quantify evidence for MD burden and suicide risk with higher precision in low and middle-income settings. These efforts will assist policy makers in framing evidence-based suicide prevention strategies and make mental health care an important part of their agendas.

1.6 Significance:

The first chapter aimed to improve relative risk estimates for MDs as risk factors for suicide. Methodological improvements were applied by incorporating a multilevel approach and unstructured variance to account for study-level and participant-level heterogeneity. Sample and methodological covariates were introduced which allowed us to better understand variation in reported risk estimates by study-level attributes. These findings may find application in the computation of population-attributable fractions (PAFs) within the Global Burden of Disease framework in order to estimate MD attributable suicide DALYs (disability-adjusted life years). This will allow us to more accurately capture the fatal burden due to MDs, over and above what is directly assigned to them within the GBD study. Differences between earlier and newly calculated PAFs and attributable suicide DALYs may provide further information on the impact of sources of measurement error. The earlier Ferrari review applied the same risk estimates across all countries, age groups, and sex which could potentially mask any differences in the distribution of attributable suicide DALYs. This current analysis for Aim I calculates adjusted RRs across sex and age with additionally analyses by income-level. Therefore, these will provide additional layers of refinement for risk estimates and subsequently attributable suicide DALYs. In summary, this chapter reported on potential methodological improvements under the broader domain of comparative epidemiological assessments that are used in the mental disorder estimation process. The next chapter will look at a novel application of a Bayesian meta-regression approach within the domain of treatment efficacy.

1.7 References:

1. Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, Abbasi-Kangevari M, Abbastabar H, Abd-Allah F, Abdelalim A, Abdollahi M. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet. 2020 Oct 17;396(10258):1204-22.

2. Vigo D, Thornicroft G, Atun R. Estimating the true global burden of mental illness. The Lancet Psychiatry. 2016;3(2):171-8.
3. Ferrari AJ, Norman RE, Freedman G, Baxter AJ, Pirkis JE, Harris MG, et al. The burden attributable to mental and substance use disorders as risk factors for suicide: findings from the Global Burden of Disease Study 2010. PloS one. 2014;9(4):e91936.
4. James SL, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet. 2018;392(10159):1789-858.
5. Wu KC-C, Chen Y-Y, Yip PS. Suicide methods in Asia: implications in suicide prevention. International journal of environmental research and public health. 2012;9(4):1135-58.
6. Ajdacic-Gross V, Weiss MG, Ring M, Hepp U, Bopp M, Gutzwiller F, et al. Methods of suicide: international suicide patterns derived from the WHO mortality database. Bulletin of the World Health Organization. 2008;86:726-32.
7. Inskip H, Harris C, Barraclough B. Lifetime risk of suicide for affective disorder, alcoholism and schizophrenia. The British Journal of Psychiatry. 1998;172(1):35-7.
8. Harris EC, Barraclough B. Suicide as an outcome for mental disorders. A meta-analysis. British journal of psychiatry. 1997;170(3):205-28.
9. Cavanagh JT, Carson AJ, Sharpe M, Lawrie SM. Psychological autopsy studies of suicide: a systematic review. Psychological medicine. 2003;33(3):395-405.
10. Harris C, Barraclough B. Excess mortality of mental disorder. The British journal of psychiatry. 1998;173(1):11-53.
11. Li Z, Page A, Martin G, Taylor R. Attributable risk of psychiatric and socio-economic factors for suicide from individual-level, population-based studies: A systematic review. Social Science and Medicine. 2011;72(4):608-16.
12. Walker ER, McGee RE, Druss BG. Mortality in mental disorders and global disease burden implications: a systematic review and meta-analysis. JAMA psychiatry. 2015;72(4):334-41.
13. Rotenstein LS, Ramos MA, Torre M, Segal JB, Peluso MJ, Guille C, et al. Prevalence of Depression, Depressive Symptoms, and Suicidal Ideation Among Medical Students: A Systematic Review and Meta-Analysis. Jama. 2016;316(21):2214-36.
14. da Silva Costa L, Alencar ÁP, Neto PJN, dos Santos MdSV, da Silva CGL, Pinheiro SdFL, et al. Risk factors for suicide in bipolar disorder: a systematic review. Journal of affective disorders. 2015;170:237-54.
15. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Annals of internal medicine. 2009;151(4):264-9.
16. Organization WH. The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines: Geneva: World Health Organization; 1992.
17. Association AP. Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR). Washington DC: American Psychiatric Assocation. 2000.
18. Fantom N, Serajuddin U. The World Bank's classification of countries by income. The World Bank; 2016 Jan 4.
19. Brent DA. The psychological autopsy: methodological considerations for the study of adolescent suicide. Suicide and Life-Threatening Behavior. 1989 Mar;19(1):43-57.
20. Lester D, Stack S. Bias resulting from the choice of sample and results of cross-national analyses of suicide rates. Quality and Quantity. 1989 Jun 1;23(2):221-3.
21. Assink M, Wibbelink CJ. Fitting three-level meta-analytic models in R: A step-by-step tutorial. The Quantitative Methods for Psychology. 2016;12(3):154-74.
22. Cheung MW-L. Modeling dependent effect sizes with three-level meta-analyses: a structural equation modeling approach. Psychological Methods. 2014;19(2):211.

23. Baujat B, Mahé C, Pignon JP, Hill C. A graphical method for exploring heterogeneity in meta-analyses: application to a meta-analysis of 65 trials. Statistics in medicine. 2002;21(18):2641-52.
24. (2019) RCT. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2019.
25. Anderberg J, Bogren M, Mattisson C, Brådvik L. Long-term suicide risk in anxiety—the Lundby Study 1947–2011. Archives of Suicide Research. 2016;20(3):463-75.
26. Athey A, Overholser J, Bagge C, Dieter L, Vallender E, Stockmeier CA. Risk-taking behaviors and stressors differentially predict suicidal preparation, non-fatal suicide attempts, and suicide deaths. Psychiatry Research. 2018;270:160-7.
27. Brent DA, Baugher M, Bridge J, Chen T, Chiappetta L. Age-and sex-related risk factors for adolescent suicide. Journal of the American Academy of Child & Adolescent Psychiatry. 1999;38(12):1497-505.
28. Conwell Y, Duberstein PR, Hirsch JK, Conner KR, Eberly S, Caine ED. Health status and suicide in the second half of life. International Journal of Geriatric Psychiatry. 2010;25(4):371-9.
29. De Leo D, Draper BM, Snowdon J, Kõlves K. Suicides in older adults: A case–control psychological autopsy study in Australia. Journal of Psychiatric Research. 2013;47(7):980-8.
30. Giupponi G, Innamorati M, Baldessarini RJ, De Leo D, de Giovannelli F, Pycha R, et al. Factors associated with suicide: Case-control study in South Tyrol. Comprehensive Psychiatry. 2018;80:150-4.
31. Hirokawa S, Kawakami N, Matsumoto T, Inagaki A, Eguchi N, Tsuchiya M, et al. Mental disorders and suicide in Japan: A nation-wide psychological autopsy case–control study. Journal of Affective Disorders. 2012;140(2):168-75.
32. Holmstrand C, Bogren M, Mattisson C, Brådvik L. Long-term suicide risk in no, one or more mental disorders: The Lundby Study 1947–1997. Acta Psychiatrica Scandinavica. 2015;132(6):459-69.
33. Lasserre AM, Marti-Soler H, Strippoli MPF, Vaucher J, Glaus J, Vandeleur CL, et al. Clinical and course characteristics of depression and all-cause mortality: A prospective population-based study. Journal of Affective Disorders. 2016;189:17-24.
34. Lesage AD, Boyer R, Grunberg F, Vanier C, Morissette R, Ménard-Buteau C, et al. Suicide and mental disorders: a case-control study of young men. The American journal of psychiatry. 1994.
35. Niu L, Jia C, Ma Z, Wang G, Yu Z, Zhou L. Validating the Geriatric Depression Scale with proxy-based data: A case-control psychological autopsy study in rural China. Journal of Affective Disorders. 2018;241:533-8.
36. Oguzhanoglu NK, Ugurlu TT, Acar K, Atesci F. A psychological and social perspective on completed suicides in western Anatolia, Turkey: A case-control psychological autopsy study. Dusunen Adam. 2018;31(4):331-9.
37. Page A, Morrell S, Hobbs C, Carter G, Dudley M, Duflou J, et al. Suicide in young adults:Psychiatric and socio-economic factors from a case-control study. BMC Psychiatry. 2014;14(1).
38. Séguin M, Renaud J, Lesage A, Robert M, Turecki G. Youth and young adult suicide: A study of life trajectory. Journal of Psychiatric Research. 2011;45(7):863-70.
39. Shaffer D, Gould MS, Fisher P, Trautman P, Moreau D, Kleinman M, et al. Psychiatric diagnosis in child and adolescent suicide. Archives of general psychiatry. 1996;53(4):339-48.
40. Shibre T, Hanlon C, Medhin G, Alem A, Kebede D, Teferra S, et al. Suicide and suicide attempts in people with severe mental disorders in Butajira, Ethiopia: 10 year follow-up of a population-based cohort. BMC Psychiatry. 2014;14(1).

41. Tong Y, Phillips MR. Cohort-specific risk of suicide for different mental disorders in China. British Journal of Psychiatry. 2010;196(6):467-73.
42. Waern M. Alcohol dependence and misuse in elderly suicides. Alcohol and alcoholism. 2003 May 1;38(3):249-54.
43. Zhang J, Wieczorek W, Conwell Y, Tu XM, Wu BY, Xiao S, et al. Characteristics of young rural Chinese suicides: a psychological autopsy study. Psychological medicine. 2010;40(4):581-9.
44. Zhou L, Wang G, Jia C, Ma Z. Being left-behind, mental disorder, and elderly suicide in rural china: A case–control psychological autopsy study. Psychological Medicine. 2018.
45. Chesney E, Goodwin GM, Fazel S. Risks of all-cause and suicide mortality in mental disorders: a meta-review. World psychiatry. 2014;13(2):153-60.
46. Hawton K, Comabella CCI, Haw C, Saunders K. Risk factors for suicide in individuals with depression: A systematic review. Journal of Affective Disorders. 2013;147(1-3):17-28.
47. Darvishi N, Farhadi M, Haghtalab T, Poorolajal J. Alcohol-related risk of suicidal ideation, suicide attempt, and completed suicide: A meta-analysis. PLoS ONE. 2015;10(5).
48. Bentley KH, Franklin JC, Ribeiro JD, Kleiman EM, Fox KR, Nock MK. Anxiety and its disorders as risk factors for suicidal thoughts and behaviors: A meta-analytic review. Clinical psychology review. 2016;43:30-46.
49. Baxter A, Scott K, Vos T, Whiteford H. Global prevalence of anxiety disorders: a systematic review and meta-regression. Psychological medicine. 2013;43(5):897-910.
50. Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. The Lancet. 2013;382(9904):1575-86.
51. Whiteford HA, Ferrari AJ, Degenhardt L, Feigin V, Vos T. The global burden of mental, neurological and substance use disorders: an analysis from the Global Burden of Disease Study 2010. PloS one. 2015;10(2):e0116820.
52. Dowdy E, Furlong M, Raines TC, Bovery B, Kauffman B, Kamphaus RW, et al. Enhancing school-based mental health services with a preventive and promotive approach to universal screening for complete mental health. Journal of Educational and Psychological Consultation. 2015;25(2-3):178-97.
53. Kline E, Schiffman J. Psychosis risk screening: A systematic review. Schizophrenia Research. 2014;158(1-3):11-8.
54. Siu AL, Bibbins-Domingo K, Grossman DC, Baumann LC, Davidson KW, Ebell M, et al. Screening for depression in adults: US Preventive Services Task Force recommendation statement. Jama. 2016;315(4):380-7.
55. Ali G-C, Ryan G, De Silva MJ. Validated screening tools for common mental disorders in low and middle income countries: a systematic review. PloS one. 2016;11(6):e0156939.
56. Guntuku SC, Yaden DB, Kern ML, Ungar LH, Eichstaedt JC. Detecting depression and mental illness on social media: an integrative review. Current Opinion in Behavioral Sciences. 2017;18:43-9.
57. Bruffaerts R, Demyttenaere K, Hwang I, Chiu W-T, Sampson N, Kessler RC, et al. Treatment of suicidal people around the world. The British Journal of Psychiatry. 2011;199(1):64-70.
58. Mann JJ, Apter A, Bertolote J, Beautrais A, Currier D, Haas A, et al. Suicide prevention strategies: a systematic review. Jama. 2005;294(16):2064-74.
59. Baxter AJ, Patton G, Scott KM, Degenhardt L, Whiteford HA. Global epidemiology of mental disorders: what are we missing? PloS one. 2013;8(6):e65514.
60. Khan MM. Suicide prevention and developing countries. Journal of the Royal Society of Medicine. 2005 Oct;98(10):459-63.

61. Fleischmann A, De Leo D. The World Health Organization's report on suicide: a fundamental step in worldwide suicide prevention.
62. Gili M, Castellví P, Vives M, de la Torre-Luque A, Almenara J, Blasco MJ, Cebrià AI, Gabilondo A, Pérez-Ara MA, Miranda-Mendizabal A, Lagares C. Mental disorders as risk factors for suicidal behavior in young people: A meta-analysis and systematic review of longitudinal studies. Journal of affective disorders. 2019 Feb 15;245:152-62.
63. Arango C, Díaz-Caneja CM, McGorry PD, Rapoport J, Sommer IE, Vorstman JA, McDaid D, Marín O, Serrano-Drozdowskyj E, Freedman R, Carpenter W. Preventive strategies for mental health. The Lancet Psychiatry. 2018 Jul 1;5(7):591-604.
64. Zalsman G, Hawton K, Wasserman D, van Heeringen K, Arensman E, Sarchiapone M, Carli V, Höschl C, Barzilay R, Balazs J, Purebl G. Suicide prevention strategies revisited: 10-year systematic review. The Lancet Psychiatry. 2016 Jul 1;3(7):646-59.
65. Viechtbauer W. Conducting meta-analyses in R with the metafor package. Journal of statistical software. 2010;36(3):1-48.
66. Barber RM, Fullman N, Sorensen RJ, Bollyky T, McKee M, Nolte E. & Murray, CJL (2017). Healthcare Access and Quality Index based on mortality from causes amenable to personal health care in 195 countries and territories, 1990–2015: A novel analysis from the Global Burden of Disease Study 2015. The Lancet.;390(10091):231-66.
67. San Too L, Spittal MJ, Bugeja L, Reifels L, Butterworth P, Pirkis J. The association between mental disorders and suicide: a systematic review and meta-analysis of record linkage studies. Journal of affective disorders. 2019 Dec 1;259:302-13.
68. Suokas J, Suominen K, Isometsä E, Ostamo A, Lönnqvist J. Long-term risk factors for suicide mortality after attempted suicide-Findings of a 14-year follow-up study. Acta Psychiatr Scand 2001: 104: 117-121

1.8 Tables and Figures

Covariate	Definition and Reference Levels	Details
Age	mid-point of age range	MD prevalence and corresponding suicide risk are known to vary with age and by sex (Whiteford et al. 2015). Therefore, they are important sources of variation to consider in our analyses
Percent Female	Continuous covariate representing the proportion of females in study sample (Ranges from 0 to 1)	
Follow-up time	Follow-up time in years	Suicide is a relatively rare outcome compared to other causes of death. Therefore, the duration of follow-up may impact the number of study-reported outcomes
Response rate	Proportion of sample remaining after loss to follow-up/dropout	The response rate provides important information about possible selection bias in the sample.
Disorder	GBD mental disorder categories (Reference: MDD)	These are the primary risk factors for suicide being assessed. MDD was chosen as the reference since it was the most commonly assessed disorder among selected studies.
Estimate adjustment	Indicator for whether or not effect size has been adjusted for potential confounders such as individual demographics, socioeconomic status, family psychiatric history, etc. Reference: Adjusted for potential confounders	Study-reported effect sizes may be adjusted for potential confounders that are known to influence the MD-suicide association. These may be different from (and typically lower) than unadjusted effect sizes. Therefore, our analyses examine variation in suicide risk by testing this methodological covariate.

Psychological Autopsy (PA) Method	Indicator for whether or not data was collected using psychological autopsy – which involves collecting data from all available sources such as family informants, medical records, and healthcare providers. (Reference: PA not used)	This covariate was assessed because the psychological autopsy method involves data collection from informants and therefore is susceptible to biases in measurement of psychopathology, event recall, choice of appropriate control groups, etc (Brent 1989).
Study design	Prospective (Reference) or retrospective design	We expected variation in study quality and effect sizes based on the choice of study design. Therefore, this covariate was included to examine if study design influenced pooled RRs (Brent 1989).
Sampling type	Random/other methods (multistage, cluster sampling)	We tested this covariate because we expected studies using random sampling to have less biased samples than studies using other methods (Lester and Stack 1989).
SDI	Sociodemographic Index value (SDI): A summarized metric of a location's socio-demographic development on a scale of 0 (lowest) to 1 (highest). More details on the composition of the SDI can be found in GBD 2015 publication.	Higher SDI and HAQI are associated with better health outcomes and lower premature mortality. Therefore, these were tested in the model to see if they had an impact on suicide risk.
HAQI	Health Access Quality Index value: A summarized metric of healthcare access and quality on a scale of 0 (worst) to 100 (best). More details on the construction of the HAQI can be found in elsewhere (Barber et al. 2015).	
High income locations	Locations that are classified as high-income as per World Bank income classification (Reference) vs other locations (Fantom et al. 2016)	High-income locations are known to have better health outcomes and lower premature mortality than low and middle-income countries. Therefore, we tested this variable to see if it influenced resulting suicide risk estimates.

Table 1. List of Covariates

Study	Region	Disorders included in Analysis
Anderberg et al. 2010	Western Europe	Anxiety Disorders
Athey et al. 2018	North America	Anxiety Disorders, Schizophrenia
Brent et al. 1999	North America	Anxiety Disorders
Conwell et al. 2010	North America	Major Depressive Disorders, Anxiety Disorders, Schizophrenia
De Leo et al. 2013	Australasia	Anxiety Disorders, Major Depressive disorders, Schizophrenia
Giupponi et al. 2018	Western Europe	Bipolar disorders, Major Depressive Disorders, Schizophrenia
Hirokawa et al. 2012	Asia-Pacific	Major Depressive Disorder, Anxiety Disorders, Schizophrenia, Dysthymia

Holmstrand et al. 2015	Western Europe	Major Depressive Disorders, Schizophrenia, Anxiety Disorders
Lasserre et al. 2016	Western Europe	Major Depressive Disorders
Lesage et al. 1994	North America	Major Depressive Disorder, Schizophrenia, Anxiety Disorders, Bipolar disorder
Niu et al. 2018	East Asia	Major Depressive Disorder
Oguzhanoglu et al. 2018	North Africa and Middle East	Major Depressive disorders, Schizophrenia
Page et al. 2014	Australasia	Anxiety Disorders
Seguin et al. 2011	North America	Anxiety disorders
Shaffer et al.1996	North America	Major Depressive Disorders, Dysthymia, Schizophrenia, Bipolar Disorder, Anxiety Disorders,
Shibre et al. 2014	Sub-Saharan Africa	Schizophrenia, Bipolar Disorder, Major Depressive Disorders
Tong et al. 2010	East Asia	Anxiety Disorders, Schizophrenia
Waern et al. 2003	Western Europe	Major Depressive Disorders, Anxiety disorders
Zhang et al. 2010	East Asia	Anxiety Disorders, Bipolar Disorders, Dysthymia, Major Depressive disorders, Schizophrenia
Zhou et al. 2018	East Asia	Anxiety Disorders, Bipolar Disorders, Dysthymia, Major Depressive disorders, Schizophrenia

Table 2. List of Selected Studies

Covariates		**Coefficient**	**P Value**
Cochran's test for residual heterogeneity		$QE_{(59)}$ = 177·98	<0·0001
QM-Test of moderators		$QM_{(9)}$ = 68·96	<0·0001
Age	NA	0·0053	0·0094
Both	11 studies/39 observations	Reference category	
Percent Female	Male: 7 studies/18 observations	-0·0346	0·7518
	Female: 6 studies/12 observations		
Disorder			
Major Depressive Disorders	14 studies/ 17 observations	Reference category	
Anxiety Disorders	17 studies/23 observations	-0·446	<0·0001
Bipolar disorder	7 studies/7 observations	-0·234	0·0217
Dysthymia	4 studies/4 observations	-0·62	0·0011
Schizophrenia	13 studies/18 observations	-0·245	0·0006
Estimate adjustment			
Adjusted for confounders	5 studies/15 observations	Reference category	
Unadjusted for confounders	15 studies/54 observations	-0·35	0·0672
Study Design			
Prospective design	4 studies/10 observations	Reference category	
Retrospective design	16 studies/ 59 observations	-0·932	0·0035

Use of Psychological Autopsy			
Psychological autopsy not used	2 studies/4 observations	Reference category	
Psychological autopsy used	18 studies/65 observations	-0·0008	0·99

Table 3. Meta-regression Model Coefficients by Covariate

Disorder	Both	Males	Females
Major Depressive Disorder	7·64 [4·3, 13·58]	7·78 [4·34, 13·93]	7·51 [4·18, 13·51]
Dysthymia	4·11 [2·09, 8·09]	4·18 [2·12, 8·26]	4·04 [2·02, 8·06]
Anxiety Disorders	4·89 [2·76, 8·69]	4·98 [2·78, 8·91]	4·81 [2·68, 8·64]
Bipolar Disorder	6·05 [3·38, 10·83]	6·15 [3·4, 11·13]	5·94 [3·29, 10·75]
Schizophrenia	5·98 [3·33, 10·72]	6·09 [3·73, 10·98]	5·88 [3·24, 10·66]

Table 4. Predicted Relative Risks for Suicide

RRs adjusted for sex, age, disorder, study design, estimate adjustment, and psych-autopsy method

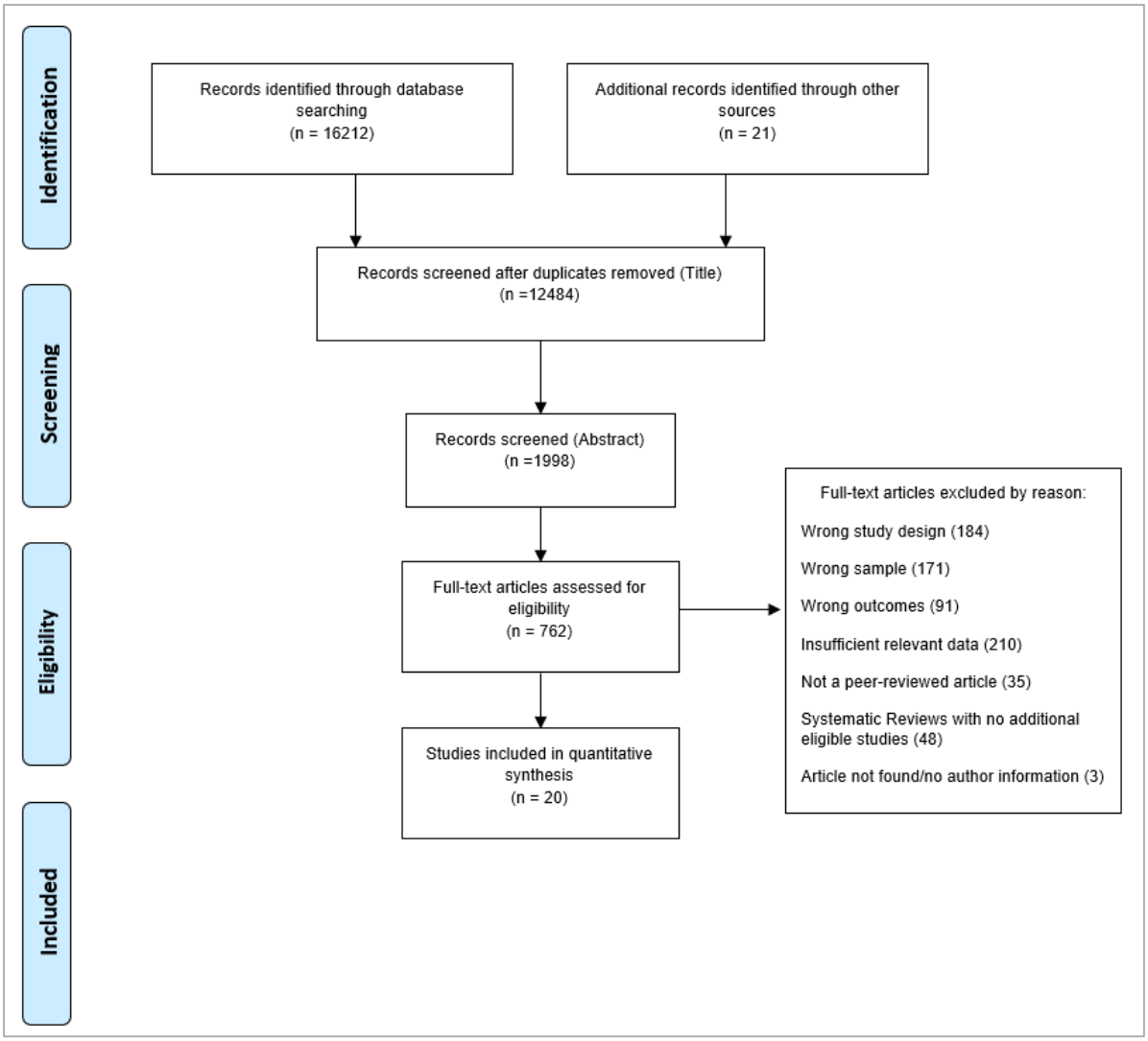


Figure 1. PRISMA Flowchart for Study Screening and Selection

1.9 Appendix:

Search String:

((((((((((((("suicid*"[Title/Abstract] OR "self-harm"[Title/Abstract] OR "suicid*"[MeSH Terms] OR "self-harm"[MeSH Terms]))))

AND
("cohort"
OR "case control"
OR "case-control"
OR "autopsy")
AND
(("mood disorder*" [Title/Abstract]
OR "depress*"[Title/Abstract]
OR "dysthymi*"[Title/Abstract]
OR "bipolar"[Title/Abstract]
OR "manic"[Title/Abstract]
OR "mania"[Title/Abstract]
OR "Mood disorders"[MeSH Terms]
OR "Depressive disorders"[MeSH Terms]
OR "Depressive disorder, Major"[MeSH Terms]
OR "Bipolar disorder"[MeSH Terms]
OR "Dysthymic Disorders"[MeSH Terms]
OR "generalized anxiety disorder*" [Title/Abstract]
OR "post traumatic stress disorder*" [Title/Abstract]
OR "posttraumatic stress disorder*" [Title/Abstract]
OR "traumatic stress disorder*" [Title/Abstract]
OR "anxiety disorder*" [Title/Abstract]
OR "Anxiety Disorders"[Mesh: No Expansion]
OR "Schizo*"[Title/Abstract]
OR "Psychosis"[Title/Abstract]
OR "Psychotic"[Title/Abstract]
OR "Psychotic disorder"[MeSH Terms]
OR "Psychotic disorders"[MeSH Terms]
OR "Schizophrenia"[Mesh: No Expansion]
OR "Personality disorder"[Mesh: No Expansion]
OR "Anorexia*" [Title/Abstract]
OR "Bulimia*" [Title/Abstract]
OR "Eating disorder*" [Title/Abstract]
OR "binge eating disorder" [Title/Abstract]
OR "Binge eating disorder"[MeSH Terms]
OR "Anorexia nervosa"[MeSH Terms]
OR "Bulimia nervosa"[MeSH Terms]
OR "Eating disorder"[MeSH Terms]
OR ("alcohol*"[Title/Abstract] AND ("abus*"[Title/Abstract] OR "dependen*"[Title/Abstract] OR "misus*"[Title/Abstract] OR "addict*"[Title/Abstract]))
OR ("drug"[Title/Abstract] OR "substance"[Title/Abstract]) AND ("abus*"[Title/Abstract] OR "dependen*"[Title/Abstract] OR "addict*"[Title/Abstract] OR "misus*"[Title/Abstract])

OR "Alcohol-Related Disorder"[Mesh: No Expansion]
OR "Alcohol-Induced Disorder"[Mesh: No Expansion]
OR "Fetal alcohol spectrum disorder" [MeSH Terms]
OR "Fetal alcohol disorder" [MeSH Terms]
OR "Alcoholism"[MeSH Terms]
OR "Amphetamine-related disorders" [MeSH Terms]
OR "Cocaine-related disorders" [MeSH Terms]
OR "Opioid-related disorders" [MeSH Terms]
OR "Substance-related disorders"[Mesh: No Expansion]
OR "opioid*"[Title/Abstract]
OR "heroin"[Title/Abstract]
OR "Cocaine"[Title/Abstract]
OR "Marijuana"[Title/Abstract]
OR "Cannabis"[Title/Abstract]
OR "Amphetamin*"[Title/Abstract]
OR "Methamphetam*"[Title/Abstract])))))))))))

Inclusion/Exclusion Criteria:

1. **Study Representativeness**: Studies using samples that are demographically representative of the community, city, region, or country of interest will be included. Therefore, estimates of suicide risk obtained from samples of minority groups or clinical trials will not be included.

2. **Timeframe**: Studies published between 2010 and 2019 (June) are included.

3. **Case Definitions**:
   i) Suicide: Defined as cases meeting the ICD-10 cause of death codes for intentional self-inflicted poisoning or injury (X60–X84)
   ii) Mental Disorders: Studies that adhere to the definition of mental disorders as determined by the International classification of diseases (ICD-10) or the Diagnostic and Statistical Manual of Mental Disorders (DSM-III, IV, V) will be included. Risk estimates for suicide derived from subthreshold cases will not be included.

4. **Risk of Suicide** attributable to Mental Disorders: Studies must report effect sizes clearly quantifying the risk of suicide that can be attributable to one or more mental disorders or provide enough data to calculate effect sizes. This review will exclude studies that only report the risk of suicide attributable to causes unrelated to mental disorders.

Data Extraction:

- Information pertaining to study methodology, sample, cases, and effect size values will be extracted from data sources that satisfy our inclusion criteria.

- Aggregated estimates of risk of suicide attributed to mental disorders will be extracted if reported or calculated if sufficient data are available for calculations.

- If studies report suicide risk estimates by disorder type, severity, age-group, sex-group, or any other category, this information was extracted for further disaggregation of suicide risk where possible.

- If studies report odds ratios, these will be converted to relative risks using methods described in Barendregt 2010 (ref).

- If studies do not report mental disorder prevalence in the entire sample, year-location-sex specific prevalence estimates from the Global Burden of Disease will be used to calculate relative risks.

- An estimate of uncertainty (either a standard error or 95% uncertainty interval) around the risk estimate will be extracted if reported, or calculated using the formula:

$$SE = \sqrt{2.1\left(\frac{P(1-P)}{N}\right)}$$

Where N=sample size, P= proportion/quantity of interest

MD categories:

MDs are categorized as per the definitions the GBD 2017 Years Lived with Disability paper.[4] Due to the low number of eligible studies, estimates could not be made for eating disorders, conduct disorders, attention-deficit hyperactivity, and personality disorders.

Multilevel Meta-regression Approach:

The multi-level meta-regression approach used in this study is an extension of the standard meta-analytic approach with higher-level effects and is formally notated as follows:

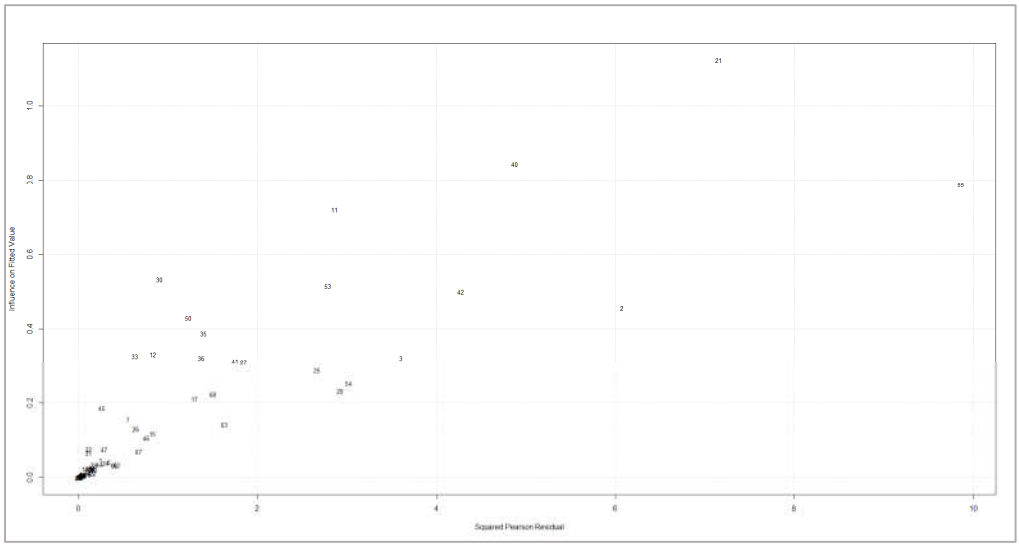
$$\hat{\theta}_{ij} = \beta_0 + \zeta_{(2)ij} + e_{ij}$$

Where $\hat{\theta}_{ij}$ is the estimator of the true effect size in the ith study and jth cluster β_0 is the average population effect, $\zeta_{(2)ij}$ represents study-level heterogeneity, and e_{ij} is the sampling error. This method is discussed in more detail in Assink 2016 and Cheung 2014.[2,5]

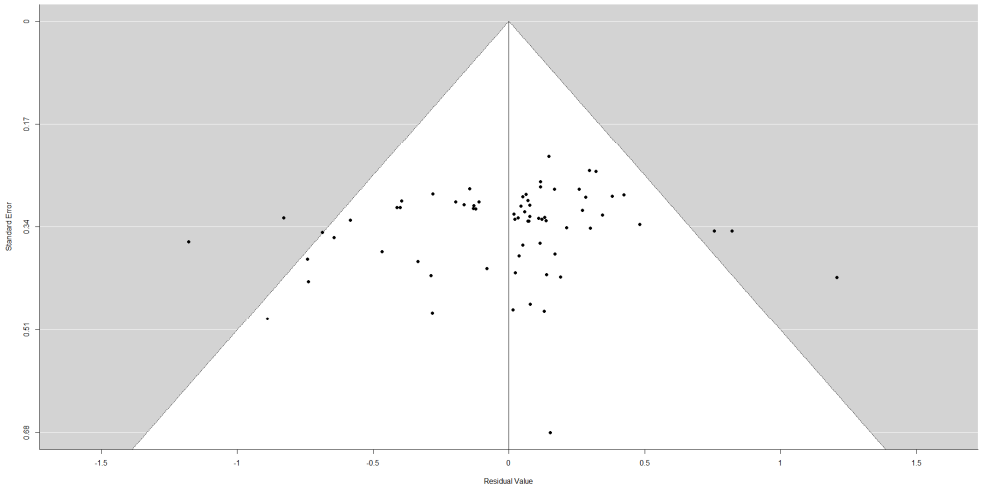
Distribution of Total Variance:

Distribution of total variance was implemented in R as per demonstrated in Harrer et al., 2019.[3]

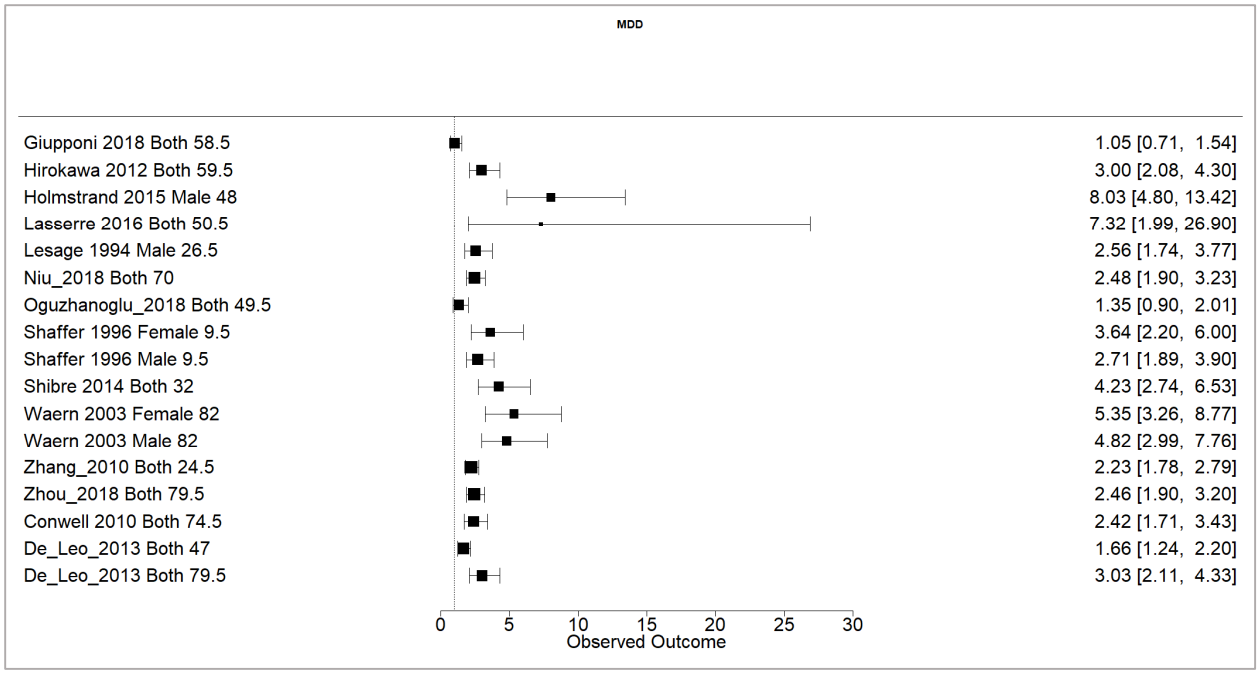
Baujat Plot for sensitivity analyses[1]:
Exponentiated intercept from standard meta-regression before outlier removal: 6·18 [3·83, 9·99]
Exponentiated intercept from standard meta-regression after outlier removal: 6·16 [4·21, 8·99]

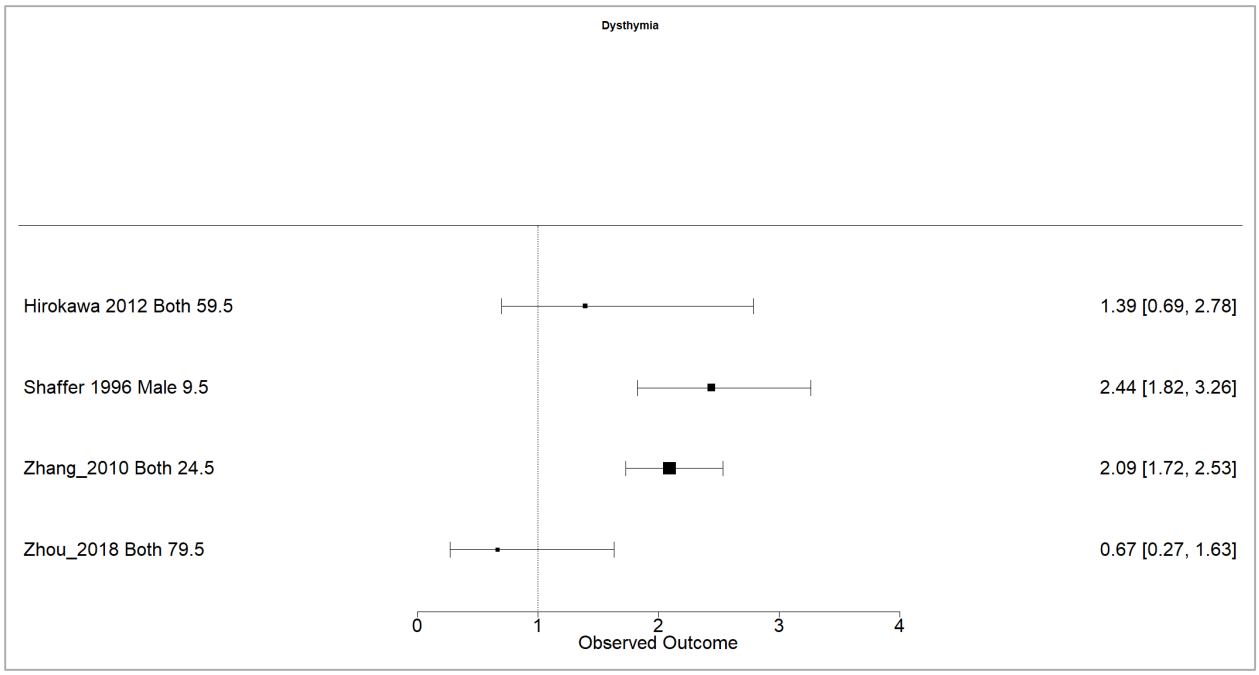


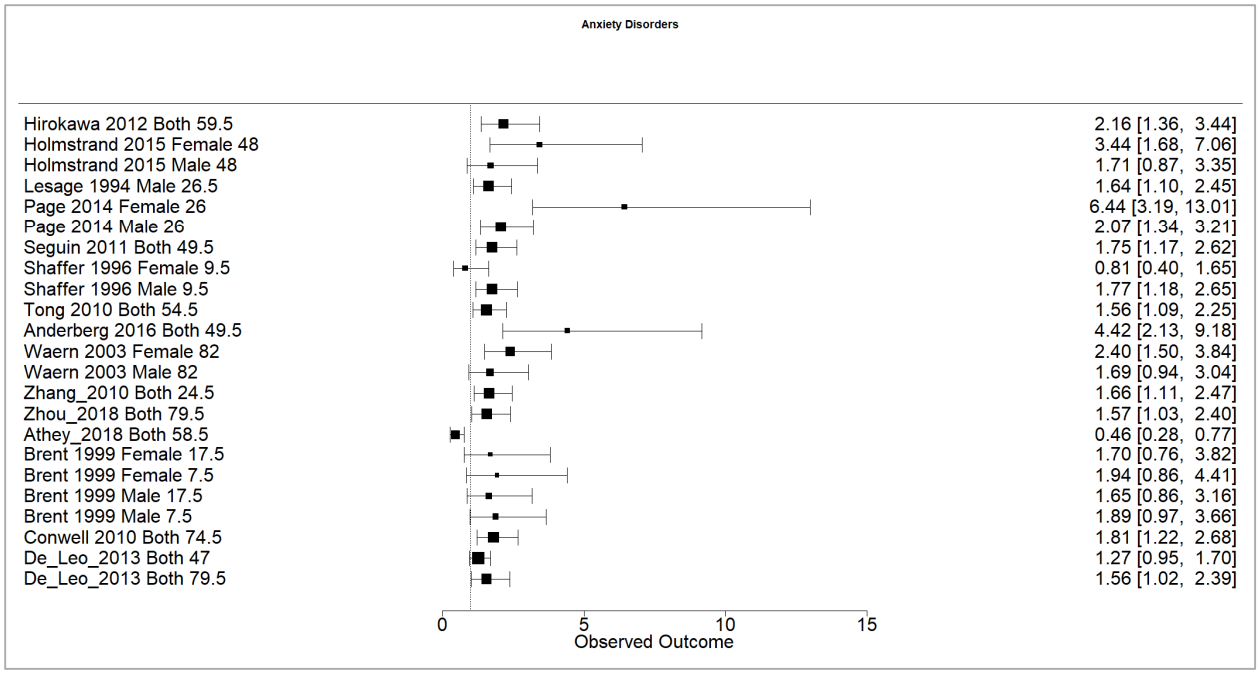
Funnel Plot to assess risk of bias:

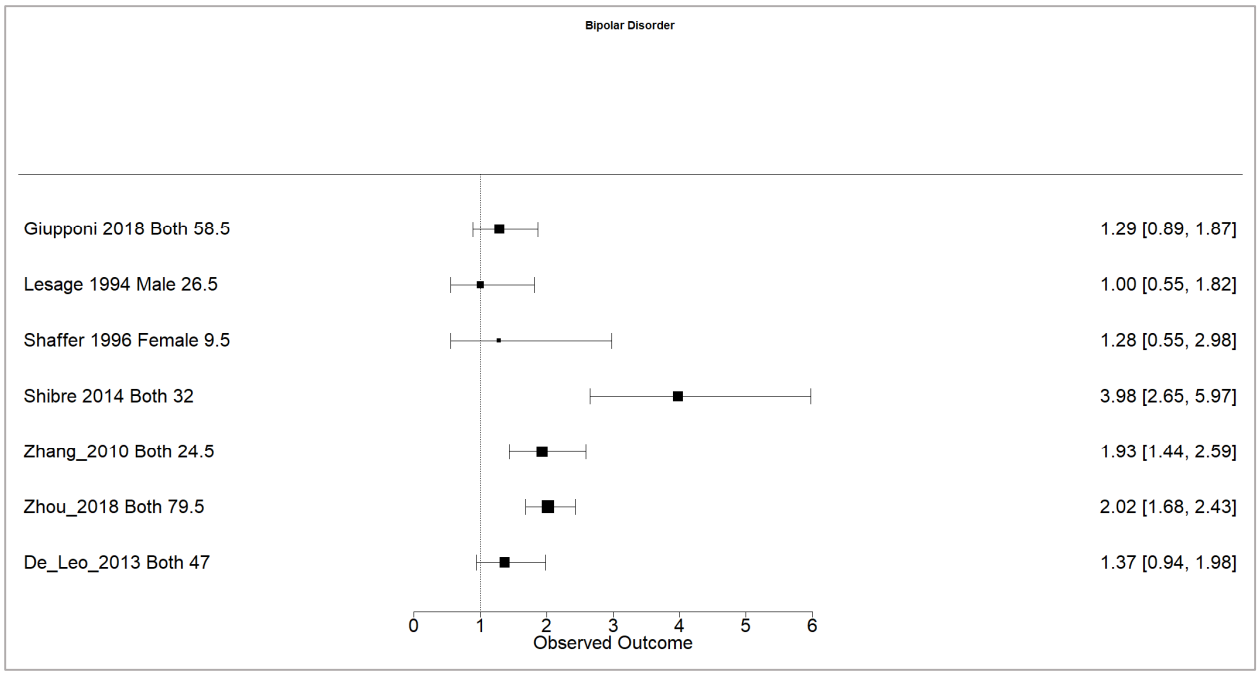


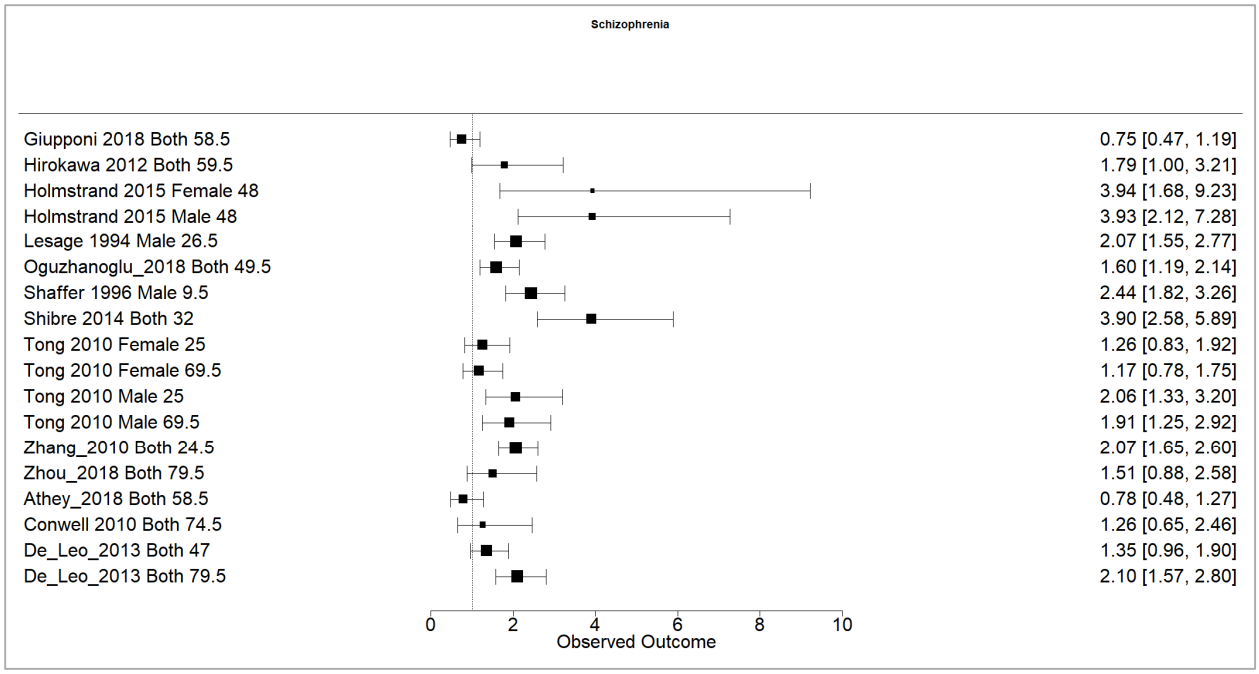
Relative Risk Forest Plots by Disorder (Labeled by Author, Year, Sex, and Mid-point of age range):


MDD
Giupponi 2018 Both 58.5
Hirokawa 2012 Both 59.5
Holmstrand 2015 Male 48
Lasserre 2016 Both 50.5
Lesage 1994 Male 26.5
Niu_2018 Both 70
Oguzhanoglu_2018 Both 49.5
Shaffer 1996 Female 9.5
Shaffer 1996 Male 9.5
Shibre 2014 Both 32
Waern 2003 Female 82
Waern 2003 Male 82
Zhang_2010 Both 24.5
Zhou_2018 Both 79.5
Conwell 2010 Both 74.5
De_Leo_2013 Both 47
De_Leo_2013 Both 79.5
1.05 [0.71, 1.54]
3.00 [2.08, 4.30]
8.03 [4.80, 13.42]
7.32 [1.99, 26.90]
2.56 [1.74, 3.77]
2.48 [1.90, 3.23]
1.35 [0.90, 2.01]
3.64 [2.20, 6.00]
2.71 [1.89, 3.90]
4.23 [2.74, 6.53]
5.35 [3.26, 8.77]
4.82 [2.99, 7.76]
2.23 [1.78, 2.79]
2.46 [1.90, 3.20]
2.42 [1.71, 3.43]
1.66 [1.24, 2.20]
3.03 [2.11, 4.33]
0
5
10
15
20
25
30
Observed Outcome


Dysthymia
Hirokawa 2012 Both 59.5
Shaffer 1996 Male 9.5
Zhang_2010 Both 24.5
Zhou_2018 Both 79.5
1.39 [0.69, 2.78]
2.44 [1.82, 3.26]
2.09 [1.72, 2.53]
0.67 [0.27, 1.63]
0
1
2
3
4
Observed Outcome


Anxiety Disorders
Hirokawa 2012 Both 59.5
Holmstrand 2015 Female 48
Holmstrand 2015 Male 48
Lesage 1994 Male 26.5
Page 2014 Female 26
Page 2014 Male 26
Seguin 2011 Both 49.5
Shaffer 1996 Female 9.5
Shaffer 1996 Male 9.5
Tong 2010 Both 54.5
Anderberg 2016 Both 49.5
Waern 2003 Female 82
Waern 2003 Male 82
Zhang_2010 Both 24.5
Zhou_2018 Both 79.5
Athey_2018 Both 58.5
Brent 1999 Female 17.5
Brent 1999 Female 7.5
Brent 1999 Male 17.5
Brent 1999 Male 7.5
Conwell 2010 Both 74.5
De_Leo_2013 Both 47
De_Leo_2013 Both 79.5
2.16 [1.36, 3.44]
3.44 [1.68, 7.06]
1.71 [0.87, 3.35]
1.64 [1.10, 2.45]
6.44 [3.19, 13.01]
2.07 [1.34, 3.21]
1.75 [1.17, 2.62]
0.81 [0.40, 1.65]
1.77 [1.18, 2.65]
1.56 [1.09, 2.25]
4.42 [2.13, 9.18]
2.40 [1.50, 3.84]
1.69 [0.94, 3.04]
1.66 [1.11, 2.47]
1.57 [1.03, 2.40]
0.46 [0.28, 0.77]
1.70 [0.76, 3.82]
1.94 [0.86, 4.41]
1.65 [0.86, 3.16]
1.89 [0.97, 3.66]
1.81 [1.22, 2.68]
1.27 [0.95, 1.70]
1.56 [1.02, 2.39]
0
5
10
15
Observed Outcome


Bipolar Disorder
Giupponi 2018 Both 58.5
Lesage 1994 Male 26.5
Shaffer 1996 Female 9.5
Shibre 2014 Both 32
Zhang_2010 Both 24.5
Zhou_2018 Both 79.5
De_Leo_2013 Both 47
1.29 [0.89, 1.87]
1.00 [0.55, 1.82]
1.28 [0.55, 2.98]
3.98 [2.65, 5.97]
1.93 [1.44, 2.59]
2.02 [1.68, 2.43]
1.37 [0.94, 1.98]
0
1
2
3
4
5
6
Observed Outcome



Variance-Covariance Matrices for varying correlation thresholds:

We tested varying correlation thresholds (ρ = *0.25, 0.5, 0.75, 0.99*) to examine the impact on covariance in the dataset. The following tables comprise the variance covariance matrices for each of these correlation thresholds:

Correlation = 0.25

Covariate	Intercept	Age	Percent Female	Anxiety disorders	Bipolar disorder	Dysthymia	Schizophrenia
1	0.0831694	0.0000482	-0.0055802	-0.0030207	-0.0040220	-0.0001996	-0.0012211
2	0.0000482	0.0000042	-0.0000017	0.0000050	-0.0000339	-0.0000898	-0.0000021
3	-0.0055802	-0.0000017	0.0123723	-0.0001103	-0.0006613	0.0009476	0.0001629
4	-0.0030207	0.0000050	-0.0001103	0.0051977	0.0029406	-0.0009951	0.0019281
5	-0.0040220	-0.0000339	-0.0006613	0.0029406	0.0094631	-0.0055033	0.0015205
6	-0.0001996	-0.0000898	0.0009476	-0.0009951	-0.0055033	0.0368400	0.0030690
7	-0.0012211	-0.0000021	0.0001629	0.0019281	0.0015205	0.0030690	0.0042281
8	-0.0231218	0.0000486	-0.0013755	0.0006285	-0.0050962	0.0038317	-0.0001160
9	-0.0554192	-0.0001067	0.0019352	-0.0017063	0.0070746	-0.0033448	-0.0005314
10	-0.0021308	0.0000117	0.0006285	0.0004509	-0.0058256	0.0106377	0.0004919

Correlation = 0.5

Covariate	Intercept	Age	Percent Female	Anxiety disorders	Bipolar disorder	Dysthymia	Schizophrenia
1	0.0778663	0.0000585	-0.0057532	-0.0030823	-0.0044041	0.0019100	-0.0012163
2	0.0000585	0.0000038	-0.0000055	0.0000021	-0.0000362	-0.0000627	0.0000008
3	-0.0057532	-0.0000055	0.0126136	-0.0001132	-0.0006252	0.0005122	0.0000847
4	-0.0030823	0.0000021	-0.0001132	0.0045612	0.0029797	-0.0005641	0.0019597
5	-0.0044041	-0.0000362	-0.0006252	0.0029797	0.0085869	-0.0059201	0.0016278
6	0.0019100	-0.0000627	0.0005122	-0.0005641	-0.0059201	0.0384992	0.0026841
7	-0.0012163	0.0000008	0.0000847	0.0019597	0.0016278	0.0026841	0.0034740
8	-0.0203466	0.0000391	-0.0013901	0.0003805	-0.0045179	0.0030941	-0.0002791
9	-0.0528397	-0.0001038	0.0020393	-0.0008540	0.0068540	-0.0049906	-0.0003878
10	-0.0011584	0.0000060	0.0006355	0.0001722	-0.0055156	0.0110591	0.0003603

Correlation = 0.75

Covariate	Intercept	Age	Percent Female	Anxiety disorders	Bipolar disorder	Dysthymia	Schizophrenia
1	0.075774	0.000080	-0.006503	-0.003318	-0.004679	0.000647	-0.001550
2	0.000080	0.000004	-0.000007	-0.000004	-0.000032	0.000017	0.000004
3	-0.006503	-0.000007	0.012861	-0.000189	-0.000899	-0.000074	0.000055
4	-0.003318	-0.000004	-0.000189	0.003397	0.003768	-0.000644	0.000963
5	-0.004679	-0.000032	-0.000899	0.003768	0.009942	-0.001148	0.001957
6	0.000647	0.000017	-0.000074	-0.000644	-0.001148	0.024031	0.005021
7	-0.001550	0.000004	0.000055	0.000963	0.001957	0.005021	0.003408
8	-0.023953	0.000016	-0.000554	-0.000427	-0.005414	-0.001048	-0.001463
9	-0.046311	-0.000101	0.001448	0.001141	0.007358	-0.000368	0.001448
10	-0.002013	0.000003	0.001356	-0.001020	-0.005998	-0.000059	-0.001260

Correlation = 0.99

Covariate	Intercept	Age	Percent Female	Anxiety disorders	Bipolar disorder	Dysthymia	Schizophrenia
1	0.0784749	0.0001071	-0.0055100	-0.0012050	-0.0060333	0.0096453	0.0003516
2	0.0001071	0.0000028	-0.0000156	-0.0000140	-0.0000376	0.0000153	-0.0000139
3	-0.0055100	-0.0000156	0.0128874	-0.0000908	-0.0003763	0.0002479	0.0000224
4	-0.0012050	-0.0000140	-0.0000908	0.0034791	0.0004326	-0.0015109	-0.0008240
5	-0.0060333	-0.0000376	-0.0003763	0.0004326	0.0095542	-0.0051794	0.0021789
6	0.0096453	0.0000153	0.0002479	-0.0015109	-0.0051794	0.0165730	0.0026977
7	0.0003516	-0.0000139	0.0000224	-0.0008240	0.0021789	0.0026977	0.0018833
8	-0.0200492	-0.0000240	-0.0020584	-0.0006679	-0.0050803	0.0012128	-0.0012327

9	-0.0556621	-0.0000706	0.0024510	-0.0002857	0.0106815	-0.0104197	0.0011423
10	0.0060437	-0.0000373	0.0004447	0.0001848	-0.0091698	0.0109821	-0.0003580

References:

1. Baujat B, Mahé C, Pignon JP, Hill C. A graphical method for exploring heterogeneity in meta-analyses: application to a meta-analysis of 65 trials. Statistics in medicine. 2002 Sep 30;21(18):2641-52.Cheung
2. Assink M, Wibbelink CJ. Fitting three-level meta-analytic models in R: A step-by-step tutorial. The Quantitative Methods for Psychology. 2016 Oct 1;12(3):154-74.
3. Harrer, M., Cuijpers, P., Furukawa, T.A, & Ebert, D. D. (2019). Doing Meta-Analysis in R: A Hands-on Guide. DOI: 10.5281/zenodo.2551803.
4. James SL, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet. 2018;392(10159):1789-858.
5. Cheung MW-L. Modeling dependent effect sizes with three-level meta-analyses: a structural equation modeling approach. Psychological Methods. 2014;19(2):211.

2. Chapter 2: Assessing the Efficacy of Pharmacological and Therapeutic Interventions for Major Depressive Disorder: An application of a novel network meta-analytic method (MR-BRT

2.1 Introduction:

The Global Burden of Disease 2019 study ranks Major depressive disorder (MDD) as the fifth leading cause of disability globally (higher than other commonly known chronic conditions such as type 2 diabetes and ischemic stroke). When compared by sex, MDD accounted for 3.7% [2.8%, 4.8%] of total years lived with disability (YLDs) in men and 4.8% [3.7%, 6.2%] of total YLDs in women [1]. MDD is also known to be a major risk factor for fatal outcomes such as suicide. The risk of suicide among people with MDD is approximately 7.6 [4.3, 13.6] times that of people without MDD [2].

The Global Burden of Disease reports the most comprehensive analysis of the burden of MDD via a thorough review of all available and usable epidemiological data sources [1]. The burden of MDD is quantified via the estimation of disability-adjusted life years (DALYs) which comprise both years lived with disability (YLDs) and years of life lost (YLLs). MDD is modeled as a non-fatal condition in the GBD and therefore comprises YLDs only. Given the magnitude of non-fatal disability accounted for by MDD, it is important to accurately estimate the burden of this disability experienced globally. This estimation process requires good quality epidemiological data and information on the distribution of severity of MDD. Severity distributions within the GBD analysis summarize the range of functional health loss across prevalent cases of MDD and are incorporated

within the estimation of YLDs. In GBD 2019 this was expressed as the proportion of MDD cases that are symptomatic, mild, moderate, or severe [1]. At present, globally representative data on the severity of MDD is lacking [3]. Severity distributions used in the GBD are derived from a single survey from Australia, the Australian National Survey of Mental Health and Wellbeing of Adults (NSMHWB) 1997 [4,5]. This survey captured the prevalence of multiple mental and physical disorders along with health status information as measured by the Short Form 12-item questionnaire (SF-12) [6]. The GBD study estimates the prevalence of MDD in the past month and the NSMHWB was the only survey with diagnoses and SF-12 questions pertaining to severity in the same timeframe. Therefore, the distribution of severity for one-month diagnoses was used from this survey. Although this approach allows us to arrive at some estimate of the severity distribution of MDD, the severity distribution is only representative of Australia in 1997. This severity distribution is then applied to all locations in GBD without accounting for variation in treatment efficacy or quality by geography or time [3]. It is reasonable to expect that this variation in treatment efficacy may influence the severity proportions for MDD. In other words, the disability burden due to MDD may be influenced by the efficacy of available treatment options. Current methods of estimation in use may then potentially underestimate the global burden of MDD – as MDD severity will likely be less severe in high-income countries (including Australia) with access to treatments with known efficacy.

To facilitate more precise assessment of severity within the GBD framework for MDD, it is first important to compare existing treatment options to obtain pooled treatment effects by treatment types that may be used to improve estimation methods for the severity distribution of MDD. Several pharmacological, physical, and psychosocial interventions can effectively treat MDD [7-10]. A review of 198 studies on seven types of psychological interventions found that each of these interventions were superior to a waitlist control condition with moderate to large effects on depressive symptoms in adults [7]. A review of trials on 21 antidepressants found that all antidepressants were more effective than placebo in the acute treatment of MDD in adults [8]. However, existing reviews are limited because they only provide evidence on efficacy for specific treatments being compared directly (for example, Treatment A compared against Treatment B, and B against Treatment C, but no evidence on comparisons between A and C). XXX Existing efficacy reviews typically summarize treatment effects from published psychiatric trials of specific treatment comparisons to report a pooled treatment effect. Hetrick et. al reviewed 10 studies comparing combined antidepressants and benzodiazepines to antidepressants alone. They also reviewed evidence from 19 trials for newer generation antidepressants compared to placebo for treating depressive disorders in children and adolescents [11]. Similarly, other reviews report a summarized treatment effect size based on select treatment pairs [12-14]. However, individual pairwise direct comparisons commonly reported in trials may not provide comparative efficacy estimates for all known treatment types. This potentially limits our understanding of efficacy across a more comprehensive set of treatments beyond those directly compared. In order to address this limitation, a network meta-analysis (NMA) approach is used to generate indirect evidence for treatment efficacy not otherwise directly estimated. Subsequently, a pooled effect size for the entire treatment ‘network’ can be generated using both direct and indirect evidence. Cipriani and colleagues’ most recent review employs the network meta-analytic approach to analyze multiple treatment pairs and generate evidence for efficacy across direct and indirect treatment comparisons [8]. An earlier review by Zhou et al. on psychotherapies for depression in children and young adults also uses NMA within a Bayesian framework [15]. Although

this approach allows us to gain higher precision by considering all available evidence for treatment efficacy, several important limitations arise. First, studies examining the efficacy of therapeutic interventions use control conditions such as care as usual that may not entirely simulate a condition of no access to treatment. Second, many earlier studies report on a select group of treatments for specific samples (example, samples with severe MDD only or with specific medical conditions) that provide only partial evidence on treatment efficacy. Third, studies that do use an NMA approach may not account for between-study heterogeneity that may account for variation in reported treatment effect sizes. In this analysis, we estimate pooled treatment effect sizes for MDD using a network meta-analytic approach. We attempt to address the first limitation discussed above by exploring the use of waitlist as a control condition for therapies (more details in the methods section). We attempt to address the second and third limitations by using a novel Bayesian network meta-analytic approach that allows for all treatment classes to be combined into a single network and incorporates between-study heterogeneity in estimation of uncertainty [16] ..

2.2 Methods:

Data Sources:
We made use of existing datasets on treatment efficacy for pharmacological and psychological interventions [7,8,10]. The search strategy and data collection procedures are described elsewhere [7, 17,18]. Briefly, these datasets comprised data extracted from existing systematic reviews of randomized controlled trials (RCTs) published through 2019 to examine the comparative efficacy of interventions for MDD.

Psychological Interventions:
RCTs on psychological interventions were obtained via a comprehensive literature search in PubMed, PsycINFO, Embase, the Cochrane Central Register of Controlled Trials and the World Health Organization Afro library. RCTs comparing psychological interventions to each other or to a control condition such as waitlist or care as usual were used for this analysis. Studies that recruited participants either from community, clinical, or other settings (including systematic screening or general medical settings) were included. Studies were excluded if they did not use depression as an inclusion criterion or if treatment was not aimed at depression, were maintenance studies aimed at preventing relapse or maintaining outcomes or did not report sufficient information to calculate an effect size.

Pharmacological Interventions:
RCTs on pharmacological interventions were obtained via a comprehensive search of databases including the Cochrane Central Register of Controlled Trials, CINAHL, Embase, LILACS database, MEDLINE, MEDLINE In-Process, PsycINFO, the websites of regulatory agencies such as the Food and Drug Administration (FDA) in the USA, the Medicines and Healthcare products Regulatory Agency in the UK, the European Medicines Agency, (EMA) in the European Union, the Medicines Evaluation Board in the Netherlands, the Medical Products Agency in Sweden, the Pharmaceuticals and Medical Devices Agency (PMDA) in Japan, and the Therapeutic Goods Administration (TGA) in Australia. International registers for published and unpublished, double-blind, randomized controlled trials were also searched. Placebo-controlled and head-to-head trials of 21 antidepressants used for the acute treatment of adults (≥18 years old and of

both sexes) with MDD diagnosed according to standard operationalized criteria [19, 20] were included. Quasi-randomized trials and trials that were incomplete or included 20% or more of participants with bipolar disorder, psychotic depression, or treatment-resistant depression; or patients with a serious concomitant medical illness were excluded. RCTS with participants diagnosed with postpartum depression were excluded as the profile of postpartum depression is clinically different from major depression [17].

In addition to the above, we made use of five additional studies comparing cognitive behavioral therapy (CBT) to pill placebo used by the Furukawa and colleagues [10]. We collaborated with the principal investigators of these studies to ensure that the datasets were complete and comprehensive.

Data Extraction:
Datasets compiled from selected RCTs included information on mean depression symptom scores and standard deviations for treatment and comparison arms, type of treatment and comparison used, symptom scales used to assess depressive symptoms, and additional sample characteristics including sample size, follow-up duration, and remission thresholds. Treatments were grouped into distinct broader classes using the USA Food and Drug Administration drug classification index [21]. Pharmacological treatments were grouped into seven classes: atypicals, noradrenaline reuptake inhibitors, noradrenergic and specific serotonergic antidepressants, serotonin modulators, serotonin-norepinephrine reuptake inhibitor (SNRIs), selective serotonin reuptake inhibitor (SSRIs), and tricyclics. Placebo was used as the control condition for pharmacological treatments. Distinct comparisons between intervention and comparison classes were used for this analysis. A full list of specific drug names and the corresponding classes are provided in Table 1 in the appendix.

For psychological interventions, treatment classes used for this analysis included cognitive behavioral therapy (CBT), counseling, care as usual (CAU), and other psychological interventions. Waitlist (WL) was chosen as the control condition used for this analysis. Although some studies may use CAU as the control condition, in this analysis, it was important to reflect a control condition that best reflects ‘no treatment’. It is plausible that participants randomized to CAU may be able to access treatment elsewhere and this may not accurately reflect a scenario where they receive no access to treatment [18]. Waitlist participants may expect to receive treatment and therefore refrain from seeking treatment elsewhere if they are randomized to this trial arm. Therefore, we chose to use Waitlist as the control condition for psychological interventions.

Analytic Approach:
The standardized mean difference (SMD) in depression symptom scores between each treatment and comparison pair was calculated using the ‘escalc’ function available in the metafor package in R. Binary variables were created for each intervention class. Funnel plots were used to assess publication bias in the data for antidepressants. SMDs were modeled in a network meta-analysis using the Meta Regression: Bayesian, Regularized Trimmed (MR-BRT) framework [16]. MR-BRT incorporates between-study heterogeneity into the estimated effect. It also uses a trimmed maximum likelihood estimator which excludes outlier effects based on their contributions to the likelihood function. Trimming

was set to 10% of the observed effect sizes. Pooled SMDs by each intervention class were estimated and reported in the results section below.

2.3 Results:

A total of 366 studies on pharmacological interventions and 241 studies on psychological interventions comprising 839 data points were used. Among pharmacological interventions, most studies reported on comparisons with SSRIs, SNRIs, and atypical antidepressants. Among therapeutic interventions, most studies reported on comparisons with cognitive behavioral therapy, followed by other categories. The funnel plot showing evidence of publication bias in Figure 1 below represents treatment effect sizes and standard errors from studies comparing antidepressant interventions to placebo. The red circles in Figure 1 represent effect sizes that were identified as outliers and trimmed by MR-BRT. The pooled SMD for each treatment class is presented in Table 1. All treatment effects were statistically significant ($p < 0.05$). Among drug intervention classes, the largest effects were observed for Tricyclics [-0.89 (-1.35, -0.43)], followed by SNRIs [-0.86 (-1.26, -0.42)], and Noradrenergic and specific serotonergic reuptake inhibitors [-0.83, (-1.32, -0.33)]. Among therapy intervention classes, the largest treatment effects were observed for CBT [-0.85, (-0.92, -0.78)], followed by other psychological interventions [-0.76, (-0.89, -0.64)], and counseling [-0.73, (-0.86,
-0.6)].

Treatment Class	Pooled SMD [95% UIs]
Atypical	-0.67 [-1.11, -0.22]
Noradrenaline reuptake inhibitor	-0.63 [-1.12, -0.14]
Noradrenergic and specific serotonergic antidepressants	-0.83 [-1.32, -0.33]
Serotonin modulators	-0.67 [-1.21, -0.12]
Selective Serotonin Reuptake Inhibitor (SSRI)	-0.73 [-1.15, -0.3]
Serotonin-norepinephrine reuptake Inhibitor (SNRI)	-0.86 [-1.29, -0.42]
Tricyclics	-0.89 [-1.35, -0.43]
CBT	-0.85 [-0.92, -0.78]
Counseling	-0.73 [-0.86, -0.6]
Other Psychological interventions	-0.76 [-0.89, -0.64]
Care as Usual	-0.29 [-0.38, -0.2]

2.4 Discussion:

This analysis demonstrates an application of the network meta-analytic framework using MR-BRT to a comprehensive dataset on treatment efficacy estimates obtained from existing systematic reviews of treatment studies on drug and psychotherapeutic interventions for MDD. For drug classes, the largest treatment effects were found for

tricyclics, SNRIs, and noradrenergic and specific serotonergic antidepressants, followed by others. For psychotherapeutic classes, the largest treatment effects were found for CBT, other psychological interventions, and counseling.

These findings are largely consistent with estimates obtained in previous network meta-analytic review studies [7-15]. Based on this analysis, the treatment efficacy for the most effective antidepressants (when compared to pill placebo) and therapies (when compared to waitlist) are relatively comparable. . CBT is highly beneficial in comparison to other psychological interventions as well as compared to many classes of antidepressants in use. The coefficient estimate obtained for the pill placebo condition was -0.45 [-0.84, -0.05] which indicates that it is an effective control condition (and more effective than CAU). Our results show a low but significant effect size for CAU versus waitlist condition. For studies included in this dataset, participants in the CAU group are assumed to have access to some form of regular routine care. The significant effect size may be indicative of some benefit of receiving usual care as opposed to being waitlisted – which depending on the trial setting, may mean receiving no treatment at all and/or deferring treatment in expectation of receiving care. A 2014 study on psychological interventions by Furukawa and colleagues found that those in the no-treatment group had a superior response compared to those in the waitlisted group [22]. This was attributed to the possibility that no-treatment group participants could theoretically seek care elsewhere whereas the waitlist group may delay care in the expectation of receiving treatment after the study, thereby contributing to possible worsening of symptoms [22]. While no-treatment and waitlist groups are not exactly equivalent, our finding conforms with the initial premise that waitlist most closely resembles a scenario of zero access to treatment and is therefore a suitable choice for a control condition for therapies in this analysis.

This analysis was constrained by lack of data on other relevant characteristics that may impact treatment effects such as sample characteristics, depression severity, or geographic coverage. Therefore, we were unable to examine these attributes analytically. Exploring variation in treatment effects by type of outcome measure used to score symptoms of depression was of particular interest; variation in symptom scale items may contribute to differences in how MDD symptoms are captured. However, the data included here were extracted using an algorithm that prioritized using effect sizes derived using scores from the Hamilton Rating Scale for Depression above effect sizes scored using other scales (Montgomery-Asberg Depression Rating Scale or MADRS, Beck Depression Inventory, and others). Therefore, this data collection algorithm may bias the data towards a few commonly used symptom scales as was the case with this dataset. There were very few intervention-comparison pairs for some individual drugs, broader intervention comparisons had to be used so that the model was sufficiently powered. Future efforts in this area will benefit from more detailed datasets that may provide more granular analyses of attributes of interest [18].

We examined the data on antidepressants to determine the presence of publication bias. Publication bias exists if the chances of a study getting published depends on whether or not the results are statistically significant [23]. Therefore, limiting the literature search to published trials only may lead to potential overestimation of the treatment effect sizes. Figure 1 below shows that these data may be susceptible to publication bias as evidenced by the asymmetrical distribution of treatment effect sizes. It also shows treatment effect sizes that were detected as outliers and trimmed using MR-BRT which may reduce the impact of existing publication bias. The distribution of effect sizes in

Figure 1 is largely consistent with the publication bias shown in the funnel plot reported in the most recent publication (in 2018) by our external collaborators [18]. Although the antidepressants dataset included data from unpublished trials (61 studies accounting for 108 data points), it is nonetheless predominantly informed by published data which may contribute to publication bias. Therefore, antidepressant treatment effect sizes reported here should be treated with caution in the context of publication bias.

Findings reported here provide comparative evidence for both therapies and antidepressants that may guide clinical service providers in determining which combinations of treatment approaches may be most effective and feasible. Although psychological interventions such as CBT have been shown to be highly effective, they may be complex to administer and potentially require extensive investments in training that may not be available in all settings. Recent adaptations to create scalable psychological interventions for use in low resource settings somewhat mitigate this challenge [24]. Our analysis included efficacy estimates on a wide range of antidepressants, all of which may not be readily available in many countries (especially the more recent antidepressants). Therefore, evidence on older classes of drugs such as tricyclics may be useful for many resource-constrained settings and can potentially contribute to minimally adequate treatment for MDD.

Our analysis incorporates several improvements over existing work in this area. First, we explore the use of waitlist as a control condition for therapies instead of the more commonly used CAU condition. Second, we leverage the use of the novel Bayesian meta-regression tool (MR-BRT) to incorporate efficacy data on all known pharmacological and therapeutic interventions for MDD in a single network. Third, we incorporate random effects to account for between-study heterogeneity which allows for a more accurate estimate of uncertainty accompanying estimated pooled treatment effects.

Findings from this analysis may serve as important inputs in the severity estimation process for MDD within the Global Burden of Disease framework. Current severity estimation methods apply the same severity distribution derived from a single survey conducted in Australia in 1997 to all locations without accounting for variation in treatment efficacy. While this may be representative for high-income countries, it is not generalizable to the rest of the world. Furthermore, it is plausible that the efficacy of treatments will influence the burden of MDD. This analysis provides pooled estimates of treatment efficacy from a comprehensive systematic review that is more representative than the one NSMHWB survey, that may be used to adjust severity distributions for MDD. Another important component that will better inform the MDD severity distribution is treatment coverage which will be covered subsequently in Chapter 3.

2.5 Conclusions:

This analysis presents an initial application of the novel MR-BRT framework to data on treatment efficacy for MDD. These findings provide comparative evidence for both therapies and antidepressants that may be useful to clinical service providers. Future efforts may incorporate attributes such as MDD assessment methods, sample characteristics, depression severity, and trial setting that may additionally inform variation in treatment efficacy to obtain estimates with greater accuracy and precision.

2.6 Significance of outputs from Chapter 2:

This chapter explored the application of a novel Bayesian meta-regression framework to estimate the efficacy of pharmacological and therapeutic interventions for MDD. We extend upon previous work by reporting pooled treatment effects for all known treatment classes in a single network along with incorporating between-study heterogeneity. We also make use of waitlist as a control condition for therapies which may better reflect a scenario of no access to treatment compared to care as usual or other control conditions used before. The outputs from Chapter 2 along with treatment rates (covered in Chapter 3) may find useful applications in re-estimating the severity distribution for MDD. Severity distributions summarize the range of functional health loss across prevalent cases of MDD. This is usually expressed as the proportion of MDD cases as asymptomatic, mild, moderate, or severe. The impact of treatment efficacy on the severity distribution of MDD is of particular interest because the GBD study currently applies the same severity distribution derived from surveys in high-income settings (such as Australia) to all locations without accounting for variation in MDD treatment efficacy and access. The treatment effects estimated in Chapter 2 provide a larger dataset of efficacy estimates that may be used to re-estimate severity distributions for MDD.

Treatment rates and projected gaps in treatment coverage are another important input in addition to treatment efficacy that may better inform the severity distribution of MDD. In Chapter 3, we will turn to an examination of treatment rates for MDD. The concluding section on the significance of the outputs for Chapter 3 will then offer a cohesive discussion of how both treatment effects and treatment coverage sequentially explored in this thesis, may be used as inputs to improve methods for MDD severity estimation.

2.7 References:

1. Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, Abbasi-Kangevari M, Abbastabar H, Abd-Allah F, Abdelalim A, Abdollahi M. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet. 2020 Oct 17;396(10258):1204-22.
2. Moitra M, Santomauro D, Degenhardt L, Collins PY, Whiteford H, Vos T, Ferrari A. Estimating the risk of suicide associated with mental disorders: A systematic review and meta-regression analysis. Journal of psychiatric research. 2021 Mar 2.
3. Wyper GM, Assuncao R, Fletcher E, Gourley M, Grant I, Haagsma JA, Hilderink H, Idavain J, Lesnik T, Von der Lippe E, Majdan M. The increasing significance of disease severity in a burden of disease framework. Scandinavian Journal of Public Health. 2021 Jul 2:14034948211024478.
4. Burstein R, Fleming T, Haagsma J, Salomon JA, Vos T, Murray CJ. Estimating distributions of health state severity for the global burden of disease study. Population health metrics. 2015 Dec 1;13(1):31.

5. Salomon JA, Vos T, Hogan DR, Gagnon M, Naghavi M, Mokdad A, Begum N, Shah R, Karyana M, Kosen S, Farje MR. Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010. The Lancet. 2012 Dec 15;380(9859):2129-43
6. Statistics AB. National survey of mental health and wellbeing: summary of results. Canberra, Australian Bureau of Statistics. 2007.
7. Barth J, Munder T, Gerger H, Nüesch E, Trelle S, Znoj H, Jüni P, Cuijpers P. Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis. Focus. 2016 Apr;14(2):229-43.
8. Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ, Ogawa Y, Leucht S, Ruhe HG, Turner EH, Higgins JP, Egger M. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. Focus. 2018 Oct;16(4):420-9.
9. Lynch D, Laws KR, McKenna PJ. Cognitive behavioural therapy for major psychiatric disorder: does it really work? A meta-analytical review of well-controlled trials. Psychological medicine. 2010 Jan;40(1):9-24.
10. Furukawa TA, Weitz ES, Tanaka S, Hollon SD, Hofmann SG, Andersson G, Twisk J, DeRubeis RJ, Dimidjian S, Hegerl U, Mergl R. Initial severity of depression and efficacy of cognitive–behavioural therapy: individual-participant data meta-analysis of pill-placebo-controlled trials. The British Journal of Psychiatry. 2017 Mar;210(3):190-6.
11. Hetrick SE, McKenzie JE, Cox GR, Simmons MB, Merry SN. Newer generation antidepressants for depressive disorders in children and adolescents. Cochrane Database of Systematic Reviews. 2012(11).
12. Komossa K, Depping AM, Gaudchau A, Kissling W, Leucht S. Second-generation antipsychotics for major depressive disorder and dysthymia. Cochrane Database of Systematic Reviews. 2010(12).
13. Churchill R, Moore TH, Furukawa TA, Caldwell DM, Davies P, Jones H, Shinohara K, Imai H, Lewis G, Hunot V. 'Third wave'cognitive and behavioural therapies versus treatment as usual for depression. Cochrane Database of Systematic Reviews. 2013(10).
14. Peng L, Min S, Wei K, Ziemann-Gimmel P. Different regimens of intravenous sedatives or hypnotics for electroconvulsive therapy (ECT) in adult patients with depression. Cochrane database of systematic reviews. 2014(4).
15. Zhou X, Hetrick SE, Cuijpers P, Qin B, Barth J, Whittington CJ, Cohen D, Del Giovane C, Liu Y, Michael KD, Zhang Y. Comparative efficacy and acceptability of psychotherapies for depression in children and adolescents: A systematic review and network meta-analysis. World psychiatry. 2015 Jun;14(2):207-22.
16. Zheng P, Barber R, Sorensen RJD, et al. Trimmed constrained mixed effects models: formulations and algorithms. Journal of Computational and Graphical Statistics 2021; 1–13.
17. Furukawa TA, Salanti G, Atkinson LZ, Leucht S, Ruhe HG, Turner EH, Chaimani A, Ogawa Y, Takeshima N, Hayasaka Y, Imai H. Comparative efficacy and acceptability of first-generation and second-generation antidepressants in the acute treatment of major depression: protocol for a network meta-analysis. BMJ open. 2016 Jul 1;6(7):e010919.
18. Cuijpers P, Reijnders M, Karyotaki E. Psychological treatment of depression: An updated meta-analytic database of randomized trials. https://evidencebasedpsychotherapies.shinyapps.io/metapsy/Submitted for publication. 2018.
19. Organization WH. The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines: Geneva: World Health Organization; 1992.

20. Association AP. Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR). Washington DC: American Psychiatric Assocation. 2000.
21. U.S. Food and Drug Administration. General Drug Categories (2015). https://www.fda.gov/drugs/investigational-new-drug-ind-application/general-drug-categories
22. Furukawa TA, Noma H, Caldwell DM, Honyashiki M, Shinohara K, Imai H, Chen P, Hunot V, Churchill R. Waiting list may be a nocebo condition in psychotherapy trials: a contribution from network meta-analysis. Acta Psychiatrica Scandinavica. 2014 Sep;130(3):181-92.
23. Dechartres, Agnes, Ignacio Atal, Carolina Riveros, Joerg Meerpohl, and Philippe Ravaud. 2018. "Association Between Publication Characteristics and Treatment Effect Estimates: A Meta-Epidemiologic Study." Annals of Internal Medicine 169 (6): 385–93.
24. World Health Organization. Scalable psychological interventions for people in communities affected by adversity: a new area of mental health and psychosocial work at WHO. 2017.

2.8 List of Tables and Figures

Treatment Class	Pooled SMD [95% UIs]
Atypical	-0.67 [-1.11, -0.22]
Noradrenaline reuptake inhibitor	-0.63 [-1.12, -0.14]
Noradrenergic and specific serotonergic antidepressants	-0.83 [-1.32, -0.33]
Serotonin modulators	-0.67 [-1.21, -0.12]
Selective Serotonin Reuptake Inhibitor (SSRI)	-0.73 [-1.15, -0.3]
Serotonin-norepinephrine reuptake Inhibitor (SNRI)	-0.86 [-1.29, -0.42]
Tricyclics	-0.89 [-1.35, -0.43]

CBT	-0.85 [-0.92, -0.78]
Counseling	-0.73 [-0.86, -0.6]
Other Psychological interventions	-0.76 [-0.89, -0.64]
Care as Usual	-0.29 [-0.38, -0.2]

Table 1. Pooled SMDs and 95% UIs by treatment class

Note: Treatment classes compared against control conditions of pill placebo (drugs) or waitlist (therapies); CBT = Cognitive behavioral therapy. Values reported here represent pooled SMDs between treatment class and comparison group. Pooled SMDs that are further away from zero (i.e., more negative) represent a higher treatment effect compared to those that are closer to zero (i.e., less negative)

Drug Class	Drug Name
Atypical	agomelatine
	bupropion
	trazodone
Noradrenaline reuptake inhibitor	reboxetine
Noradrenergic and specific serotonergic antidepressants	mirtazapine
Serotonin modulators	nefazodone
serotonin-norepinephrine reuptake Inhibitor (SNRI)	duloxetine
	milnacipran
	venlafaxine
Selective Serotonin Reuptake Inhibitor (SSRI)	citalopram
	fluoxetine
	fluvoxamine
	paroxetine
	sertraline
	vilazodone
	vortioxetine
Tricyclics	amitriptyline
	clomipramine

Table 2. List of Drug Classes and Drug Names

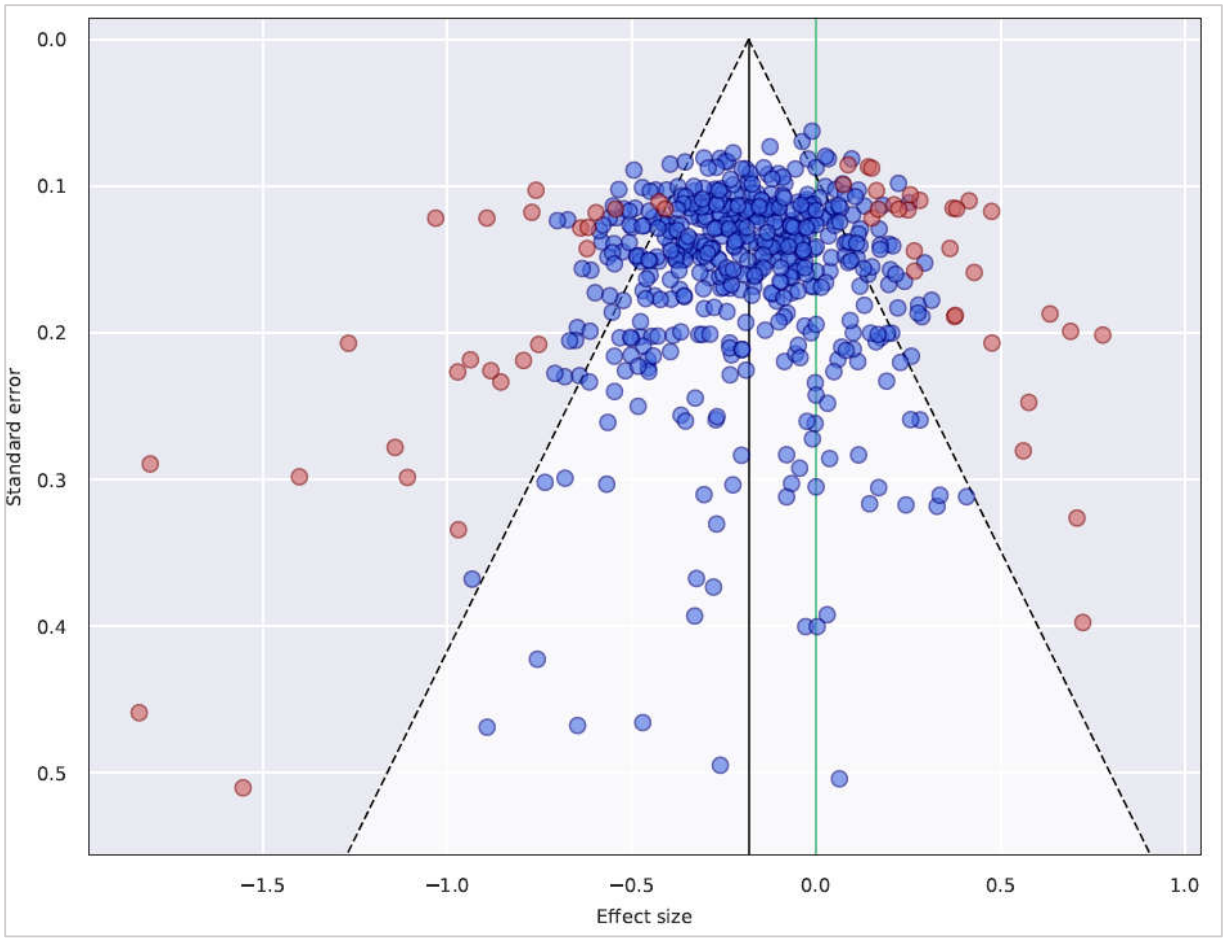


Figure 1. Funnel Plot of treatment effect sizes and standard errors from studies comparing antidepressant interventions to placebo. [Red circles represent outliers identified by MR-BRT]

Chapter 3: Modeling global treatment rates for major depressive disorder in 83 countries from 2000-2019: A systematic review and Bayesian meta-regression analysis.

[Note: The contents of this chapter have been submitted as a publication for review to PLOS Medicine (2021)]

3.1 Introduction

Depressive disorders are highly prevalent and disabling. The Global Burden of Disease Study 2019 (GBD 2019) estimated that major depressive disorder (MDD) and dysthymia were jointly responsible for 46.9 million disability-adjusted life years (DALYs) globally in 2019, with each DALY equivalent to a healthy year of life lost to the disability caused by depressive disorders [1]. When benchmarked against a total of 369 diseases and injuries, depressive disorders were the 13th leading cause of overall burden and the 7th leading cause of non-fatal burden, globally [1, 2]. The impact of depressive disorders also goes beyond the disability and mortality captured by the DALY. People with depressive disorders, caregivers, employers, and governments must manage the associated reductions in work productivity and increased reliance on state health and welfare services [3]. Depressive disorders are also known to be major risk factors for fatal outcomes such as suicide [4, 5].

Effective and efficient treatment strategies are available for depressive disorders and consist of pharmacotherapy, psychological, and social interventions [6]. In recent years, there have been signs of increasing global commitment to prioritize mental health and reduce the burden imposed by severe forms of mental disorders such as MDD. In 2013, the World Health Assembly adopted the Comprehensive Mental Health Action Plan which was extended to 2030 at the 72nd World Health Assembly [7, 8]. Amongst the global targets set are for member states to increase service coverage for severe mental disorders by 20% by the year 2030 and to routinely collect information on key mental health indicators such as disorder prevalence and treatment. In 2015, the United Nations General Assembly passed the 2030 Agenda for sustainable development, which for the first time identified the promotion of mental health and well-being, and the prevention and treatment of substance abuse, as health priorities within the global development agenda [9].

These national and global advances suggest increasing commitment by governments to reduce the negative effects of mental and substance use disorders. However, despite depression being a major cause of disability, treatment rates for depression are remarkably low. An estimated 7% to 28% of those with depression receive appropriate care and treatment [3]. Previous reviews on depression treatment rates have found wide geographic variation by WHO region with gaps in treatment ranging from 45.4% in Europe to 67% in the African region and 70.2% in the Eastern Mediterranean region [10]. There also exists variation by resource setting in the quality of care received. The proportion of those receiving minimally adequate care (MAT) ranged from 22.4% in high-income countries to 3.7% in lower-middle income countries [11].

Health information systems in many countries are not designed to routinely collect data on key mental health indicators (such as treatment coverage) from which the extent of any progress can be measured [12]. As an alternative, we can turn to population-representative epidemiological surveys to estimate the treated and untreated prevalence of depressive disorders as an indicator of treatment gaps. Whilst epidemiological surveys capturing data on service use for individuals with depressive disorders exist, efforts to assemble and critically evaluate the data for a representative global summary of treatment rates are outdated or do not capture all available information [13]. Furthermore, most existing reviews rely on a descriptive summarization of treatment rates or gaps, without accounting for variation in study methodology that may potentially contribute to heterogeneity in the existing evidence thereby resulting in imprecise estimates.

In this paper, we sought to update the 2004 study by Kohn and collaborators who undertook a literature review of population surveys of mental and substance use disorders for data on the proportion of individuals receiving care [10]. Treatment gap here referred to the difference between the proportion of the individuals within a given population with a mental disorder (i.e.,

total prevalent cases) and the proportion of these individuals who received treatment for that disorder (i.e. treated prevalent cases). More recent data on the prevalence and treatment rates of depressive disorders can now be used to derive treatment gaps estimates that consider: (1) the increased availability of data for some regions; (2) recent health reforms that may have occurred in some countries that influence treatment rates; and (3) changes in the use of some interventions that could affect treatment rates. In this study, we update the knowledge base on the gaps in treatment coverage for MDD. We conducted a systematic literature review to identify the existing literature on the global treatment rates of MDD. Potential sources of heterogeneity were analytically explored and accounted for to generate predicted treatment rates. These were combined with population-representative prevalence estimates derived by GBD 2019 to estimate treatment gaps for MDD.

3.2 Methods:

Case Definitions: This study focused on MDD as defined by the Diagnostic and Statistical Manual of Mental Disorders (DSM) or the equivalent International Classification of Diseases (ICD) diagnostic criteria [14, 15] According to the DSM, MDD is an episodic disorder characterized by at least one major depressive episode (MDE). A diagnosis of an MDE is characterized by meeting five out of nine criteria including the core symptoms of depressed mood and/or loss of interest causing clinically significant impairment in the main areas of functioning. To meet the threshold for a diagnosis of MDD, depressed mood must be experienced mostly all day and every day for a minimum period of two weeks.

Treatment Rates: Treatment rates were defined as the proportion of cases of MDD that received treatment for the disorder. Types of treatment were classified into categories used in previous studies by Thornicroft et al. (2017) and Wang et al. (2007) [11, 16]. These classifications were used to ensure consistent cross-national comparisons of the multiple sectors from which people may receive treatment. The treatment type categories used for this analysis are listed in Table 1 below:

Data sources and search strategy: The estimation of treatment gaps required data on prevalence and treatment rates for depressive disorders. Prevalence data came from work undertaken as part of the GBD 2019 study (see appendix for further details). A systematic literature review of the treatment rates data was undertaken as part of the present study.

Disorder specific treatment rates: A new systematic review was conducted to capture information on treatment rates, using methods that would ensure estimates were comparable with the GBD 2019 literature review and analysis of prevalence data. Searches were performed in two online scholarly databases Embase and PubMed from 2000 – mid-2020 including keywords such as “depres*” OR “dysthymia” AND “service OR care” AND “utilization” (See Appendix for full search strings). An additional search of all data sources used to estimate the prevalence of MDD in the GBD was conducted to ensure all relevant data sources were screened (see more details on GBD data below). Both reviews adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [17]. Search strings and the PRISMA checklist are provided in the Appendix. For each of the two literature searches, potential data sources were assessed for inclusion through a title, abstract and full text search, respectively. MM, AJF, JL, and KJ (see acknowledgement) conducted both systematic reviews. This systematic review is registered with PROSPERO (ID: CRD42020212552)

The primary metric of interest for the review of treatment rates was the proportion of individuals from general population surveys, meeting criteria for MDD that received treatment for their disorder. Studies were included if treatment rates pertaining to the last 12 months or less were reported directly or if sufficient information was available to calculate this along with 95% uncertainty intervals (UIs). Studies were included if they made use of population-based surveys that were representative of communities, countries, or regions under study. Studies were included if they used DSM 3, 4 or 5 or ICD 9 or ICD-10 criteria to diagnose cases of MDD. Only data collected between 2000 and 2020 were included. Earlier samples were excluded given that changes in the available treatment and service systems between countries have likely evolved over time. Therefore, treatment rates from years earlier than 2000 may not be representative of the current state of treatment coverage and service quality [18] . To maximize data availability, studies that reported on the current prevalence of 'depressive disorders' (comprising both MDD and dysthymia) or 'mood disorders' were included and reported on separately (see appendix). Studies were excluded if they exclusively reported on non-representative samples (example: inpatient samples, incarcerated samples, populations without fixed residences, populations that are racial or ethnic minorities in the study location) or used symptom scales to assess for the presence of depressive symptoms not meeting diagnostic thresholds within the DSM and ICD. In addition to our primary literature search, a grey literature search was also conducted as part of the review of GBD data to identify datasets from the World Health Surveys (WHS). We made an exception to the symptom scale exclusion criteria for data from this survey, which captured data on both depressive symptoms and service use across 70 countries in order to maximize data inclusion and geographic representativeness. To improve comparability of WHS treatment rates with other treatment rates in the dataset, WHS estimates were adjusted upwards towards the level they would have been had they been based on a formal diagnosis of MDD (See appendix for further details).

Data Collection and Processing: Data from included studies were extracted using a Microsoft Excel template that ensured that minimum amount of information was extracted from each study. Study characteristics that were extracted included location, study setting, methodological design, urbanicity (mixed/rural/urban), years, and MDD assessment tools used. Sample characteristics that were extracted included age, sex, response rate, treatment type, and sample size (See Table 2 for a full list of study characteristics extracted). Studies were classified by income level according to the World Bank Country and Lending Groups and GBD super regions [1, 19]. If studies reported multiple treatment rates (e.g., stratified by age or sex), the estimate for each was extracted. Similarly, if studies reported multiple treatment rates by severity of MDD, these were extracted and analyzed separately. Estimates for treatment rates were stratified into categories of treatment types described above. Treatment rate estimates were Freeman-Tukey double arcsine transformed to ensure all treatment rates and their uncertainty intervals remained between 0% and 100% [20, 21]. Pooled estimates were then back-transformed into natural number space and reported in the results section (see next section for more details).

Statistical Analysis:

Our primary regression analysis was restricted to data on MDD. We modeled MDD treatment rates as a function of covariates listed in Table 2 above using a Meta Regression: Bayesian, Regularized Trimmed (MR-BRT) framework to estimate pooled treatment rates adjusted for parameters of interest [22]. Parameters that suggested statistically significant differences in treatment rates were retained for meta-analyses reported in the results below. We used fixed effects for selected covariates and random effects for studies chosen a priori to account for between-study variation. Due to considerable heterogeneity and sparsity of data by select covariates, we chose to analyze data on disorder severity, MAT definitions, dysthymia, and mood disorders, and treatment rates by year as part of our supplementary analyses (see

appendix). Our final model included treatment type, income group, age, sex, type of assessment tool, and response rate as significant predictors of treatment rates. The MR-BRT model was used to obtain predicted treatment rates adjusted for these covariates (Tables 5 and 6).

Predicted treatment rates for MDD by age, sex, and super-region were estimated from the model and used to calculate treatment gaps as 1 – Treatment Rate. We estimated uncertainty for our analyses at the 1000 draws level. This involved taking 1000 samples of the predicted treatment rate estimates to generate 1000 samples of the treatment gap distribution. Estimates of projected treatment gaps were computed using the mean estimate across 1000 draws and the 95% uncertainty intervals (UIs) are determined on the basis of the 25^{th} and 975^{th} percentile values across a total of 1000 draws. Projected treatment gaps by age, sex, and GBD super-region for health service use and mental health service use are reported in the results section below as our dataset had the most data for these select treatment types.

GBD prevalence data were multiplied by treatment rates to calculate the number of treated and corresponding untreated cases of MDD. Uncertainty was estimated at the 1000 draws level described above. The generated 95% uncertainty reflected the main sources of sampling uncertainty from both the prevalence and treatment rates.

3.3 Results:

Study characteristics

We identified 325 data points from 146 studies reporting on treatment rates for MDD from 83 countries. The literature search and data sources are summarized in the appendix (Appendix sections 1-5). Table 3 summarizes the number of available datapoints for each of the six treatment types by income group. Figure 1. shows the global availability of relevant studies on treatment rates for MDD.

Regression Results for MDD data:

The coefficients from our regression model are presented below (Table 4). Our final model included treatment type, World Bank income group, age, sex, type of assessment tool, and response rate as significant predictors of treatment rates. Treatment rates for any service use, any health service use, and general health service use were significantly higher than those for mental health service use (reference category). Treatment rates for MAT and non-health service use were significantly lower than those for mental health service use. Age was associated with an increase in MDD treatment rates. Treatment rates were higher for females compared to males. Response rate was associated with an increase in treatment rates. The use of non-diagnostic assessment tools was associated with a decrease in treatment rates. Treatment rates from upper-middle income locations were significantly lower than those from high-income locations (reference category). Treatment rates from low and lower-middle income locations were significantly lower than those from high and upper-middle income locations.

Predicted treatment rates by service type and income group and adjusted for covariates are reported in Table 5. Treatment rates were the highest for any service use, followed by any health service use, and other treatment types. The lowest treatment rates were observed for MAT and non-health service use. Treatment rates across all service types were the highest for high-income locations compared to upper-middle and low and lower-middle income locations. For high-income locations, treatment rates ranged from 59% [34%, 82%] for any service use to 16% [2%, 38%] for any non-health service use. For upper-middle income locations, treatment rates ranged from 39% [16%, 65%] for any service use to 4% [<0.1%, 22%] for non-health

service use. For low and lower-middle income countries, treatment rates ranged from 31% [9%, 57%] for any service use to 1% [<0.1%, 15%] for any non-health service use.

We also modeled treatment rates as a function of the covariates described above along with GBD super-region (instead of income group). The regression coefficients from this model are presented in the appendix. Table 6 shows overall predicted treatment rates by service type and GBD super-region. Figures 2 and 3 show projected treatment gaps disaggregated by age, sex, and GBD super-region for health service use and mental health service use.

3.4 Discussion:

Our systematic review identified 325 datapoints from 83 countries on the treatment rates of MDD. From this dataset we characterized the patterns of service use for MDD across 83 countries. Treatment rates were modeled as a function of service type, location, age, sex, diagnostic tool, and response rate. Treatment rates for any service use, any health service use, and any general health service use were significantly higher than mental health service use. Although mental health service use is traditionally considered the most ideal for MDD, the higher treatment rates of broader categories of any, health, or general health service use indicates the importance of these types of services used to treat MDD given the lack of specialized mental health services in many countries. These findings are also largely consistent with WHO recommendations on treatment of mental disorders within general or primary health care settings for all countries [7].

Age was associated with an increase in treatment rates. This is consistent with earlier findings that older age is typically associated with greater use of treatment services. Treatment rates were higher in females than males. This is also consistent with trends found in other studies that females are perhaps better at detecting and seeking out care for emotional problems than males [23]. Response rate and type of diagnostic tool were important to incorporate in our main model as methodological features that may impact population-based analyses. Higher response rates indicate the sample is likely to be more representative of the broader population and the results therefore more generalizable to the whole population compared to lower response rates which are more susceptible to bias. High response rates being associated with greater treatment coverage may indicate that large-scale sample recruitment efforts may succeed in including those who generally have poorer health statuses with a higher need and use of services, and who might be non-responders for studies with low response rates that fail to adequately capture a more representative sample from the target population [24, 25]. We adjusted for type of diagnostic tool used to account for any difference between studies that use CIDI or other equivalent diagnostic instruments, and our WHS estimates which were adjusted prior to analyses. Treatment rate estimates from studies using CIDI or other diagnostic MDD assessment tools were found to be higher compared to WHS estimates that were obtained using symptom scales. This is consistent with our sensitivity analyses where treatment rates from studies using diagnostic assessment tools were found to be higher than those obtained from studies using symptom scales as was the case with WHS estimates. WHS treatment rates estimates obtained using symptom scales likely include samples with subthreshold MDD as well which may reflect lower treatment rates.

LMICs had significantly lower treatment rates compared to high-income locations. With up to 75% of individuals with depressive disorders residing in low- and middle-income countries, this indicates that a substantial proportion of people with depression globally do not access any health related services for depression care. Our findings echo the importance of existing calls

for the prioritization of mental health in national health agendas. Many countries featured in our review still lack the mental health policy, legislation, or resources to guide their mental health programs and services [7, 26]. Our findings also showed that even in high income countries where treatment rates are comparatively higher, the majority of individuals receiving care for depressive disorders, failed to receive a level of care that is consistent with practice guideline recommendations. Only a small minority of individuals with depressive disorders accessed treatment in the specialized mental health care system or received minimally adequate treatment.

In this paper, we analyzed rates of minimally adequate treatment by categorizing study-reported definitions of MAT as either "stringent", i.e., requiring some combination of at least 8 visits to a mental health professional and at least 30 days of prescribed antidepressant use or "non-stringent", i.e., having lower threshold for mental health visits and medication use, and deemed additional types of service use as adequate treatment. Treatment rates for stringent definitions of MAT were lower than those for non-stringent definitions of MAT. This indicates that the high thresholds of treatment requirements considered minimally adequate may not be feasible to attain for all those who need treatment [6]. Moreover, stringent definitions of MAT were most commonly found in studies originating from high-income countries (e.g., Canada, Finland, USA, and Spain). Many countries may be unable to deliver stringent MAT especially if mental health treatment is provided within primary or general healthcare settings by trained healthcare providers instead of specialized mental health professionals. In countries where attaining MAT may pose a challenge, alternative interventions (e.g., The Friendship Bench program in Zimbabwe) that leverage community and non-specialized human resources and ensure appropriate levels of treatment intensity have shown promise [27–30].

The findings presented here and elsewhere indicate that access to care for MDD needs to be improved. However, the specifics of which components of care increase 'access' still need to be systematically explored. Dedicated mental health services, institutes, and hospital units along with community-based care exist in many HICs. However, mental healthcare institutions in LMICs are likely under-resourced and overburdened with treatment for competing, acute health conditions. Therefore, examining access alone is not enough. The proportion of people who receive sufficient care once they enter treatment is still difficult to estimate from the current literature. MAT is also difficult to quantify because not everyone who meets criteria for MDD will need or want care. Therefore, it is important to consider not simply the presence of services, but what behavioral or environmental drivers such as awareness (or lack thereof) of effectiveness of interventions, income or food insecurity, and other more immediate stressors that may impact contact with and adherence to treatment. A recent paper using WMHS data estimated a 90% gap in effective treatment with lack of utilization and inadequate quality or adherence being critical bottlenecks [13]. While this provides an important decomposition of elements of treatment coverage and quality, it is also important to consider variation in real-world treatment settings and variation in MAT thresholds that impact treatment rates as shown in this analysis. Treatment gaps for MDD also need to consider gaps in psychosocial and physical health care [31]. Psychosocial interventions have been shown to be highly effective in symptom reduction and physical health care is important to include considering the high and often untreated physical comorbidity and premature mortality that accompanies MDD and other mental disorders [32].

Efforts to close the depression treatment gap would also need significant boosts in funding allocations Global health financing has historically been prioritized for malaria, HIV/AIDs, and tuberculosis – which are some of the leading causes of disability and mortality in many LMICs. However, financing for mental health is still far from adequate. In 2019, development assistance

for health (DAH) for non-communicable diseases (which includes mental disorders) for SDG 3 targets was $0.7 billion for 135 low and middle income countries – which is less than 2% of the total estimated DAH in 2019 of $40.6 billion [33]. Therefore, it is important to align funding priorities with epidemiological shifts in LMICs that are likely to be accompanied by an increase in non-communicable disease burden including mental disorders. A global return on investment analysis by Chisholm and colleagues showed that scaling up effective treatment for depression and anxiety disorders leads to 43 million extra years of healthy life and a net present economic value of $310 billion between 2016 to 2030 [3]. In addition to differences in treatment gaps between countries, it is also important to consider subnational variation in the magnitude of these gaps. For instance, it is plausible that urban areas may have higher treatment rates (although our analysis did not find any significant differences by urbanicity possibly due to lack of sufficient data informing urban vs rural estimates). Unobserved factors such as regional differences in health system performance, personal and public healthcare spending, regional socioeconomic inequality, history of regional instability, and ongoing epidemiological shifts in disease burden may well play a role in within-country treatment gaps for MDD. While this paper did not explicitly analyze subnational estimates, it is important to recognize that there may be regional differences in MDD treatment. Subsequently within-country funding priorities should take these differences into account. Future work in analyzing subnational variation in treatment rates may also contribute to our understanding of how MDD burden at the subnational level can be influenced by access to care.

The Comprehensive Mental Health Action Plan 2013–2020 adopted by the World Health Assembly was recently revised and extended through 2030 to include an updated set of indicators. Of particular significance is a newly added indicator to quantify the proportion of people with depression who are using services during the past 12 months [7, 8, 26]. The presence of an indicator to track treatment use among those with depression specifically may serve as an important impetus for regular data monitoring and tracking for treatment coverage. In the United Nations Sustainable Development Goals, mental health was for the first time explicitly recognized within the concept of Universal Health Coverage [34]. It is clear that providing effective services for people with depression, integrated into general health services, care for HIV or maternal and child health, is a vital element of basic healthcare provisions [35, 36]. As we now have evidence for effective and feasible interventions suitable for low-, middle- and high-income countries, we call upon national and international organizations to make firm and time-bound commitments to make adequate resources available for scaling up the provision of mental health services so that ‘no one is left behind’. This is particularly pertinent during the current COVID-19 pandemic, which has been accompanied by a substantial increase in the prevalence of depression and a simultaneous decrease in access to services in many countries [38].

The analyses conducted were limited by lack of high quality data on service use. Most studies originated from high-income countries largely located in North America and Western Europe. However, LMICs (mostly in Sub-Saharan Africa and South Asia) that comprise approximately 70% of the world’s population and 80.9% of MDD non-fatal disability globally, accounted for only 22% of studies on MDD in this dataset. Despite the significant disparities in available data, the available evidence indicates that the treatment gap for MDD is consistently wide across most locations. There was considerable variation in treatment rates across countries, suggesting that resources available for MDD not only continue to be scarce but unequally distributed across the globe, and far from commensurate to the prevalence of depressive disorders [37].

Some additional limitations are important to note. First, only 22% of the studies originated from low or lower-middle income countries with very few datapoints by treatment type. Therefore, we

were unable to generate estimates by type of treatment beyond the location groupings used. This is also reflected in the large uncertainty bounds accompanying estimates for low and lower-middle income countries and therefore these should be interpreted with caution. Second, given the nature of our systematic review, we had to rely on definitions for service use set by each individual data source. Definitions for what comprise minimally adequate treatment in particular varied widely by data source and highlighted the lack of consistency in the literature (and even within the World Mental Health Survey collaboration) in how this concept should be defined. Thornicroft and collaborators restricted their definition of MAT to those "receiving either pharmacotherapy (at least 1 month of a medication, and 4 visits to any type of medical doctor) or psychotherapy (at least 8 visits with any professional including religious or spiritual advisor, social worker or counsellor)". However, the extent to which these thresholds of care should be considered a practice guideline recommendation is unclear as depression exists on a severity continuum with more intensive treatment needed for depression of higher severity. Third, there were insufficient data by location and year to appropriately analyze changes in MDD treatment rates over time for all locations (see appendix). Nonetheless, this study provides the most comprehensive review of MDD treatment rates to date. We improve upon earlier work by applying updated modeling methods that better capture heterogeneity in the data and account for bias that may be contributed by study-level characteristics. Findings from this study may contribute to future work in modeling potentially avoidable burden of MDD in varying scenarios of treatment coverage.

3.5 Conclusion:

Treatment coverage for MDD continues to be low globally. . It is important to prioritize better and frequent collection of treatment coverage data over time and the availability of appropriate care for MDD and examine barriers and facilitators of treatment.

Acknowledgements:

Janni Leung and Kara Jaeschke contributed to the original data search and extraction. Dan Chisholm contributed to discussions on the implications of the findings.

3.6 Significance of Outputs from Chapter 3:

In Chapter 3, we report findings from an updated systematic review on treatment rates using data from 2000 through 2019 and 83 countries. We include carefully selected studies that used diagnostic criteria to assess MDD and included samples that were representative of their given locations. We consider and analytically quantify sample and methodological covariates that may contribute to variation in reported treatment rates and model these in MR-BRT to produce predicted treatment rates by location and six treatment types. Although there were insufficient data to allow for a more detailed exploration of minimally adequate treatment (MAT), we briefly explore variation by stringent and non-stringent care thresholds of what is considered adequate care which provide some indication of feasibility of desired MAT particularly in low and middle-income countries. Overall, this work provides the most comprehensive evidence on treatment coverage estimates that highlight the need to prioritize MDD treatment in many parts of the world.

In the previous chapter, we mentioned that outputs on treatment efficacy along with outputs on treatment coverage from the current chapter may find useful applications in re-estimating the severity distribution for MDD within the GBD study. We also discussed the limitations of current methods that may be addressed using more informative data on treatment efficacy and treatment coverage provided in Chapters 2 and 3. We now offer a broad, conceptual framework of how treatment efficacy and treatment coverage may be used to re-estimate the severity distribution of MDD. Current GBD methods map Short Form-12 scores to disability weights and derive severity distributions. This process is documented in more detail in Burstein et al. and Salomon et al [39, 40]. However, most data sources are from high-income settings and are not fully representative of severity in low and middle-income settings. Therefore, more informative pooled treatment efficacy estimates informed by a comprehensive set of 607 studies on interventions for MDD may be used to adjust the severity distribution such that they better reflect the impact of treatment efficacy. These pooled treatment efficacy estimates may be further enhanced by being adjusted for treatment coverage (inputs from comprehensive systematic review in Chapter 3) that will incorporate information on access to treatment. Therefore, pooled treatment efficacy estimates adjusted for treatment coverage may be used to adjust the severity distribution such that they better reflect the impact of both treatment efficacy and treatment coverage. Extensions of this work may include grading this new distribution of severity by health access quality index (HAQI) by location and year. Additional extensions may include quantifying the proportion of depression burden that has (i) been averted by treatment, (ii) remains untreated, and importantly – (iii) is potentially avoidable under assumptions of routine and best healthcare functioning scenarios.

The paragraph above offers some potential insights into the relevance of work in Chapters 2 and 3 to re-estimating the severity distribution for MDD. Beyond the GBD framework, these findings may serve as useful contributions to the broader fields of mental health epidemiology by empirically addressing methodological improvements to comparative epidemiological assessments, and the evidence on treatment efficacy and coverage for MDD. These domains are closely connected and complementary in their efforts to improve global mental health estimates. Chapter 4 offers a summary of key findings that result from this work and their relevance to future areas of research in these areas.

3.7 References:

[1] Vos T, Lim SS, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet* 2020; 396: 1204–1222.

[2] Vigo D, Thornicroft G, Atun R. Estimating the true global burden of mental illness. *The Lancet Psychiatry* 2016; 3: 171–178.

[3] Chisholm D, Sweeny K, Sheehan P, et al. Scaling-up treatment of depression and anxiety: a global return on investment analysis. *The Lancet Psychiatry* 2016; 3: 415–424.

[4] Moitra M, Santomauro D, Degenhardt L, et al. Estimating the risk of suicide associated with mental disorders: A systematic review and meta-regression analysis. *Journal of psychiatric research*.

[5] Ferrari AJ, Norman RE, Freedman G, et al. The burden attributable to mental and substance use disorders as risk factors for suicide: findings from the Global Burden of Disease Study 2010. *PloS one* 2014; 9: e91936.

[6] Patel V, Chisholm D, Parikh R, et al. Addressing the burden of mental, neurological, and substance use disorders: key messages from Disease Control Priorities. *The Lancet* 2016; 387: 1672–1685.

[7] Saxena S, Funk M, Chisholm D. World health assembly adopts comprehensive mental health action plan 2013–2020. *The Lancet* 2013; 381: 1970–1971.

[8] World Health Organization. *Extracts from document EB144/2019/REC/1 for consideration by the Seventy-second World Health Assembly*. 2019.

[9] Izutsu T, Tsutsumi A, Minas H, et al. Mental health and wellbeing in the Sustainable Development Goals. *The Lancet Psychiatry* 2015; 2: 1052–1054.

[10] Kohn R, Saxena S, Levav I, et al. The treatment gap in mental health care. *Bulletin of the World health Organization* 2004; 82: 858–866.

[11] Thornicroft G, Chatterji S, Evans-Lacko S, et al. Undertreatment of people with major depressive disorder in 21 countries. *The British Journal of Psychiatry* 2017; 210: 119–124.

[12] World Health Organization. Health in 2015: from MDGs, millennium development goals to SDGs, sustainable development goals.

[13] Vigo D, Haro JM, Hwang I, et al. Toward measuring effective treatment coverage: critical bottlenecks in quality-and user-adjusted coverage for major depressive disorder. *Psychological medicine* 2020; 1–11.

[14] American Psychiatric Association AP. Diagnostic and statistical manual of mental disorders (DSM-IV).

[15] WHO. ICD-10 Version:2019. *World Health Organization*.

[16] Wang PS, Angermeyer M, Borges G, et al. Delay and failure in treatment seeking after first onset of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World psychiatry* 2007; 6: 177.

[17] Moher D, Liberati A, Tetzlaff J, et al. Academia and clinic annals of internal medicine preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annu Intern Med* 2009; 151: 264–269.

[18] Marcus SC, Olfson M. National trends in the treatment for depression from 1998 to 2007. *Archives of general psychiatry* 2010; 67: 1265–1273.

[19] Fantom N, Serajuddin U. *The World Bank's classification of countries by income*. The World Bank, 2016.

[20] Freeman MF, Tukey JW. Transforma-tions Related to the Angular and Square Root. Memo-randum Report 24. *Statistical Research Group, Princeton University*.

[21] Miller JJ. The inverse of the Freeman–Tukey double arcsine transformation. *The American Statistician* 1978; 32: 138.

[22] Zheng P, Barber R, Sorensen RJD, et al. Trimmed constrained mixed effects models: formulations and algorithms. *Journal of Computational and Graphical Statistics* 2021; 1–13.

[23] Wang PS, Angermeyer M, Borges G, et al. Delay and failure in treatment seeking after first onset of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World psychiatry* 2007; 6: 177.

[24] Knudsen AK, Hotopf M, Skogen JC, et al. The health status of nonparticipants in a population-based health study: the Hordaland Health Study. *American journal of epidemiology* 2010; 172: 1306–1314.

[25] Galea S, Tracy M. Participation rates in epidemiologic studies. *Annals of epidemiology* 2007; 17: 643–653.

[26] World Health Organization. Mental health atlas 2017. Geneva: World Health Organization; 2018. *World Health Organization WHO MiNDbank http://www who int/mental_health/mindbank/en*.

[27] Chibanda D, Bowers T, Verhey R, et al. The Friendship Bench programme: a cluster randomised controlled trial of a brief psychological intervention for common mental disorders delivered by lay health workers in Zimbabwe. *International journal of mental health systems* 2015; 9: 1–7.

[28] Mendenhall E, de Silva MJ, Hanlon C, et al. Acceptability and feasibility of using non-specialist health workers to deliver mental health care: stakeholder perceptions from the PRIME district sites in Ethiopia, India, Nepal, South Africa, and Uganda. *Social science & medicine* 2014; 118: 33–42.

[29] Thornicroft G, Deb T, Henderson C. Community mental health care worldwide: current status and further developments. *World Psychiatry* 2016; 15: 276–286.

[30] Hoeft TJ, Fortney JC, Patel V, et al. Task-sharing approaches to improve mental health care in rural and other low-resource settings: a systematic review. *The Journal of rural health* 2018; 34: 48–62.

[31] Pathare S, Brazinova A, Levav I. Care gap: a comprehensive measure to quantify unmet needs in mental health. *Epidemiology and psychiatric sciences* 2018; 27: 463–467.

[32] Liu NH, Daumit GL, Dua T, et al. Excess mortality in persons with severe mental disorders: a multilevel intervention framework and priorities for clinical practice, policy and research agendas. *World psychiatry* 2017; 16: 30–40.

[33] Micah AE, Su Y, Bachmeier SD, et al. Health sector spending and spending on HIV/AIDS, tuberculosis, and malaria, and development assistance for health: progress towards Sustainable Development Goal 3. *The Lancet* 2020; 396: 693–724.

[34] Votruba N, Thornicroft G, Group FS. Sustainable development goals and mental health: learnings from the contribution of the FundaMentalSDG global initiative. *Global Mental Health*; 3.

[35] Collins PY, Insel TR, Chockalingam A, et al. Grand challenges in global mental health: integration in research, policy, and practice. *PLoS Med* 2013; 10: e1001434.

[36] Shidhaye R, Lund C, Chisholm D. Closing the treatment gap for mental, neurological and substance use disorders by strengthening existing health care platforms: strategies for delivery and integration of evidence-based interventions. *International journal of mental health systems* 2015; 9: 1–11.

[37] Saxena S, Thornicroft G, Knapp M, et al. Resources for mental health: scarcity, inequity, and inefficiency. *The lancet* 2007; 370: 878–889.

[38] World Health Organization. The impact of COVID-19 on mental, neurological and substance use services: results of a rapid assessment.

[39] Burstein R, Fleming T, Haagsma J, Salomon JA, Vos T, Murray CJ. Estimating distributions of health state severity for the global burden of disease study. Population health metrics. 2015 Dec 1;13(1):31.

[40] Salomon JA, Vos T, Hogan DR, Gagnon M, Naghavi M, Mokdad A, Begum N, Shah R, Karyana M, Kosen S, Farje MR. Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010. The Lancet. 2012 Dec 15;380(9859):2129-43

[41] Patten SB. Recall bias and major depression lifetime prevalence. Social Psychiatry and Psychiatric Epidemiology. 2003 Jun;38(6):290-6.

3.8 Tables and Figures:

Treatment Type	Definition
Any service use	Studies that reported treatment rates without differentiating between health and non-health sectors
Health service use	Services offered within the health sector
General health service use	Services provided by primary care doctors, other general medical doctors, nurses, or other health professionals not within the mental health sector
Mental health service use	Services provided by psychiatrists, psychologists, other mental health professionals in any setting, social workers or counsellors in a mental health specialty setting or use of a mental health hotline
Non-health service use	Services outside of the health sector. This includes service provided by spiritual or religious advisers, chiropractors, traditional

	healers, participation in internet support groups, and self-help groups
Minimally adequate treatment (MAT)	Treatment which was potentially minimally adequate according to evidence-based guidelines. Due to the variation in the definition of MAT between studies, we chose to group definitions as being "stringent" or "non-stringent". Stringent MAT was defined as receiving 8 or more visits to any service sector for psychotherapy or 4 or more visits to any service sector and at least 30 days of pharmacotherapy or its nearest equivalent. Non-stringent MAT was defined as requiring fewer visits and days of medication use than the stringent definition of MAT. These groups best reflected the variation observed in the reported definitions of MAT

Table 1. Treatment types and corresponding definitions.

Parameter	Definition
Disorder	As reported by the study: Major Depressive Disorder or Dysthymia or Depressive Disorders or Mood disorders. (The main analyses focus on MDD only)
Country	As reported by study
World Bank Income Group	High-income (ref) /Upper-middle/Lower-middle/Low-income
Year	Mid-point of duration between start and end years of study period
Age	Mid-point of age range of sample reported by study
Percent Female	Percentage of study sample that comprised female participants
Treatment Type	Any service/health service/general health service/mental health service/non-health service
MDD Assessment tool	CIDI/other diagnostic instruments (ref) or symptom-scale assessed WHS estimates
Recall Period of Treatment	12 months (ref) or less
Response Rate	Proportion of sample contacted that provided data for the study
Sample size	Total number of study participants
Urbanicity	Information on urban, rural, or mixed setting of study location
MAT	As defined by study; Categorized as lenient, moderate, or stringent definition
Disorder severity	As reported by study (mild, moderate, or severe)

Table 2. List of Parameters and Definitions

Treatment Type	Low and lower middle	Upper-middle	High Income Countries	Total
Any general health service	1	2	32	35
Any health service	27	14	45	86
Any mental health service	6	13	79	98
Any non-health service	1	3	22	26
Any service use	2	13	27	42
MAT	1	9	28	38

Table 3. Number of datapoints for MDD by Treatment Type and Income group

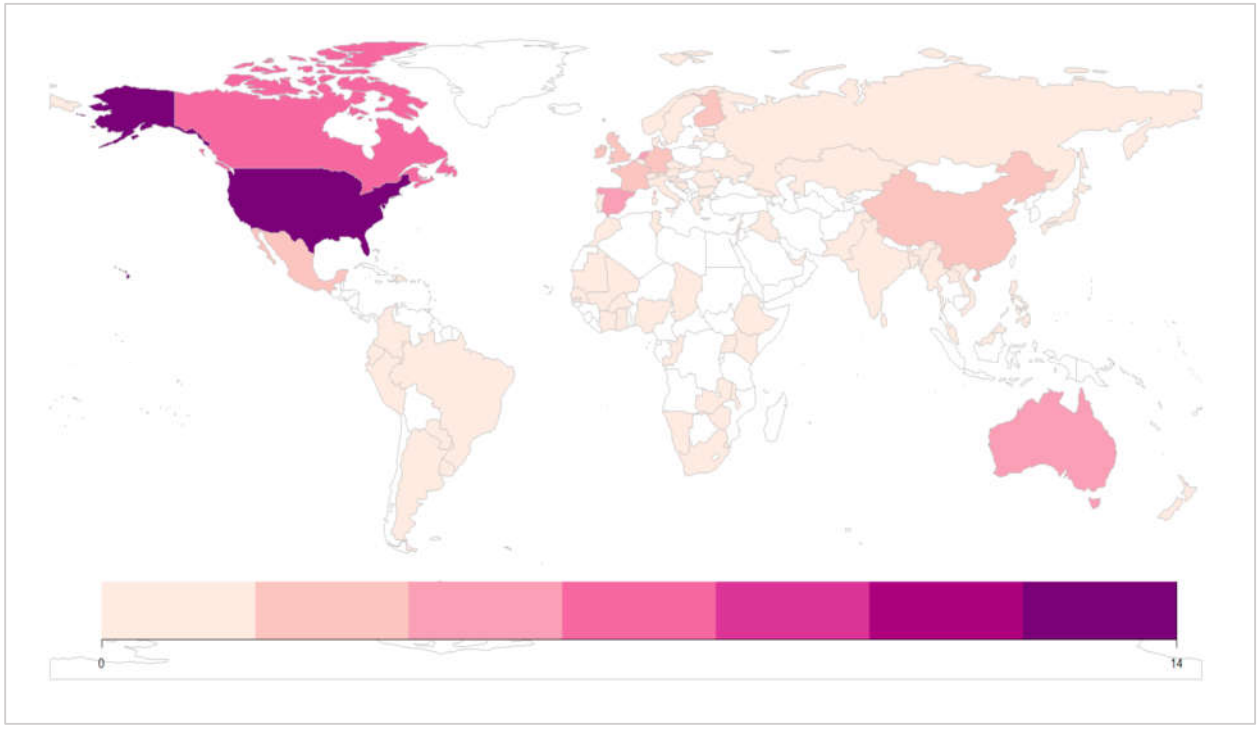


Figure 1. Number of studies on MDD treatment coverage by country

Covariate	Parameter Estimate [95% UI]
Intercept [a]	0.567 [0.543, 0.591] ***
Treatment Type	
Any service use	0.3 [0.286, 0.315] ***
General health service use	0.049 [0.036, 0.062] *
Minimally adequate treatment (MAT)	-0.112 [-0.126, -0.098] *
Health service use	0.196 [0.179, 0.213] ***
Non-health service use	-0.153 [-0.167, -0.139] ***
Sample characteristics	
Age	0.001 [0.0006, 0.002] *
Percent female	0.032 [0.016, 0.048] *

World Bank Income Group	
Upper-middle-income	-0.201 [-0.281, -0.121] ***
Low and lower-middle income	-0.287 [-0.379, -0.195] ***
Methodological Covariates	
Response rate	0.17 [0.028, 0.313] *
Type of assessment tool	-0.082 [-0.159, -0.005] *

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

[a] Intercept represents mental health service use.

Table 4. Regression Coefficients and 95% uncertainty intervals for MDD treatment rates modeled as a function of select covariates: treatment type (ref = mental health service use), income group (ref= high-income), age (ref = mid-point of sample ~ 50 years), sex (ref = both), response rate (ref = 100%), and type of assessment tool (ref = CIDI/equivalent diagnostic tool).

Treatment Type	High-Income	Upper-middle Income	Lower middle and Low Income
Any service use	0.59 [0.34, 0.82]	0.39 [0.16, 0.65]	0.31 [0.09, 0.57]
Health service use	0.49 [0.24, 0.73]	0.29 [0.09, 0.56]	0.21 [0.04, 0.46]
General health service use	0.35 [0.13, 0.59]	0.17 [0.02, 0.4]	0.11 [0.0002, 0.31]
Mental health service use	0.3 [0.09, 0.54]	0.13 [0.01, 0.36]	0.08 [<0.001, 0.28]
Non-health service use	0.16 [0.02, 0.38]	0.04 [<0.001, 0.22]	0.01 [<0.001, 0.15]
Minimally Adequate Treatment	0.2 [0.04, 0.42]	0.07 [<0.001, 0.25]	0.03 [<0.001, 0.18]

Table 5: Predicted treatment rates by Income Group

Treatment Type	Southeast Asia, East Asia & Oceania	Central Europe, Eastern Europe & Central Asia	High Income	Latin America & Caribbean	North Africa & Middle East	South Asia	Sub-Saharan Africa
Any service use	0.29 [0.05, 0.61]	0.54 [0.25, 0.82]	0.58 [0.29, 0.85]	0.45 [0.16, 0.74]	0.28 [0.06, 0.57]	0.28 [0.05, 0.6]	0.28 [0.06, 0.57]
Health service use	0.2 [0.01, 0.49]	0.44 [0.16, 0.74]	0.47 [0.2, 0.76]	0.34 [0.09, 0.64]	0.19 [0.02, 0.46]	0.2 [0.01, 0.48]	0.19 [0.02, 0.46]
General health service use	0.12 [<0.01, 0.38]	0.33 [0.09, 0.63]	0.36 [0.11, 0.66]	0.24 [0.03, 0.52]	0.12 [<0.01, 0.34]	0.12 [<0.01, 0.37]	0.12 [<0.01, 0.34]
Mental health service use	0.08 [<0.01, 0.31]	0.27 [0.05, 0.55]	0.3 [0.07, 0.59]	0.19 [0.01, 0.44]	0.08 [<0.01, 0.28]	0.08 [<0.01, 0.3]	0.08 [<0.01, 0.28]

Non-health service use	0.03 [<0.01, 0.17]	0.14 [<0.01, 0.38]	0.16 [0.01, 0.41]	0.08 [<0.01, 0.28]	0.02 [<0.01, 0.14]	0.03 [<0.01, 0.16]	0.02 [<0.01, 0.14]
Minimally Adequate Treatment	0.04 [<0.01, 0.2]	0.18 [0.01, 0.44]	0.2 [0.02, 0.47]	0.11 [<0.01, 0.33]	0.04 [<0.01, 0.18]	0.04 [<0.01, 0.2]	0.04 [<0.01, 0.18]

Table 6. Predicted treatment rates by GBD Super-region

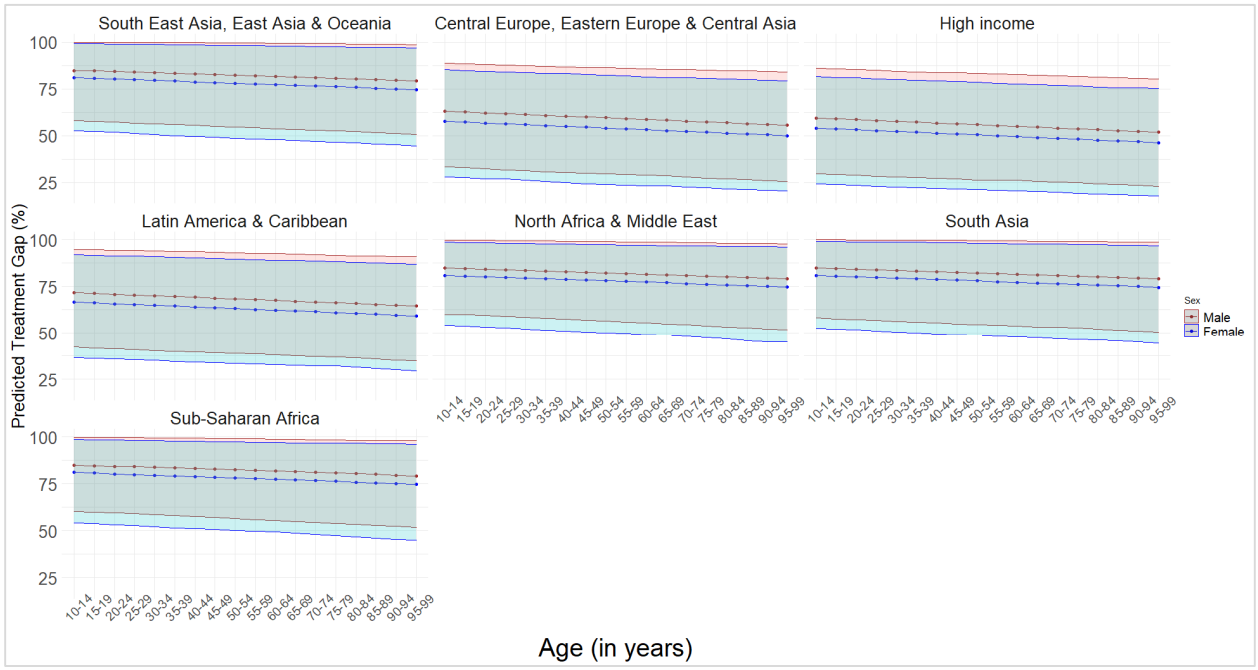


Figure 2. Predicted Treatment Gap for any health service use by age, sex, and GBD super-region

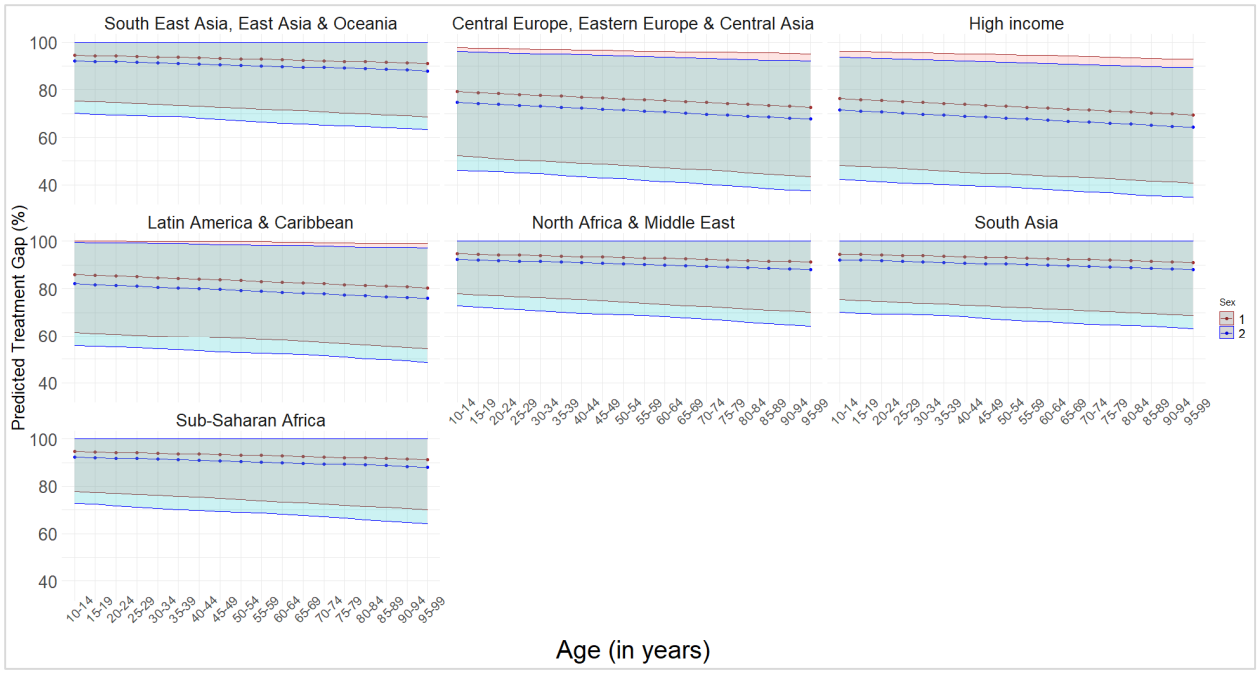


Figure 3. Predicted Treatment Gap for any mental health service use by age, sex, and GBD super-region.

3.9 Appendix:

Systematic Review Search Strings:

PubMed:

Treatment rates systematic review:

- ((((“Mental Health Services” AND “Mental Disorders” AND (Epidemiology OR “Health Care Surveys”))OR ((((((“Mental Health Services/ utilization" OR “unmet needs” OR Untreated OR (Treatment AND rate*) OR (Treatment AND gap)) AND (epidemiolog* OR population*)) AND (("Mental health" OR "Mental disorder*" OR " Mental illness*" OR "Psychiatric disorder*" OR "Psychiatric illness*") OR ("Depress*" OR "Dysthymi*" OR "Mood disorder*" OR "Affective disorder*") OR Bipolar OR Schizophrenia OR (("Substance" OR "Addictive" OR "Addiction" OR "Drug dependence" OR "Cannabi*" OR "Hallucinogen*" OR "Opioid*" OR "Heroin*" OR "Amphetamine*" OR "Cocaine*" OR "Marijuana*") AND ("Use" OR "Dependence" OR "disorder")) OR ("Anxiety" OR "Obsessive compulsive" OR "Agoraphobia" OR "Social phobia" OR "Phobia" OR "Post-traumatic stress" OR "Posttraumatic stress" OR "PTSD" OR "Panic disorder")))))) AND "humans"[Filter]) NOT "clinical trial"[Filter]) NOT ((cancer[sb] OR AIDS[sb])))

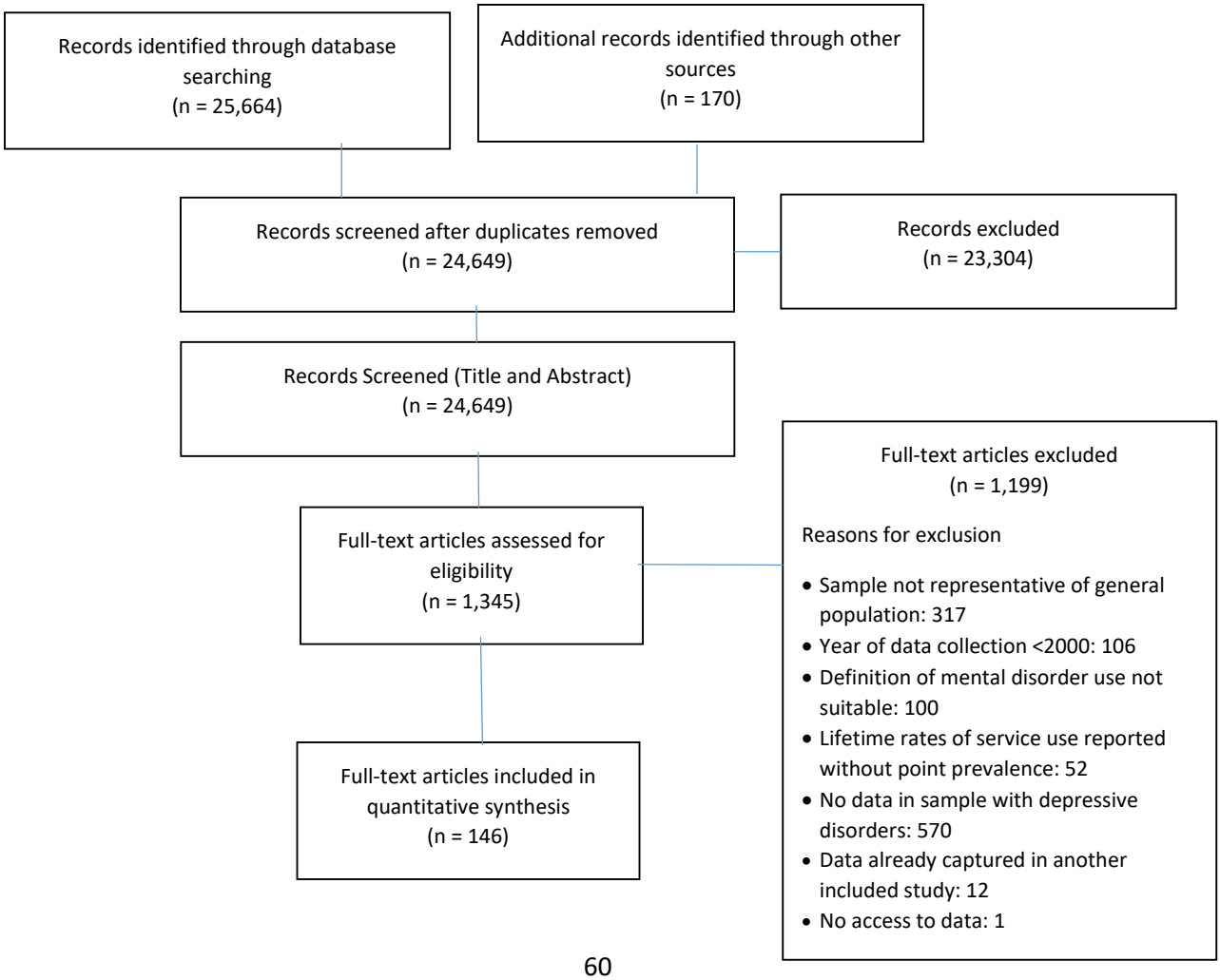
Embase:

- (major AND depression:ab,ti OR depression:ab,ti OR depres*:ab,ti OR dysthymia:ab,ti OR mental:ab,ti OR behavioral:ab,ti) AND ((mental AND health AND service:ab,ti OR health) AND service:ab,ti OR care:ab,ti OR service) AND utilization:ab,ti

Prevalence systematic review:

- ((((((((((((((((depress*[Title/Abstract]) OR dysthymi*[Title/Abstract]) OR bipolar[Title/Abstract]) OR manic[Title/Abstract]) OR mania[Title/Abstract]) OR "mood disorders"[Title/Abstract]) OR "mood disorder"[Title/Abstract]) OR mood disorders[MeSH Terms]) OR depressive disorders[MeSH Terms]) OR depressive disorders, major[MeSH Terms]) OR bipolar disorders[MeSH Terms]) OR dysthymic disorders[MeSH Terms] AND ((((((((((((((prevalen*[Title/Abstract]) OR mortality[Title/Abstract]) OR death*[Title/Abstract]) OR inciden*[Title/Abstract]) OR recurren*[Title/Abstract]) OR remission[Title/Abstract]) OR duration[Title/Abstract]) OR remit*[Title/Abstract]) OR epidemiolog*[Title/Abstract]) OR prevalence[MeSH Terms]) OR mortality[MeSH Terms]) OR incidence[MeSH Terms]) OR recurrence[MeSH Terms])

PRISMA flowchart:



Data sources on MDD prevalence from the Global Burden of Disease Study:

Disorder-specific prevalence rates. GBD 2019 estimated the prevalence of MDD and dysthymia by age, sex, year, and location as part of their analysis of non-fatal burden. The GBD search strategy has been reported elsewhere and is summarized here [1]. Briefly, prevalence estimates were derived from an analysis of epidemiological population survey data obtained from comprehensive systematic reviews of the literature reporting on the prevalence, incidence, remission, duration, and excess mortality associated with MDD and dysthymia. The literature search involved examining the peer-reviewed literature (via PubMed, PsychInfo, and Embase) between 1980 and 2019, and obtaining other relevant data sources from the grey literature or through expert consultation up to 2019. To meet criteria for inclusion, studies reporting prevalence must have: defined a case of MDD or dysthymia using diagnostic classifications proposed in the DSM or ICD; involved/recruited a sample representative of the community, region, or country under study (i.e., samples of minority groups, or those derived from hospital records were not accepted); and reported prevalence within the past year or less. Lifetime prevalence estimates were not accepted as they are more prone to recall bias [41].

WHS Data sources:

WHS surveys were implemented by WHO to monitor adult health outcomes and health systems between 2002 and 2004 in countries chosen to represent all world regions. Samples were representative of the general population and probabilistically selected. For each country, unit record data were aggregated to estimate the proportion of cases likely meeting diagnostic criteria for MDD and the proportion of these probable MDD cases accessing any health service in the last two weeks. The survey item relating to service use identified individuals taking medication or receiving other professional treatment for their symptoms of depression in the last two weeks. Treatment was defined as attending psychological therapy or counselling session with a trained health professional during the last two weeks. The survey items relating to depression in the WHS captured the majority but not all symptoms required for a full-diagnosis of MDD according to the DSM or ICD. As such, the estimated treatment rates likely pertained to a combination of individuals with MDD and sub-threshold MDD.

We conducted a sensitivity analysis to examine the difference between other treatment rates available in our dataset where MDD was defined according to full DSM/ICD criteria and those derived from the WHS. There were three countries (Australia, Spain, France) for which we had comparable WHS and diagnostic based treatment rates for ‘any health service’. The pooled ratio of other diagnostic-based treatment rates: WHS treatment rates was 1.4 (95% confidence interval: 1.3-1.6). This suggested that surveys capturing a full diagnosis of MDD produced treatment rates that were on average 1.4 times higher than WHS surveys. We maximized data inclusion by including the WHS treatment rates in our estimation of treatment gaps. To improve comparability of WHS treatment rates with other treatment rates in the dataset, WHS estimates were adjusted upwards towards the level they would have been had they been based on a full-diagnosis of MDD using our estimated ratio. The diagnostic algorithm used to identify probable cases of MDD within the WHS, the analysis to estimate the pooled ratio of other diagnostic-based treatment rates: WHS

treatment rates, and a comparison of the adjusted and unadjusted WHS estimates are shown below.

The DSM diagnostic criteria for MDD stipulates that five (or more) of the following symptoms need to be present during the same 2-week period and represent a change from previous functioning: at least one of the symptoms is either (1) depressed mood or (2) loss of interest or pleasure.

- Depressed mood most of the day, nearly every day, as indicated by either subjective report (e.g., feels sad, empty, hopeless) or observation made by others (e.g., appears tearful).
- Markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day (as indicated by either subjective account or observation).3. Significant weight loss when not dieting or weight gain (e.g., a change of more than 5% of body weight in a month), or decrease or increase in appetite nearly every day.
- Insomnia or hypersomnia nearly every day.
- Psychomotor agitation or retardation nearly every day (observable by others, not merely subjective feelings of restlessness or being slowed down).
- Fatigue or loss of energy nearly every day.
- Feelings of worthlessness or excessive or inappropriate guilt (which may be delusional) nearly every day (not merely self-reproach or guilt about being sick).
- Diminished ability to think or concentrate, or indecisiveness, nearly every day (either by subjective account or as observed by others).
- Recurrent thoughts of death (not just fear of dying), recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide.

The symptoms above need to have caused clinically significant distress or impairment in social, occupational, or other important areas of functioning. Additionally, the episode is not attributable to the physiological effects of a substance or to another medical condition. Table 1 summarizes the diagnostic algorithm used to identify probable cases of major depressive disorder within each World Health Survey (WHS).

Table 1: DSM-IV-TR criteria for major depressive disorder with corresponding World Health Survey item

DSM-IV-TR criteria for major depressive disorder (single episode)	WHS Item	WHS item description
First determine if all of the following apply;		
Not a mixed episode (e.g. bipolar disorder)	N/A	N/A
Symptoms cause clinically significant distress or impairment in social, occupational or other important areas of concern	N/A	N/A
Not due to direct effect of a substance	N/A	N/A

Not accounted for by bereavement unless continuous for over 2 months or severe functional impairment, morbid preoccupation with worthlessness, psychotic symptoms or psychomotor retardation	N/A	N/A
If all of the above is true, must have one or both of these symptoms;		
Present for the same 2 week period	Q6031	*"Was this period [of sadness/loss of interest/low energy] for more than 2 weeks?"*
Depressed mood most of the day and nearly every day, self-reported or observed by others	Q6028	*"Have you had a period lasting several days when you felt sad, empty or depressed?"*
Or		
Markedly diminished interest or pleasure in all, or almost all, activities on most days, self-reported or reported by others	Q6029	"*Have you had a period lasting several days when you lost interest in most things you usually enjoy such as hobbies, personal relationships or work?"*
Must have either one or both of the above symptoms plus 3 or 4 of these to make a total of 5 or more symptoms;		
Significant weight loss (not due to dieting) or gain (e.g. 5% change in one month); or decrease or increase in appetite nearly every day	Q6033	*"During this period, did you lose your appetite?"*
Insomnia or hypersomnia nearly every day	Q2080	*"Overall in the last 30 days, how much of a problem did you have with sleeping, such as falling asleep, waking up frequently during the night or waking up too early in the morning?"*
Psychomotor agitation or retardation nearly every day, observable by others	Q6034	*"During this period, did you notice any slowing down in your thinking?"*
Fatigue or loss of energy nearly every day	Q6030	"*Have you had a period lasting several days when you have been feeling your energy decreased or that you are tired all the time?"*
Feelings of worthlessness or excessive or inappropriate guilt nearly every day; May be delusional; Not	N/A	N/A

merely self-reproach or guilt about being sick		
Diminished ability to think or concentrate, or indecisiveness, nearly every day (self-reported or observed by others)	Q2050	*"Overall in the last 30 days, how much difficulty did you have with concentrating or remembering things?"*
Recurrent thoughts of death, recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide		

Note: DSM-IV-TR: Diagnostic and Statistical Manual of Mental Disorders Fourth Edition, Test Revision; WHS: World Health Survey

A participant was counted as a probable case if major depressive disorder if Q6031 = 1 AND ((Min (Q6028, Q6029) = 1) AND Min (Q6033, Q2080 Q6034, Q6030, Q2050) =>4) OR (Min (Q6028, Q6029) = 2) AND Min (Q6033, Q2080, Q6034, Q6030, Q2050) =>3))

- For Q6028, Q6029, Q6030, Q6031, Q6033, and Q6034, a value of 5 ("No") was recoded as 0.
- For Q6025 and Q6028, a value =>3 was recoded as 1, and a value < 3 was recoded as 0.

The WHS item related to any health service utilization asked the following question;

> "*Have you taken medication or received other professional treatment for your depression symptoms? The purpose of this question is to find out if the respondent is currently under treatment for depression. An answer of "yes" is appropriate only if the respondent was taking any medications for depression during the last two weeks, or received some kind of treatment such as attending a psychological therapy or counselling session (often referred to as "therapy" or "psychotherapy") with a trained health professional during the last two weeks"*

There were three countries (Australia, Spain, France) for which we had comparable WHS and diagnostic based treatment rates for 'any health service'. These are summarised in Table 2. The pooled ratio of other diagnostic-based treatment rate: WHS treatment rates was 1.4 (95% confidence interval: 1.3-1.6).

Table 2: Comparable World Health Survey and diagnostic based treatment rates for 'any health service' in the dataset

WHS data-source						Corresponding diagnostic data-source			
Country	**Sex**	**Year**	**Age range**	**Treatment rate**	**Source**	**Year**	**Age range**	**Treatment rate**	**Source**
Australia	F	2003-2003	18-100	34.8%	(1)	2007-2007	16-85	62.2%	(2)
Australia	M	2003-2003	18-100	36.6%	(1)	2007-2007	16-85	42.9%	(2)
Spain	F	2002-2003	18-96	48.8%	(3)	2001-2002	18-100	60.8%	(4)
Spain	M	2002-2003	18-96	35.8%	(3)	2001-2002	18-100	54.2%	(4)
France	P	2003-2003	18-93	40.8%	(5)	2005-2005	20-100	60.4%	(6)

Note: WHS: World Health Survey; P: Person; F: Female; M: Male

The estimated pooled ratio suggested that surveys capturing a full diagnosis of MDD produced treatment rates that were approximately 1.4 times higher than WHS surveys. We maximized data inclusion by including the WHS treatment rates in our estimation of treatment gaps and adjusting them upwards towards the level they would have been had they been based on a full-diagnosis of MDD using our estimated pooled ratio. The adjusted and un-adjusted WHS estimates are summarized in Table 3.

Table 3: Adjusted and un-adjusted WHS treatment rates and gaps for major depressive disorder

			Un-adjusted estimates		Adjusted estimates	
Country	**Year**	**Age range**	**Treatment rate % (standard error)**	**Treatment gap % (95% confidence interval)**	**Treatment rate % (standard error)**	**Treatment gap % (95% confidence interval)**
Australia	2003-2003	18-100	35.4 (3.6)	64.6 (57.2-71.7)	50.2 (5.8)	49.8 (38.4-60.9)
Austria	2003-2003	20-93	38.1 (10.6)	61.9 (40.4-82.1)	54 (15.3)	46 (18.5-75.7)
Bangladesh	2003-2003	18-100	2.6 (0.6)	97.4 (96.1-98.4)	3.7 (0.9)	96.3 (94.3-97.9)
Belgium	2003-2003	17-100	37.1 (5.8)	62.9 (50.8-74.2)	52.7 (8.7)	47.3 (31.1-65)
Bosnia and Herzegovina	2003-2003	18-85	35 (5.6)	65 (53.5-75.3)	49.6 (8.4)	50.4 (34.2-66.8)
Brazil	2003-2003	18-96	25.9 (1.7)	74.1 (70.6-77.4)	36.8 (3.1)	63.2 (57.2-69.3)
Burkina Faso	2002-2003	18-100	8.4 (1.7)	91.6 (87.9-94.8)	11.9 (2.5)	88.1 (82.9-92.5)
Chad	2003-2003	18-100	6 (1.2)	94 (91.4-96.1)	8.6 (1.8)	91.4 (87.8-94.5)
China	2002-2002	18-96	1 (1.2)	99 (95.7-100)	1.5 (1.7)	98.5 (94.1-100)
Comoros	2003-2003	16-86	8.3 (2.7)	91.7 (86.1-96.2)	11.8 (3.9)	88.2 (79.9-94.7)
Cote d'Ivoire	2003-2003	18-96	8.7 (2.6)	91.3 (85.6-95.9)	12.3 (3.8)	87.7 (79.1-94.2)

Croatia	2003-2003	18-94	44.8 (6.3)	55.2 (42.7-67.6)	63.5 (9.6)	36.5 (18.3-56.9)
Czech Republic	2002-2003	18-95	10.6 (4.5)	89.4 (78.6-96.5)	15.1 (6.4)	84.9 (70.3-95.1)
Denmark	2003-2003	17-93	53.1 (8.8)	46.9 (30.3-64)	75.3 (13.2)	24.7 (5.1-53.8)
Dominican Republic	2003-2003	18-100	4 (1.1)	96 (93.5-97.9)	5.7 (1.6)	94.3 (90.8-96.9)
Ecuador	2003-2003	15-100	16.3 (2.5)	83.7 (78.4-88.4)	23.1 (3.8)	76.9 (69.2-83.9)
Estonia	2003-2003	17-93	14.6 (4.4)	85.4 (76.3-93)	20.6 (6.3)	79.4 (66.2-90.2)
France	2003-2003	18-93	40.8 (5.1)	59.2 (49.5-69)	57.8 (7.9)	42.2 (27.2-58.6)
Georgia	2003-2003	18-100	13.4 (2.5)	86.6 (81.3-91.3)	19 (3.6)	81 (73.6-87.8)
Germany	2004-2004	18-95	54.5 (8.7)	45.5 (28.7-62.9)	77.3 (13)	22.7 (3.7-52.4)
Ghana	2003-2003	18-100	2.8 (1)	97.2 (95-98.8)	4 (1.4)	96 (93-98.3)
Greece	2003-2003	18-96	11 (3.7)	89 (81-95.1)	15.5 (5.3)	84.5 (73.3-92.9)
India	2003-2003	18-100	4.7 (0.7)	95.3 (93.8-96.6)	6.6 (1.1)	93.4 (91.1-95.4)
Ireland	2003-2003	18-93	42.4 (8.7)	57.6 (41-74.3)	60.2 (12.8)	39.8 (15.9-65.6)
Italy	2003-2003	18-97	39.6 (6.7)	60.4 (46.8-72.8)	56.2 (10)	43.8 (24.9-64.4)
Kazakhstan	2002-2003	18-90	5.5 (1.9)	94.5 (90.3-97.7)	7.8 (2.7)	92.2 (86.2-96.5)
Kenya	2004-2004	18-100	4.3 (1.2)	95.7 (93.2-97.7)	6.1 (1.7)	93.9 (90.3-96.7)
Laos	2003-2003	18-100	23.6 (5.4)	76.4 (65.1-86.1)	33.4 (7.9)	66.6 (51-81.1)
Latvia	2003-2003	18-90	32.4 (6.6)	67.6 (54.1-79.3)	45.9 (9.7)	54.1 (35.6-73.3)
Luxembourg	2003-2003	18-91	34.7 (7.7)	65.3 (49.2-79.2)	49.3 (11.3)	50.7 (30-71.6)
Malawi	2003-2003	18-100	0	100	0	100
Malaysia	2003-2003	18-100	8.9 (2.8)	91.1 (85.2-95.6)	12.6 (4)	87.4 (79-94.2)
Mali	2003-2003	15-100	4.8 (1.9)	95.2 (91.2-98.1)	6.7 (2.6)	93.3 (87.3-97.3)
Mauritania	2003-2003	18-100	24.9 (3.6)	75.1 (67.6-82.1)	35.4 (5.5)	64.6 (53.6-74.7)
Mauritius	2003-2003	18-94	22.5 (2.3)	77.5 (72.4-81.9)	31.9 (3.7)	68.1 (60.8-75.2)
Morocco	2003-2003	18-100	5.6 (0.8)	94.4 (92.7-95.9)	7.9 (1.2)	92.1 (89.5-94.2)
Myanmar	2003-2003	18-100	7.3 (3.9)	92.7 (83.1-98.4)	10.4 (5.5)	89.6 (76.2-97.9)
Namibia	2003-2003	18-100	19.8 (3.1)	80.2 (73.9-86.2)	28.1 (4.6)	71.9 (62.4-80.2)
Netherlands	2004-2004	18-85	26.1 (4.7)	73.9 (64.4-82.4)	37.1 (6.9)	62.9 (49.5-75.8)
Norway	2003-2003	18-94	21.5 (6.4)	78.5 (64.3-89.4)	30.5 (9.2)	69.5 (50.6-85.7)
Pakistan	2003-2004	18-100	9.9 (1.3)	90.1 (87.2-92.5)	14 (2)	86 (81.9-89.9)
Paraguay	2002-2003	18-96	12 (2)	88 (83.8-91.6)	17 (2.9)	83 (76.7-88.3)
Philippines	2003-2003	18-100	8.7 (1.6)	91.3 (87.9-94.1)	12.4 (2.4)	87.6 (82.6-91.9)
Portugal	2003-2003	18-95	53.6 (3.9)	46.4 (39-54.5)	76 (6.9)	24 (11.7-39.6)
Republic of Congo	2003-2003	16-86	28.2 (2.6)	71.8 (66.5-76.7)	39.9 (4.3)	60.1 (51.6-68.3)
Russia	2003-2003	18-95	10.4 (1.9)	89.6 (85.6-93.1)	14.7 (2.8)	85.3 (79.3-90.1)
Senegal	2003-2003	18-100	7 (2.6)	93 (87.2-97.1)	9.9 (3.7)	90.1 (81.8-96.2)
Slovakia	2003-2003	18-100	5.9 (3)	94.1 (86.9-98.4)	8.4 (4.3)	91.6 (82.4-97.7)
Slovenia	2003-2003	18-92	18.8 (6.9)	81.2 (66.7-92.9)	26.6 (9.9)	73.4 (52-90.3)
South Africa	2002-2003	18-100	15.1 (4.1)	84.9 (76.1-91.6)	21.4 (5.9)	78.6 (66.1-88.4)
Spain	2002-2003	18-96	45.8 (2.3)	54.2 (49.7-58.5)	65 (4.8)	35 (26-44.6)
Sri Lanka	2003-2003	18-100	6.9 (2.2)	93.1 (88.1-96.7)	9.8 (3.2)	90.2 (83.2-95.5)
Swaziland	2003-2003	18-100	32.5 (3.4)	67.5 (60.9-74.2)	46 (5.5)	54 (43.5-64.6)
Sweden	2003-2003	18-92	23.1 (4.9)	76.9 (66.9-85.7)	32.8 (7.2)	67.2 (53.3-80.8)
Tunisia	2003-2003	18-100	10.6 (1.5)	89.4 (86.1-92.3)	15.1 (2.3)	84.9 (80.2-89.3)

Ukraine	2002-2003	18-96	9.5 (2.2)	90.5 (85.4-94.3)	13.5 (3.3)	86.5 (79.7-92)
United Arab Emirates	2003-2003	18-91	3.1 (2.2)	96.9 (91.1-99.7)	4.4 (3.1)	95.6 (88.1-99.5)
United Kingdom	2004-2004	18-100	41.2 (5.3)	58.8 (48.5-69.2)	58.4 (8.2)	41.6 (26.3-58.2)
Uruguay	2002-2003	18-100	25.3 (4.2)	74.7 (65.6-82.2)	35.9 (6.3)	64.1 (52.3-75.9)
Vietnam	2002-2003	18-96	12.6 (8.3)	87.4 (66.9-98.3)	17.9 (11.8)	82.1 (54.1-97.9)
Zambia	2003-2003	18-100	5.1 (1.5)	94.9 (91.3-97.5)	7.3 (2.2)	92.7 (87.9-96.3)
Zimbabwe	2003-2003	18-100	5.4 (2.2)	94.6 (89.9-98.2)	7.6 (3.1)	92.4 (85.1-97.4)

Study characteristics:

Treatment rates for dysthymia, depressive disorders combined, and mood disorders are reported separately. The level of service use investigated varied considerably between studies. Only a few countries (largely high-income countries) had treatment rates available for what we considered to be the most informative group i.e. aggregated treatment rates for any health service as well as either mental health service or minimally adequate treatment. The majority of countries had aggregated treatment rates available for any health service only, with large parts of the world with no data available. Treatment rates for non-health based services were generally reported in combination with rates of any service use, except for Singapore where one study focused non-health based services for MDD.

MDD Treatment Rates:
Any service use: There were 42 data-points pertaining to overall treatment rates of any service, from which we estimated treatment gaps for MDD, irrespective of whether the service was in the health or non-health sector. Treatment rates for the utilization of any service ranged from 79.6% [73.5, 85.7] in Australia to 15.7% [9.1, 22.3] in Mexico.

Any health service use: Treatment rates for any health service use were lowest in Malawi, China, Bangladesh, Ghana, and the UAE, and the highest in Denmark, Portugal, Australia, Germany, and Latvia. Treatment rates for any health service utilization ranged from 87.7% [83.5, 91.8] in Latvia to 0% (indicating no treatment) in Malawi. Figure 4 presents treatment gaps for MDD estimated for any heath service utilization and location. Although there were many parts of the world with missing data, we found a statistically significant effect by income group in our analyses. Treatment rates were significantly lower in middle and low income countries compared to high income countries.

Any general health service use: There were 35 data-points available on treatment rates of general health services, (i.e., services provided by general practitioners or other medical doctor). Treatment rates for any general health service use ranged from 80.7% [75.7, 85.7] in Latvia to 8.6% [2.9, 14.3] in Finland.

Any mental health service use: Treatment rates for any mental health service utilization ranged from 85.1% [47.4, 100] in Australia to 0% (Indicating no treatment) in Laos PDR for MDD. Pooled estimates reported in the paper by World Bank income level showed significant differences in treatment gaps between income groups.

Any non-health service use: Treatment rates for non-health services tended to be low. These ranged from 55.1% [46.2, 64.1] in Australia to 2.5 % [0.08, 4.3] in Canada for MDD.

Minimally adequate treatment: There were 38 data-points available on rates of minimally adequate treatment. The majority of estimates also came from high-income countries (28 data points from high-income countries; 10 data points from middle-income countries; 0 data points from low income countries). The criteria used to

define minimally adequate treatment varied considerably between studies and therefore these data were also analyzed separately. The key sources of variation included: (1) Whether or not a visit duration was specified; and (2) How 'psychotherapy' or 'psychological treatment' was operationalized, either according to the type of provider seen or therapy received. Based on the available definitions reported by studies that reported on MAT, we pooled rates for MAT definitions that were either considered stringent or non-stringent.

Table of Coefficients by GBD super-region

Covariate	Parameter Estimate [95% UI]
Intercept [a]	0.56 [0.54, 0.59] *
Treatment Type	
Any service use	0.3 [0.28, 0.31] *
General health service use	0.07 [0.06, 0.08] *
Minimally Adequate Treatment (MAT)	-0.12 [-0.13, -0.1] *
Any health service use	0.19 [0.17, 0.21] *
Any non-health service use	-0.17 [-0.18, -0.16] *
Sample characteristics	
Age	0.001 [0.0003, 0.0016] *
Percent Female	0.06 [0.04, 0.07] *
GBD Super Region	
Central Europe, Eastern Europe & Central Asia	-0.04 [-0.12, 0.04]
Latin America and Caribbean	-0.14 [-0.24, -0.04] *
North Africa, Middle East, and Sub-Saharan Africa	-0.31 [-0.39, -0.22] *
South Asia	-0.31 [-0.41, -0.2] *
Methodological Covariates	
Response rate	0.22 [0.08, 0.36] *
Type of diagnostic tool	-0.07 [-0.14, 0.001]

[a] Intercept represents mental health service use in High Income locations

*Statistically significant ($p < 0.05$)

Table 5. Regression Coefficients and 95% uncertainty intervals for MDD treatment rates modeled as a function of select covariates (treatment type, age, sex, location, response rate, and type of diagnostic tool)

Supplementary Results:

Treatment rates disaggregated by MDD severity vary for those with mild, moderate, or severe MDD and were therefore analyzed separately as part of the supplementary analyses. Minimally adequate treatment rates were also expected to vary by scope of MAT definitions. Therefore, analyses with MAT definitions were also conducted as part of our supplementary analyses (see below).

Dysthymia, depressive disorders, and mood disorders were expected to have different prevalence and treatment profiles and were therefore reported separately in our supplementary analyses (see appendix). For MDD data, initial model iterations were tested with year as a covariate to assess changes in treatment rates over time. However, there were very few countries (particularly in the low income group) with data spanning enough time points. Therefore, we chose to examine a few select countries (USA, Canada, Netherlands, and Australia) with relatively more data points by year to examine trends in treatment rates over time. For years with missing data, a spline interpolation with low degree polynomials was used to impute treatment coverage in order to better assess trends over a complete time series.

Treatment by Disorder severity: There were 55 datapoints for MDD from high-income countries. There were 23 estimates available for severe disorders, and 16 estimates for disorders moderate and mild severity respectively. Treatment rates were higher for all treatment types for severe MDD compared to moderate and mild MDD (See Fig 1).

Treatment by Year: We examined treatment rates by year for the United States, Canada, Netherlands, and Australia between 2002 and 2016. Based on data from the United States, estimated treatment rates increased between 35.5% in 2002 to 63.1% in 2007. Estimated treatment rates by year for the remaining countries are in Fig 2.

Treatment rates by MAT definition type: There were 38 datapoints for minimally adequate treatment for MDD. There were 12 datapoints with stringent definitions of MAT and 26 datapoints with non-stringent definitions of MAT. We examined variation in MAT rates by income group and definition stringency. The small number of data points precluded further analysis by other covariates. Overall, MAT rates were higher in high income locations compared to middle and low-income locations and MAT rates for stringent care thresholds were lower compared to those for non-stringent thresholds. In high income locations, the pooled MAT rate was 31.6% [9.1%, 54.5%] for stringent thresholds and 35.1% [12.8%, 57.1%] for non-stringent thresholds. In middle and low income locations, the pooled MAT rate was 11.95% [<1%, 35.2%] for stringent thresholds and 15.4% [<1%, 38.1%] for non-stringent thresholds.

Treatment rates for Dysthymia, Depressive Disorders, and Mood Disorders:

Dysthymia Treatment Rates: There were 11 studies (31 observations) reporting treatment rates for dysthymia. Data were from 9 high income and upper-middle income countries. Pooled treatment rates for dysthymia were higher for high-income countries 38.9% [29, 49.2] compared to upper-middle income countries 25.8% [3.7, 56].

Depressive Disorders Treatment Rates: There were 11 studies (43 observations) reporting treatment rates for depressive disorders as an aggregated group. Data from high-income countries originated from Western Europe and the United States. Upper-middle income countries included China, Mexico, Russia, and South Africa. Lower-middle income countries included Ghana and India. Differences in treatment rates by income group were similar to those for MDD and dysthymia. Pooled treatment rates for depressive disorders were significantly higher for high-income locations at 18.6% [15.5, 21.9] compared to upper-middle income at 13.7% [3.7, 28.6] and lower-middle income locations at 15.9% [4.1, 12.3]

Mood Disorders Treatment Rates: There were 35 studies (125 observations) reporting treatment rates for mood disorders. High income countries in this dataset included Australia, Belgium, Czech Republic, Israel, Japan, New Zealand, Spain, and the United States. Upper-middle income countries in this dataset included China, Argentina, Brazil, Colombia, Guatemala, Lebanon, Mexico, Peru, and South Africa. There were twenty four studies that used data from the World Mental Health surveys. Sensitivity analyses did not reveal any significant differences between treatment rates from WMHS compared to other studies (p-value = 0.81). The pooled treatment rate for mood disorders were 32.1% [27.1, 37.3] for high income countries and 13.8% [5.6, 25] for upper-middle income countries.

Figure 1. Treatment rates by MDD severity and treatment type

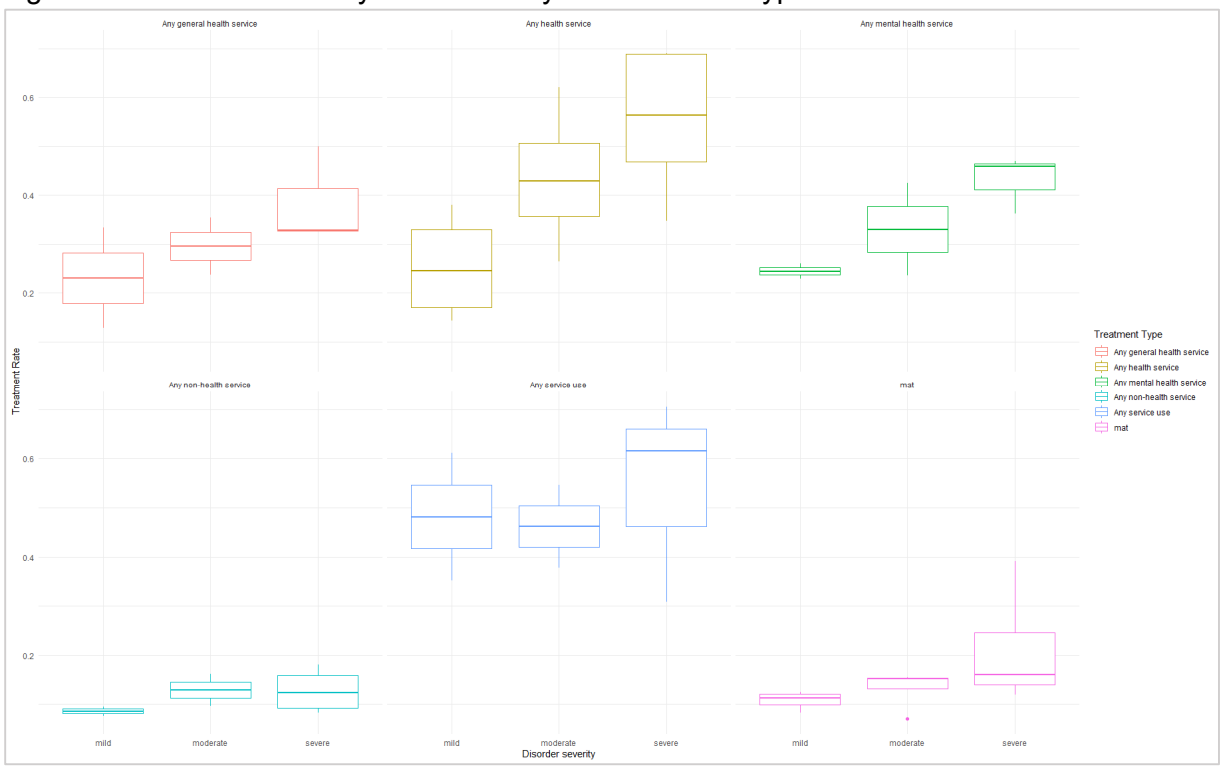
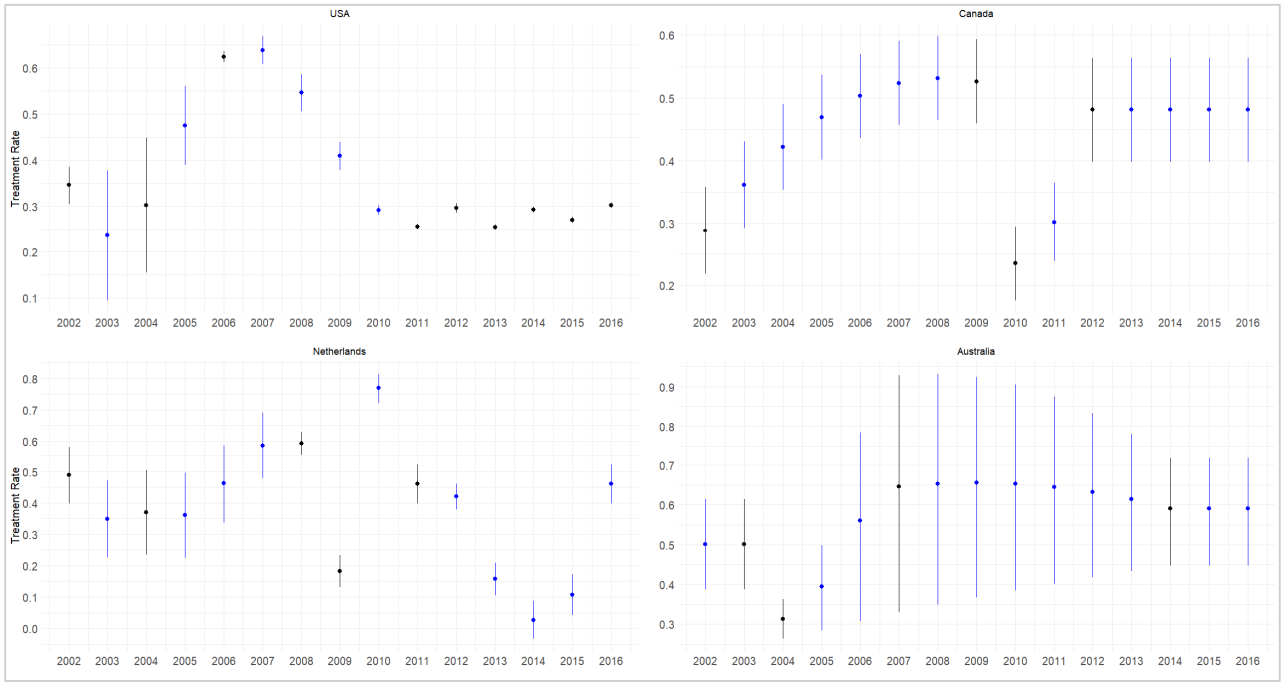


Figure 2: MDD Treatment rates by year; USA, Canada, Netherlands, Australia; 2002-2016



Note: Blue = interpolated; Black = Observed.

4. Discussion

The process of estimating the burden of mental disorders is challenging due to several reasons including sparsity of data from many parts of the world, methodological constraints, and accurate information on the efficacy of existing interventions and treatment coverage for common mental disorders such as major depressive disorder. To this end, this thesis aimed to examine three main areas of the mental disorder burden estimation process, namely comparative epidemiological assessments, efficacy of existing interventions, and examining variation in treatment coverage.

The most comprehensive study of mental disorders as risk factors for suicide was conducted by Ferrari and colleagues in 2014. The work in Chapter one improves upon the Ferrari 2014 study by updating the evidence base and analytically incorporating relevant sample and methodological characteristics that may contribute to variation in suicide risk. The main improvements included adopting a hierarchical approach to incorporate between-study random effects, inclusion of relevant covariates, and the use of more rigid study selection procedures. These updated findings highlight the importance of study characteristics and how they may contribute to variation in reported relative risks. Death by suicide and associated risk factors can be inherently challenging to empirically study in many settings, which may then lead to bias in reported results. This review may provide insight on analyzing and interpreting synthesized evidence by incorporating these sources of bias. The main limitation was the lack of sufficient studies that met our rigorous selection criteria therefore precluding our ability to provide estimates on a wider range of MDs and incorporating additional covariates in our final model.

Chapter two includes an application of the MR-BRT framework to data on treatment efficacy for major depressive disorders (MDD). A network meta-analytic approach was used to estimate pooled standardized mean differences by intervention class. An analysis of variation in treatment efficacy by type of outcome measure used to assess symptoms of depression was also conducted. Our reported efficacy estimates show that CBT is a highly effective intervention and that the pill placebo is a powerful control condition with its own significant treatment effect. We were limited in our ability to incorporate information on trial and sample characteristics that may inform variation in treatment efficacy estimates due to lack of these data for this work.

In addition to treatment efficacy, treatment coverage is another important input needed to improve the severity estimation process for MDD. Chapter three presents an updated systematic review of treatment coverage for MDD using evidence from eighty-three countries from 2000 to 2019. Data were sourced from a literature review as well as from the World Health Surveys (WHS) and World Mental Health surveys (WMHS). A Bayesian meta-regression framework (MR-BRT) is used to examine heterogeneity in the existing published data and estimates treatment coverage by types of service including health services, mental health services, and minimally adequate treatment. Sample and methodological characteristics that may influence treatment coverage estimates were explored analytically. Treatment rates by service type and location and corresponding treatment gaps by age group, sex, and location were estimated. Our findings showed wide disparities in treatment coverage by type of treatment and resource-setting. Sample and study method attributes were also found to be important predictors of treatment rates. We were limited by the number of available studies by location with most evidence originating from high-income countries.

4.1 Future Research Areas:

The work presented in this thesis have several important implications in measurement of population mental health. Findings from chapter one may improve population-attributable-fraction (PAF) estimation by incorporating relative risks that are adjusted for methodological covariates. Subsequently, these relative risk estimates may also find applications in calculating estimates of mental disorder attributable suicide DALYs (disability-adjusted life years). This will allow us to examine any shifts in fatal burden due to mental disorders (assigned to them as a direct cause) and overall burden of mental disorders. Differences between earlier and

newly calculated PAFs and attributable suicide DALYs may provide further information on the impact of sources of measurement error. The earlier Ferrari review applied the same risk estimates across all countries, age groups, and sex which could potentially mask any differences in the distribution of attributable suicide DALYs. The adjusted RRs across sex and age with additional analyses by income-level in Chapter one may provide additional layers of refinement for risk estimates and subsequently attributable suicide YLLs.

Treatment efficacy analyses conducted as part of Aim II may find useful applications in re-estimating the severity distribution for MDD and quantifying the proportion of MDD burden that may be potentially avoidable under routine and best-functioning healthcare scenarios. Current methods in mental disorder burden estimation incorporate data on the severity of disorders. Severity of disorders in turn is affected by the effectiveness of interventions. Therefore, the outputs from chapter two may find applications in re-estimating the severity of depression.

Future analyses of MDD treatment coverage (Aim III) may incorporate more detailed and geographically representative data on adequacy of care. Most current knowledge of minimally adequate treatment is informed by data from high-income locations. Better data on MAT from LMICs could consider what MAT treatment rates may look like in LMICs and how that subsequently affects depression burden for those locations. As detailed earlier, efforts to improve mental disorder burden estimates need to account for treatment efficacy and geographic variation in treatment use. These are important to consider as they may influence the range of functional health loss due to mental disorders that is experienced across populations. In other words, depression burden may be amenable to quality and use of healthcare services. Therefore, an important application of the work in Chapter three may be to quantify the amount of depression burden that can be averted or prevented by treatment.

This work relies on studies that collect primary data from research participants. Primary data collection efforts are very often accompanied by several constraints such as sampling procedures, participant response rate and loss to follow-up, measurement approaches, choice of study design, and others. These constraints are indicative of study quality and potential bias that may exist in study-reported data. These subsequently contribute to variation in study-reported metrics that are then detected in meta-analytic assessments of the accumulated evidence. The work presented here explains why specific study elements were quantified and accounted for in the analyses along with the rationale for these approaches. For instance, we explain why a prospective study design is considered “gold-standard” and therefore used as a reference category in our analyses in Chapter 1. We explain why we only include studies that use clinical diagnostic instruments to assess MDs and the limitations of estimates reported by studies that use symptom scales in both Chapters 1 and 3. Methodological details such as these may serve as useful guidelines for researchers collecting primary data on MDs and their treatment while developing study protocols to ensure the highest possible study quality as well as documenting constraints and limitations encountered during the research study period while writing up results for publication. This may in turn increase the clarity and usability of research outputs for both systematic review researchers as well as clinical decision-makers.

This thesis addresses several important and emerging areas of work in population-level measurement of mental disorders. Ongoing reviews and syntheses of evidence such as those presented in the chapters above are needed in order to arrive at precise estimates of the global burden of mental disorders that may be of use to healthcare providers, researchers, and policy makers to target specific areas of need and improvement in mental health.

5. Acknowledgements:

I have many to thank for their support throughout my time in the PhD program.

I would like to thank IHME and UW for financial and administrative support. Thank you to the many UW librarians who have helped me source countless papers for my work. Thank you to the UW community and our PhD cohort. I have thoroughly enjoyed being part of such a vibrant and encouraging space and sharing so many experiences together.

I would like to thank Dr. Angela Micah for being a very understanding and supportive supervisor at IHME. I have also benefited from many classes taken at UW. In particular, I am grateful for learning from Chris Adolph in his excellent class on maximum likelihood estimation. And thank you to Ali Rowhani-Rahbar for leading a wonderful introductory series in epidemiology and also for serving as the GSR on my committee.

I could not have hoped for a better dissertation committee. Thank you to Theo for your valuable advice and feedback. Thank you to Damian for providing much-needed support and guidance for many important components of my work. Thank you to Alize for helping shape and guide my work, for taking the time to check in regularly, and providing excellent feedback to many manuscript drafts. And lastly, a very big thank you to my Chair Dr. Pamela Y. Collins for graciously agreeing to take on the task of chairing my committee. I have learned a lot from you in our many discussions and can only hope to be as inspiring and empathetic a mentor and as impactful in the global mental health space as you are.

Thank you to my friends near and far. I hope to enjoy many more hikes, boardgame sessions, and happy hours in person soon. Thank you, Diya, for always being there to hear me out and being so generous with your time, patience, and advice. I look forward to many more fun times with you and Michael (and all our meme swaps of cats, capybaras, and much more).

It would be remiss of me to not acknowledge the extraordinary privilege I have of belonging to a remarkably progressive family in Kolkata, India that prioritizes women's education and autonomy. I stand on their shoulders now and always. Thank you to my brother Deep, my aunt Nayantara, and extended family in India for cheering me on and always having my back. A special thank you to my 93-year-old grandmother Chitra Palchoudhuri who loves asking me questions about what I am working on and for her unconditional affection. Thank you to my father Soumyo Moitra for inspiring me to pursue doctoral studies. Thank you to my mother Sreoshi Moitra for being a such a strong pillar of emotional support and for providing much-needed encouragement and enthusiasm for my work.

And lastly, none of this would have been possible without Pradeep. Thank you for being such an enduring partner through graduate school and many more adventures. I have been inspired daily as I watched you complete your doctoral studies and now have been fortunate to have you by my side and hold space for me through many hard and happy days. Thank you for your unfailing support, empathy, the good and not-so-good jokes, and for all the love.