Comparison of major depression diagnostic classification probability using the SCID, CIDI, and MINI diagnostic interviews among women in pregnancy or postpartum: An individual participant data meta-analysis

Brooke Levis1,2 | Dean McMillan3 | Ying Sun1 | Chen He1,2 | Danielle B. Rice1,4 |
Ankur Krishnan1 | Yin Wu1,5 | Marlinee Azar1,2 | Tatiana A. Sanchez1 |
Matthew J. Chiovitti1 | Parash Mani Bhandari1,2 | Dipika Neupane1,2 |
Nazanin Saadat1 | Kira E. Riehm1,6 | Mahrurkh Imran1 | Jill T. Boruff7 |
Pim Cuijpers8 | Simon Gilbody3 | John P.A. Ioannidis9 | Lorie A. Kloda10 |
Scott B. Patten11,12,13 | Ian Shrier1,2,14 | Roy C. Ziegelstein15 | Liane Comeau16 |
Nicholas D. Mitchell17,18 | Marcello Tonelli19 | Simone N. Vigod20 | Franci Aceti21 |
Rubén Alvarado22 | Cosme Alvarado-Equivel23 | Muideen O. Bakare24,25 |
Jacqueline Barnes26 | Cheryl Tatano Beck27 | Karola Bindt28 | Philip M. Boyce29,30 |
Adomas Bunevicius31 | Tiago Castro e Couto32 | Linda H. Chaudron33 |
Humberto Correa34 | Felippe Pinheiro de Figueiredo35 | Valsamma Eapen36,37,38 |
Michelle Fernandes39,40 | Barbara Figueiredo41 | Jane R.W. Fisher42 |
Lluisa Garcia-Esteve43,44,45 | Lisa Giardinelli46 | Nadine Helle28 |
Louise M. Howard47,48 | Dina Sami Khalifa49,50,51 | Jane Kohlhoff36,37,52 |
Laima Kusminskas53 | Zoltán Kozinszky54 | Lorenzo Lelli46 |
Angeliki A. Leonardou55 | Beth A. Lewis56 | Michael Maes57,58 | Valentina Meuti21 |
Sandra Nakic Rados59 | Purificacion Navarro Garcia43,60 | Daisuke Nishi61,62 |
Daniel Okitundu Luwa E-Andjafo63 | Emma Robertson-Blackmore64 |
Tamsen J. Rochat65,66 | Heather J. Rowe42 | Bonnie W.M. Siu67 |
Alkistis Skalkidou68 | Alan Stein69,70 | Robert C. Stewart71,72 | Kuan-Pin Su73,74 |
Inger Sundstrom-Poromaa68 | Meri Tadinac75 | S. Darius Tandon76 | Iva Tendoi41 |
Papaavi Thiagayson77,78,79 | Annamaria Toereki80 | Anna Torres-Gimenez43,44,45 |
Thach D. Tran42 | Kylee Trevillion47 | Katherine Turner81 |
Johann M. Vega-Dienstmaier82 | Karen Wynter42,83 | Kimberly A. Yonkers84,85,86 |

Received: 14 May 2019 Revised: 5 August 2019 Accepted: 16 August 2019
DOI: 10.1002/mpr.1803

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.


Int J Methods Psychiatr Res. 2019:e1803. wileyonlinelibrary.com/journal/mpr 1 of 11
https://doi.org/10.1002/mpr.1803
Abstract

Objectives: A previous individual participant data meta-analysis (IPDMA) identified differences in major depression classification rates between different diagnostic interviews, controlling for depressive symptoms on the basis of the Patient Health Questionnaire-9. We aimed to determine whether similar results would be seen in a different population, using studies that administered the Edinburgh Postnatal Depression Scale (EPDS) in pregnancy or postpartum.
Methods: Data accrued for an EPDS diagnostic accuracy IPDMA were analysed. Binomial generalised linear mixed models were fit to compare depression classification odds for the MINI International Neuropsychiatric Interview (MINI), Composite International Diagnostic Interview (CIDI), and Structured Clinical Interview for DSM (SCID), controlling for EPDS scores and participant characteristics.

Results: Among fully structured interviews, the MINI (15 studies, 2,532 participants, 342 major depression cases) classified depression more often than the CIDI (3 studies, 2,948 participants, 194 major depression cases; adjusted odds ratio [aOR] = 3.72, 95% confidence interval [CI] [1.21, 11.43]). Compared with the semistructured SCID (28 studies, 7,403 participants, 1,027 major depression cases), odds with the CIDI (interaction aOR = 0.88, 95% CI [0.85, 0.92]) and MINI (interaction aOR = 0.95, 95% CI [0.92, 0.99]) increased less as EPDS scores increased.

Conclusion: Different interviews may not classify major depression equivalently.

KEYWORDS
depressive disorders, diagnostic interviews, Edinburgh Postnatal Depression Scale, individual participant data meta-analysis, major depression
INTRODUCTION

Among diagnostic interviews for classifying major depression in research, semistructured interviews, such as the Structured Clinical Interview for DSM (SCID; First, 1995), are designed to be administered by clinically trained professionals, who may insert unscripted queries and use judgement to decide whether symptoms are present. Fully structured interviews, such as the Composite International Diagnostic Interview (CIDI; Robins et al., 1988), are completely scripted and can be administered by lay interviewers. The Mini International Neuropsychiatric Interview (MINI; Lecrubier et al., 1997; Sheehan et al., 1997) is a very brief fully structured interview that was designed for rapid administration and intended to be overinclusive.

The different diagnostic interviews are typically considered equivalent for major depression classification in research (Hewitt, Gilbody, Brealey, et al., 2009; Manea, Gilbody, & McMillan, 2012; Mitchell, Meader, & Symonds, 2010; Moriarty, Gilbody, McMillan, & Manea, 2015). However, a recent individual participant data meta-analysis (IPDMA) of 57 studies (17,158 participants) from diverse settings that controlled for participant characteristics and depressive symptom severity on the basis of EPDS scores. Finally, we tested whether differences in classification rates between interviews were associated with depressive symptom severity.

METHODS

We used data accrued for an IPDMA on the diagnostic accuracy of the EPDS, which is the most commonly used depression screening tool for women in pregnancy or postpartum (Hewitt et al., 2009). The IPDMA was registered in PROSPERO (CRD42015024785), a protocol was published (Thombs et al., 2015), and results were reported following PRISMA-DTA (McInnes et al., 2018) and PRISMA-IPD (Stewart et al., 2015) reporting guidelines.

Identification of eligible studies

For the main IPDMA, data sets from articles in any language were eligible for inclusion if (a) they included diagnostic classification for current major depressive disorder (MDD) or major depressive episode (MDE) using any version of Diagnostic and Statistical Manual of Mental Disorders (DSM; American Psychiatric Association [APA], 1987; APA, 1994; APA, 2000) or International Classification of Diseases (ICD; World Health Organization, 1992) criteria on the basis of a validated semistructured or fully structured interview; (b) they included EPDS scores; (c) the diagnostic interview and EPDS were administered within 2 weeks of each other because DSM and ICD criteria specify that symptoms must have been present in the last 2 weeks; (d) participants were women aged ≤18 years who were not recruited from youth or college settings; and (e) participants were not recruited from psychiatric settings or because they were identified as having symptoms of depression because screening is done to identify previously unrecognised cases. For the present study, we only included studies that assessed major depression using the SCID, CIDI, and MINI because there were only three studies that used other interviews.

Data sets where not all participants were eligible were included if primary data allowed selection of eligible participants. For defining major depression, we considered MDD or MDE on the basis of the DSM or ICD. If more than one was reported, we prioritised MDE over MDD, because screening would attempt to detect depressive symptom severity on the basis of EPDS scores. Finally, we tested whether differences in classification rates between interviews were associated with depressive symptom severity.
episodes and further interview would determine if the episode is related to MDD or bipolar disorder, and DSM over ICD.

### 2.2 Search strategy and study selection

A medical librarian searched Medline, Medline In-Process & Other Non-Indexed Citations and PsycINFO via OvidSP, and Web of Science via ISI Web of Knowledge from inception to June 10, 2016, using a peer-reviewed search strategy (Methods S1: PRESS, 2016). We also reviewed reference lists of relevant reviews and queried contributing authors about non-published studies. Search results were uploaded into RefWorks (RefWorks-COS, Bethesda, MD, USA). After deduplication, unique citations were uploaded into DistillerSR (Evidence Partners, Ottawa, Canada) for storing and tracking search results.

Two investigators independently reviewed titles and abstracts for eligibility. If either deemed a study potentially eligible, full-text review was done by two investigators, independently, with disagreements resolved by consensus, consulting a third investigator when necessary. A translator was consulted for determining the eligibility of one Chinese article.

### 2.3 Data extraction, contribution, and synthesis

Authors of eligible data sets were invited to contribute de-identified primary data. We emailed corresponding authors of eligible primary studies at least three times, as necessary. If we did not receive a response, we emailed co-authors and attempted to contact corresponding authors by phone.

Diagnostic interview used as the reference standard and country were extracted from published reports by two investigators independently, with disagreements resolved by consensus. Countries were categorised as “very high,” “high,” or “low–medium” development on the basis of the United Nation’s Human Development Index, a statistical composite index that includes indicators of life expectancy, education, and income (United Nations, 2019). Participant-level data provided in data sets included age, pregnancy status (pregnant vs. postpartum), EPDS scores, and major depression status.

Individual participant data were converted to a standard format and synthesised into a single data set with study-level data. We compared published participant characteristics and diagnostic accuracy results with results from raw data sets and resolved any discrepancies in consultation with the original investigators. For the present study, we restricted our data to participants with complete data for all variables included in our analyses. Then, for studies that collected data at multiple time points, we restricted our data to the time point with the most participants. If there was a tie, we selected the time point with the largest number of major depression cases.

### 2.4 Statistical analyses

To isolate the association between diagnostic assessment method and major depression classification, we estimated binomial generalised linear mixed models with a logit link function. All analyses controlled for depressive symptom severity (continuous EPDS scores), age (continuous), country Human Development Index (very high, high, or low–medium), and pregnant versus postpartum status. Given that each study only administered one diagnostic interview, these covariates were included in analyses to account for their potential influence on major depression classification. Covariates were chosen a priori on the basis of their potential influence on major depression classification as well as their availability across primary studies. To account for correlation between subjects within the same primary study, a random intercept was fit for each primary study. Fixed slopes were estimated for EPDS score, diagnostic interview, age, Human Development Index, and pregnant versus postpartum status.

We estimated generalised linear mixed models to compare major depression classification odds for MINI versus CIDI, CIDI versus SCID, and MINI versus SCID. We then fit additional models including an interaction between interview and EPDS score. All analyses were run in R using the glmer function within the lme4 package.

### 3 RESULTS

Of 3,418 unique titles and abstracts identified from the database search, 3,097 were excluded after title and abstract review and 226 were excluded after full text review, leaving 95 eligible articles with data from 64 unique participant samples, of which 45 (70% of data sets; 70% of participants) contributed data (Figure 1). Reasons for exclusion for the articles excluded at the full-text level are given in Table S1. In addition, authors of included studies contributed data from an additional eligible study that was not identified in the search, for a total of 46 data sets. Characteristics of included studies and eligible studies that did not provide data sets are shown in Table S2. In total, 12,759 participants (1,553 [12%] with major depression) were included; none of whom were included in the previous PHQ-9 analysis (Levis et al., 2018).

Of the 46 total included studies, there were 28 SCID studies (7,279 participants, 14% major depression), 3 CIDI studies (2,948 participants, 7% major depression), and 15 MINI studies (2,532 participants, 14% major depression; Table 1). Seventeen of the 28 SCID studies described the SCID as having been administered by clinically trained professionals.

As shown in Figure 2 and Table S3, for all interviews, the proportion with major depression generally increased as EPDS scores increased.

Model coefficients for each analysis are shown in Table S4. Among fully structured interviews, controlling for EPDS scores, the MINI was more likely to classify major depression than the CIDI (adjusted odds ratio [aOR] = 3.72; 95% confidence interval [CI] [1.21, 11.43]). The CIDI and MINI tended to classify major depression less often than the SCID, but there was high uncertainty in estimates (aOR for CIDI vs. SCID = 0.34, 95% CI [0.09, 1.34]; aOR for MINI vs. SCID = 0.91, 95% CI [0.43, 1.94]).
As EPDS scores increased, the probability of diagnosis increased more for the MINI than for the CIDI (interaction aOR = 1.07, 95% CI [1.03, 1.12]) but increased less for both the CIDI and MINI than for the SCID (interaction aOR for CIDI = 0.88, 95% CI [0.85, 0.92]; interaction aOR for MINI = 0.95, 95% CI [0.92, 0.99]).

4 | DISCUSSION

We compared depression classification across diagnostic interviews in studies that administered the EPDS with women in pregnancy or postpartum, controlling for participant characteristics and depressive symptom severity on the basis of EPDS scores. Among fully structured interviews, odds of major depression were substantially higher for the MINI than the CIDI. As depressive symptom severity increased, the probability of diagnosis increased more for the MINI than for the CIDI. There were no definitive differences in classification odds between the CIDI and SCID and between the MINI and SCID, but, as EPDS scores increased, likelihood of classification increased less for the CIDI and MINI than for the SCID. Results were similar to those of our previous study that assessed depressive symptom severity in diverse patient groups with the PHQ-9 (Levis et al., 2018). In that study, on the basis of subgroup analyses by PHQ-9 scores, we found that the CIDI classified fewer participants with high-level depressive symptoms as depressed than the SCID. Due to limited numbers of participants and major depression cases for each interview across EPDS scores in the present study, we were unable to conduct subgroup analyses based on EPDS scores. However, our interaction analyses were generally consistent with previous findings.

There are limitations to consider. First, we were unable to obtain primary data for 19 of 64 eligible data sets identified in our search (30% of data sets; 30% of participants). Second, only three included studies used the CIDI, one of which had only one major depression case. Third, across interviews, there were few participants with high EPDS scores and few major depression cases with low EPDS scores. For the CIDI, data were sparse across EPDS scores. Notwithstanding,
the previous PHQ-9 study and the present study have used samples many times the size of other studies that have attempted to compare diagnostic interviews for major depression (Anthony et al., 1985; Booth, Kirchner, Hamilton, Harrell, & Smith, 1998; Brugha, Jenkins, Taub, Meltzer, & Bebbington, 2001; Hesselbrock, Stabenau, Hesselbrock, Mirkin, & Meyer, 1982; Jordanova, Wickramesinghe, Gerada, & Prince, 2004). Fourth, residual confounding may exist. We were only able to consider variables collected in the original investigations, and the included study-level variables may not apply uniformly to all participants in a study. Finally, not all SCID studies described interviewer qualifications. It is possible that use of untrained interviewers may have reduced performance differences across interviews.

5 | CONCLUSION

The previous PHQ-9 IPDMA found that different diagnostic interviews may not be equivalent for major depression classification. In the present study, we observed similar patterns. The CIDI and MINI were designed as less resource-intensive options that can be administered by research staff without diagnostic skills, but they may misclassify major depression in substantial numbers of patients compared with the SCID. The findings of both the previous and present IPDMAs suggest that different interviews may not classify major depression equivalently and should be combined in meta-analyses with caution.

DATA ACCESSIBILITY

Requests to access data should be made to the corresponding author at brett.thombs@mcgill.ca.

DECLARATION OF INTEREST STATEMENT

All authors have completed the ICMJE uniform disclosure form and declare no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous 3 years with the following exceptions: Dr. Patten reports grants from Hotchkiss Brain Institute/Pfizer Competition, outside the submitted work. Dr. Tonelli declares that he has received a grant from Merck Canada, outside the submitted work. Dr. Vigod declares that she receives royalties for her Postpartum Depression Screening Scale published by Western Psychological Services. Dr. Boyce declares that he receives grants and personal fees from Servier, grants from Lundbeck, and personal fees from AstraZeneca, all outside the submitted work. Dr. Howard declares that she has received personal fees from NICE Scientific Advice, outside the submitted work. Dr. Nishi declares that he has received personal fees from Sumitomo Dainippon Pharma Co., Ltd., outside the submitted work. Dr. Sundström-Poromaa declares that she has served on advisory boards and acted as invited speaker at scientific meetings for MSD, Novo Nordisk, Bayer Health Care, and Lundbeck A/S. Dr. Yonkers declares that she receives royalties from UpToDate, outside the submitted work.

ACKNOWLEDGEMENTS

This study was funded by the Canadian Institutes of Health Research (CIHR, KRS-140994). Ms. Levis was supported by a CIHR Frederick Banting and Charles Best Canada Graduate Scholarship Doctoral Awards. Ms. Rice was supported by a Vanier Canada Graduate Scholarship. Dr. Wu was supported by an Utting Postdoctoral Fellowship from the Jewish General Hospital, Montreal, Quebec, Canada. Ms. Azar was supported by a Fonds de recherche...
du Québec - Santé (FRQS) Masters Training Award. Mr. Bhandari was supported by a studentship from the Research Institute of the McGill University Health Centre. The primary study by Alvarado et al. was supported by the Ministry of Health of Chile. The primary study by Barnes et al. was supported by a grant from the Health Foundation (1665/608). The primary study by Beck et al. was supported by the Patrick and Catherine Weldon Donaghue Medical Research Foundation and the University of Connecticut Research Foundation. The primary study by Helle et al. was supported by the Werner Otto Foundation, the Kroschke Foundation, and the Feindt Foundation. Prof. Robertas Bunevicius, MD, PhD (1958–2016) was principal investigator of the primary study by Bunevicius et al. but passed away and was unable to participate in this project. The primary study by Couto et al. was supported by the National Counsel of Technological and Scientific Development (CNPq; Grant 444254/2014-5) and the Minas Gerais State Research Foundation (FAPEMIG; Grant APQ-01954-14). The primary study by Chaudron et al. was supported by a grant from the National Institute of Mental Health (Grant K23 MH64476). The primary study by Figueira et al. was supported by the Brazilian Ministry of Health and by the National Counsel of Technological and Scientific Development (CNPq; Grant 403433/2004-5). The primary study by de Figueiredo et al. was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo. The primary study by Tisot et al. was supported by the Swiss National Science Foundation (Grant 32003B 125493). The primary study by Fernandes et al. was supported by grants from the Child: Care Health and Development Trust and the Department of Psychiatry, University of Oxford, Oxford, UK, and by the Ashok Ranganathan Bursary from Exeter College, University of Oxford. Dr. Fernandes was supported by a National Institute for Health Research (NIHR) academic clinical fellowship. The primary study by Tendas et al. was supported under the project POCI/SAU-ESP/56397/2004 by the Operational Program Science and Innovation 2010 (POCI 2010) of the Community Support Board III and by the European Community Fund FEDER. The primary study by Fisher et al. was supported by a grant under the Invest to Grow Scheme from the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs. The primary study by Garcia-Esteve et al. was supported by Grant 7/98 from the Ministerio de Trabajo y Asuntos Sociales, Women’s Institute, Spain. The primary study by Howard et al. was supported by the NIHR under its Programme Grants for Applied Research Programme (Grants RP-PG-1210-12002 and RP-DG-1108-10012) and by the South London Clinical Research Network. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care. The primary study by Phillips et al. was supported by a scholarship from the National Health and Medical and Research Council (NHMRC). The primary study by Roomruangwong et al. was supported by the Ratchadhaphiseksomphot Endowment Fund 2013 of Chulalongkorn University (CU-56-457-HR). The primary study by Nakić Radoš et al. was supported by the Croatian Ministry of Science, Education, and Sports. The primary study by Navarro et al. was supported by Grant 13/00 from the Ministry of Work and Social Affairs, Institute of Women, Spain. The primary study by Usuda et al. was supported by Grants-in-Aid for Young Scientists (A) from the Japan Society for the Promotion of Science (primary investigator: Daisuke Nishi, MD, PhD) and by an Intramural Research Grant for Neurological and Psychiatric Disorders from the National Center of Neurology and Psychiatry, Japan. Dr. Robertson-Blackmore was supported by a Young Investigator Award from the Brain and Behavior Research Foundation and NIMH Grant K23MH080290. The primary study by Rochat et al. was supported by grants from University of Oxford (HQS035), the Tuixen Foundation (9940), the Wellcome Trust (082384/Z/07/Z and 071571), and the American Psychological Association. Dr. Rochat receives salary support from a Wellcome Trust Intermediate Fellowship (211374/Z/18/Z). The primary study by Rowe et al. was supported by the diamond Consortium, beyondblue Victorian Centre of Excellence in Depression and Related Disorders. The primary study by Comasco et al. was supported by funds from the Swedish Research Council (VR: 521-2013-2339, VR: 523-2014-2342), the Swedish Council for Working Life and Social Research (FAS: 2011-0627), the Marta Lundqvist Foundation (2013, 2014), and the Swedish Society of Medicine (SLS-331991). The primary study by Prenoveau et al. was supported by the Wellcome Trust (Grant 071571). The primary study by Stewart et al. was supported by Professor Francis Creed’s Journal of Psychosomatic Research Editorship fund (BA00457) administered through University of Manchester. The primary study by Su et al. was supported by grants from the Department of Health (DOH94F044 and DOH95F022) and the China Medical University and Hospital (CMU94-105, DMR-92-92, and DMR94-46). The primary study by Tandon et al. was supported by the Thomas Wilson Sanitarium. The primary study by Tran et al. was supported by the Myer Foundation who funded the study under its Beyond Australia scheme. Dr. Tran was supported by an early career fellowship from the Australian National Health and Medical Research Council. The primary study by Vega-Dienstmaier et al. was supported by Tejada Family Foundation, Inc. and Peruvian-American Endowment, Inc. The primary study by Yokners et al. was supported by a National Institute of Child Health and Human Development grant (5 R01HD045735). Drs. Benedetti and Thombs were supported by FRQS researcher salary awards.

AUTHOR CONTRIBUTIONS


SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Levis B, McMillan D, Sun Y, et al. Comparison of major depression diagnostic classification probability using the SCID, CIDI, and MINI diagnostic interviews among women in pregnancy or postpartum: An individual participant data meta-analysis. *Int J Methods Psychiatr Res.* 2019:e1803. [https://doi.org/10.1002/mpr.1803](https://doi.org/10.1002/mpr.1803)