

1 *Probability of Major Depression Classification Based on the SCID, CIDI and MINI*
2 *Diagnostic Interviews: A Synthesis of Three Individual Participant Data Meta-Analyses*

3
4 Yin Wu¹⁻³; Brooke Levis^{1,3,4}; John P. A. Ioannidis⁵; Andrea Benedetti^{3,6,7*}; Brett D. Thombs^{1-3,6-9*}; and the
5 DEPRESSion Screening Data (DEPRESSD) Collaboration.

6
7 ¹ Lady Davis Institute for Medical Research, Jewish General Hospital, Montréal, Québec, Canada;

8 ² Department of Psychiatry, McGill University, Montréal, Québec, Canada;

9 ³ Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montréal,
10 Québec, Canada;

11 ⁴ Centre for Prognosis Research, School of Primary, Community and Social Care, Keele University,
12 Staffordshire, UK;

13 ⁵ Departments of Medicine, Health Research and Policy, Biomedical Data Science, and Statistics, Stanford
14 University, Stanford, California, USA;

15 ⁶ Respiratory Epidemiology and Clinical Research Unit, McGill University Health Centre, Montréal,
16 Québec, Canada;

17 ⁷ Department of Medicine, McGill University, Montréal, Québec, Canada; Biomedical Ethics Unit, McGill
18 University, Montréal, Québec, Canada;

19 ⁸ Department of Psychology, McGill University, Montréal, Québec, Canada;

20 ⁹ Department of Educational and Counselling Psychology, McGill University, Montréal, Québec, Canada.

21 * Co-senior authors

22
23 Short title: Comparison of Diagnostic Interviews for Major Depression

24
25 **Corresponding authors:**

26 Brett D. Thombs, PhD

27 Jewish General Hospital

28 4333 Cote Ste Catherine Road

29 Montreal, Quebec H3T 1E4

30 Tel: (514) 340-8222 ext. 25112

31 E-mail: brett.thombs@mcgill.ca

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

32

33 Andrea Benedetti, PhD

34 Centre for Outcomes Research & Evaluation, Research Institute of the McGill University Health Centre

35 5252 Boulevard de Maisonneuve

36 Montréal, Quebec, H4A 3S5, Canada

37 Tel: (514) 934-1934 ext. 32161;

38 E-mail: andrea.benedetti@mcgill.ca

39

40 **Keywords:**

41 depressive disorders, diagnostic interviews, individual participant data meta-analysis, major depression,

42 classification

43

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

44 **ABSTRACT**

45 **Objective:** To compare odds of major depression classification based on the Structured Clinical Interview
46 for DSM (SCID), the Composite International Diagnostic Interview (CIDI), and the Mini International
47 Neuropsychiatric Interview (MINI).

48 **Methods:** We included and standardized data from three individual participant data meta-analysis (IPDMA)
49 databases, which included primary studies with depressive symptom scores from the Patient Health
50 Questionnaire-9, Edinburgh Postnatal Depression Scale, or Hospital Anxiety and Depression Scale –
51 Depression subscale plus diagnostic interview-based major depression status. For each IPDMA, separately,
52 we fit binomial generalized linear mixed models to compare adjusted odds ratios (aORs) of (1) major
53 depression classification, controlling for depression symptom severity and participant characteristics, and
54 (2) the interaction between interview and symptom severity. Next, we synthesized results using
55 DerSimonian-Laird random-effects meta-analysis.

56 **Results:** In total, 69,405 participants (7,574 [11%] with major depression) from 212 studies were included.
57 Controlling for symptom severity and participant characteristics, the MINI (74 studies; 25,749 participants)
58 classified major depression more often than the SCID (108 studies; 21,953 participants; aOR [95% CI] =
59 1.46 [1.11-1.92]). Classification odds for the CIDI (30 studies; 21,703 participants) and SCID did not differ
60 overall (aOR [95% CI] =1.19 [0.79, 1.75]), but as screening scores increased, aOR increased less for the
61 CIDI than the SCID (interaction aOR [95% CI] = 0.64 [0.52-0.80]).

62 **Conclusions:** Compared to the SCID, the MINI classified major depression more often. Odds of depression
63 classification with the CIDI increased less as symptom levels increased. Interpretation of research that uses
64 diagnostic interviews to classify depression should consider interview characteristics.

65

66 INTRODUCTION

67 In mental health research, diagnostic interviews are used to classify disorders in a manner consistent
68 with standard classification systems and replicable across studies [1-4]. There are important differences,
69 however, in the designs of commonly used interviews. Semi-structured interviews are designed for
70 administration by trained professionals with diagnostic experience; evaluators can interject queries and use
71 their clinical judgment to determine whether symptoms are present and significant [1-3]. The Structured
72 Clinical Interview for DSM (SCID) [4] is the most commonly used semi-structured interview in depression
73 research [5-7]. Fully structured interviews, in contrast, are designed for lay interviewer administration to
74 reduce the cost of clinician-administered interviews. They are completely scripted, and evaluators cannot
75 provide additional explanations or rephrase questions; minimal judgment is involved. They are intended to
76 maximize reliability but may reduce validity [8]. The Composite International Diagnostic Interview (CIDI)
77 [8] is the most commonly used fully structured interview for depression research [5-7]. The Mini
78 International Neuropsychiatric Interview (MINI) [9,10], also common in depression research, is a very brief
79 fully structured interview, originally described by its developers as a screening interview and intended to be
80 over-inclusive [10].

81 Despite their differences, semi-structured interviews, fully structured interviews of conventional
82 length, and abbreviated alternatives such as the MINI are usually treated as equivalent. For instance, meta-
83 analyses of depression screening tool accuracy typically pool primary study results without consideration of
84 reference standards [11-17]. Until recently, however, only several small studies, each with 61 depression
85 cases or fewer, compared classification by different diagnostic interviews [2,18-23]. Recently, three
86 individual participant data meta-analyses (IPDMA) compared odds of major depression classification
87 between different diagnostic interviews, controlling for depression symptom severity scores and participant
88 characteristics [5-7]. Those included an IPDMA with 17,158 participants from 57 primary studies that used
89 the Patient Health Questionnaire-9 (PHQ-9) to control for depression symptom severity [5], 12,759 women

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

90 in pregnancy or postpartum from 46 studies that used the Edinburgh Postnatal Depression Scale (EPDS) [6],
91 and 15,856 participants from 73 studies that used the depression subscale of the Hospital Anxiety and
92 Depression Scale (HADS-D) [7]. Results suggested that, compared to semi-structured interviews (e.g.,
93 SCID) [4], the CIDI may classify more people with relatively low-level symptoms as depressed but fewer
94 people with higher symptom levels. The MINI appeared to classify major depression in more people across
95 the symptom spectrum. There was important imprecision in results, however, including wide confidence
96 intervals (CIs) around estimates.

97 Our objective was to synthesize results from three separate IPDMAs datasets to and compare the most
98 commonly used diagnostic interviews for major depression, the SCID, CIDI, and MINI to determine (1) if
99 odds ratios for major depression classification using the CIDI and MINI differ from the SCID, controlling
100 for depression symptom severity and participant characteristics, and (2) if there is an interaction between the
101 interview and depressive symptom level that would suggest that differences in classification odds are
102 associated with symptom levels.

103 **MATERIALS AND METHODS**

104 We conducted a two-stage evidence synthesis. We first conducted IPDMAs in the PHQ-9, EPDS, and
105 HADS datasets, separately, by fitting models with and without interaction terms for depressive symptom
106 severity in each dataset, separately. Second, we pooled estimates from the results of the three IPDMAs.

107 **Inclusion Criteria for the Included Datasets**

108 For the PHQ-9, EPDS, and HADS-D IPDMAs, datasets from articles in any language were eligible
109 for inclusion if (1) they included diagnostic classification for current Major Depressive Disorder or Major
110 Depressive Episode using Diagnostic and Statistical Manual of Mental Disorders [24-27] or International
111 Classification of Diseases [28] criteria based on a validated semi-structured or fully structured interview; (2)
112 they included PHQ-9, EPDS, or HADS-D scores; (3) the diagnostic interview and depression screening test
113 were administered within two weeks of each other; and (4) participants were ≥ 18 years, not recruited from

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

114 youth or college settings, and not recruited from psychiatric settings or because a screening test identified
115 them as having symptoms of depression [29-31]. For the EPDS, participants were women in pregnancy or
116 within 12 months postpartum [30]. In each IPDMA, datasets where not all participants were eligible were
117 included if primary data allowed selection of eligible participants [29-31]. Over 90% of all included studies
118 in the IPDMA databases used the SCID, CIDI, or MINI diagnostic interviews. Thus, for the present study,
119 as we did in the published IPDMAs of the EDPS [6] and HADS-D [7], we restricted analyses to studies that
120 used SCID, CIDI, or MINI.

121 **Search Strategy, Study Selection, Data Acquisition, and Data Extraction**

122 For more details on the search and selection processes, as well as data contribution, extraction, and
123 synthesis, please see Supplementary Method 1. For information on how the IPDMA datasets and the
124 analyses conducted in the present study deviated from our previous published IPDMAs on diagnostic
125 interview performance using the PHQ-9 [5], EPDS [6], and HADS-D [7] IPDMA databases, please see
126 Supplementary Method 2, Supplementary Method 3, and Supplementary Figure 1.

127 **Statistical Analysis**

128 *IPDMAs of PHQ-9, EPDS, and HADS-D Datasets:*

129 We initially standardized symptom severity scores in each dataset. To do this, for each measure, we
130 converted raw screening tool scores to standardized scores by Z-transformation (subtracting the mean and
131 dividing by the standard deviation of raw scores). We then meta-analyzed the PHQ-9, EPDS, and HADS
132 datasets, separately. In each dataset, we fit binomial generalized linear mixed models with a logit link
133 function to compare the adjusted odds ratio (aOR) of major depression classification for the CIDI versus the
134 SCID, the MINI versus the SCID, and, as a supplementary analysis, the MINI versus the CIDI, controlling
135 for depressive symptom levels and other participant characteristics. We adjusted for different covariates in
136 the models for each dataset, based on relevant measures. For the PHQ-9 and HADS-D datasets, as in the
137 previously published IPDMAs [5,7], we controlled for depressive symptom severity (continuous

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

138 standardized scores), age, sex, country Human Development Index (very high, high, or low-medium) [32],
139 and patient care setting (PHQ-9: primary care, outpatient specialty care, inpatient specialty care, non-
140 medical care [33]; HADS-D: outpatient care, inpatient care, non-medical care, mixed inpatient and
141 outpatient [7]). For the EPDS, we did not control for sex or patient care settings but controlled for
142 pregnancy versus postpartum status [6]. To account for the correlation between subjects within primary
143 studies in each dataset, a random intercept was fit. Fixed slopes were estimated for all covariates in each
144 model. We also fit additional models in each dataset, where we added an interaction term between interview
145 and depressive symptom severity (continuous PHQ-9, EPDS, and HADS-D standardized scores), to
146 evaluate whether any differences in aOR of major depression classification were associated with depression
147 symptom severity.

148 *Synthesis of IPDMA Results:*

149 To synthesize results from the three IPDMAs, we pooled estimates of the aOR for each comparison
150 (CIDI versus SCID, MINI versus SCID, MINI versus CIDI) and the aOR for the interaction of interview
151 and depression symptom severity in each comparison, along with 95% CIs. We used DerSimonian-Laird
152 random effects meta-analysis to pool the aORs [34]. Heterogeneity was examined using the I_2 statistic based
153 on log aORs [35]. Because some studies were included in both the PHQ-9 and HADS-D IPDMAs, as a
154 sensitivity analysis, we re-analyzed results after removing those studies.

155 All analyses were conducted in R (R version R 3.5.1 and R Studio version 1.1.463) [36,37] using the
156 glmer function within the lme4 package [38] and the rma function within the metafor package [39].

157 **RESULTS**

158 In total, 69,405 participants (7,574 [11%] with major depression) were included in the three individual
159 IPDMAs (Table 1). Of the 212 included primary studies, the SCID was used in 108 studies (21,953
160 participants, 14% major depression), the CIDI in 30 studies (21,703 participants, 7% major depression), and
161 the MINI in 74 studies (25,749 participants, 12% major depression). Mean (standard deviation) of raw

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

162 screening tool scores, prior to standardization, were 4.99 (5.26) for the PHQ-9, 6.98 (5.58) for the EPDS,
163 and 5.16 (4.07) for the HADS-D. Characteristics of individual primary studies are available in
164 Supplementary Table 1 with details for PHQ-9 update in Supplementary Method 1. There were 13 studies
165 that were included in both the PHQ-9 and HADS-D datasets, including 2,383 (6%) participants in the PHQ-
166 9 IPDMA and 2,349 participants (15%) in the HADS-D IPDMA. There was no overlap between the EPDS
167 and the PHQ-9 or HADS-D IPDMAs.

168 Estimates of aORs of major depression classification by diagnostic interview, controlling for
169 depressive symptom severity and other participant characteristics, individually and pooled, are reported in
170 Table 2. Overall odds of major depression classification did not differ for the CIDI versus the SCID (aOR
171 1.19, 95% CI = 0.79 to 1.75) in the full model that included the interaction term, but there was a significant
172 interaction between the CIDI and depressive symptom severity; as screening tool scores increased, odds of
173 major depression classification increased less for the CIDI than for the SCID (interaction aOR = 0.64, 95%
174 CI = 0.52 to 0.80). As shown in Figure 1, participants with lower depressive symptom severity were more
175 likely to be classified with major depression with the CIDI compared to the SCID, but the opposite was true
176 with greater symptom severity. Compared to the SCID, the MINI classified major depression more often
177 (aOR 1.45; 95% CI = 1.08 to 1.93), controlling for depressive symptom severity and participant
178 characteristics. There was no apparent interaction between symptom levels and odds of classification
179 (interaction aOR = 0.95, 95% CI = 0.78 to 1.15). See Figure 2.

180 Trends of the probability of major depression classification by reference standards for individual
181 IPDMAs are presented in Supplementary Figures 2-4. There was minimal between-IPDMA heterogeneity in
182 overall aORs for the comparison of the CIDI versus the SCID and the MINI versus the SCID in models
183 without the interaction term ($I_2 = 11\%$ and 0% , respectively) and including the interaction term ($I_2 = 0\%$ and
184 0% , respectively). However, there was substantial between-IPDMA heterogeneity of interaction aORs for
185 both comparisons ($I_2 = 82\%$ and 82%). See Table 2.

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

186 In the comparison of the MINI versus the CIDI, the MINI was more likely to classify participants as
187 having major depression than the CIDI (aOR = 2.05; 95% CI = 1.36 to 2.10), controlling for depressive
188 symptom levels and other participant characteristics. As screening tool scores increased, the odds of major
189 depression classification increased more for the MINI than for the CIDI (interaction aOR = 1.48, 95% CI =
190 1.36 to 1.60). Heterogeneity was low for aORs with and without the interaction term, and interaction aORs
191 ($I_2 = 0\%$, 0% , and 0%).

192 In the individual IPDMAs, some results from the EPDS dataset appeared to diverge from those
193 generated in the PHQ-9 and HADS-D datasets. However, the number of studies and cases included in the
194 EPDS dataset for the CIDI and MINI were smaller than any other combination of screening tool and
195 diagnostic interview. See Table 1.

196 As a sensitivity analysis, we removed the 13 datasets that were included in both the PHQ-9 and
197 HADS-D IPDMAs and re-ran all analyses. Results were similar (see Supplementary Table 2).

198 **DISCUSSION**

199 There were two main findings. First, overall odds of major depression classification did not differ
200 between the fully structured CIDI and the semi-structured SCID. However, adjusting for depressive
201 symptom levels and participant characteristics, odds of major depression classification with the CIDI
202 increased significantly less than for the SCID as depressive symptom levels increased. This suggests that,
203 compared to the SCID, the CIDI is relatively more likely to classify individuals with subthreshold or mild
204 depressive symptoms and relatively less likely to classify people with more severe symptoms. Second,
205 participants evaluated with the MINI were significantly more likely to be classified as having major
206 depression compared to those assessed with the SCID, independent of symptom severity. Between-study
207 heterogeneity was low for models without the interaction term, but higher for models with interaction terms.
208 Estimates from the EPDS IPDMA appeared to diverge somewhat from the PHQ-9 and HADS-D IPDMAs.

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

209 This may have been related to the small numbers of studies and major depression cases for the CIDI and
210 MINI among studies that used the EPDS.

211 Our findings appear to be consistent with characteristics of the different types of diagnostic
212 interviews. The MINI was designed as a screening interview and described by its developers as over-
213 inclusive in classifying psychiatric disorders [10]. For the CIDI, the lack of sensitivity to different levels of
214 depressive symptoms severity may be because the CIDI assesses symptoms in the last 12 months and over
215 the lifetime, then probes to determine if those symptoms are currently present using only a single question.
216 In contrast, the SCID and the MINI specifically assess symptoms in the past two weeks. In addition, the
217 CIDI is much more complicated than the MINI or the SCID. It includes complex branches and is scored
218 using algorithms subject to calibration, which may influence how well diagnoses map onto DSM criteria.
219 This could lead to error at all symptom levels, which would result in more people classified at lower
220 symptom severity levels and fewer at higher levels.

221 Results were generally consistent with limited evidence from small studies that previously directly
222 compared depression classification by administering semi- and fully structured diagnostic interviews to the
223 same participants. In two studies that examined general population samples with low prevalence, fully
224 structured interviews classified major depression substantially more frequently than semi-structured
225 interviews [2,20]. On the other hand, in a study of participants in inpatient alcohol treatment, where
226 symptom severity would be expected to be higher, depression classification likelihood was similar with
227 semi-structured and fully structured interviews [22].

228 Our findings have important implications for research, including clinical trials, prognostic and risk
229 factor studies, diagnostic accuracy studies, and prevalence studies. Concerns have been raised about the
230 degree to which antidepressant trials are generalizable to real-world clinical practice [40]. Based on our
231 findings, the method used to classify depression status is also an important consideration. If used to
232 determine trial eligibility, the CIDI may not identify some participants who would be eligible based on the

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

233 SCID, whereas both CIDI and MINI may include some participants who would not be eligible based on the
234 SCID, which could reduce the ability to detect treatment effects and further limit applicability to
235 participants in practice who meet diagnostic criteria. Differences in classifying participants could similarly
236 reduce the ability to identify potential associations between risk factors and depression. In diagnostic test
237 accuracy studies, depression screening tool accuracy has been shown to differ across reference standards
238 [33,41,42]. In studies of major depression prevalence, the MINI will overestimate compared to the SCID,
239 whereas with the CIDI, relative prevalence will depend on the underlying distribution of depressive
240 symptoms.

241 Our findings, which are contrary to the common belief that different reference standards can be
242 treated equivalently in mental health research, provide evidence that different approaches are needed [43].
243 Ideally, researchers would use semi-structured interviews, such as the SCID, which are designed to replicate
244 diagnostic procedures as closely as possible, to establish diagnostic status. However, this is not always
245 feasible due to the resources required, including highly trained staff. Future studies are needed to develop
246 models to calibrate weights of major depression classification based on different reference standards that
247 could facilitate synthesis of results using different diagnostic interviews. Meanwhile, in selecting a
248 diagnostic interview for use in research, investigators should consider advantages and disadvantages of
249 different interviews, including performance characteristics and resources required. In published studies,
250 authors should comment on potential implications of the type of diagnostic interview that was used. Users
251 of research, including clinicians, should be aware that results from studies that use the CIDI or MINI may
252 differ from what would be found using semi-structured interviews, which are designed to replicate
253 diagnostic procedures as closely as possible. It is also important to underline that from a clinimetric
254 perspective [44-46], assessment of diagnostic status alone is not sufficient, but that rating tools and self-
255 report questionnaires are needed to characterize symptom severity and the specific nature of experienced
256 symptoms.

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

257 A strength of the present study was the inclusion of 69,405 participants with 7,574 (11%) major
258 depression cases from 212 studies. This allowed us to overcome limitations of previous IPDMAs and
259 generate more precise estimates. A second strength was that data within each included dataset were
260 standardized in terms of definitions of major depression classification, eligibility criteria, and variables. A
261 limitation to consider is that for included IPDMAs, we could not obtain primary data for 28 of 117 eligible
262 PHQ-9 studies (24% of eligible studies, 17% of eligible participants), 19 of 64 EPDS studies (30% of
263 eligible studies, 30% of eligible participants), and 47 of 116 HADS-D studies (41% of eligible studies, 29%
264 of eligible participants). A second is that we used standardized scores instead of raw depression symptom
265 scores, which required making the assumption that a standard deviation change in scores was equivalent
266 across different screening tools. Third, because only three estimates were pooled, our ability to estimate
267 heterogeneity and explore possible causes was limited. Fourth, some studies were included in both the
268 PHQ-9 and HADS-D IPDMAs. However, a sensitivity analysis showed that results were similar when these
269 studies were removed. Fifth, we examined the SCID, CIDI, and MINI, because we did not have access to
270 enough studies to include other diagnostic interviews. It is unclear to what degree our findings would
271 generalize to other diagnostic interviews. Finally, our study did not include a head-to-head comparison of
272 interviews from a randomized controlled trial or by administering different interviews to all participants. It
273 is unlikely, however, that such a study would be feasible with a large enough sample to draw conclusions
274 with confidence. Our study design, despite its limitations, overcame this barrier.

275 To conclude, the semi-structured SCID was designed to replicate diagnostic standards and procedures
276 as closely as possible. By synthesizing results from three large IPDMAs, we found that the most commonly
277 used fully structured diagnostic interviews to classify major depression, the CIDI and MINI, did not
278 perform equivalently to the SCID. The CIDI is not as responsive as the SCID to different levels of reported
279 depressive symptoms, and the MINI identifies more cases across the spectrum of depressive symptom
280 levels. Researchers should carefully consider the advantages and disadvantages of using these diagnostic

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

281 interviews, and findings from studies based on the CIDI or the MINI should be interpreted considering how
282 their performance deviates from that of the SCID.

283

284 **Acknowledgements:** Not applicable

285

286 **Author Contributions:** YW, BL, JPAI, AB, and BDT were responsible for the study conception and
287 design. BDT contributed a primary dataset that was included in this study. YW, BL, and BDT contributed to
288 data extraction and coding for the meta-analysis. YW, BL, AB, JPAI, and BDT contributed to data analysis
289 and interpretation. YW, AB and BDT contributed to drafting the manuscript. All authors provided a critical
290 review and approved the final manuscript. AB and BDT are the guarantors; they had full access to all the
291 data in the study and take responsibility for the integrity of the data and the accuracy of the data analyses.

292

293 **Group Author Contributions:** The DEPRESSD Collaboration includes collaborators who contributed:

294 To data extraction, coding, and synthesis: Ying Sun, Lady Davis Institute for Medical Research,
295 Jewish General Hospital, Montréal, Québec, Canada; Chen He, Lady Davis Institute for Medical Research,
296 Jewish General Hospital, Montréal, Québec, Canada; Ankur Krishnan, Lady Davis Institute for Medical
297 Research, Jewish General Hospital, Montréal, Québec, Canada; Parash Mani Bhandari, Lady Davis Institute
298 for Medical Research, Jewish General Hospital, Montréal, Québec, Canada; Dipika Neupane, Lady Davis
299 Institute for Medical Research, Jewish General Hospital, Montréal, Québec, Canada; Zelalem Negeri, Lady
300 Davis Institute for Medical Research, Jewish General Hospital, Montréal, Québec, Canada; Mahrukh Imran,
301 Lady Davis Institute for Medical Research, Jewish General Hospital, Montréal, Québec, Canada; Danielle
302 B. Rice, Lady Davis Institute for Medical Research, Jewish General Hospital, Montréal, Québec, Canada;
303 Kira E. Riehm, Lady Davis Institute for Medical Research, Jewish General Hospital, Montréal, Québec,
304 Canada; Nazanin Saadat, Lady Davis Institute for Medical Research, Jewish General Hospital, Montréal,
305 Québec, Canada; Marleine Azar, Lady Davis Institute for Medical Research, Jewish General Hospital,
306 Montréal, Québec, Canada; Alexander W. Levis, Lady Davis Institute for Medical Research, Jewish
307 General Hospital, Montréal, Québec, Canada; Tatiana A. Sanchez, Lady Davis Institute for Medical

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

308 Research, Jewish General Hospital, Montréal, Québec, Canada; Matthew J. Chiovitti, Lady Davis Institute
309 for Medical Research, Jewish General Hospital, Montréal, Québec, Canada; Xin Wei Yan, Lady Davis
310 Institute for Medical Research, Jewish General Hospital, Montréal, Québec, Canada.

311 Via the design and conduct of database searches: Jill Boruff, McGill University, Montréal, Québec,
312 Canada; Lorie A. Kloda, Concordia University, Montréal, Québec, Canada.

313 As members of the DEPRESSD Steering Committee, including conception and oversight of
314 collaboration: Pim Cuijpers, Vrije Universiteit, Amsterdam, the Netherlands; Simon Gilbody, University of
315 York, Heslington, York, UK; Dean McMillan, University of York, Heslington, York, UK; Scott B. Patten,
316 University of Calgary, Calgary, Alberta, Canada; Ian Shrier, McGill University, Montréal, Québec, Canada;
317 Roy C. Ziegelstein, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

318 As knowledge user consultants: Liane Comeau, International Union for Health Promotion and
319 Health Education, École de santé publique de l'Université de Montréal, Montréal, Québec, Canada;
320 Nicholas D. Mitchell, Department of Psychiatry, University of Alberta, Edmonton, Alberta, Canada;
321 Marcello Tonelli, Department of Medicine, University of Calgary, Calgary, Alberta, Canada; Simone N.
322 Vigod, Women's College Hospital and Research Institute, University of Toronto, Toronto, Ontario, Canada;
323 Melissa Henry, Lady Davis Institute for Medical Research, Jewish General Hospital, Montréal, Québec,
324 Canada; Zahinoor Ismail, Hotchkiss Brain Institute & O'Brien Institute for Public Health, Calgary, Alberta,
325 Canada; Carmen G. Loiselle, Lady Davis Institute for Medical Research, Jewish General Hospital,
326 Montréal, Québec, Canada.

327 By contributing included datasets: Dickens H. Akena, Department of Psychiatry, Makerere
328 University College of Health Sciences, Kampala, Uganda; Samir Al-Adawi, Department of Behavioural
329 Medicine, College of Medicine & Health Sciences, Sultan Qaboos University, Muscat, Oman; Sultan H.
330 Alamri, King Abdulaziz University, Abdullah Sulayman, Jeddah, Makkah, Saudi Arabia; Rubén Alvarado,
331 School of Public Health, Faculty of Medicine, Universidad de Chile, Santiago, Chile; Cosme Alvarado-

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

332 Esquivel, Laboratorio de Investigación Biomédica, Facultad de Medicina y Nutrición, Avenida Universidad,
333 Dgo, Mexico; Dagmar Amtmann, Department of Rehabilitation Medicine, University of Washington,
334 Seattle, Washington, USA; Bruce Arroll, Department of General Practice and Primary Health Care,
335 University of Auckland, New Zealand; Liat Ayalon, Louis and Gabi Weisfeld School of Social Work, Bar
336 Ilan University, Ramat Gan, Israel; Muideen O. Bakare, Child and Adolescent Unit, Federal
337 Neuropsychiatric Hospital, Enugu, Nigeria; Hamid R. Baradaran, Endocrine Research Center, Institute of
338 Endocrinology and Metabolism, Iran University of Medical Sciences, Tehran, Iran; Jacqueline Barnes,
339 Department of Psychological Sciences, Birkbeck, University of London, UK; Amar D. Bavle, Department
340 of Psychiatry, Rajarajeswari Medical College and Hospital, Bengaluru, Karnataka, India; Cheryl Tatano
341 Beck, University of Connecticut School of Nursing, Mansfield, Connecticut, USA; Anna Beraldi, Kbo-
342 Lech-Mangfall-Klinik Garmisch-Partenkirchen, Klinik für Psychiatrie, Psychotherapie & Psychosomatik,
343 Lehrkrankenhaus der Technischen Universität München, Munich, Germany; Charles N. Bernstein ,
344 University of Manitoba IBD Clinical and Research Centre, Winnipeg, Manitoba, Canada; Arvin Bhana,
345 Centre for Rural Health, School of Nursing and Public Health, College of Health Sciences, University of
346 KwaZulu-Natal, Durban, South Africa; Carola Bindt, Department of Child and Adolescent Psychiatry,
347 University Medical Center Hamburg-Eppendorf, Germany; Charles H. Bombardier, Department of
348 Rehabilitation Medicine, University of Washington, Seattle, Washington, USA; Philip M. Boyce, Discipline
349 of Psychiatry, Westmead Clinical School, Sydney Medical School, University of Sydney, Sydney,
350 Australia; Natalie Büel-Drabe, Department of Psychiatry and Psychotherapy, University Hospital Zürich,
351 Zürich Switzerland; Ryna Imma Buji, Department of Psychiatry, Hospital Mesra Bukit Padang, Sabah,
352 Malaysia; Adomas Bunevicius, Neuroscience Institute, Lithuanian University of Health Sciences,
353 Kaunas ,Lithuania; Dr. Jurate Butnorienė, PhD, who did the data collection and analysis as part of her PhD
354 thesis for the primary study by Butnorienė et al., passed away and was unable to participate in this project;
355 Dr. Robertas Bunevicius, MD, PhD (1958-2016) was the Principal Investigator of the primary studies by

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

356 Butnoriene et al. and Bunevicius et al, but passed away and was unable to participate in this project; Peter
357 Butterworth, Centre for Research on Ageing, Health and Wellbeing, Research School of Population Health,
358 The Australian National University, Canberra, Australia; Gregory Carter, Centre for Brain and Mental
359 Health Research, University of Newcastle, New South Wales, Australia; Marcos H. Chagas, Department of
360 Neurosciences and Behavior, Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto,
361 Brazil; Juliana C. N. Chan, Department of Medicine and Therapeutics, Prince of Wales Hospital, The
362 Chinese University of Hong Kong, Hong Kong Special Administrative Region, China; Lai Fong Chan,
363 Department of Psychiatry, National University of Malaysia, Kuala Lumpur, Malaysia; Linda H. Chaudron,
364 Departments of Psychiatry, Pediatrics, Obstetrics and Gynecology, School of Medicine and Dentistry,
365 University of Rochester, Rochester, NY, USA; Chih-Ken Chen, Community Medicine Research Center,
366 Keelung Chang Gung Memorial Hospital and Chang Gung University College of Medicine, Keelung,
367 Taiwan; Rushina Cholera, Department of Pediatrics, Duke University, Durham, North Carolina, USA;
368 Kerrie Clover, Centre for Brain and Mental Health Research, University of Newcastle, New South Wales,
369 Australia; Ronán M. Conroy, Royal College of Surgeons in Ireland Division of Population Health Sciences,
370 Dublin, Ireland; Aaron Conway, Lawrence S. Bloomberg Faculty of Nursing, University of Toronto,
371 Toronto, Canada; Yeates Conwell, Department of Psychiatry, University of Rochester Medical Center,
372 Rochester, New York, USA; Humberto Correa, Faculty of Medicine, Universidade Federal de Minas
373 Gerais. Belo Horizonte, MG, Brazil; Tiago Castro e Couto, Federal University of Uberlândia, Brazil; Daniel
374 Cukor, Rogosin Institute, New York, New York, USA; Eli Dabscheck, The Alfred Hospital, Prahran, VIC,
375 Australia; Federico M. Daray, Institute of Pharmacology, School of Medicine, University of Buenos Aires,
376 Argentina; Felipe Pinheiro de Figueiredo, Department of Neurosciences and Behavior, Ribeirão Preto
377 Medical School, Brazil; Janneke M. de Man-van Ginkel, Julius Center for Health Sciences and Primary
378 Care, Department of Nursing Science, University Medical Center Utrecht – University Utrecht, Utrecht, the
379 Netherlands; Crisanto Diez-Quevedo, Servei de Psiquiatria, Hospital Germans Trias i Pujol, Badalona,

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

380 Spain; Elles Douven, Alzheimer Center Limburg and School for Mental Health and Neuroscience
381 (MHeNs), Department of Psychiatry and Neuropsychology, Maastricht University, Maastricht, the
382 Netherlands; Marina G. Downing, School of Psychological Sciences, Monash University, Melbourne VIC,
383 Australia; Valsamma Eapen, School of Psychiatry, University of New South Wales, Kensington, Australia;
384 Jesse R. Fann, Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle,
385 Washington, USA; Anthony Feinstein, University of Toronto, Toronto, Ontario, Canada; Panagiotis P.
386 Ferentinos, National and Kapodistrian University of Athens, 2nd Department of Psychiatry, Attikon General
387 Hospital, Athens, Greece; Michelle Fernandes, Faculty of Medicine, Department of Paediatrics, University
388 of Southampton, Southampton and Nuffield Department of Women's & Reproductive Health, University of
389 Oxford, UK; Sally Field, Perinatal Mental Health Project, Alan J Flisher Centre for Public Mental Health,
390 Department of Psychiatry and Mental Health, University of Cape Town, Cape Town, South Africa; Barbara
391 Figueiredo, School of Psychology, University of Minho, Portugal; Felix H. Fischer, Department of
392 Psychosomatic Medicine, Center for Internal Medicine and Dermatology, Charité - Universitätsmedizin
393 Berlin, Germany; Jane R. W. Fisher, School of Public Health and Preventive Medicine, Monash University,
394 Melbourne, Australia; Alastair J. Flint, University Health Network, Toronto, Ontario, Canada; Maiko
395 Fujimori, Section of Psychological Science, Division of Health Care Research, Center for Public Health
396 Sciences, National Cancer Center, Tokyo, Japan; Daniel S. S. Fung, Department of Developmental
397 Psychiatry, Institute of Mental Health, Singapore; Pamela Gallagher, School of Psychology, Dublin City
398 University, Dublin, Ireland; Milena Gandy, The Department of Psychology, Macquarie University, Sydney,
399 Australia; Lluïsa Garcia-Esteve, Perinatal Mental Health Unit CLINIC-BCN. Institut Clínic de
400 Neurociències, Hospital Clínic, Barcelona, Spain; Emily C. Garman, Alan J Flisher Centre for Public
401 Mental Health, Department of Psychiatry and Mental Health, University of Cape Town; Bizu Gelaye,
402 Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, Massachusetts, USA;
403 Leila Gholizadeh, Faculty of Health, University of Technology Sydney, Sydney, Australia; Lisa Giardinelli,

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

404 Psychiatry Unit, Department of Health Sciences, University of Florence, Firenze, Italy; Lorna J. Gibson,
405 Tropical Epidemiology Group, Faculty of Epidemiology and Population Health, London School of Hygiene
406 and Tropical Medicine, London, UK; Felicity Goodyear-Smith, Department of General Practice and
407 Primary Health Care, University of Auckland, New Zealand; Luigi Grassi, Institute of Psychiatry,
408 Department of Biomedical and Specialty Surgical Sciences, University of Ferrara, Ferrara, Italy; Eric P.
409 Green, Duke Global Health Institute, Duke University, Durham, North Carolina, USA; Catherine G.
410 Greeno, School of Social Work, University of Pittsburgh, Pittsburgh, Pennsylvania, USA; Brian J. Hall,
411 Global and Community Mental Health Research Group, Department of Psychology, Faculty of Social
412 Sciences, University of Macau, Macau Special Administrative Region, China; Liisa Hantsoo, Department of
413 Psychiatry & Behavioral Sciences , The Johns Hopkins University School of Medicine, Baltimore,
414 Maryland; Emily E. Haroz, Center For American Indian Health, Department of International Health, Johns
415 Hopkins Bloomberg School of Public Health; Martin Härter, Department of Medical Psychology,
416 University of Hamburg, Hamburg, Germany; Ulrich Hegerl, Department of Psychiatry, Psychosomatics and
417 Psychotherapy, Goethe-Universität Frankfurt, Germany; Nadine Helle, Department of Child and Adolescent
418 Psychiatry, University Medical Center Hamburg-Eppendorf, Germany; Leanne Hides, School of
419 Psychology, University of Queensland, Brisbane, Queensland, Australia; Stevan E. Hobfoll, STAR-Stress,
420 Anxiety and Resilience Consultants, Chicago, Illinois, USA; Simone Honikman, Perinatal Mental Health
421 Project, Alan J Flisher Centre for Public Mental Health, Department of Psychiatry and Mental Health,
422 University of Cape Town; Louise M. Howard, Institute of Psychiatry, Psychology & Neuroscience, King's
423 College London, London, UK; Marie Hudson, Lady Davis Institute for Medical Research, Jewish General
424 Hospital, Montréal, Québec, Canada; Thomas Hyphantis, Department of Psychiatry, Faculty of Medicine,
425 School of Health Sciences, University of Ioannina, Greece; Masatoshi Inagaki, Department of Psychiatry,
426 Faculty of Medicine, Shimane University, Shimane, Japan; Josef Jenewein, Clinic Zugersee, Center for
427 Psychiatry and Psychotherapie, Oberwil-Zug, Switzerland; Hong Jin Jeon, Department of Psychiatry,

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

428 Depression Center, Samsung Medical Center, Sungkyunkwan University School of Medicine, Suwon,
429 South Korea; Nathalie Jetté, Departments of Neurology and Population Health Science and Policy, Icahn
430 School of Medicine at Mount Sinai, New York, New York, USA; Monika Keller, Division of Psycho-
431 oncology, Department of General Internal Medicine and Psychosomatics, University Hospital Heidelberg,
432 Germany; Dina Sami Khalifa, Faculty of Health Sciences, Ahfad University for Women, Omdurman,
433 Sudan; Mohammad E. Khamseh, Endocrine Research Center, Institute of Endocrinology and Metabolism,
434 Iran University of Medical Sciences, Tehran, Iran; Kim M. Kiely, School of Psychology, The University of
435 New South Wales, and Neuroscience Research Australia (NeuRA), Sydney, Australia; Sung-Wan Kim,
436 Department of Psychiatry, Chonnam National University Medical School, Republic of Korea; Marie
437 Kjærgaard, Endocrinology Research Group, Medical Clinic, University Hospital of North Norway, Norway;
438 Sebastian Köhler, Department of Psychiatry and Neuropsychology, School for Mental Health and
439 Neuroscience, Maastricht University, Maastricht, The Netherlands; Jane Kohlhoff, School of Psychiatry,
440 University of New South Wales, Kensington, Australia; Brandon A. Kohrt, Department of Psychiatry and
441 Behavioral Sciences, The George Washington University, Washington, DC, USA; Zoltán Kozinszky,
442 Department of Obstetrics and Gynecology, Danderyd Hospital, Stockholm, Sweden; Laima Kusminskas,
443 Private Practice, Hamburg, Germany; Yunxin Kwan, Department of Psychological Medicine, Tan Tock
444 Seng Hospital, Singapore; Femke Lamers, Department of Psychiatry, Amsterdam Public Health Research
445 Institute, Amsterdam UMC, Amsterdam, the Netherlands; Maria Asunción Lara, Instituto Nacional de
446 Psiquiatría Ramón de la Fuente Muñiz. San Lorenzo Huipulco, Tlalpan, México D. F. Mexico; Lorenzo
447 Lelli, Psychiatry Unit, Department of Health Sciences, University of Florence, Firenze, Italy; Angeliki A.
448 Leonardou, First Department of Psychiatry, Women's Mental Health Clinic, Athens University Medical
449 School, Athens, Greece; Holly F. Levin-Aspenson, Department of Psychology, University of Notre Dame,
450 Notre Dame, Indiana, USA; Manote Lotrakul, Department of Psychiatry, Faculty of Medicine, Ramathibodi
451 Hospital, Mahidol University, Bangkok, Thailand; Sonia R. Loureiro, Department of Neurosciences and

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

452 Behavior, Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, Brazil; Bernd Löwe,
453 Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-
454 Eppendorf, Hamburg, Germany; Nagendra P. Luitel, Research Department, TPO Nepal, Kathmandu, Nepal;
455 Crick Lund , Alan J Flisher Centre for Public Mental Health, Department of Psychiatry and Mental Health,
456 University of Cape Town; Michael Maes, Department of Psychiatry, Faculty of Medicine, Chulalongkorn
457 University, Bangkok, Thailand; Ruth Ann Marrie, Departments of Medicine and Community Health
458 Sciences, Max Rady College of Medicine, Rady Faculty of Health Sciences, University of Manitoba,
459 Winnipeg, Manitoba, Canada; Laura Marsh, Baylor College of Medicine, Houston and Michael E. DeBakey
460 Veterans Affairs Medical Center, Houston, Texas, USA; Rocio Martin-Santos, Department of Psychiatry
461 and Psychology, Hospital Clinic, IDIBAPS, CIBERSAM, Barcelona, Spain; Brian P. Marx, National Center
462 for PTSD at VA Boston Healthcare System, Boston, MA, USA; Loreto Massardo, Centro de Biología
463 Celular y Biomedicina, Facultad de Medicina y Ciencia, Universidad San Sebastián. Santiago, Chile;
464 Yutaka Matsuoka, Division of Health Care Research, Center for Public Health Sciences, National Cancer
465 Center, Tokyo, Japan; Anja Mehner, Department of Medical Psychology and Medical Sociology, University
466 of Leipzig, Germany; Valentina Meuti, Department of Neurology and Psychiatry, Sapienza University of
467 Rome, Rome, Italy; Ioannis Michopoulos, 2nd Department of Psychiatry, Attikon General Hospital,
468 National and Kapodistrian University of Athens, Athens, Greece; Laurent Misery, Department of
469 Dermatology, University Hospital of Brest, Brest, France; Sherina Mohd Sidik, Cancer Resource &
470 Education Centre, and Department of Psychiatry, Faculty of Medicine and Health Sciences, Universiti Putra
471 Malaysia, Serdang, Selangor, Malaysia; Tiago N. Munhoz, Post-graduate Program in Epidemiology,
472 Federal University of Pelotas, Pelotas, RS, Brazil; Kumiko Muramatsu, Department of Clinical Psychology,
473 Graduate School of Niigata Seiryō University, Niigata, Japan; Sandra Nakić Radoš, Department of
474 Psychology, Catholic University of Croatia, Zagreb, Croatia; Juliet E. M. Nakku, Butabika National
475 Referral Teaching Hospital; Laura Navarrete, Department of Epidemiology and Psychosocial Research,

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

476 Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz, Ciudad de México, México; Purificación
477 Navarro García, Perinatal Mental Health Unit CLINIC-BCN. Institut Clínic de Neurociències, Hospital
478 Clínic, Barcelona, Spain; Ricard Navines, Department of Psychiatry and Psychology, Hospital Clinic,
479 IDIBAPS, CIBERSAM, Barcelona, Spain; Daisuke Nishi, Department of Mental Health, Graduate School
480 of Medicine, The University of Tokyo, Japan; Meaghan L. O'Donnell, Phoenix Australia, Carlton VIC,
481 Australia; Daniel Okitundu Luwa E-Andjafono, Unité de Neuropsychologie, Département de Neurologie,
482 Centre Neuro-psycho-pathologique, Faculté de Médecine, Université de Kinshasa, République
483 Démocratique du Congo; Flávia L. Osório, Department of Neurosciences and Behavior, Ribeirão Preto
484 Medical School, University of São Paulo, Ribeirão Preto, Brazil; Ahmet Öztürk, Bezmialem Vakif
485 University, Istanbul, Turkey; Jurate Peceliuniene, Vilnius University Faculty of Medicine, Clinic of Internal
486 Diseases, Family Medicine and Oncology, Vilnius, Lithuania; Brian W. Pence, Department of
487 Epidemiology, Gillings School of Global Public Health, The University of North Carolina at Chapel Hill,
488 Chapel Hill, North Carolina, USA; Philippe Persoons, Department of Adult Psychiatry, University Hospitals
489 Leuven, Leuven, Belgium; Angelo Picardi, Centre for Behavioural Sciences and Mental Health, Italian
490 National Institute of Health, Rome, Italy; Luis Pintor, Consultation Liaison Psychiatry Unit. Hospital
491 Clínic de Barcelona, Barcelona, Spain; Jennie L. Ponsford, School of Psychological Sciences, Monash
492 University, Melbourne VIC, Australia; Stephanie L. Pugh, NRG Oncology Statistics and Data Management
493 Center, Philadelphia, PA, USA; Terence J. Quinn, Institute of Cardiovascular & Medical Sciences,
494 University of Glasgow, Glasgow, Scotland; Elmars Rancans, Department of Psychiatry and Narcology,
495 Riga Stradins University, Latvia; Sujit D. Rathod, Department of Population Health, London School of
496 Hygiene and Tropical Medicine, London, United Kingdom; Silje E. Reme, Department of psychology,
497 Faculty of Social Sciences, University of Oslo, Oslo, Norway; Katrin Reuter, Group Practice for
498 Psychotherapy and Psycho-oncology, Freiburg, Germany; Emma Robertson-Blackmore, Halifax Health,
499 Graduate Medical Education, Daytona Beach, FL. USA; Tamsen J. Rochat, Developmental Pathways to

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

500 Health Research Unit, Faculty of Health Sciences, University of Witwatersrand, South Africa; Alasdair G.
501 Rooney, Division of Psychiatry, Royal Edinburgh Hospital, University of Edinburgh, Edinburgh, Scotland;
502 Heather J. Rowe, School of Public Health and Preventive Medicine, Monash University, Melbourne,
503 Australia; Roberto Sánchez-González, Department of Psychiatry. Institut de Neuropsiquiatria i Addiccions,
504 Centre Emili Mira, Parc de Salut Mar. Barcelona. Spain; Iná S. Santos, Post-graduate Program in
505 Epidemiology, Federal University of Pelotas, Pelotas, RS, Brazil; Miranda T. Schram, Department of
506 Internal Medicine, Maastricht University Medical Center, Maastricht, The Netherlands; Marcelo L.
507 Schwarzbald, Department of Internal Medicine, Federal University of Santa Catarina, Florianópolis, Santa
508 Catarina, Brazil; Vesile Senturk Cankorur, Ankara University Faculty of Medicine Psychiatry Department,
509 Ankara, Turkey; Juwita Shaaban, Department of Family Medicine, School of Medical Sciences, Universiti
510 Sains Malaysia, Kelantan, Malaysia; Louise Sharpe, School of Psychology, The University of Sydney,
511 Sydney NSW, Australia; Eileen H. Shinn, Department of Behavioral Science, University of Texas M. D.
512 Anderson Cancer Center, Houston, Texas, USA; Abbey Sidebottom, Allina Health, Minneapolis,
513 Minnesota, USA; Sébastien Simard, Département des sciences de la santé, Université du Québec à
514 Chicoutimi (UQAC), Québec, Canada; Adam Simning, Department of Psychiatry, University of Rochester
515 Medical Center, Rochester, New York, USA; Susanne Singer, University Medical Centre Mainz, Institute of
516 Medical Biostatistics, Epidemiology and Informatics, Mainz, Germany; Bonnie W. M. Siu, Department of
517 Psychiatry, Castle Peak Hospital, Hong Kong SAR, China; Alkistis Skalkidou, Department of Women's
518 and Children's Health, Uppsala University, Uppsala, Sweden; Lena Spangenberg, Department of Medical
519 Psychology and Medical Sociology, University of Leipzig, Germany; Lesley Stafford, Centre for Women's
520 Mental Health, Royal Women's Hospital, Parkville, Australia; Alan Stein, Department of Psychiatry,
521 University of Oxford, Oxford, UK; Robert C. Stewart, Division of Psychiatry, University of Edinburgh,
522 Edinburgh, UK; Jon Stone, University of Edinburgh, Edinburgh, UK; Kuan-Pin Su, Tainan Municipal An-
523 Nan Hospital & College of Medicine, China Medical University, Taichung, Taiwan; Serge Sultan,

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

524 Université de Montréal, Québec, Canada; Inger Sundström-Poromaa, Department of Women's and
525 Children's Health, Uppsala University, Uppsala, Sweden; Sharon C. Sung, Department of Child &
526 Adolescent Psychiatry, Institute of Mental Health, Singapore; Keiko Suzuki, Department of General
527 Medicine, Asahikawa University Hospital, Asahikawa, Hokkaido, Japan; Meri Tadinac, Department of
528 Psychology, Faculty of Humanities and Social Sciences, University of Zagreb, Croatia; Pei Lin Lynnette
529 Tan, Department of Psychological Medicine, Tan Tock Seng Hospital, Singapore; S. Darius Tandon,
530 Feinberg School of Medicine, Northwestern University, Chicago, IL, USA; Martin Taylor-Rowan, Institute
531 of Cardiovascular and Medical Science, University of Glasgow, Glasgow, Scotland; Antonio L. Teixeira,
532 University of Texas Health Science Center at Houston, Houston, Texas, USA; Iva Tendais, School of
533 Psychology, University of Minho, Portugal; Pavaani Thiagayson, Institute of Mental Health, Singapore;
534 Istvan Tiringier, Pécs University, Medical School, Institute of Behavioral Sciences, Pécs, Hungary;
535 Annamária Töreki, Department of Emergency, University of Szeged, Hungary; Anna Torres-Giménez,
536 Perinatal Mental Health Unit CLINIC-BCN. Institut Clínic de Neurociències, Hospital Clínic, Barcelona,
537 Spain; Thach D. Tran, School of Public Health and Preventive Medicine, Monash University, Melbourne,
538 Australia; Kylee Trevillion, Institute of Psychiatry, Psychology & Neuroscience, King's College London,
539 London, UK; Ka-Yee Tung, Kwai Chung Hospital, Hong Kong SAR, China; Alyna Turner, School of
540 Medicine and Public Health, University of Newcastle, New South Wales, Newcastle, Australia; Katherine
541 Turner, Epilepsy Center-Child Neuropsychiatry Unit, ASST Santi Paolo Carlo, San Paolo Hospital, Milan,
542 Italy; Christina M. van der Feltz-Cornelis, Department of Health Sciences, HYMS, University of York,
543 York, UK; Thandi van Heyningen, Division of Epidemiology & Biostatistics, School of Public Health &
544 Family Medicine, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa; Henk
545 C. van Weert, Department General Practice, Institute Public Health, Amsterdam Universities Medical
546 Centers, Amsterdam, the Netherlands; Johann M. Vega-Dienstmaier, Facultad de Medicina Alberto
547 Hurtado, Universidad Peruana Cayetano Heredia. Lima, Perú; Paul A. Vöhringer, Department of Psychiatry

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

548 and Mental Health, Clinical Hospital, Universidad de Chile, Santiago, Chile; Lynne I. Wagner, Department
549 of Social Sciences and Health Policy, Wake Forest School of Medicine, Wake Forest University, Winston-
550 Salem, North Carolina, USA; Mark Walterfang, Neuropsychiatry Unit, Royal Melbourne Hospital,
551 Melbourne, Australia; Jian Li Wang, University of Ottawa Institute of Mental Health Research, Ottawa,
552 Canada; Wenzheng Wang, Shanghai Mental Health Center, Shanghai Jiao Tong University School of
553 Medicine, Shanghai, China; Liang-Jen Wang, Department of Child and Adolescent Psychiatry, Kaohsiung
554 Chang Gung Memorial Hospital and Chang Gung University College of Medicine, Kaohsiung, Taiwan;
555 Jennifer White, Department of Physiotherapy, School of Primary and Allied Health Care, Monash
556 University, Melbourne, Australia; Dana K. Wong, School of Psychology & Public Health, La Trobe
557 University, Melbourne, Australia; Karen Wynter, School of Nursing and Midwifery, Deakin University,
558 Melbourne, Australia; Mitsuhiko Yamada, Department of Neuropsychopharmacology, National Institute of
559 Mental Health, National Center of Neurology and Psychiatry, Ogawa-Higashi, Kodaira, Tokyo, Japan;
560 Kimberly A. Yonkers, Department of Psychiatry, Yale School of Medicine, New Haven, Connecticut, USA;
561 Qing Zhi Zeng, Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine,
562 Shanghai, China; Yuying Zhang, Department of Medicine and Therapeutics, Prince of Wales Hospital, The
563 Chinese University of Hong Kong, Hong Kong Special Administrative Region, China.

564

565 **Funding Sources:** This study was funded by the Canadian Institutes of Health Research (KRS-134297,
566 PCG-155468, PJT-162206, KRS-140994, KRS-144045). Drs. Wu and Levis were supported by Fonds de
567 recherche du Québec – Santé (FRQS) Postdoctoral Training Fellowships. Dr. Benedetti was supported by a
568 FRQS researcher salary award.

569

570 **Disclosure Statement:** All authors have completed the ICJME uniform disclosure form and declare no
571 support from any organization for the submitted work; no financial relationships with any organizations that

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

572 might have an interest in the submitted work in the previous three years. All authors declare no other
573 relationships or activities that could appear to have influenced the submitted work. No funder had any role
574 in the design and conduct of the study; collection, management, analysis, and interpretation of the data;
575 preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

576
577 **Statement of Ethics:** The authors have no ethical conflicts to disclose.

578
579 **Transparency Declaration:** The manuscript's guarantor affirms that this manuscript is an honest, accurate,
580 and transparent account of the study being reported; that no important aspects of the study have been
581 omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been
582 explained.

583
584 **Data Sharing:** Requests to access data should be made to the corresponding authors.

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

REFERENCES

1. Brugha TS, Bebbington PE, Jenkins R. A difference that matters: comparisons of structured and semi-structured psychiatric diagnostic interviews in the general population. *Psychol Med.* 1999;29:1013-1020.
2. Brugha TS, Jenkins R, Taub N, Meltzer H, Bebbington PE: A general population comparison of the Composite International Diagnostic Interview (CIDI) and the Schedules for Clinical Assessment in Neuropsychiatry (SCAN). *Psychol Med.* 2001;31:1001-1013.
3. Nosen E, Woody SR: Chapter 8: Diagnostic Assessment in Research: In *Handbook of research methods in abnormal and clinical psychology*. Edited by McKay D. Thousand Oaks, CA, Sage, 2008, pp 109-124.
4. First MB: *Structured Clinical Interview for the DSM (SCID)*. John Wiley & Sons, Inc. 1995.
5. Levis B, Benedetti A, Riehm KE, et al. Probability of major depression diagnostic classification using semi-structured vs. fully structured diagnostic interviews. *Br J Psychiatry.* 2018;212:377-385
6. 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;28:e1803.
7. Wu Y, Levis B, Sun Y, et al: Probability of major depression diagnostic classification based on the SCID, CIDI and MINI diagnostic interviews controlling for Hospital Anxiety and Depression Scale – Depression subscale scores: an individual participant data meta-analysis of 73 primary studies. *J Psychosom Res.* 2020;129:109892.
8. Robins LN, Wing J, Wittchen HU, et al: The Composite International Diagnostic Interview: an epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Arch Gen Psychiatry.* 1988;45:1069-1077.

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

- 609 9. Lecrubier Y, Sheehan DV, Weiller E, et al: The Mini International Neuropsychiatric Interview (MINI).
610 A short diagnostic structured interview: reliability and validity according to the CIDI. *Eur Psychiatry*.
611 1997;12:224-231.
- 612 10. Sheehan DV, Lecrubier Y, Sheehan KH, et al. The validity of the Mini International Neuropsychiatric
613 Interview (MINI) according to the SCID-P and its reliability. *Eur Psychiatry*. 1997;12:232-241.
- 614 11. Gilbody S, Sheldon T, House A. Screening and case-finding instruments for depression: a meta-
615 analysis. *CMAJ*. 2008;178:997-1003.
- 616 12. Moriarty AS, Gilbody S, McMillan D, Manea L. Screening and case finding for major depressive
617 disorder using the Patient Health Questionnaire (PHQ-9): a meta-analysis. *Gen Hosp Psychiatry*.
618 2015;37:567-576.
- 619 13. Manea L, Gilbody S, McMillan D. Optimal cut-off score for diagnosing depression with the Patient
620 Health Questionnaire (PHQ-9): a meta-analysis. *CMAJ*. 2012;184:E191-E196.
- 621 14. Manea L, Gilbody S, Hewitt C, et al: Identifying depression with the PHQ-2: a diagnostic meta-
622 analysis. *J Affect Disord*. 2016;203:382-395.
- 623 15. Hewitt CE, Gilbody SM, Brealey S, et al. Methods to identify postnatal depression in primary care: an
624 integrated evidence synthesis and value of information analysis. *Health Technol Assess*. 2009;13:1-
625 145,147-230.
- 626 16. Mitchell AJ, Meader N, Symonds P. Diagnostic validity of the Hospital Anxiety and Depression Scale
627 (HADS) in cancer and palliative settings: a meta-analysis. *J Affect Disord*. 2010;126:335-348.
- 628 17. Rice DB, Kloda LA, Shrier I, Thombs BD: Reporting completeness and transparency of meta-analyses
629 of depression screening tool accuracy: a comparison of meta-analyses published before and after the
630 PRISMA statement. *J Psychosom Res*. 2016;87:57-69.
- 631 18. Lu J, Huang YQ, Liu ZR, Cao XL. Validity of Chinese version of the composite international
632 diagnostic interview-3.0 in psychiatric settings. *Chin Med J*. 2015;128:2462-2466.

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

- 633 19. Gelaye B, Williams MA, Lemma S, et al. Diagnostic validity of the composite international diagnostic
634 interview (CIDI) depression module in an East African population. *Int J Psychiat Med.* 2013;46:387-
635 405.
- 636 20. Anthony JC, Folstein M, Romanoski AJ, et al. Comparison of the lay Diagnostic Interview Schedule
637 and a standardized psychiatric diagnosis: experience in eastern Baltimore. *Arch Gen Psychiatry.*
638 1985;42:667-675.
- 639 21. Booth BM, Kirchner JA, Hamiltonc G, Harrell R, Smith GR. Diagnosing depression in the medically
640 ill: validity of a lay-administered structured diagnostic interview. *J Psychiatr Res.* 1998;32:353-360.
- 641 22. Hesselbrock V, Stabenau J, Hesselbrock M, Mirkin P, Meyer R. A comparison of two interview
642 schedules: the Schedule for Affective Disorders and Schizophrenia-Lifetime and the National Institute
643 for Mental Health Diagnostic Interview Schedule. *Arch Gen Psychiatry.* 1982;39:674-677.
- 644 23. Jordanova V, Wickramesinghe C, Gerada C, Prince M. Validation of two survey diagnostic interviews
645 among primary care attendees: a comparison of CIS-R and CIDI with SCAN ICD-10 diagnostic
646 categories. *Psychol Med.* 2004;34:1013-1024.
- 647 24. Diagnostic and statistical manual of mental disorders: DSM-III 3rd ed, revised. Washington, DC:
648 American Psychiatric Association 1987.
- 649 25. Diagnostic and statistical manual of mental disorders: DSM-IV 4th ed. Washington, DC: American
650 Psychiatric Association 1994.
- 651 26. Diagnostic and statistical manual of mental disorders: DSM-IV 4th ed, text revised. Washington, DC:
652 American Psychiatric Association 2000.
- 653 27. Diagnostic and statistical manual of mental disorders: DSM-V 5th ed. Washington, DC: American
654 Psychiatric Association 2013.
- 655 28. The ICD-10 Classifications of Mental and Behavioural Disorder: Clinical Descriptions and Diagnostic
656 Guidelines. Geneva: World Health Organization 1992.

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

- 657 29. Thombs BD, Benedetti A, Kloda LA, et al. The diagnostic accuracy of the Patient Health
658 Questionnaire-2 (PHQ-2), Patient Health Questionnaire-8 (PHQ-8), and Patient Health Questionnaire-9
659 (PHQ9) for detecting major depression: protocol for a systematic review and individual patient data
660 meta-analyses. *Syst Rev.* 2014;3:124.
- 661 30. Thombs BD, Benedetti A, Kloda LA, et al: Diagnostic accuracy of the Edinburgh Postnatal Depression
662 Scale (EPDS) for detecting major depression in pregnant and postnatal women: protocol for a
663 systematic review and individual patient data meta-analyses. *BMJ Open.* 2015;5:e009742.
- 664 31. Thombs BD, Benedetti A, Kloda LA, et al: Diagnostic accuracy of the Depression subscale of the
665 Hospital Anxiety and Depression Scale (HADS-D) for detecting major depression: protocol for a
666 systematic review and individual patient data meta-analyses. *BMJ Open.* 2016;6:e011913.
- 667 32. United Nations Development Programme: Human Development Report 2019.
668 <http://hdr.undp.org/sites/default/files/hdr2019.pdf>. Accessed January 14, 2020.
- 669 33. Levis B, Benedetti A, Thombs BD, and the DEPRESSion Screening Data (DEPRESSD) Collaboration:
670 Accuracy of Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression:
671 individual participant data meta-analysis. *BMJ.* 2019;365:11476.
- 672 34. Deeks JJ, Higgins JP, Altman DG: Chapter 9: Analysing data and undertaking meta-analyses. In
673 *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*.
674 Edited by Higgins JP, Green S, editors. The Cochrane Collaboration, 2011.
- 675 35. Higgins JP, Thompson SG, Deeks JJ, Altman DG: Measuring inconsistency in meta-analyses. *BMJ.*
676 2003;327:557-560.
- 677 36. Team RC: R: a language and environment for statistical computing. Vienna, Austria, R Foundation for
678 Statistical Computing, 2018.
- 679 37. Team R: RStudio: integrated development for R. Boston, MA, RStudio, Inc., 2015, pp 639-640.

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

- 680 38. Bates D, Machler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw.*
681 2014;67:1-48.
- 682 39. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw.* 2010;36:1-48.
- 683 40. Zimmerman M, Balling C, Chelminski I, Dalrymple K. Have treatment studies of depression become
684 even less generalizable? Applying the inclusion and exclusion criteria in placebo-controlled
685 antidepressant efficacy trials published over 20 years to a clinical sample. *Psychother Psychosom.*
686 2019;88:165-70.
- 687 41. Wu Y, Levis B, Riehm KE, et al: Equivalency of the diagnostic accuracy of the PHQ-8 and PHQ-9: a
688 systematic review and individual participant data meta-analysis. *Psychol Med* 2019;12:1-13.
- 689 42. He C, Levis B, Riehm KE, et al. The accuracy of the Patient Health Questionnaire-9 algorithm for
690 screening to detect major depression: An individual participant data meta-
691 analysis. *Psychother Psychosom.* 2020;89:25-37.
- 692 43. Fava GA. The decline of pluralism in medicine: dissent is welcome. *Psychother Psychosom.* 2020;89:1-
693 5.
- 694 44. Fava GA, Tomba E, Sonino N. Clinimetrics: the science of clinical measurements. *Int J Clin Pract.*
695 2012;66:11-15.
- 696 45. Fava GA, Carrozzino D, Lindberg L, Tomba E. The clinimetric approach to psychological assessment:
697 a tribute to Per Bech, MD (1942-2018). *Psychother Psychosom.* 2018;87:321-326.
- 698 46. Tomba E, Bech P. Clinimetrics and clinical psychometrics: macro- and micro-analysis. *Psychother*
699 *Psychosom.* 2012;81:333-343.

700

701 **FIGURE LEGENDS**

702
703 **Figure 1.** Comparison of major depression classification odds of the Composite International Diagnostic
704 Interview (CIDI) versus the Structured Clinical Interview for DSM (SCID)

705
706 The figure presents the aOR of major depression classification for the CIDI compared to the SCID for
707 primary studies based on the PHQ-9, EPDS, and HADS-D and pooled estimates at standardized scores of -
708 1, 0, 1, 2 and 3. The standardized scores of -1, 0, 1, 2 and 3 are approximately equal to scores of 0, 5, 10, 16
709 and 21 on the PHQ-9 (SD = 5.26); 1, 7, 13, 18 and 24 on the EPDS (SD = 5.58); and 1, 5, 9, 13 and 17 on
710 the HADS-D (SD = 4.07). We present standardized scores from -1 to 3, because raw scores corresponding
711 to standardized scores below -1 or above 3 would be negative or beyond the maximum scores of the
712 included screening tools.

713
714 Abbreviations: EPDS: Edinburgh Postnatal Depression Scale; HADS-D: Depression subscale of Hospital
715 Anxiety and Depression Scale; META: Pooled estimates from the synthesis meta-analysis. PHQ-9: Patient
716 Health Questionnaire-9.

COMPARISON OF DIAGNOSTIC INTERVIEWS FOR MAJOR DEPRESSION

719 **Figure 2.** Comparison of major depression classification odds of the Mini International Neuropsychiatric
720 Interview (MINI) vs. the SCID considering the interaction between depressive symptom severity and the
721 MINI

722

723 The figure presents the aOR of major depression classification for the MINI compared to the SCID for
724 primary studies based on the PHQ-9, EPDS, and HADS-D and pooled estimates at standardized scores of -
725 1, 0, 1, 2 and 3. The standardized scores of -1, 0, 1, 2 and 3 are approximately equal to scores of 0, 5, 10, 16
726 and 21 on the PHQ-9 (SD = 5.26); 1, 7, 13, 18 and 24 on the EPDS (SD = 5.58); and 1, 5, 9, 13 and 17 on
727 the HADS-D (SD = 4.07). We present standardized scores from -1 to 3, because raw scores corresponding
728 to standardized scores below -1 or above 3 would be negative or beyond the maximum scores of the
729 included screening tools.

730

731 Abbreviations: EPDS: Edinburgh Postnatal Depression Scale; HADS-D: Depression subscale of Hospital
732 Anxiety and Depression Scale; META: Pooled estimates from the synthesis meta-analysis. PHQ-9: Patient
733 Health Questionnaire-9.

734