

1 **TITLE PAGE**

2 **Title:** Neurofilament light and glial fibrillary acidic protein in mood and anxiety disorders: A
 3 systematic review and meta-analysis

4 **Authors (incl order)**

Full name	Highest academic degree	Affiliation	Email address
Matthew JY Kang	MBBS, FRANZCP	Neuropsychiatry Centre, Royal Melbourne Hospital Department of Psychiatry, University of Melbourne	Matthew.kang1@unimelb.edu.au
Jasleen Grewal	MBBS	Alfred Mental and Addiction Health, Alfred Health Neuropsychiatry Centre, Royal Melbourne Hospital	Jas.grewal@alfred.org.au
Dhamidhu Eratne	MBChB, FRANZCP	Neuropsychiatry Centre, Royal Melbourne Hospital Department of Psychiatry, University of Melbourne	deratne@unimelb.edu.au ORCID: 0000-0002-3226-7645
Charles Malpas	PhD	Melbourne School of Psychological Sciences, University of Melbourne Department of Medicine, Royal Melbourne Hospital, University of Melbourne	charles.malpas@unimelb.edu.au
Wei-Hsuan Chiu	BBiomed(Hons), GDipPsych(Adv)	The Neuropsychiatry Centre, The Royal Melbourne Hospital Department of Psychiatry, The University of Melbourne	whchiu@student.unimelb.edu.au ORCID: 0000-0002-5066-2999
Kasper Katisko	MD, PhD	Institute of Clinical Medicine – Neurology, University of Eastern Finland, Kuopio, Finland	kasper.katisko@uef.fi
Eino Solje	MD, PhD	Institute of Clinical Medicine – Neurology, University of Eastern Finland, Kuopio, Finland	eino.solje@uef.fi

		Neuro Center – Neurology, Kuopio University Hospital, Kuopio, Finland	
Alexander F Santillo 0000-0001-9717-0820	MD, PhD	Department of Clinical Sciences, Clinical Memory Research Unit, Faculty of Medicine, Lund University, Malmö, Sweden	alexander.frizell_santillo@med.lu.se
Philip B. Mitchell	MBBS, MD	Discipline of Psychiatry & Mental Health, Faculty of Medicine and Health, UNSW, Sydney	phil.mitchell@unsw.edu.au
Malcolm Hopwood	MD, FRANZCPsych	Ramsay Clinic Albert Road & Department of Psychiatry, University of Melbourne	mhopwood@unimelb.edu.au
Dennis Velakoulis	MBBS, DMedSci	Neuropsychiatry Centre, Royal Melbourne Hospital Department of Psychiatry, University of Melbourne	Dennis.Velakoulis@mh.org.au

5

6 **Corresponding Author**

7 Dr. Matthew JY Kang

8 The Royal Melbourne Hospital Mental Health Services & Neuropsychiatry Centre

9 The RMH Elizabeth Street

10 635 Elizabeth Street | Melbourne 3000

11 P: 03 9138 2850 | F: 03 7024 7004 | M: 0437 631 101

12 Matthew.Kang@mh.org.au | thermh.org.au

13 **Word count:** 2998 words

14 **Funding and sponsors**

15 The MiND Study is generously funded from the following bodies:

16 - 2020-2024 NHMRC Ideas Grant GNT1185180: The Markers in Neuropsychiatric Disorders
17 study.

18 - MRFF

19 - Ramsay Health Research Foundation Translation Challenge (2023-2024)

20 Matthew Kang is supported by the Research Training Program Scholarship (stipend) from the
21 Department of Psychiatry, University of Melbourne with contributions from the Australian
22 Commonwealth Government and the Ramsay Health Research Foundation Translation
23 Challenge.

24 The funders had no role in the design or conduct of this study.

25 **Conflicts of interest:** None from MK. PBM has received remuneration from Janssen (Australia)
26 for advisory board membership and lectures in the last 3 years.

27

28 **Acknowledgements**

29 We would like to thank Dr Mu-Hong Chen, Dr Leila Simani, and Dr Wen-Yin Chen for their
30 assistance in providing their published data separately to be included in the meta-analysis. We
31 are also grateful for the correspondence from the following researchers who assisted us in
32 clarifying questions about their published data: Dr Anja Fernqvist, Dr Sterre de Boer, Dr
33 Naghmeh Nikkheslat, Prof Carmine Pariante, Dr. Pedro J. Serrano Castro, Dr Johanna
34 Wallensten, Dr. Carol Van Hulle, Dr Mark Miller, Dr Daniel Alcolea, Dr. Suzanne Schindler,
35 and Dr. Andrew Aschenbrenner.

36 Abstract

37 **Background:** Neurofilament light chain (NfL) and glial fibrillary acidic protein (GFAP) are
38 biomarkers of neuronal injury measurable in cerebrospinal fluid (CSF) and blood. Despite their
39 potential as diagnostic tests for neurodegenerative disorders, it is unclear how they behave in
40 mood and anxiety disorders. We conducted a systematic review and meta-analysis to investigate
41 whether NfL and GFAP concentrations were altered in adults with mood and anxiety disorders
42 compared to healthy controls.

43
44 **Methods:** The study was prospectively registered on PROSPERO (CRD42023434617). We
45 followed the PRISMA guidelines, searched PubMed, Web of Science, PsycINFO, MEDLINE
46 and Embase up to the 31/05/2023, and assessed relevant studies and their risk of bias. The
47 primary outcome was the standardised mean difference (SMD) and 95% confidence interval
48 (95% CI) of NfL and GFAP concentrations, which was pooled using a random-effects model
49 adopting the restricted maximum likelihood estimator.

50
51 **Results:** Twenty-one studies met inclusion criteria, comprising of 2327 individuals (695 major
52 depression, 502 bipolar disorder, and 1130 controls). When we compared people with major
53 depression and controls, there was no difference in NfL (SMD = 0.29; 95% CI: -0.10, 0.68) nor
54 GFAP (SMD = 0.47; 95% CI: -0.74, 1.68). In people with bipolar disorder, NfL was
55 significantly elevated compared to controls (SMD = 0.58; 95% CI: 0.16, 0.99). However, the
56 subgroup analysis including more sensitive assay kits (blood Simoa and CSF ELISA), found no
57 significant difference (SMD = 0.40; 95% CI: -0.04, 0.85). Only one study studied GFAP in
58 bipolar disorder. No studies explored NfL nor GFAP concentrations in anxiety disorders.

59
60 **Discussion:** We found that NfL and GFAP concentrations were not elevated in depression. In
61 bipolar disorder, NfL concentration was elevated, though not in the sensitivity analysis. Our
62 study informs clinicians about how to interpret these emerging biomarkers in determining
63 whether a person's symptoms are caused by a neurodegenerative or mood disorder.

64

65 Introduction

66 Recent developments in blood and cerebrospinal fluid (CSF) biomarkers in neurological
67 conditions have offered exciting prospects for enhancing diagnostic assessments. Neurofilament
68 light chain (NfL), a cytoskeletal protein of neurons, has emerged as a sensitive biomarker for
69 assessing neuronal injury (1). Neurodegenerative processes lead to the release of NfL into the
70 CSF, which subsequently pass through to the bloodstream at lower concentrations (2). Although
71 initial studies were limited to measuring NfL in CSF, technological advances have facilitated
72 precise quantification of NfL in blood (3). In particular, blood NfL measured using single-
73 molecule array digital immunoassay (Simoa) correlate well with CSF NfL (4). Numerous studies
74 have shown that NfL, in both CSF and blood, is a valuable marker for diagnosing, assessing
75 disease severity, and predicting prognosis in a diverse range of neurodegenerative conditions,
76 including Alzheimer's disease and frontotemporal lobar degeneration (5–7). NfL can also
77 distinguish neurodegenerative disorders from primary psychiatric disorders within memory
78 clinics and neuropsychiatry services (8–10). More recently, glial fibrillary acidic protein
79 (GFAP), a component of glial cell cytoskeleton, has emerged as a sensitive measure of
80 astrogliosis and neuroinflammation, demonstrating superior prognostic accuracy to NfL in
81 multiple sclerosis (11,12).

82
83 Given the potential clinical utility of these biomarkers in determining whether an individual's
84 undifferentiated presentation with psychiatric, cognitive and/or neurological symptoms is due to
85 a neurological or a psychiatric cause, it is crucial to understand how NfL or GFAP behave in
86 primary mood or anxiety disorders. If NfL and GFAP levels remain within normal ranges in
87 primary mood and anxiety disorders, which are the two most prevalent mental disorders (13),
88 such findings would support the use of these biomarkers in distinguishing neurodegenerative
89 conditions at diagnostic assessments. Conversely, should they be elevated in primary mood and
90 anxiety disorders, this may provide insights into their neurobiological underpinnings.

91
92 One of the earliest investigations into NfL and GFAP in psychiatric disorders involved a cohort
93 of elderly women who remained dementia-free during a 10-year follow-up period (14). Those
94 with major depression exhibited significantly higher concentrations of baseline CSF NfL
95 compared to the rest of the group (major depression mean: 427 ± 186 vs control mean: 277 ± 186

96 ng/L), whilst there was no difference in the GFAP concentrations (major depression mean: $946 \pm$
97 196 vs control: 887 ± 308 ng/L). Subsequent studies comparing blood NfL and GFAP in mood
98 disorders to healthy controls have yielded mixed findings, either finding elevated concentrations
99 in mood disorders (15–17), or no difference (6,18).

100
101 More recent studies have examined the relationship between NfL or GFAP with clinical
102 outcomes in mood disorders. The concentration of GFAP was higher in people with major
103 depression and correlated with the severity of depressive symptoms (17). NfL was found to be
104 negatively correlated with cognitive processing speed in people with major depression compared
105 to healthy controls (19). Several theories have been proposed to explain why NfL or GFAP may
106 be elevated in mood disorders, including astroglia dysfunction (17), vascular pathology (14), and
107 neuroinflammatory processes (20).

108
109 Given these conflicting findings in this area of major clinical and scientific importance, we have
110 undertaken a systematic review and meta-analysis of studies investigating NfL and GFAP
111 concentrations in people with mood (depression and bipolar disorder) and anxiety disorders
112 compared to healthy controls. Secondly, we sought to identify any associations between the
113 severity of mood and anxiety disorders with NfL and GFAP.

114

115 **Methods**

116 This systematic review was developed using the Preferred Reporting Items for Systematic
117 Reviews and Meta-Analysis (PRISMA) guidelines (21). The protocol was prospectively
118 registered with PROSPERO (CRD42023434617).

119

120 **Eligibility criteria**

121 All eligible studies met the following criteria: (1) adult patients diagnosed with either mood
122 (depression and bipolar disorder) or anxiety disorder according to DSM or ICD (22–25); and (2)
123 measured NfL or GFAP in plasma, serum or CSF measured by Simoa,
124 electrochemiluminescence method (ECL) or enzyme linked immunosorbent assay (ELISA).
125 Case-series, meeting abstracts, reviews and meta-analyses were excluded. For articles with
126 overlapped samples, only the report with largest sample size was included.

127

128 Exclusions criteria were: (1) mood or anxiety disorders due to general medical conditions
129 including dementia and stroke, and (2) co-morbid neurological disorders including multiple
130 sclerosis and dementia.

131

132 **Search strategy**

133 Electronic searches were conducted *via* the Web of Science, PubMed, MEDLINE(OVID),
134 Embase and PsycINFO from the inception of the databases through to 31st May 2023. The
135 search terms included (“neurofilament light” *OR* “NfL”) *AND* (“glial fibrillary acidic protein”
136 *OR* “GFAP”) *AND* (“depressi*” *OR* ”bipolar*” *OR* “anxiet*” *OR* “mood disorder*”). Studies
137 were required to be in English. We have described our full search strategy in the supplementary
138 file.

139

140 **Study Selection and Data Collection**

141 We used Cochrane’s Covidence, a web-based systematic review manager, for study selection and
142 quality analysis. One investigator (MK) screened the title and abstracts. Two investigators (MK
143 and JG) independently reviewed the main reports and supplementary materials for their
144 eligibility. We resolved discrepancies by consensus and involved a third senior investigator
145 (MH) when required.

146

147 The primary outcome was the mean difference of NfL and GFAP in people with mood and
148 anxiety disorders. The mean and standard deviation (SD) of NfL and GFAP was collected in the
149 different cohorts. If this was not available, authors were contacted for the data. If the authors did
150 not respond, the following methods were used to convert the data according to Cochrane
151 Handbook guidelines (26): 1) median, interquartile ranges and minimum/maximum values were
152 converted using an estimation formula (27,28), 2) mean and SD of subgroups (i.e. male and
153 female) were combined using Cochrane’s formula (29), and 3) studies that only presented the
154 results in plots were estimated using WebPlotDigitizer (30).

155

156 Our secondary outcomes were as follows: 1) mood state and severity (measured by a
157 standardised rating scale), 2) duration of mood/anxiety disorder, and 3) cognitive impairment

158 (measured by a standardised rating scale). We also collected known covariates of NfL and GFAP
159 if they were reported, including age and weight (31).

160

161 **Quality of individual studies**

162 Two investigators (MK and JG) independently assessed the quality of included studies using the
163 Newcastle-Ottawa Scale (NOS) quality assessment tool (32). NOS contains a total of 8 small
164 items, and evaluates three aspects: selection, comparability and outcome. Studies were classified
165 as “poor quality” (0-2 points), “average quality” (3-5 points), and “good or high quality” (6-9
166 points).

167

168 **Statistical analysis**

169 All statistical analyses were performed using R version 4.2.2 (2022-10-31) and the meta package
170 (33). Meta-analyses were separately conducted for individual psychiatric diagnoses (i.e. major
171 depression and bipolar disorder). For each diagnosis, we calculated standardised mean
172 differences (SMDs) of each biomarker (NfL and GFAP) between psychiatric disorders and
173 controls to be able to pool the different types of assay kits. Of note, Random effects meta-
174 analysis models, rather than fixed effects models, were selected a priori based on the theoretical
175 absence of a consistent ‘true effect’ across the studies. The restricted maximum likelihood
176 estimator was adopted. Ninety-five percent confidence intervals (95% CI) and prediction
177 intervals (PI) were calculated and visualised using forest plots. A PI quantifies the likely range in
178 which a new study’s effect size will fall into, assuming that this future study is of a similar nature
179 to those meta-analysed. We assumed a two-sided $P < 0.05$ to indicate statistical significance.

180

181 Where there were a sufficient number of studies, sensitivity analyses were performed to ensure
182 any findings were consistent with the main analyses: type of biomarker analysis (i.e. CSF ELISA
183 and serum/plasma Simoa; blood Simoa) and studies that used age-matched controls. We intended
184 to analyse further subgroups including duration of psychiatric illness and age at diagnosis,
185 however there was an insufficient number of studies.

186

187 We performed meta regression on the severity of mood symptoms as measured by the scales
188 used in the study (i.e. Hamilton’s Depression Rating Scale; HDRS and Montgomery-Asberg

189 Depression Rating Scale; MADRS). We could not perform other meta regression analyses
190 including variables such as cognition, duration of psychiatric illness and number of mood
191 disorder episodes due to the limited numbers of studies.

192
193 As a post-hoc analysis to include studies (all of which used Simoa assay kits) that did not have
194 any control groups, we used a robust mixed-means model to compare the mean concentration of
195 NfL and GFAP between cases (depression or bipolar) and controls.

196
197 We assessed for publication bias using funnel plots of the effect sizes of mood disorders for NfL
198 and GFAP. We performed meta-analyses only where there were at least five studies to ensure
199 acceptable statistical power to make an inference. I^2 and τ^2 statistics were used to quantify
200 between-study heterogeneity in the meta-analyses.

201

202 **Results**

203 **Study selection**

204 The systematic search identified 2,890 unique studies, of which 110 were selected for full-text
205 review. Twenty-one studies met our inclusion criteria for the meta-analysis, as shown in Figure
206 1. Twelve studies investigated major depression, five bipolar disorder, four reported data on
207 both, and no studies reported on other types of mood disorders nor anxiety disorders.

208
209 The main reason that some full-text articles were excluded was due to studies either not reporting
210 or collecting individual psychiatric diagnoses. A number of studies used a less rigorous process
211 to collect diagnoses (self-report, file review and diagnosis based on a cut-off score from a mood
212 rating scale) (34). Four studies had an insufficient sample size ($n < 10$) (9,35,36). Some studies
213 used indirect measures of NfL and GFAP, including extracellular vesicles, mRNA, IgG
214 autoantibodies and urine (37–40).

215

216 **Study characteristics**

217 Final sample data were obtained for 2327 individuals, of which 695 had major depression, 502
218 bipolar disorder, and 1130 healthy controls. Two studies (41,42) did not have a healthy control
219 group, so were only included in the additional mixed-model meta-analyses. Seventeen studies

220 analysed blood samples, three used CSF, and Knorr et al's study (16) investigated both CSF and
221 blood. The majority of the studies (75%) were published after 2020, reflecting the relatively
222 recent availability of Simoa and ELISA technology to measure NfL and GFAP in blood. (Tables
223 1-2)

224
225 Although Ashton et al. (2021) reported that the sample size of their control (cognitively
226 unimpaired and CSF amyloid- β negative) and depression groups were 130 and 37 respectively,
227 this conflicted with their supplementary data source table that reported 115 and 51 values
228 respectively. We chose to perform the meta-analysis with the latter values, as they matched the
229 sex proportions reported in their main table. Bavato et al. (2021) drew on 295 healthy controls
230 and referenced their earlier work (Barro et al., 2018, n=258) for the characteristics of the healthy
231 controls, which is included in this meta-analysis instead.

232

233 **Quality of studies**

234 The quality of studies was considered at least moderate across studies. The area of low quality
235 was in the domain of comparability due to the inequalities in age and sex between study groups,
236 as summarised in the supplementary eTable 1.

237

238 **Synthesis of results**

239 *Major depression*

240 *Neurofilament light chain*

241 **All studies (Figure 2a):** There was no significant difference in the concentration of NfL in
242 people with major depression compared to healthy controls (SMD = 0.29; 95% CI: -0.10, 0.68).
243 The prediction interval for this pooled mean difference was very wide (PI: -1.11, 1.69) with
244 significant heterogeneity ($I^2 = 90\%$; $\tau^2 = 0.37$). A regression-based Egger's test did not show
245 any evidence of small study effects (p-value = 0.35) though on visual inspection of the funnel
246 plot (supplementary eFigure 1), Ashton et al's study's was an outlier (6). This study had a
247 significant age difference between the depression and control groups (mean difference 31.7
248 years).

249 **In the subgroup analysis of studies that included age-matched controls (Figure 2b),** NfL was
250 significantly higher in major depression, though with a wide PI (SMD = 0.44; 95% CI: 0.11,

251 0.77; PI: -0.61, 1.50). **The subgroup analysis model which only included data using CSF**
252 **ELISA and blood Simoa (Figure 2c)**, which are more accurate measures of NfL (2,4), found no
253 significant difference (SMD = -0.03, 95% CI: -0.53, 0.47). The heterogeneity for both subgroup
254 analysis models were high ($I^2 = 79%$, $\tau^2 = 0.17$ and $I^2 = 57%$, $\tau^2 = 0.07$ respectively).
255 Similarly, the subgroup analysis model including six age-matched studies that used CSF ELISA
256 and Simoa also found no significant difference (SMD = 0.11 95% CI: -0.32, 0.55).

257
258 A random-effects meta-regression model which included all NfL studies suggested that mood
259 severity did not influence NfL. In the subgroup mixed-model mean analysis of studies that only
260 used blood Simoa, we found no significant difference in NfL levels.

261
262 **Glial fibrillary acidic protein**
263 GFAP (Figure 2d) was not significantly different between people with major depression and
264 healthy controls (SMD = 0.47; 95% CI: -0.74, 1.68; PI: -3.09, 4.03), with significant
265 heterogeneity between the studies ($I^2 = 92%$, $\tau^2 = 0.53$). There was insufficient data to perform
266 additional meta-regressions including cognition and GFAP.

267

268 *Bipolar disorder*

269 *Neurofilament light chain*

270 **All studies (Figure 3a)**: The mean concentration of NfL in people with bipolar disorder was
271 significantly higher than healthy control (SMD = 0.58; 95% CI: 0.16, 0.99). However, the
272 prediction interval for this pooled mean difference was wide (PI: -0.67, 1.82) with significant
273 heterogeneity ($I^2 = 84.0%$, $\tau^2 = 0.24$). As Knorr et al's study (2022) reported both CSF ELISA
274 and plasma Simoa samples, we performed a sensitivity analysis where we excluded the Simoa
275 dataset, which found NfL was still significantly higher in bipolar disorder compared to healthy
276 controls (SMD = 0.62; 95% CI: 0.14, 1.09), There was no asymmetry on the funnel plot
277 (supplementary eFigure 1), and we could not use Egger's test due to insufficient number of
278 studies.

279 **The subgroup analysis model with age-matched studies (Figure 3b)** demonstrated significantly
280 higher NfL in bipolar disorder (SMD = 0.50; 95% CI: 0.06, 0.94; PI: -0.74, 1.74). **For the**
281 **subgroup analysis including only CSF ELISA and blood Simoa studies (Figure 3c)**, there was

282 no significant difference in the mean concentration of NfL between bipolar disorder and healthy
283 controls (SMD = 0.40; 95% CI: -0.04, 0.85) with a wide prediction interval (PI: -0.80, 1.61).
284 Both models were heterogenous ($I^2 = 82.0\%$, $\tau^2 = 0.22$ and $I^2 = 81.8\%$, $\tau^2 = 0.18$
285 respectively). The subgroup analysis including six age-matched studies which used CSF ELISA
286 and Simoa also found no significant difference in mean NfL between bipolar disorder and
287 healthy controls. (SMD = 0.27 95% CI: -0.12, 0.65). We performed a subgroup mixed-model
288 mean analysis of blood Simoa studies which showed no difference in NfL concentrations.

289

290 **Glial fibrillary acidic protein**

291 Only one study (17) studied GFAP in people with bipolar disorder (n=11), which found GFAP
292 concentrations were similar in bipolar disorder compared to healthy controls (125 ± 50 vs $147 \pm$
293 61 pg/mL).

294

295 **Discussion**

296 Our systematic review found that NfL and GFAP concentrations were not elevated in people
297 with major depression compared to healthy controls. NfL was elevated in people with bipolar
298 disorder, though the difference was no longer significant with subgroup analysis using more
299 sensitive assay kits (CSF ELISA and blood Simoa). We identified no studies examining NfL or
300 GFAP in people with anxiety disorders. To the best of our knowledge, this is the most
301 comprehensive systematic review of these biomarkers in people with primary mood and anxiety
302 disorders.

303

304 Our findings align with the existing literature supporting the use of these biomarkers in the
305 differentiation between mood disorders and neurodegenerative disorders (2,5,6). The complex
306 relationship between mood disorders and neurodegenerative disorders, which can present with
307 overlapping symptoms or co-occur, necessitates careful assessment. These findings provide
308 reassurance to clinicians that people with primary mood disorders are unlikely to have very
309 elevated concentrations of NfL and GFAP, as seen in the majority of neurodegenerative
310 disorders where there often is marked elevation (for example, 1.86-fold in Alzheimer's disease,
311 3.91-fold in frontotemporal dementia; 44,45).

312

313 In individuals with mood disorders, we noted variable instances where NfL concentration was
314 elevated compared to healthy controls, indicating that neuronal injury or neuroinflammation
315 might be present in a specific subset. The limitations of the available data, including detailed
316 clinical and treatment information, prevented us from being able to investigate this further. We
317 plan to conduct a ‘mega-analysis’ by pooling the raw data from the studies included. This
318 approach will enable us to perform statistical adjustments for known covariates including as age,
319 sex and weight/BMI, which were not possible in the meta-analysis. Moreover, identifying
320 specific clusters of symptomatology that correlate with NfL/GFAP may offer valuable insights
321 into the neurodegenerative and neuroinflammatory phenotypes within mood disorders. In turn,
322 opens a new paradigm of diagnosis and treatment for these patients based on biomarker data as
323 well as reported symptoms, in line with the RDOC Framework (56).

324
325 Despite the widespread prevalence and burden of anxiety disorders (13), no prior research has
326 explored NfL and GFAP in individuals with primary anxiety disorders. This is an important gap
327 in the literature that this systematic review has identified for the use of NfL and GFAP in clinical
328 practice.

329
330 The review process faced several limitations. The evidence was constrained by the quality and
331 heterogeneity of included studies. Controls were not consistently age- and sex-matched (6,18),
332 and some studies lacked control subjects (41,42,51). The limited number of studies and data
333 precluded adequate adjustment for known covariates of NfL and GFAP, such as age, sex,
334 weight/BMI, and renal function (2). This also hindered exploration of factors like mood disorder
335 states (euthymic vs depressed) or effects of psychotropic medications (46). Furthermore, to
336 harmonise the dataset, some of the meta-analyses relied on estimates derived from raw data,
337 which may have introduced potential inaccuracies in the true estimates. Despite our attempts to
338 include as many articles as possible by searching important databases, grey literature, as well as
339 snowballing for other relevant articles, it is possible that some studies may have been
340 inadvertently omitted.

341

342

343 **Conclusions**

344

345 Our primary analysis indicated that NfL and GFAP levels were not significantly elevated in
346 major depression, while in bipolar disorder, NfL was elevated compared to controls. Subgroup
347 analysis using more sensitive assays revealed no notable differences. These findings suggest that
348 NfL and GFAP elevations in mood disorders are minimal, contrasting with the large elevations
349 in neurodegenerative disorders. This supports the use of these biomarkers in clinical settings to
350 distinguish psychiatric from neurodegenerative causes. The practicality of blood-based
351 biomarker measurements offers valuable diagnostic insights to answer the common clinical
352 question faced by patients, carers and clinicians.

353

354 **References**

- 355 1. Gaetani L, Blennow K, Calabresi P, Filippo MD, Parnetti L, Zetterberg H. Neurofilament
356 light chain as a biomarker in neurological disorders. *J Neurol Neurosurg Psychiatry*. 2019
357 Aug 1;90(8):870–81.
- 358 2. Arslan B, Zetterberg H. Neurofilament light chain as neuronal injury marker – what is
359 needed to facilitate implementation in clinical laboratory practice? *Clinical Chemistry and*
360 *Laboratory Medicine (CCLM)*. 2023 Jun 27;61(7):1140–9.
- 361 3. Khalil M, Teunissen CE, Otto M, Piehl F, Sormani MP, Gattringer T, et al. Neurofilaments
362 as biomarkers in neurological disorders. *Nature Reviews Neurology*. 2018 Oct;14(10):577–
363 89.
- 364 4. Alagaratnam J, von Widekind S, De Francesco D, Underwood J, Edison P, Winston A, et al.
365 Correlation between CSF and blood neurofilament light chain protein: a systematic review
366 and meta-analysis. *BMJ Neurol Open*. 2021 Jun 16;3(1):e000143.
- 367 5. Benkert P, Meier S, Schaedelin S, Manouchehrinia A, Yaldizli Ö, Maceski A, et al. Serum
368 neurofilament light chain for individual prognostication of disease activity in people with
369 multiple sclerosis: a retrospective modelling and validation study. *The Lancet Neurology*.
370 2022 Mar;21(3):246–57.
- 371 6. Ashton NJ, Janelidze S, Al Khleifat A, Leuzy A, van der Ende EL, Karikari TK, et al. A
372 multicentre validation study of the diagnostic value of plasma neurofilament light. *Nat*
373 *Commun*. 2021 Jun 7;12(1):3400.
- 374 7. Eratne D, Loi SM, Li QX, Stehmann C, Malpas CB, Santillo A, et al. Cerebrospinal fluid
375 neurofilament light chain differentiates primary psychiatric disorders from rapidly
376 progressive, Alzheimer’s disease and frontotemporal disorders in clinical settings.
377 *Alzheimer’s & Dementia*. 2022;18(11):2218–33.
- 378 8. Eratne D, Kang M, Malpas C, Simpson-Yap S, Lewis C, Dang C, et al. Plasma
379 neurofilament light in behavioural variant frontotemporal dementia compared to mood and
380 psychotic disorders. *Aust N Z J Psychiatry*. 2023 Jul 21;00048674231187312.
- 381 9. Kang MJ, Eratne D, Dobson H, Malpas CB, Keem M, Lewis C, et al. Cerebrospinal fluid
382 neurofilament light predicts longitudinal diagnostic change in patients with psychiatric and
383 neurodegenerative disorders. *Acta Neuropsychiatrica*. 2023 Apr 28;1–36.
- 384 10. Willems EAJ, Scheltens P, Teunissen CE, Vijverberg EGB. A neurologist’s perspective on
385 serum neurofilament light in the memory clinic: a prospective implementation study. *Alz*
386 *Res Therapy*. 2021 Dec;13(1):101.
- 387 11. Abdelhak A, Foschi M, Abu-Rumeileh S, Yue JK, D’Anna L, Huss A, et al. Blood GFAP as
388 an emerging biomarker in brain and spinal cord disorders. *Nat Rev Neurol*. 2022
389 Mar;18(3):158–72.

- 390 12. Meier S, Willemse EAJ, Schaedelin S, Oechtering J, Lorscheider J, Melie-Garcia L, et al.
391 Serum Glial Fibrillary Acidic Protein Compared With Neurofilament Light Chain as a
392 Biomarker for Disease Progression in Multiple Sclerosis. *JAMA Neurology*. 2023 Mar
393 1;80(3):287–97.
- 394 13. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12
395 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the
396 Global Burden of Disease Study 2019. *The Lancet Psychiatry*. 2022 Feb 1;9(2):137–50.
- 397 14. Gudmundsson P, Skoog I, Waern M, Blennow K, Zetterberg H, Rosengren L, et al. Is there a
398 CSF biomarker profile related to depression in elderly women? *Psychiatry Research*.
399 2010;176(2–3):174–8.
- 400 15. Hviid CVB, Benros ME, Krogh J, Nordentoft M, Christensen SH. Serum glial fibrillary
401 acidic protein and neurofilament light chain in treatment-naïve patients with unipolar
402 depression. *Journal of Affective Disorders*. 2023 Oct 1;338:341–8.
- 403 16. Knorr U, Simonsen AH, Jensen CS, Zetterberg H, Blennow K, Akhøj M, et al. Alzheimer’s
404 disease related biomarkers in bipolar disorder - A longitudinal one-year case-control study. *J*
405 *Affect Disord*. 2022 Jan 15;297:623–33.
- 406 17. Steinacker P, Al Shweiki MR, Oeckl P, Graf H, Ludolph AC, Schönfeldt-Lecuona C, et al.
407 Glial fibrillary acidic protein as blood biomarker for differential diagnosis and severity of
408 major depressive disorder. *Journal of Psychiatric Research*. 2021 Dec;144:54–8.
- 409 18. Huang MC, Chen CH, Liu TH, Chung AN, Liu YL, Quednow BB, et al. Comorbidity of
410 ketamine dependence with major depressive disorder increases the vulnerability to
411 neuroaxonal pathology. *Journal of Psychiatric Research*. 2023 Feb;158:360–4.
- 412 19. Bavato F, Cathomas F, Klaus F, Gütter K, Barro C, Maceski A, et al. Altered neuroaxonal
413 integrity in schizophrenia and major depressive disorder assessed with neurofilament light
414 chain in serum. *Journal of Psychiatric Research*. 2021 Aug;140:141–8.
- 415 20. Michel M, Fiebich BL, Kuzior H, Meixensberger S, Berger B, Maier S, et al. Increased
416 GFAP concentrations in the cerebrospinal fluid of patients with unipolar depression. *Transl*
417 *Psychiatry*. 2021 Jun;11(1):308.
- 418 21. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The
419 PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021
420 Mar 29;n71.
- 421 22. American Psychiatric Association, American Psychiatric Association. *Diagnostic and*
422 *statistical manual of mental disorders: DSM-IV-TR*. 4th ed., text revision. Washington, DC:
423 American Psychiatric Association; 2000. 943 p.
- 424 23. American Psychiatric Association, American Psychiatric Association, editors. *Diagnostic*
425 *and statistical manual of mental disorders: DSM-5*. 5th ed. Washington, D.C: American
426 Psychiatric Association; 2013. 947 p.

- 427 24. Sheehan DV. The Mini-International Neuropsychiatric Interview (M.I.N.I.): The
428 Development and Validation of a Structured Diagnostic Psychiatric Interview for DSM-IV
429 and ICD-10. *J Clin Psychiatry*. 1998;
- 430 25. World Health Organization. International statistical classification of diseases and related
431 health problems (11th ed.) [Internet]. 2019. Available from: <https://icd.who.int/>
- 432 26. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. *Cochrane*
433 *Handbook for Systematic Reviews of Interventions* [Internet]. 1st ed. Wiley; 2019 [cited
434 2024 Jan 12]. Available from:
435 <https://onlinelibrary.wiley.com/doi/book/10.1002/9781119536604>
- 436 27. Luo D, Wan X, Liu J, Tong T. Optimally estimating the sample mean from the sample size,
437 median, mid-range, and/or mid-quartile range. *Stat Methods Med Res*. 2018 Jun;27(6):1785–
438 805.
- 439 28. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from
440 the sample size, median, range and/or interquartile range. *BMC Medical Research*
441 *Methodology*. 2014 Dec 19;14(1):135.
- 442 29. Higgins J, Li T, Deeks J. Chapter 6: Choosing effect measures and computing estimates of
443 effect [Internet]. Vol. *Cochrane Handbook for Systematic Reviews of Interventions* version
444 6.0. Cochrane; 2019 [cited 2023 Dec 24]. Available from:
445 <https://training.cochrane.org/handbook/current/chapter-06>
- 446 30. Rohatgi A. *WebPlotDigitizer* [Internet]. Pacifica, California, USA; 2022 [cited 2023 Oct 31].
447 Available from: <https://automeris.io/WebPlotDigitizer>
- 448 31. Simrén J, Andreasson U, Gobom J, Suarez Calvet M, Borroni B, Gillberg C, et al.
449 Establishment of reference values for plasma neurofilament light based on healthy
450 individuals aged 5–90 years. *Brain Communications*. 2022 Jul 4;4(4):fcac174.
- 451 32. Wells G, Shea B, O’Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle -Ottawa
452 scale (NOS) for assessing the quality if nonrandomized studies in meta-analyses [Internet].
453 2011 [cited 2024 Jan 12]. Available from:
454 http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp
- 455 33. Balduzzi S, Rucker G, Schwarzer G. How to perform a meta-analysis with R: a practical
456 tutorial. *Evid Based Mental Health*. 2019 Nov;22(4):153–60.
- 457 34. Serrano-Castro PJ, Garzón-Maldonado FJ, Casado-Naranjo I, Ollero-Ortiz A, Mínguez-
458 Castellanos A, Iglesias-Espinosa M, et al. The cognitive and psychiatric subacute impairment
459 in severe Covid-19. *Sci Rep*. 2022 Mar 3;12(1):3563.
- 460 35. Fernqvist A, Alexiou E, Zetterberg H, Howner K, Nilsson T, Andiné P. Plasma
461 neurofilament light chain protein is not increased in forensic psychiatric populations: a pilot
462 study. *Front Psychiatry*. 2023 May 4;14:1176266.

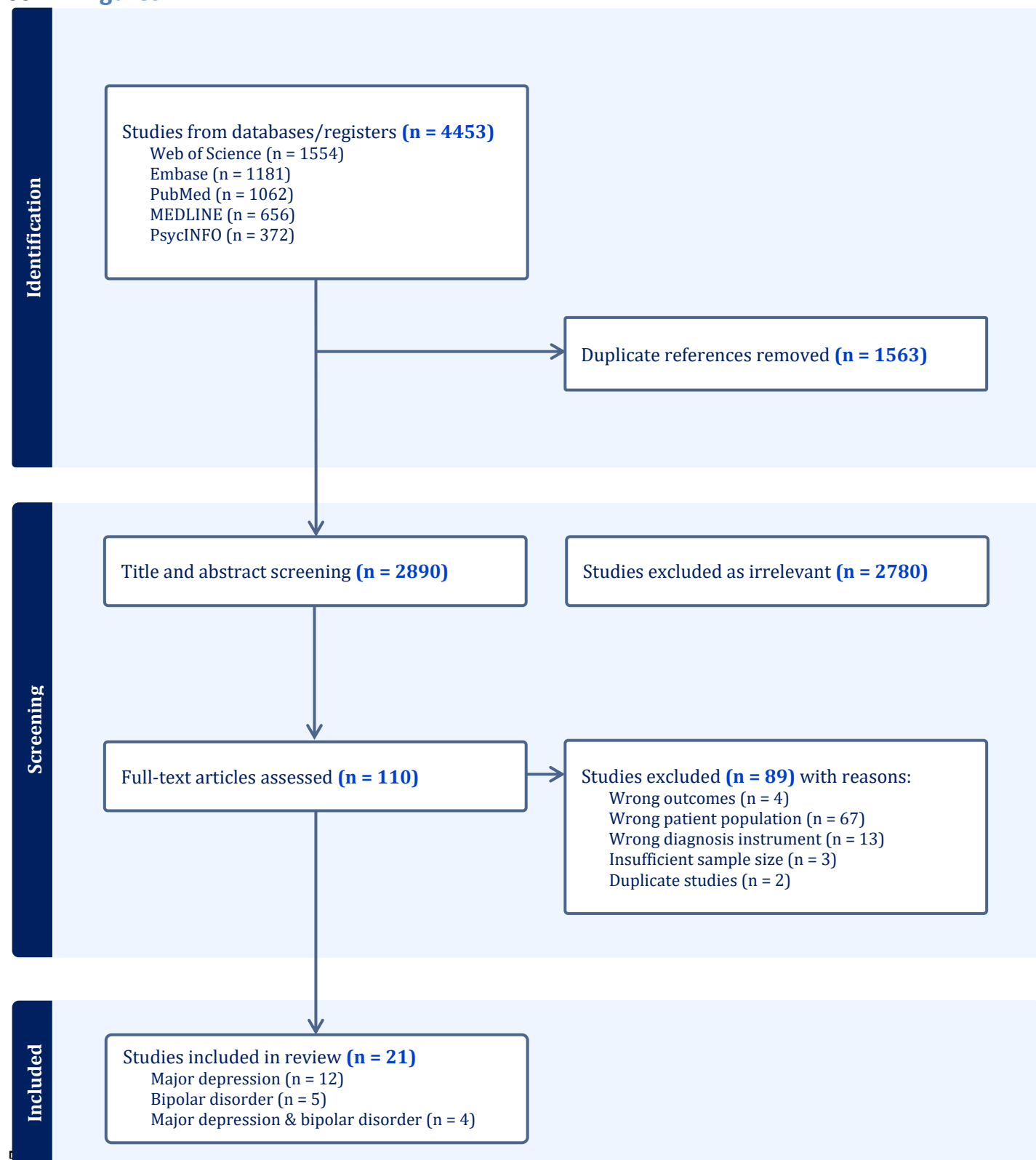
- 463 36. Zachrisson OCG, Balldin J, Ekman R, Naesh O, Rosengren L, Ågren H, et al. No evident
464 neuronal damage after electroconvulsive therapy. *Psychiatry Research*. 2000 Oct;96(2):157–
465 65.
- 466 37. Ferensztajn-Rochowiak E, Tarnowski M, Samochowiec J, Michalak M, Ratajczak MZ,
467 Rybakowski JK. Increased mRNA expression of peripheral glial cell markers in bipolar
468 disorder: The effect of long-term lithium treatment. *European Neuropsychopharmacology*.
469 2016 Sep;26(9):1516–21.
- 470 38. Poletaev AB, Morozov SG, Gnedenko BB, Zlunikin VM, Korzhenevsky DA. Serum Anti-
471 SIOOb, Anti-GFAP and Anti-NGF Autoantibodies of IgG Class in Healthy Persons and
472 Patients with Mental and Neurological Disorders. *Autoimmunity*. 2000 Jan;32(1):33–8.
- 473 39. van Engelen MPE, Heijst H, Willemse EAJ, Oudega ML, Vermunt L, Scheltens P, et al.
474 Urine as matrix for analysis of neurofilament light chain is not suitable to distinguish
475 frontotemporal dementia from psychiatric diseases. *Brain Communications*. 2023;5(2).
- 476 40. Wallensten J, Mobarrez F, Åsberg M, Borg K, Beser A, Wilczek A, et al. Isoforms of soluble
477 vascular endothelial growth factor in stress-related mental disorders: a cross-sectional study.
478 *Sci Rep*. 2021 Aug 17;11(1):16693.
- 479 41. Katisko K, Cajanus A, Jääskeläinen O, Kontkanen A, Hartikainen P, Korhonen VE, et al.
480 Serum neurofilament light chain is a discriminative biomarker between frontotemporal lobar
481 degeneration and primary psychiatric disorders. *J Neurol*. 2020 Jan;267(1):162–7.
- 482 42. Katisko K, Cajanus A, Huber N, Jääskeläinen O, Kokkola T, Kärkkäinen V, et al. GFAP as a
483 biomarker in frontotemporal dementia and primary psychiatric disorders: diagnostic and
484 prognostic performance. *J Neurol Neurosurg Psychiatry*. 2021 Dec;92(12):1305–12.
- 485 43. Bai YM, Liu YL, Kuo HW, Tsai SJ, Hsu JW, Huang KL, et al. Procollagen type 1 N-
486 terminal propeptide, neurofilament light chain, proinflammatory cytokines, and cognitive
487 function in bipolar and major depressive disorders: An exploratory study of brain–bone axis
488 and systemic inflammation. *Journal of Psychiatric Research*. 2023 Feb;158:403–8.
- 489 44. Besse M, Belz M, Folsche T, Vogelgsang J, Methfessel I, Steinacker P, et al. Serum
490 neurofilament light chain (NFL) remains unchanged during electroconvulsive therapy. *The*
491 *World Journal of Biological Psychiatry*. 2020 Feb 7;21(2):148–54.
- 492 45. Chen MH, Liu YL, Kuo HW, Tsai SJ, Hsu JW, Huang KL, et al. Neurofilament light chain is
493 a novel biomarker for major depression and related executive dysfunction. *International*
494 *Journal of Neuropsychopharmacology*. 2022;25(2):99–105.
- 495 46. Jiang L, Shen Z, Cheng Y, Lu J, He B, Xu J, et al. Elevated serum neurofilament levels in
496 young first-episode and medication-naïve major depressive disorder patients with alterative
497 white matter integrity. *Psychiatry Research: Neuroimaging*. 2021 Nov;317:111351.

- 498 47. Ramezani M, Simani L, Fard MG, Abbaszadeh F, Shadnia S. Increased levels of
499 neurofilament light chain in suicide attempters' serum. *Translational Neuroscience*. 2022
500 Aug 5;13(1):218–23.
- 501 48. Wallensten J, Mobarrez F, Åsberg M, Borg K, Beser A, Wilczek A, et al. Plasma levels of
502 S100B and neurofilament light chain protein in stress-related mental disorders. *Sci Rep*.
503 2022 May 18;12(1):8339.
- 504 49. Aggio V, Fabbella L, Finardi A, Mazza EB, Colombo C, Falini A, et al. Neurofilaments
505 light: Possible biomarker of brain modifications in bipolar disorder. *Journal of Affective*
506 *Disorders*. 2022 Mar;300:243–8.
- 507 50. Al Shweiki MR, Steinacker P, Oeckl P, Hengerer B, Danek A, Fassbender K, et al.
508 Neurofilament light chain as a blood biomarker to differentiate psychiatric disorders from
509 behavioural variant frontotemporal dementia. *Journal of Psychiatric Research*.
510 2019;113(February):137–40.
- 511 51. Chen WY, Huang MC, Chiu CC, Cheng YC, Kuo CJ, Chen PY, et al. The interactions
512 between vitamin D and neurofilament light chain levels on cognitive domains in bipolar
513 disorder. *BJPsych open*. 2022 Nov;8(6):e207.
- 514 52. Rolstad S, Jakobsson J, Sellgren C, Ekman CJ, Blennow K, Zetterberg H, et al. Cognitive
515 performance and cerebrospinal fluid biomarkers of neurodegeneration: a study of patients
516 with bipolar disorder and healthy controls. *PloS one*. 2015;10(5):e0127100–e0127100.
- 517 53. Barro C, Benkert P, Disanto G, Tsagkas C, Amann M, Naegelin Y, et al. Serum
518 neurofilament as a predictor of disease worsening and brain and spinal cord atrophy in
519 multiple sclerosis. *Brain*. 2018 Aug 1;141(8):2382–91.
- 520 54. Bridel C, Van Wieringen WN, Zetterberg H, Tijms BM, Teunissen CE, Alvarez-Cermeño
521 JC, et al. Diagnostic Value of Cerebrospinal Fluid Neurofilament Light Protein in
522 Neurology: A Systematic Review and Meta-analysis. *JAMA Neurology*. 2019
523 Jun;76(9):1035–48.
- 524 55. Eratne D, Keem M, Lewis C, Kang M, Walterfang M, Farrand S, et al. Cerebrospinal fluid
525 neurofilament light chain differentiates behavioural variant frontotemporal dementia
526 progressors from non-progressors. *Journal of the Neurological Sciences*. 2022
527 Nov;442:120439.
- 528 56. Cuthbert BN. Research Domain Criteria (RDoC): Progress and Potential. *Curr Dir Psychol*
529 *Sci*. 2022 Apr 1;31(2):107–14.

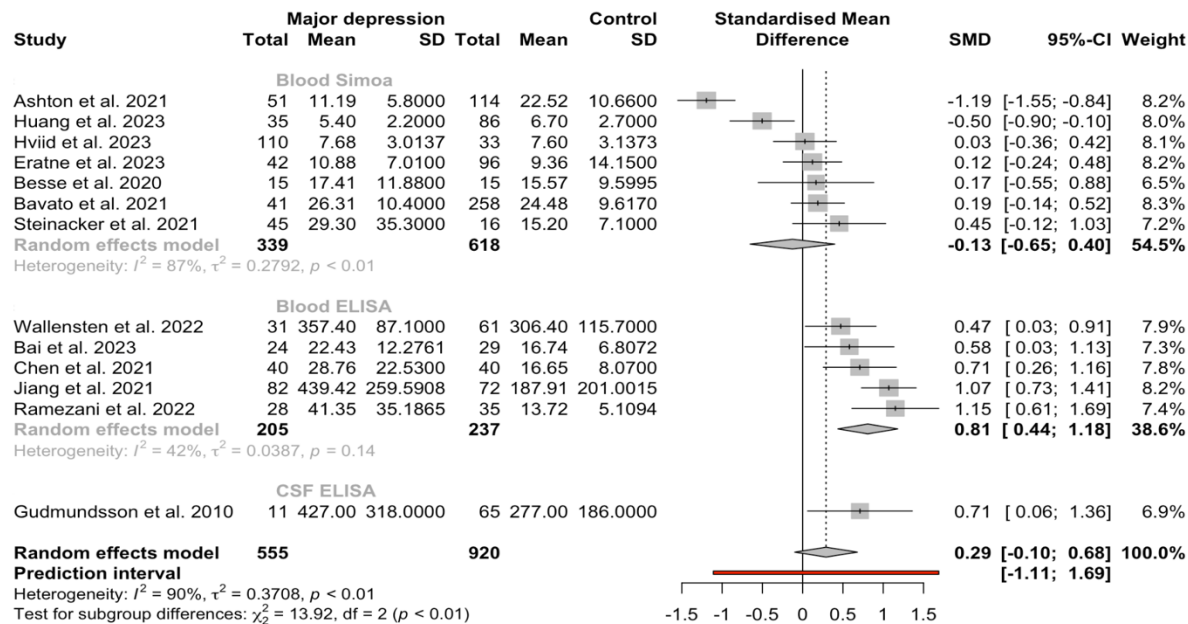
530

531

532 **Figures**



534 *Figure 1. PRISMA Diagram for study selection*



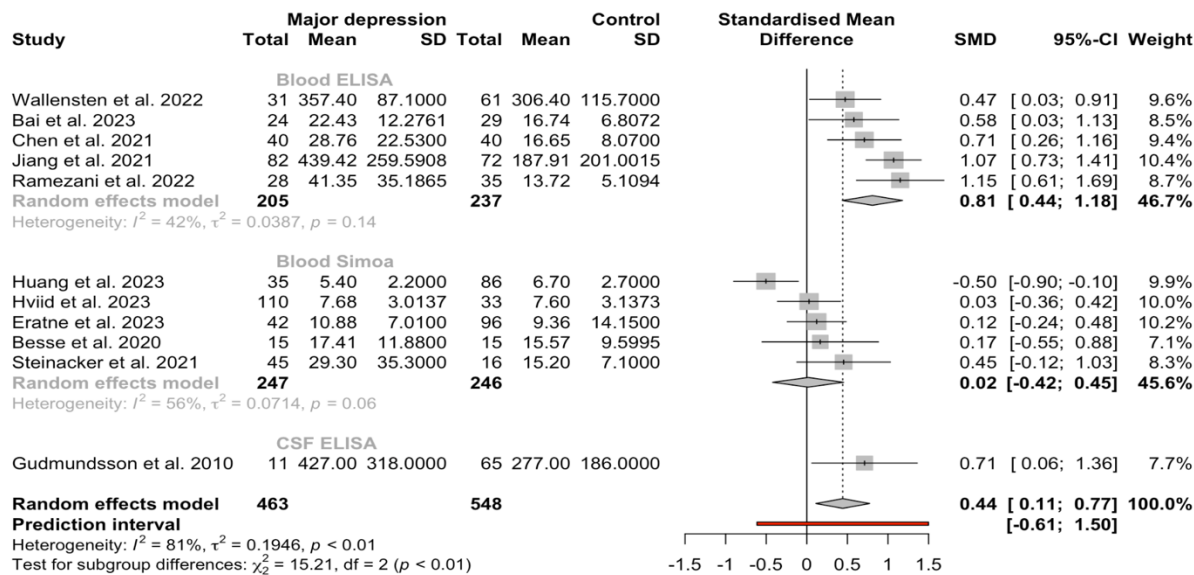
535

536 *Figure 2a. Forest plot of NfL in major depression (all studies)*

537

538

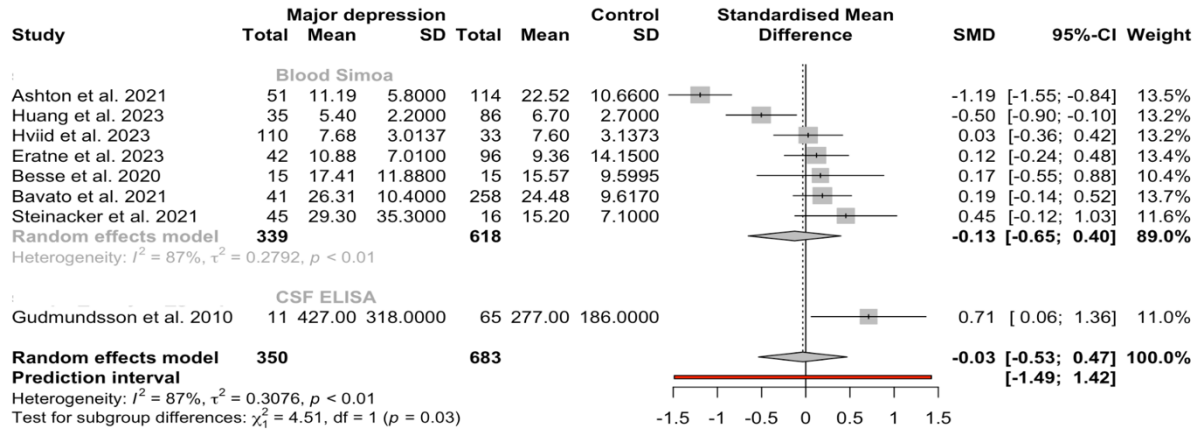
539



540

541 *Figure 2b. Forest plot of NfL in major depression (age-matched studies only)*

542



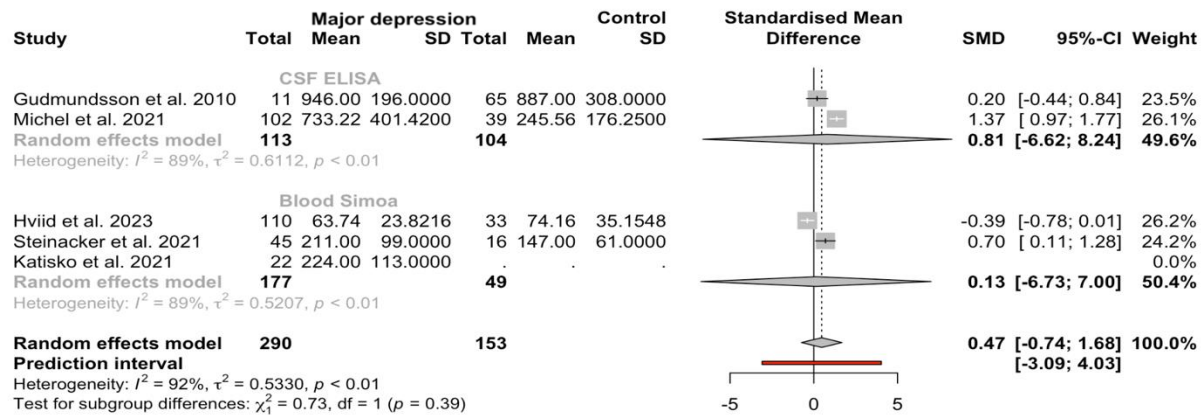
543

544 *Figure 2c. Forest plot of NfL in major depression (Simoa and CSF studies only)*

545

546

547



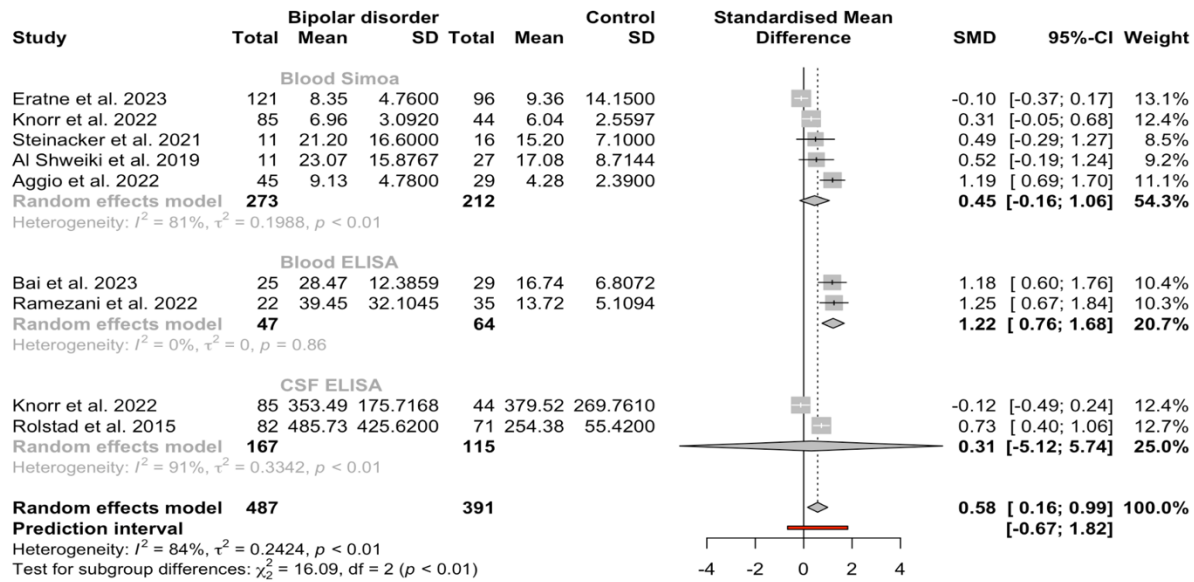
548

549 *Figure 2d. Forest plot of GFAP in major depression*

550

551

552



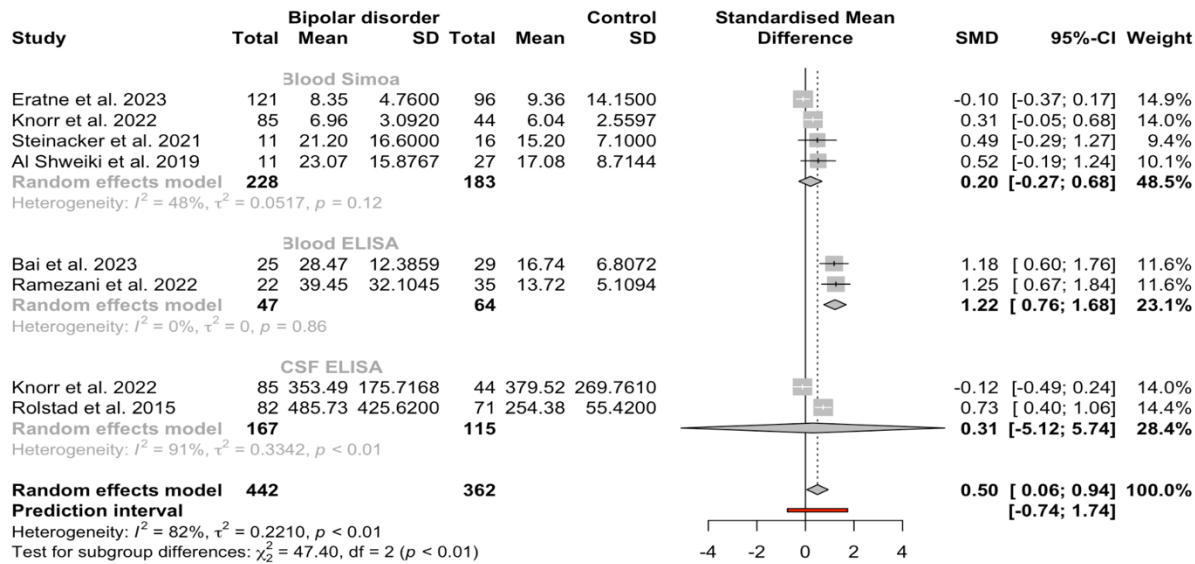
553

554 *Figure 3a. Forest plot of NfL in bipolar disorder (all studies)*

555

556

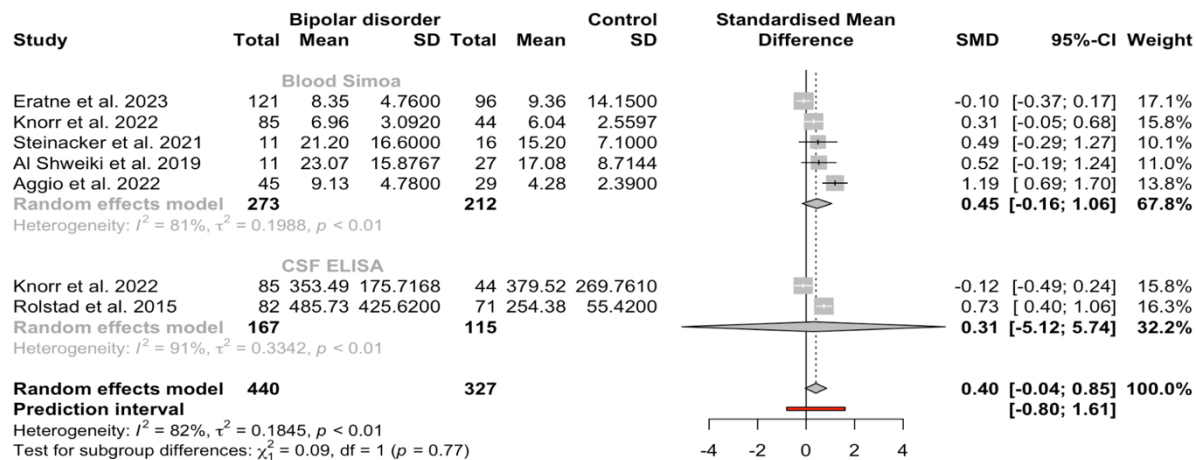
557



558

559 *Figure 3b. Forest plot of NfL in bipolar disorder (age-matched studies only)*

560



561

562 *Figure 3c. Forest plot of NfL in bipolar disorder (CSF and Simoa only)*

563

564

565

Table 1. Characteristics of studies included: NFL in major depression

Authors and year	Country	Sample Type	Analysis type	Diagnosis classification	Mood state	Sample size		Age Mean (SD)		Sex (Female %)		NFL concentration ^a		Total NOS Score
						Case	Control	Case	Control	Case	Control	Case	Control	
Ashton et al, 2021 (6)	England	Plasma	Simoa	DSM-5	Not specified	51	114	32.3 (6.9)	64.0 (14.1)	69%	55%	11.19 (5.8)	22.5 (10.7)	3
Bai et al, 2023 (43)	Taiwan	Serum	HRP ELISA	DSM-5	HDRS ≤16	24	29	30.8 (13.0)	30.9 (10.7)	67%	62%	22.4 (12.3)**	16.7 (6.8)	5
Bavato et al, 2021 (19)	Switzerland	Plasma	Simoa	DSM-5	Clinically stable	41	258	36.0 (10.9)	Unavailable	59%	NA%	26.3 (10.4)	24.5 (9.6)	5
Besse et al, 2020 (44)	Germany	Serum	Simoa	ICD-10	Moderate to severe depressive episode	15	15	49.2 (14.0)	"± 3 years"	73%	NA%	17.4 (11.88)	Mean = 15.57 95% CI*: 10.69, 20.40	5
Chen et al, 2022 (45)	Taiwan	Plasma	HRP ELISA	DSM-5	Not specified	40	40	28.3 (14.4)	28.3 (14.1)	68%	68%	28.8 (22.5)**	16.7 (8.1)	5
Eratne et al, 2023 (8)	Australia	Plasma	Simoa	DSM-IV	Depressive phase ≥ moderate severity	42	96	55.0 (12.9)	44.5 (14.4)	57%	52%	10.9 (7.0)**	9.4 (14.2)**	4
Gudmundsson et al, 2010 (14)	Sweden	CSF	ELISA Sandwich	DSM-III-R	Not specified	11	65	Combined: 73.9 (3.2)		100%	100%	427 (318)	277 (186)	5
Huang et al, 2023 (18)	Taiwan	Serum	Simoa	DSM-IV-TR	Outpatients	35	86	34.7 (10.3)	33.4 (6.2)	77%	13%	5.4 (2.2)	6.7 (2.7)	5
Hviid et al, 2023 (15)	Denmark	Serum	Simoa	DSM-IV	Treatment naive	110	33	42 (11)	39 (14)	65%	70%	7.7 (3.0)*	7.6 (3.1)*	5
Jiang et al 2021, (46)	China	Serum	ELISA Sandwich	DSM-IV	First episode, HDRS > 16	82	72	34 (11)	34 (12)	61%	49%	Median = 405.8 IQR: 281.5–625.5	Median = 143.5 IQR: 73.6–339.3	5
Katisko et al, 2020 (41)	Finland	Serum	Simoa	ICD-10	Included moderate, severe & psychotic depression	16	0	55.7 (6.7)	NA	56%	NA	12.4 (4.3)**	NA	3
Ramezani et al, 2022 (47)	Iran	Serum	ELISA	DSM-5	Suicide attempt	28	35	M: 26.28 (9.37) F: 26.61 (9.49)	M: 28.10 (8.54) F: 30.76 (5.12)	75%	71%	M: 55.42 (33.63) F: 36.66 (35.21)	M: 14.11 (4.38) F: 13.57 (5.45)	5
Steinacker et al, 2021 (17)	Germany	Serum	Simoa	DSM-5	Moderate-severe depressive episode	45	16	Median=48 Range 19-69	Median=45 Range 27-64	64%	75%	29.3 (35.3)	15.2 (7.1)	4
Wallensten et al, 2022 (48)	Sweden	Plasma	ELISA	ICD-10	Not specified	31	61	40.3 (10.8)	42.9 (9.5)	84%	85%	357.4 (87.1)	306.4 (115.7)	6

^aCSF = ng/L, plasma/serum = pg/mL; *estimates using WebPlotDigitizer; **data obtained directly from corresponding author

CI = confidence interval, F = females, HDRS = Hamilton Depression Rating Scale, M = males, MADRS = Montgomery-Asberg Depression Rating Scale, NA = not applicable, NOS = Newcastle-Ottawa Scale, YMRS = Young Mania Rating Scale

Table 2. Characteristics of studies included: NFL in bipolar disorder

Authors and year	Country	Sample Type	Analysis type	Diagnosis classification	Mood state	Sample size		Age Mean (SD)		Sex (Female %)		NFL concentration [^]		Total NOS Score
						Case	Control	Case	Control	Case	Control	Case	Control	
Aggio et al, 2022 (49)	Italy	Plasma	Simoa	DSM-IV	Depressed	45	29	48.2 (11.9)	41.8 (10.2)	78%	55%	9.1 (4.8)	4.28 (2.39)	3
Al Shweiki et al, 2019 (50)	Germany	Serum	Simoa	DSM-5	Manic, depressive and mixed states	11	27	Median = 51.4 IQR: 33.5-58.1	Median = 46.8 IQR: 39.1-54.1	36%	63%	Median = 17.8 IQR: 12.6 - 23.1	Median = 17.8 IQR: 11.3 - 19.1	4
Bai et al, 2023 (43)	Taiwan	Serum	HRP ELISA	DSM-5	YMRS ≤12	25	29	33.2 (13.2)	30.9 (10.7)	60%	62%	28.5 (12.4)**	16.7 (6.8)	5
Chen et al, 2022 (51)	Taiwan	Serum	Simoa	DSM-5	Euthymic	100		46.5 (11.7)	NA	52%	NA%	11.1 (8.1)**	NA	5
Eratne et al, 2023 (8)	Australia	Plasma	Simoa	DSM-IV	Depressive phase of at least moderate severity	121	96	44.1 (12.18)	44.48 (14.35)	62%	52%	8.4 (4.8)**	9.4 (14.2)**	5
Knorr et al, 2022 (16)	Denmark	CSF & Plasma	ELISA & Simoa	ICD-10	In remission (YMRS <=8; HDRS <= 8)	85	44	Median = 33 IQR: 26-42	Median = 30 IQR: 25-42	51%	52%	CSF: Median = 332 IQR: 246-479 Plasma: Median = 6.81 IQR: 4.97 - 9.07	CSF: Median = 354 IQR: 214-566 Plasma: Median = 5.73 IQR: 4.50 - 7.84	6
Ramezani et al, 2022 (47)	Iran	Serum	ELISA	DSM-5	Suicide attempt	22	35	M: 33.00 (15.18) F: 31.72 (13.10)	M: 28.10 (8.54) F: 30.76 (5.12)	82%	71%	M: 46.50 (38.27) F: 37.88 (31.63)	M: 14.11 (4.38) F: 13.57 (5.45)	5
Rolstad et al, 2015 (52)	Sweden	CSF	ELISA	DSM-IV	Euthymic (MADRS < 14, YMRS < 14)	82	71	38.3 (12.5)	37.8 (14.6)	59%	62%	485.7 (425.6)	254.4 (55.4)	5
Steinacker et al, 2021 (17)	Germany	Serum	Simoa	DSM-5	Manic, depressive and mixed states	11	16	Median = 48 Range: 19-69	Median = 45 Range: 27-64	27%	75%	21.2 (16.6)	15.2 (7.1)	4

[^]CSF = ng/L, plasma/serum = pg/mL; *estimates using WebPlotDigitizer; **data obtained directly from corresponding author

CI = confidence interval, F = females, HDRS = Hamilton Depression Rating Scale, M = males, MADRS = Montgomery-Asberg Depression Rating Scale, NA = not applicable, NOS = Newcastle-Ottawa Scale, YMRS = Young Mania Rating Scale

Table 3. Characteristics of studies included: GFAP in major depression and bipolar disorder

Author	Country	Sample Type	Analysis type	Diagnosis classification	Mood state	Sample size		Age Mean (SD)		Sex (Female %)		GFAP concentration [^]		Total NOS Score
						Case	Control	Case	Control	Case	Control	Case	Control	
Major depressive disorder														
Gudmundsson et al, 2010 (14)	Sweden	CSF	ELISA Sandwich	DSM-III-R	Not specified	11	65	Cases + controls: 73.9 (3.2)		100%	100%	946 (196)	887 (308)	5
Hviid et al, 2023 (15)	Denmark	Serum	Simoa	DSM-IV	Treatment naïve	110	33	42 (11)	39 (14)	65%	70%	63.7 (23.8)*	74.2 (35.2)*	5
Katisko et al, 2021 (42)	Finland	Serum	Simoa	ICD-10	Included moderate, severe and psychotic depression	22	0	55.7 (6.3)	NA	55%	NA	224 (113)**	NA	3
Michel et al, 2021 (20)	Germany	CSF	ELISA	ICD-10	Inpatients	102	39	44.2 (13.6)	34.6 (12.0)	54%	85%	733 (401)	245.6 (176.3)	5
Steinacker et al, 2021 (17)	Germany	Serum	Simoa	DSM-5	Moderate-severe depressive episode	45	16	Median = 48 Range 19-69	Median = 45 Range 27-64	64%	75%	211 (99)	147.0 (61.0)	4
Bipolar disorder														
Steinacker et al, 2021 (17)	Germany	Serum	Simoa	DSM-5	Manic, depressive and mixed states	11	16	Median = 48 Range 19-69	Median = 45 Range 27-64	27%	75%	125 (50)	147.0 (61.0)	4

[^]CSF = ng/L, plasma/serum = pg/mL; *estimates using WebPlotDigitizer; **unpublished data obtained directly from corresponding author
 F = females, M = males, NA = not applicable, NOS = Newcastle-Ottawa Scale

eMethods - Search strategy for respective databases

Database: Ovid MEDLINE(R): inception to June 27, 2023

Search Strategy:

1. Affective Symptoms/ or exp Catatonia/ or exp Depression/ or exp Mental Fatigue/ or exp Mood Disorders/ or exp Anxiety/ or exp Anxiety Disorders/ or Mania/ or Depression, Chemical/ or Cognitive Dysfunction/ or depress*.mp. or bipolar*.mp. or anxiety*.mp. or *Mental Disorders/ or psychiatric disorder*.mp. (1080981)
2. Glial Fibrillary Acidic Protein/ or Neurofilament Proteins/ (22084)
3. (neurofilament light* or glial fibrillary acidic protein*).mp. (27935)
4. 2 or 3 (32209)
5. 1 and 4 (1445)
6. (5 not rats not animals/) or (5 and humans/) (656)

Database: Web of Science Search Strategy: inception to June 27, 2023

Searches:

- 1: TS=(bipolar*) OR TS=(mania*) OR TS=(depressi*) OR TS=(anxiety*) OR TS=(psychiatr* disord*) OR TS=(mental disorder*) 1230138
- 2: TS=(neurofilament*) OR TS=(Glial fibrillary acidic protein*) OR TS=(NF-L) OR TS=(GFAP) 40397
- 3: #1 AND #2 and Article or Early Access or Proceeding Paper or Data Paper or Note or Letter (Document Types) 1554

Embase search: inception to May 31 2023

1. exp glial fibrillary acidic protein/ 36983
2. exp neurofilament/ or exp neurofilament protein/ 19626
3. exp bipolar depression/ or exp bipolar disorder/ or exp mania/ 96498
4. exp bipolar depression/ or exp bipolar disorder/ or exp mania/ 96498
5. exp anxiety/ or exp anxiety assessment/ or exp anxiety disorder/ 611058
6. 1 or 2 53764
7. 3 or 4 or 5 1070529
8. 6 and 7 1554
9. limit 8 to human 809

medRxiv preprint doi: <https://doi.org/10.1101/2024.03.07.24303938>; this version posted March 9, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY-ND 4.0 International license](#).

Pubmed: inception to May 31 2023

((bipolar* OR depress* OR anxie*)) AND (neurofilament light* OR glial fibrillary acidic protein*) 1059 hits

PSYCHInfo: Inception to May 31 2023

- 1 (Neurofilaments or neurofilament protein L or Neurofilament Proteins or Glial Fibrillary Acidic Protein or (neurofilament light* or glial fibrillary acidic protein*).mp. (4714)
- 2 exp Chronic Fatigue Syndrome/ or exp Chronic Mental Illness/ or exp Affective Disorders/ or exp Anxiety Disorders/ or exp Bipolar Disorder/ or exp Chronic Mental Illness/ or exp Affective Psychosis/ or exp Acute Stress Disorder/ or exp Adjustment Disorders/ or exp Prolonged Grief Disorder/ or exp Fatigue/ or exp Psychiatric Symptoms/ or exp Treatment Resistant Disorders/ or exp Treatment Resistant Depression/ or exp Atypical Depression/ or exp Acute Stress Disorder/ or depress*.mp. or bipolar*.mp. or anxiety*.mp. or psychiatric disorder*.mp. or mental disorder*.mp. (779119)
- 3 1 and 2 (372)
- 4 limit 7 to human (150)

eTable 1 - Newcastle-Ottawa Scale scores for the included studies

Study Author	Selection	Comparability	Outcome	Total
	- Representativeness of the cohort - Sample size justification - Ascertainment of mood or anxiety disorder	- Controls for age between case and controls - Controls for additional covariates (sex, weight, BMI, renal function) between cases and controls	- Validated measure of biomarker - Appropriate statistical analysis which is well described	
<i>Aggio et al, 2022 (1)</i>	1	0	2	3
<i>Al Shweiki et al, 2019 (2)</i>	1	1	2	4
<i>Ashton et al, 2021 (3)</i>	1	0	2	3
<i>Bai et al, 2023 (4)</i>	1	2	2	5
<i>Bavato et al, 2021 (5)</i>	1	2	2	5
<i>Besse et al, 2020 (6)</i>	1	2	2	5
<i>Chen et al, 2022b (7)</i>	1	2	2	5
<i>Chen et al, 2022a (8)</i>	1	2	2	5
<i>Chen et al, 2022b (9)</i> (bipolar)	1	2	2	5
<i>Eratne et al, 2023 (9)</i> (depression)	1	1	2	4
<i>Gudmundsson et al, 2010 (10)</i>	1	2	2	5
<i>Huang et al, 2023 (11)</i>	1	2	2	5
<i>Hviid et al, 2023 (12)</i>	1	2	2	5
<i>Jiang et al 2021, (13)</i>	1	2	2	5
<i>Katisko et al, 2020 (14)</i>	1	0	2	3
<i>Katisko et al, 2021 (15)</i>	1	0	2	3
<i>Knorr et al, 2022 (16)</i>	3	2	1	6
<i>Michel et al, 2021 (17)</i>	1	2	2	5
<i>Ramezani et al, 2022 (18)</i>	1	2	2	5
<i>Rolstad et al, 2015 (19)</i>	1	2	2	5
<i>Steinacker et al, 2021 (20)</i>	1	1	2	4
<i>Wallensten et al, 2022 (21)</i>	2	2	2	6

medRxiv preprint doi: <https://doi.org/10.1101/2024.03.07.24303938>; this version posted March 9, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY-ND 4.0 International license](https://creativecommons.org/licenses/by-nd/4.0/).

eReferences

1. Aggio V, Fabbella L, Finardi A, Mazza EB, Colombo C, Falini A, et al. Neurofilaments light: Possible biomarker of brain modifications in bipolar disorder. *Journal of Affective Disorders*. 2022 Mar;300:243–8.
2. Al Shweiki MR, Steinacker P, Oeckl P, Hengerer B, Danek A, Fassbender K, et al. Neurofilament light chain as a blood biomarker to differentiate psychiatric disorders from behavioural variant frontotemporal dementia. *Journal of Psychiatric Research*. 2019;113(February):137–40.
3. Ashton NJ, Janelidze S, Al Khleifat A, Leuzy A, van der Ende EL, Karikari TK, et al. A multicentre validation study of the diagnostic value of plasma neurofilament light. *Nat Commun*. 2021 Jun 7;12(1):3400.
4. Bai YM, Liu YL, Kuo HW, Tsai SJ, Hsu JW, Huang KL, et al. Procollagen type 1 N-terminal propeptide, neurofilament light chain, proinflammatory cytokines, and cognitive function in bipolar and major depressive disorders: An exploratory study of brain–bone axis and systemic inflammation. *Journal of Psychiatric Research*. 2023 Feb;158:403–8.
5. Bavato F, Cathomas F, Klaus F, Gütter K, Barro C, Maceski A, et al. Altered neuroaxonal integrity in schizophrenia and major depressive disorder assessed with neurofilament light chain in serum. *Journal of Psychiatric Research*. 2021 Aug;140:141–8.
6. Besse M, Belz M, Folsche T, Vogelgsang J, Methfessel I, Steinacker P, et al. Serum neurofilament light chain (NFL) remains unchanged during electroconvulsive therapy. *The World Journal of Biological Psychiatry*. 2020 Feb 7;21(2):148–54.
7. Chen WY, Huang MC, Chiu CC, Cheng YC, Kuo CJ, Chen PY, et al. The interactions between vitamin D and neurofilament light chain levels on cognitive domains in bipolar disorder. *BJPsych open*. 2022 Nov;8(6):e207.

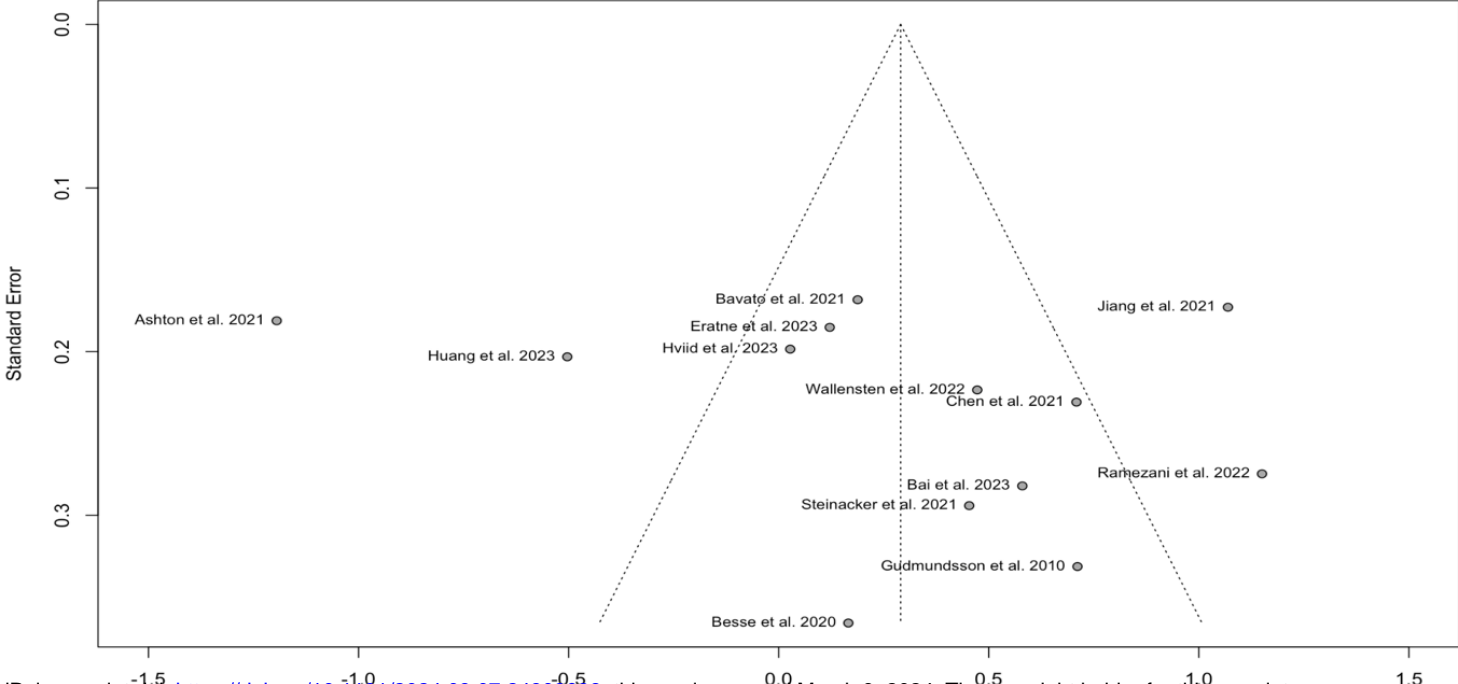
8. Chen MH, Liu YL, Kuo HW, Tsai SJ, Hsu JW, Huang KL, et al. Neurofilament light chain is a novel biomarker for major depression and related executive dysfunction. *International Journal of Neuropsychopharmacology*. 2022;25(2):99–105.
9. Eratne D, Kang M, Malpas C, Simpson-Yap S, Lewis C, Dang C, et al. Plasma neurofilament light in behavioural variant frontotemporal dementia compared to mood and psychotic disorders. *Aust N Z J Psychiatry*. 2023 Jul 21;00048674231187312.
10. Gudmundsson P, Skoog I, Waern M, Blennow K, Zetterberg H, Rosengren L, et al. Is there a CSF biomarker profile related to depression in elderly women? *Psychiatry Research*. 2010;176(2–3):174–8.
11. Huang MC, Chen CH, Liu TH, Chung AN, Liu YL, Quednow BB, et al. Comorbidity of ketamine dependence with major depressive disorder increases the vulnerability to neuroaxonal pathology. *Journal of Psychiatric Research*. 2023 Feb;158:360–4.
12. Hviid CVB, Benros ME, Krogh J, Nordentoft M, Christensen SH. Serum glial fibrillary acidic protein and neurofilament light chain in treatment-naïve patients with unipolar depression. *Journal of Affective Disorders*. 2023 Oct 1;338:341–8.
13. Jiang L, Shen Z, Cheng Y, Lu J, He B, Xu J, et al. Elevated serum neurofilament levels in young first-episode and medication-naïve major depressive disorder patients with alternative white matter integrity. *Psychiatry Research: Neuroimaging*. 2021 Nov;317:111351.
14. Katisko K, Cajanus A, Jääskeläinen O, Kontkanen A, Hartikainen P, Korhonen VE, et al. Serum neurofilament light chain is a discriminative biomarker between frontotemporal lobar degeneration and primary psychiatric disorders. *J Neurol*. 2020 Jan;267(1):162–7.
15. Katisko K, Cajanus A, Huber N, Jääskeläinen O, Kokkola T, Kärkkäinen V, et al. GFAP as a biomarker in frontotemporal dementia and primary psychiatric disorders: diagnostic and prognostic performance. *J Neurol Neurosurg Psychiatry*. 2021 Dec;92(12):1305–12.
16. Kwon J, Simonsen AH, Hye IS, et al. Zetterberg H, Blennow K, Almkvist O, et al. Multiple Alzheimer's disease related biomarkers in bipolar disorder - A longitudinal one-year case-control study. *J Affect Disord*. 2022 Jan 15;297:623–33.
17. Michel M, Fiebich BL, Kuzior H, Meixensberger S, Berger B, Maier S, et al. Increased GFAP concentrations in the cerebrospinal fluid of patients with unipolar depression. *Transl Psychiatry*. 2021 Jun;11(1):308.
18. Ramezani M, Simani L, Fard MG, Abbaszadeh F, Shadnia S. Increased levels of neurofilament light chain in suicide attempters' serum. *Translational Neuroscience*. 2022 Aug 5;13(1):218–23.
19. Rolstad S, Jakobsson J, Sellgren C, Ekman CJ, Blennow K, Zetterberg H, et al. Cognitive performance and cerebrospinal fluid biomarkers of neurodegeneration: a study of patients with bipolar disorder and healthy controls. *PloS one*. 2015;10(5):e0127100–e0127100.
20. Steinacker P, Al Shweiki MR, Oeckl P, Graf H, Ludolph AC, Schönfeldt-Lecuona C, et al. Glial fibrillary acidic protein as blood biomarker for differential diagnosis and severity of major depressive disorder. *Journal of Psychiatric Research*. 2021 Dec;144:54–8.
21. Wallensten J, Mobarrez F, Åsberg M, Borg K, Beser A, Wilczek A, et al. Plasma levels of S100B and neurofilament light chain protein in stress-related mental disorders. *Sci Rep*. 2022 May 18;12(1):8339.

medRxiv preprint doi: <https://doi.org/10.1101/2024.03.07.24303938>; this version posted March 9, 2024. The copyright holder for this preprint

(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY-ND 4.0 International license](https://creativecommons.org/licenses/by-nd/4.0/).

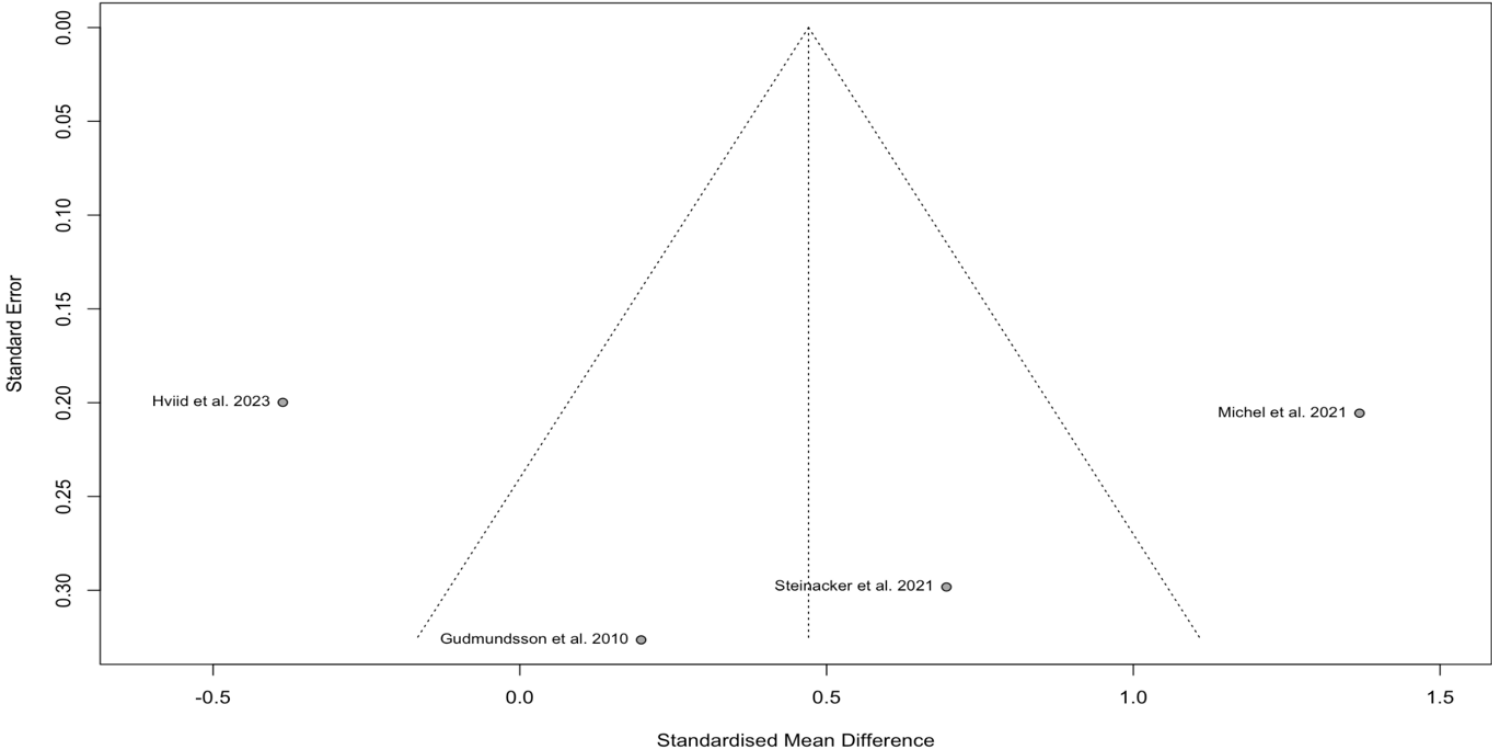
sFigure 1 - Funnel plots

a. NFL in major depression

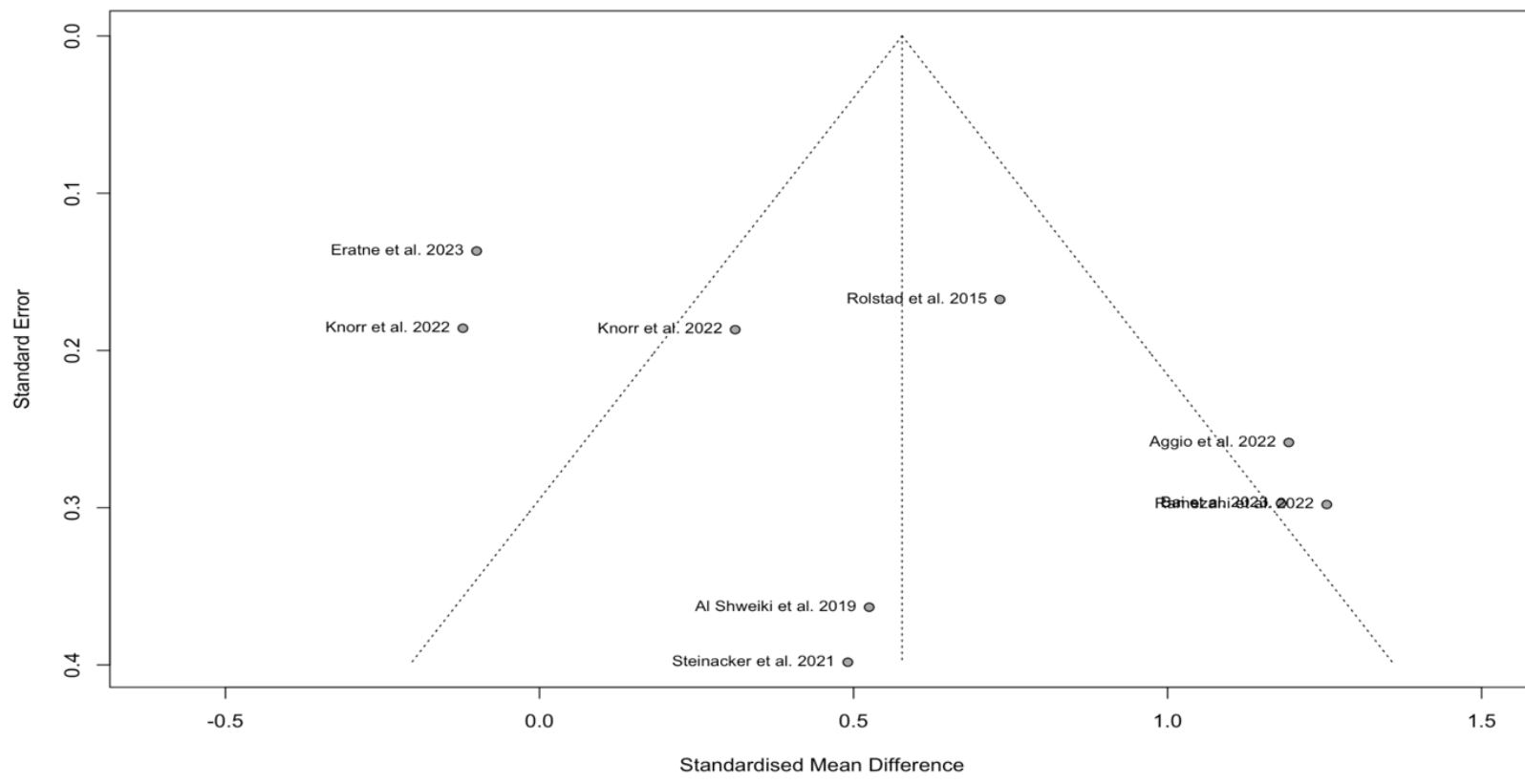


medRxiv preprint doi: <https://doi.org/10.1101/2024.03.07.24303938>; this version posted March 9, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY-ND 4.0 International license](https://creativecommons.org/licenses/by-nd/4.0/).

b. GFAP in major depression



c. NfL in bipolar disorder



medRxiv preprint doi: <https://doi.org/10.1101/2024.03.07.24303938>; this version posted March 9, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY-ND 4.0 International license](https://creativecommons.org/licenses/by-nd/4.0/).



PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Pg 1
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Pg 4-5
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Pg 6/line 89
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Pg 7/line 133
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Pg 7/line 144
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Pg 8/line 156
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Supplementary
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Pg 8/line 164
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Pg 8/line 171
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Pg 8/line 171
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Pg 8/line 171
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Pg 9/line 185
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Pg 8/line 171
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	Pg 7/line 186
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Pg 8/line 173
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Pg 9/line 182
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Pg 9/line 182
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Pg 9/line 205
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	Pg 9/line 205
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Pg 10/line 22
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Pg 9/line 201



PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Figure 1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Pg 10/line 233
Study characteristics	17	Cite each included study and present its characteristics.	Table 1-2
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Supplementary eTable 1
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Pg 11/line 264 Figures 2-3
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Supplementary eTable 1
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	Pg 11/line 263
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Pg 11/line 267
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	Pg 11/line 73
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	Supplementary eFigure 1
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Pg 11/line 268
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Pg 13/line 320
	23b	Discuss any limitations of the evidence included in the review.	Pg 14/line 355
	23c	Discuss any limitations of the review processes used.	Pg 14/line 360
	23d	Discuss implications of the results for practice, policy, and future research.	Pg 15/line 368
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Pg 7/line 143
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Pg 7/line 143
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	N/A
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Pg 2/line 14
Competing interests	26	Declare any competing interests of review authors.	Pg 2/line 25
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	N/A