

Review: predicting psychosis using blood biomarkers

1 **Prognostic models predicting transition to psychotic disorder using blood-**  
2 **based biomarkers: a systematic review and critical appraisal.**

3

4 **Authors:** Jonah F. Byrne<sup>1,2</sup>, BSc; David Mongan<sup>1,3</sup>, MB PhD; Jennifer Murphy, MSc<sup>1</sup>; Colm  
5 Healy<sup>1</sup>, PhD; Melanie Föcking<sup>1,2</sup>, PhD; Mary Cannon<sup>1,2</sup>, MB PhD; David R. Cotter<sup>1,2</sup>, MB  
6 PhD.

7 <sup>1</sup> Department of Psychiatry, Royal College of Surgeons in Ireland, Dublin, Ireland.

8 <sup>2</sup> SFI FutureNeuro Research Centre, Royal College of Surgeons in Ireland, Dublin, Ireland.

9 <sup>3</sup> Centre for Public Health, Queen's University Belfast, United Kingdom.

10

11 **Correspondence:** Jonah Byrne

12 E-mail: [jonahbyrne21@rcsi.ie](mailto:jonahbyrne21@rcsi.ie)

13 Address: Department of Psychiatry, Royal College of Surgeons in Ireland, Education and  
14 Research Centre, Beaumont Hospital, Dublin 9, Ireland

15

Review: predicting psychosis using blood biomarkers

## 16 **Abstract**

### 17 **Background**

18 Accumulating evidence suggests individuals with psychotic disorder show abnormalities in  
19 metabolic and inflammatory processes. Recently, several studies have employed blood-based  
20 predictors in models predicting transition to psychotic disorder in risk-enriched populations. A  
21 systematic review of the performance and methodology of prognostic models using blood-  
22 based biomarkers in the prediction of psychotic disorder from risk-enriched populations is  
23 warranted.

### 24 **Methods**

25 Databases (PubMed, EMBASE and PsycINFO) were searched for eligible texts from 1998 to  
26 15/05/2023 which detailed model development or validation studies. The checklist for Critical  
27 Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies  
28 (CHARMS) was used to guide data extraction from eligible texts and the Prediction Model Risk  
29 of Bias Assessment Tool (PROBAST) was used to assess risk of bias and applicability of the  
30 studies. A narrative synthesis of included studies was performed.

### 31 **Results**

32 17 eligible studies were identified: 16 eligible model development studies and one eligible  
33 model validation study. A wide range of biomarkers were assessed including nucleic acids,  
34 proteins, metabolites and lipids. The range of C-index (area under the curve) estimates  
35 reported for the models was 0.67-1.00. No studies assessed model calibration. According to  
36 PROBAST criteria, all studies were at high risk of bias in the analysis domain.

### 37 **Discussion**

38 While a wide range of potentially predictive biomarkers were identified in the included studies,  
39 most studies did not account for overfitting in model performance estimates, no studies  
40 assessed calibration, and all models were at high risk of bias according to PROBAST criteria.

Review: predicting psychosis using blood biomarkers

41 External validation of the models is needed to provide more accurate estimates of their  
42 performance. Future studies which follow the latest available methodological and reporting  
43 guidelines and adopt strategies to accommodate required sample sizes for model  
44 development or validation will clarify the value of including blood-based biomarkers in models  
45 predicting psychosis.

46

## 47 **Introduction**

### 48 **Background and Rationale**

49 Recent research in the field of early intervention in psychosis has focused on building models  
50 to predict the development of psychotic disorder [1–3]. These models have largely been  
51 developed in “clinical high-risk” populations, which include individuals showing prodromal  
52 symptoms or genetic risk combined with functional decline, as determined with validated  
53 assessment tools such as the Comprehensive Assessment of At-Risk Mental State  
54 (CAARMS) [4] and the Structured Interview for Prodromal Symptoms (SIPS) [5,6]. Meta-  
55 analytic estimates indicate 19% of individuals at clinical high-risk develop psychosis within two  
56 years [7].

57 Accumulating evidence points towards abnormalities in metabolic and inflammatory  
58 processes in individuals with psychosis [8–10] and there is some evidence to suggest that  
59 these abnormalities may precede medication use [9] or even the onset of psychosis [11,12].  
60 The prognostic value of peripheral markers over lifestyle or environmental factors, such as  
61 smoking and exercise, is unclear [13,14]. However, several studies have employed a range of  
62 blood-based predictors in models predicting transition to psychotic disorder, with several  
63 published since the last systematic reviews of models in the field [15–19]. A recent large-scale  
64 systematic review included all prediction models in psychiatry, except those using biological  
65 predictors [20]. Biological predictors have the advantage of often being more objective and  
66 precise than, for example, scores given on symptom scales. Blood biomarkers are among the

Review: predicting psychosis using blood biomarkers

67 least invasive biological parameters, with relative low-cost. There is a clear pathway for their  
68 integration into clinical practice, as they could be measured along with routine blood markers,  
69 or predictors could consist of routine blood measures. As such, a systematic review of  
70 prognostic models predicting transition to psychotic disorder using blood-based biomarkers is  
71 warranted. Furthermore, the weak standard of prediction modelling methodology in psychiatry  
72 and medicine in general has been highlighted previously, in particular regarding the need for  
73 validation and implementation of models [20–22]. Therefore, a review of methodology that  
74 takes into account the interaction between the use of blood-based biomarkers and prediction  
75 modelling may help future research in the field.

## 76 **Objectives**

77 To systematically review the performance and methodology of models predicting transition to  
78 psychotic disorder from risk-enriched populations, with a focus on those using blood-based  
79 biomarkers, to determine their potential utility in predictive models and to help guide future  
80 research in the field.

81

## 82 **Methods**

83 This review was reported using the guidelines for transparent reporting of multivariable  
84 prediction models for individual prognosis or diagnosis: systematic reviews and meta-analyses  
85 (TRIPOD-SRMA) [23].

## 86 **Eligibility Criteria**

87 Studies were eligible for inclusion if they; 1) described the development, validation or updating  
88 of a prognostic model of transition to psychotic disorder in “at-risk” (or similar psychosis risk-  
89 enriched populations) help-seeking individuals, 2) used blood-based biomarkers in the  
90 prognostic model described, 3) were published in peer-reviewed journals, 4) were published  
91 after 1998 (after the first prospective studies using clinical high-risk criteria were published;

Review: predicting psychosis using blood biomarkers

92 [24]) and 5) had full-texts available in English. The PICOTS (Population, Intervention,  
93 Comparator, Outcome, Timing, Setting) criteria were used to guide the development of  
94 eligibility criteria and can be found in the Supplementary Material.

95 Studies were excluded if they focused on psychiatric disorders other than psychotic disorders  
96 (for example studies solely focusing on prediction of depression) and if they investigated risk  
97 factors or longitudinal associations but did not report the development and/or validation of a  
98 risk prediction model. Studies that reported only on models including blood-based predictors  
99 in combination with brain or cerebrospinal fluid biomarkers were excluded as the predictive  
100 value of the blood-based biomarkers in the models would be difficult to determine precisely,  
101 and as the aim of the review was to identify models that required the measurement of blood  
102 biomarkers only, without the requirement for a further invasive and expensive procedure.  
103 Studies must have had a binary outcome of transition to psychosis to be included in the current  
104 study (studies predicting continuous outcomes were excluded, for example psychotic  
105 symptom scales). Studies of continuous outcomes were excluded for several reasons,  
106 including a) transition is the key outcome in the literature, b) the difficulty in implementing a  
107 prediction model in clinical practice without a clear diagnostic outcome, c) individual clinical  
108 scales (positive symptoms or negative symptom scales) can't define diagnosis alone, d)  
109 potentially multiple different outcomes (functioning and clinical symptom subscales) that may  
110 use multiple different non-comparable scales (e.g. the Positive And Negative Symptom Scale  
111 and the Community Assessment of Psychic Experiences).

112

### 113 **Information Sources and Search Strategy**

114 PubMed, EMBASE and PsycINFO were searched from 01/01/1998 to 15/05/2023 using the  
115 following general search strategy: psychosis risk-enrichment keywords AND transition to  
116 psychotic disorder keywords AND prediction modelling keywords AND blood-based biomarker  
117 keywords. The search strategy was developed with the use of established strategies for

Review: predicting psychosis using blood biomarkers

118 searching for predictive modelling studies [25]. The full search strategies, as formatted for  
119 each database, are included in the Supplementary Material.

120 In an attempt to find other references that may meet inclusion criteria, reference lists of  
121 relevant reviews that appeared in the databases searched were examined and forward citation  
122 searching was carried out on Google Scholar for texts eligible for full-text screening up to  
123 01/06/2023. Where clearly eligible models were detailed in conference abstracts and  
124 corresponding full-texts could not be found, we contacted the corresponding authors for  
125 information on potential unpublished full-texts.

126

### 127 **Selection Process**

128 Duplicate records were identified with guidance from previous recommendations [26] and  
129 removed. Abstracts identified by the search strategy were screened independently by JFB,  
130 DM and JM. Prediction modelling studies that clearly did not meet the eligibility criteria were  
131 excluded and the full-texts of all other studies were examined. Disagreements were resolved  
132 through discussion and or by referral to a third author (DRC).

133

### 134 **Data Collection Process**

135 Data were extracted from studies using the CHARMS checklist for Critical Appraisal and Data  
136 Extraction for Systematic Reviews of Prediction Modelling Studies [27]. Data were extracted  
137 independently by JFB, DM and JM. Disagreements were resolved through discussion or  
138 referral to a third author (CH or MF). Where several similar models were presented in studies  
139 (e.g. models with minor differences in predictors included), data pertaining to the final model  
140 as indicated by the study authors was extracted. Where study authors did not indicate a final  
141 model, data pertaining to the best performing model was extracted.

Review: predicting psychosis using blood biomarkers

142 As part of the CHARMS, we extracted two main model performance metrics: discrimination  
143 (how well the model differentiates between individuals who do and do not develop the  
144 outcome; the Concordance (C)-index [28], which is equivalent to the area-under-the-curve in  
145 the case of a binary outcome) and calibration (how well the predicted probabilities match the  
146 actual proportion of outcomes) [29], where available. Where studies reported on the added  
147 predictive value of blood-based biomarkers to models, relevant metrics and tests (such as the  
148 likelihood ratio chi-square test) of added predictive information were extracted, if available.

149

### 150 **Risk of Bias and Applicability Assessment**

151 Risk of bias assessments were carried out using the prediction model risk of bias assessment  
152 tool (PROBAST) [30,31], which assesses risk of bias in the selection of participants,  
153 measurement of predictors or outcomes and in the analysis. Using the PROBAST tool,  
154 concerns of applicability of the study to the review question were also rated in the domains  
155 selection of participants, measurement of predictors or measurement of outcomes. PROBAST  
156 ratings for risk of bias and applicability assessments can be either high, low or unclear.  
157 Assessments were carried out independently by JFB, DM and JM. Disagreements were  
158 resolved through discussion with a third author (CH). PROBAST figures were generated using  
159 the robvis package [32] in R (<https://github.com/mcguinlu/robvis>). Required sample sizes for  
160 precise estimates of model performance on external validation were calculated in R  
161 (<https://github.com/c-qu/samplesize-validation>) according to guidelines from a previous  
162 publication [33].

163

### 164 **Synthesis Methods**

165 We planned to use a narrative synthesis method [34]. We stratified performance estimates of  
166 the included model development studies based on the PROBAST signalling question “Were  
167 model overfitting and optimism in model performance accounted for?”. As per PROBAST

Review: predicting psychosis using blood biomarkers

168 guidance, model overfitting and optimism in model performance estimates were accounted for  
169 if both internal validation and shrinkage techniques were applied, and if predictor selection  
170 procedures (e.g. univariable screening or backwards selection) were included in the internal  
171 validation framework [31]. Meta-analyses were to be carried out only if a particular model had  
172 multiple external validation studies [35].

173

## 174 **Results**

### 175 **Study Selection**

176 A PRISMA flow diagram is presented in Figure 1. The database search identified 2,565  
177 records. Following removal of duplicates, 1,677 titles and abstracts were screened, and 45  
178 articles were brought forward for full-text screening. 13 records (relating to 13 studies) were  
179 included, and 32 records were excluded. Reasons for exclusion can be summarised in the  
180 following categories: not predictive modelling, no blood-based predictors used, ineligible study  
181 design (participants or outcome do not meet eligibility criteria), conference abstract only. We  
182 contacted authors of three conference abstracts which detailed clearly eligible models for  
183 further information on potential full-texts related to the abstracts. Authors of two conference  
184 abstracts confirmed that both abstracts related to a single study for which we had already  
185 obtained a peer-reviewed full-text. We did not receive a response from the authors of one  
186 conference abstract which detailed a clearly eligible model.

187 Forward citation searching of the 45 full-texts identified eight further potentially eligible full-  
188 texts, of which four were excluded (not predictive modelling) and four records (relating to four  
189 studies) were included (two of which were full-text reports of conference abstracts identified  
190 in the database search). Therefore, 17 studies were included overall.

191

192



## 193 **Study Characteristics**

194 Of the 17 included studies, 16 were prognostic model development studies [15–19,36–46] and  
195 one study was a prognostic model external validation study [47]. All studies were conducted  
196 in outpatients. There were five models developed in the Shanghai At Risk for Psychosis  
197 (SHARP) study [18,19] and the extension of that study [44–46]. There were three model  
198 development studies each for the North American Prodrome Longitudinal Study (NAPLS 2)  
199 cohort [36,38,41] and for the EU Gene-Environment Interaction (EU-GEI) study [15,16,42].  
200 Two studies were developed models in participants of the Vienna omega-3 randomised-  
201 controlled trial [37,39]. There was one model developed in the Personalised Prognostic Tools  
202 for Early Psychosis Management (PRONIA) study [17], one model developed in a Korean  
203 cohort study [40], one model developed in participants recruited from the Outreach and  
204 Support in South London high-risk service [43], and one validated in the ICAAR (Influence of  
205 Cannabis in the emergence of psychopathological symptoms in Adolescents and Adults at-  
206 Risk) study [47].

207 Mean study participant ages ranged from 15.8 years to 24.6 years. The majority of studies  
208 (16/17) defined participants at increased risk of psychosis through use of the SIPS (ten  
209 studies), CAARMS (four studies) or CAARMS-equivalent (two studies) criteria. A wide range  
210 of biomarkers were assessed including cytokines (four studies), single-nucleotide  
211 polymorphisms (SNPs; four studies), hormonal, inflammatory and metabolic-related analytes  
212 (two studies), ribonucleic acids (two studies), lipids (two studies), proteins (one study),  
213 metabolites (one study), and glutathione (one study) (Table 1).

214 Of the 16 development studies, nine internally validated model performance, of which five  
215 accounted for optimism in their performance estimates [15,17,41]. Four studies which used  
216 internal validation did not include the predictor selection process within the internal validation  
217 procedure. The remaining seven studies reported apparent performance (Table 1). Reported  
218 C-indices for logistic regression models ranged from 0.67-1.00. Reported C-indices for cox  
219 models ranged from 0.82-0.88. One study reported a balanced accuracy of 46.2%. None of

Review: predicting psychosis using blood biomarkers

220 the studies reported calibration measures or assessed clinical utility of their models. Further  
221 characteristics and performance metrics of the included studies are detailed in the  
222 Supplementary Material.

223

#### 224 **Risk of Bias and Applicability Assessments of Included Studies**

225 All studies included in the systematic review were at high overall risk of bias according to  
226 PROBAST criteria. This was related to risk of bias in the analysis domain (Figure 2;  
227 Supplementary Table 3).

228 The risk of bias due to the selection of participants was rated unclear in ten out of 17 studies  
229 (58.8%). Generally, this related to either a lack of reporting of exclusion criteria used, a lack  
230 of comparison between included and excluded participants, or a lack of reporting of how many  
231 exclusions were made due to potentially inappropriate criteria (such as abnormal levels of  
232 blood parameters). Regarding applicability concerns in this domain, for the same reasons it  
233 was unclear in ten studies (58.8%) whether participants matched the review question (Figure  
234 3).

235 Risk of bias due to the predictors or their assessment was unclear for 12 out of 17 studies  
236 (70.6%) as they did not report whether predictor assessments were made blind to the outcome  
237 data. The method of predictor assessment used in four studies (23.5%) may not match the  
238 review question, as it was unclear if the biomarker measurement methods used provide  
239 absolute quantification. A lack of standardised units for the biomarker measurement could  
240 hinder the generalisability of models. There were low applicability concerns for the remaining  
241 13 studies in this applicability domain.

242 Risk of bias introduced by the outcome or its determination was low for 11 studies which used  
243 standard measures or structured interviews. Risk of bias introduced by the outcome was  
244 unclear for six studies (35.3%), largely because insufficient information on how the outcome  
245 was determined was reported. There were generally low concerns regarding the applicability

Review: predicting psychosis using blood biomarkers

246 of the outcome, with one study rated unclear in this domain, as the timing of the outcome  
247 determination relative to collection of the blood sample was unclear.

248 All studies were at high risk of bias in the analysis domain. None of the studies reported  
249 measures of model calibration. Generally, studies did not have a sufficient number of  
250 participants with the outcome or had a low number of events per candidate predictor. 9/16  
251 development studies (56.3%) had less than one event per candidate predictor, and 14/16  
252 (87.5%) had less than three events per candidate predictor. The one model validation study  
253 had a sample size of 76. However, we calculated that with the given study outcome prevalence  
254 of 0.237 and expected external validation C-index of 0.80, the minimum recommended sample  
255 size for precise estimation of model performance metrics on external validation would be 709.  
256 [33,48,49]. 5/16 (31.3%) development studies accounted for model overfitting and optimism  
257 in performance estimates. Four of these studies used cross-validation for internal validation  
258 and penalised regression models to shrink coefficients. One of these studies used the  
259 bootstrap for internal validation and shrinkage. Other limitations included not accounting for  
260 subsampling of controls, not reporting how missing data was handled or the use of univariable  
261 analysis to select predictors for inclusion in the model.

262

### 263 **Narrative Synthesis**

264 None of the included studies assessed model calibration, a key metric for which a model must  
265 perform acceptably if it is to be used to determine whether an intervention should be offered  
266 [29]. The logistic regression model C-indices in model development studies which did not  
267 account for optimism in their performance estimates ranged from 0.81 to 1 (with one study  
268 reporting a cox model C-index of 0.82). Model development studies which accounted for  
269 optimism in their performance estimates reported C-indices from logistic regression models  
270 ranging from 0.67-0.95 (with one study reporting a cox model C-index of 0.88 and one study  
271 reporting a balanced accuracy of 46.2%).

Review: predicting psychosis using blood biomarkers

272 Three studies which adjusted their performance estimates for optimism used polygenic risk  
273 scores (PRS) as a blood-based predictors [17,41,43]. Perkins et al. [41] investigated adding  
274 PRS to an established cox regression model comprised of clinical predictors. When PRS was  
275 included as a variable together with the clinical predictors, there was no evidence from a  
276 likelihood ratio chi-square test that the PRS added predictive value. Furthermore, the C-index  
277 was unchanged for participants of non-European descent and was similar for participants of  
278 European descent (point estimate increase of 0.01). In line with this, Koutsouleris et al. [17]  
279 found that a model based on PRS variables had a similar C-index to that achieved by  
280 predictions from clinical raters, and Tavares et al. [43] found that a PRS model did not  
281 discriminate between individuals who transitioned and did not transition better than chance.

282 The final model presented by Chan et al. [47] and the “15-analyte” model in Perkins et al. [36]  
283 had two overlapping predictors, Interleukin-8 and Factor VII. While model coefficients were  
284 not available in Perkins et al. [36], the two studies presented the same direction of effect for  
285 IL-8 and opposite directions of effect for Factor VII in univariate analyses. Zhang et al. [46]  
286 and Mondelli et al [42] did not find evidence for the predictive ability of IL-8 for transition to  
287 psychotic disorder. Zhang et al. 2022 [44], Zhang et al. 2023a [45] and Mondelli et al. [42] all  
288 included different cytokine ratios in their models (IL-1 $\beta$ /IL-6, IL-2/IL-6 and IL-10/IL-6  
289 respectively) with the aim of capturing inflammatory balance. Chan et al. [47] and Mongan et  
290 al. [15] both included Alpha-2-Macroglobulin as a predictor in their final models, though  
291 coefficients were in opposite directions. No other overlapping predictors were noted between  
292 final models presented in studies.

293

## 294 **Discussion**

295 We undertook a focused systematic review of models predicting transition to psychosis with  
296 use of recent guidelines for systematically reviewing prognostic models [31,50]. Models  
297 developed with blood biomarkers require unique consideration, as any predictive benefit of

Review: predicting psychosis using blood biomarkers

298 blood biomarkers over clinical predictors must outweigh the disadvantages associated with  
299 their measurement, namely the cost or potential lack of wide accessibility. While studies  
300 included in this review described a wide range of blood biomarkers that potentially have altered  
301 concentrations preceding psychosis, the prognostic models including blood biomarkers were  
302 not developed according to the latest methodological recommendations, lacked calibration  
303 and lacked sufficient external validation. Therefore, similar to recent systematic reviews of  
304 prediction models in psychiatry [20,51,52], we did not find evidence for a model suitable for  
305 implementation into clinical practice [53].

306 As all the studies were rated at high risk of bias, we are unable to recommend a particular  
307 model to be externally validated, and it is unclear at the present time whether any specific  
308 blood biomarkers could potentially contribute to improved prediction of transition to psychosis  
309 in individuals at risk. However, evidence from three studies suggested that models including  
310 polygenic risk scores do not sufficiently outperform models based on clinical variables or  
311 clinical rater predictions. A wide range of other biomarkers were assessed in the included  
312 studies, however, going forward, the field will need to externally validate models to truly  
313 estimate their generalizability. Risk prediction models cannot be recommended for clinical  
314 practice without sufficient external validation.

315 In general, the reporting of study design and methods could be improved in the field. We  
316 recommend that future studies follow the transparent reporting of a multivariable prediction  
317 model for individual prognosis or diagnosis (TRIPOD) reporting guidelines [54] and explicitly  
318 report on participant exclusion criteria and whether predictors were assessed blind to outcome  
319 status. Blood-biomarkers in cohort studies are often assessed after the outcome is determined  
320 and therefore studies involving blood-biomarkers are at particular risk of bias in this domain.  
321 As well as blinding, studies should report whether samples were randomised prior to  
322 biomarker measurement. Recent guidance proposes block randomisation as a gold standard  
323 [55,56].

Review: predicting psychosis using blood biomarkers

324 Sixteen out of the 17 studies in this review defined their participants to be at risk of psychosis  
325 through use of SIPS and CAARMS interviews for prodromal symptoms or reduced functioning  
326 combined with genetic risk. While these established “clinical-high risk” and “at-risk mental  
327 state” constructs are risk-enriched populations with an estimated 2-year transition rate of 19%  
328 [7], they may only capture a minority (4-14%) of the population who present with a first episode  
329 of psychosis, indicating a relatively poor predictive capacity [57–59]. There have been recent  
330 calls to expand the clinical high risk paradigm such as to include individuals attending child  
331 and adolescent mental health services or presenting to emergency departments due to self-  
332 harm, as they are at similar risk of psychosis [60–62]. This “systems approach” may capture  
333 a greater proportion of first episode psychosis cases than prodromal constructs. We  
334 recommend that future studies are designed in populations with a higher predictive capacity  
335 for psychosis [62], as this may help to mitigate issues with insufficient sample sizes in the field  
336 and potential recruitment bias.

337 Some of the studies in this review excluded participants based on established psychiatric  
338 diagnoses or when a participant’s prodromal symptoms were caused by a mood or anxiety  
339 disorder. In light of studies showing that psychotic symptoms are highly prevalent in disorders  
340 of depression and anxiety [63] and that psychosis can be predicted in individuals with non-  
341 psychotic mental illnesses [64–66], these may be inappropriate exclusion criteria. On the other  
342 hand, a more nuanced assessment of clinical and biological risk factors associated with  
343 psychosis, such as minimal self [67], circadian rhythms [68] or trait-like EEG signatures [69],  
344 could be used to reduce biological heterogeneity. Reducing biological heterogeneity could  
345 complement traditional risk-enrichment approaches and allow for the identification of more  
346 replicable blood biomarkers of psychosis risk.

347 Half of the studies included in this systematic review considered over 100 candidate  
348 predictors. With the growing popularity of “omics” methods, it must be highlighted that the  
349 latest research does not suggest that data-driven methods of predictor selection involving the  
350 outcome data can alleviate overfitting in situations of high dimensionality and low sample size.

Review: predicting psychosis using blood biomarkers

351 Univariable screening has long been highlighted as problematic [70]. Multivariable selection  
352 methods such as LASSO have recently been shown to be unstable in small sample sizes or  
353 with small numbers of events [71]. Researchers should be aware that penalisation methods  
354 do not solve the issue of a small ratio of events to predictors. Sample size calculators for the  
355 development or validation of prediction models with binary or time-to-event outcomes are now  
356 easily accessible and should be utilised prior to designing studies to ensure adequate power  
357 [49,72,73].

358 One of the main limitations of the studies in the analysis domain was the lack of calibration of  
359 models. In the first instance, clinical prediction models must produce predicted event  
360 probabilities for each individual rather than binary event or non-event predictions alone - for a  
361 model to be implemented into clinical practice, a probability cut-off relating to maximum clinical  
362 benefit is required to determine whether a specific intervention should be offered or not  
363 [22,30,74]. Calibration measures how well the predicted probabilities match the observed  
364 proportion of outcomes, i.e. the accuracy of risk estimates, and models can have poor  
365 calibration even when models show good discrimination. An over- or underestimation of the  
366 probability of developing a psychotic disorder is ethically unacceptable, and would lead to  
367 inappropriately offered interventions or undertreatment [29]. Future studies in the field should  
368 assess model calibration to improve the chances of models being implemented clinically.

369 The studies included in this review had further limitations in the analysis domain. Some studies  
370 did not handle missing data to PROBAST standards. Missing biomarker data is often related  
371 to biomarker concentrations being below the limit of detection. In this case, the data should  
372 be considered missing not-at-random. PROBAST guidelines recommend multiple imputation  
373 as best practice in prediction modelling [31], and solutions combining multiple imputation with  
374 left-censored missing data have been proposed [75]. Furthermore, participant subsampling  
375 was frequent. When subsampling is necessary, researchers should weight cases and controls  
376 by the inverse of their sampling fractions in analyses [28].

Review: predicting psychosis using blood biomarkers

377 This review has several strengths. The review has benefitted from prospective registration and  
378 the use of recommended reporting guidelines [23], search strategies [25], data extraction tools  
379 [27] and risk of bias assessment tools [30]. However, there are also several limitations to this  
380 review. We were unable to perform meta-analyses as we did not identify any models that were  
381 externally validated multiple times. Meta-analyses of models with different predictors or  
382 validation approaches would not have been readily interpretable. Due to limitations of the  
383 modelling strategies in the studies and lack of external validation of models, we were not able  
384 to perform a head-to-head comparison of the performance of each of the prediction models  
385 as the performance estimates were at high risk of bias. Finally, we acknowledge that the  
386 concept of a binary “transition” to psychosis, even in the presence of assessment criteria, can  
387 be subjective or can sometimes represent small increases in the severity or frequency of  
388 symptoms [76]. This review did not examine the prediction of positive symptoms as continuous  
389 outcomes, which may be worth examining in future reviews in the field.

390 In conclusion, while there have been several studies developing models using blood-based  
391 biomarkers for prediction of transition to psychotic disorder, this review found no models that  
392 are ready for implementation in clinical practice, and the value of including blood-based  
393 biomarkers in models predicting transition to psychosis is unclear due to the high risk of bias  
394 of the eligible studies. The field of prediction modelling is rapidly progressing and it should be  
395 noted that new methodological recommendations have been made since the majority of the  
396 studies in this review were published [71,72]. Future studies should aim to follow the latest  
397 available reporting guidelines, assess model calibration, internally and repeatedly externally  
398 validate models, and adopt strategies to accommodate minimum required sample sizes in  
399 order to maximise potential clinical benefits and outcomes for patients.

400

#### 401 **Registration and Protocol**

402 This systematic review was prospectively registered with PROSPERO, CRD42022302047.



Review: predicting psychosis using blood biomarkers

## 403 **Funding**

404 D.R.C., M.F. and J.F.B. are supported by a Wellcome Flagship Innovations Award (IMPETUS  
405 - 220438Z/20/Z). D.R.C. is supported by a Health Research Board Investigator Led Project  
406 Grant (ILP-POR-2017-039), M.C. is supported by a European Research Council Consolidator  
407 Award (iHEAR 724809), C.H. is supported by Health Research Board Investigator Led Project  
408 Grant (ILP-PHR-2019-009), D.M. is a Fellow on the Irish Clinical Academic Training (ICAT)  
409 Programme which is supported by the Wellcome Trust and the Health Research Board (Grant  
410 Number 203930/B/16/Z), the Health Service Executive National Doctors Training and  
411 Planning and the Health and Social Care, Research and Development Division, Northern  
412 Ireland. S.R.S is supported by the Health Research Board (HRB) under grant number  
413 HRB/HRA/PHR/2015-1293. D.R.C and J.F.B are supported in part by Science Foundation  
414 Ireland (SFI) under Grant Number 16/RC/3948 and co-funded under the European Regional  
415 Development Fund and by FutureNeuro industry partners.

## 416 **Conflict of Interest**

417 D.M., M.C. and D.R.C. report a patent pending (UK Patent Application No. 1919155.0,  
418 “Biomarkers to predict psychosis”). C.H., M.F., J.M. and J.F.B. report no financial relationships  
419 with commercial interests.

420

## 421 **References**

- 422 1. Riecher-Rössler A, Studerus E. Prediction of conversion to psychosis in individuals with an at-  
423 risk mental state: a brief update on recent developments. *Current Opinion in Psychiatry*.  
424 2017;30:209–219.
- 425 2. Sanfelici R, Dwyer DB, Antonucci LA, Koutsouleris N. Individualized Diagnostic and Prognostic  
426 Models for Patients With Psychosis Risk Syndromes: A Meta-analytic View on the State of the  
427 Art. *Biological Psychiatry*. 2020;88:349–360.
- 428 3. Montemagni C, Bellino S, Bracale N, Bozzatello P, Rocca P. Models predicting psychosis in  
429 patients with high clinical risk: a systematic review. *Frontiers in Psychiatry*. 2020;11:223.

Review: predicting psychosis using blood biomarkers

- 430 4. Yung AR, Yung AR, Pan Yuen H, MCGorry PD, Phillips LJ, Kelly D, et al. Mapping the onset of  
431 psychosis: the comprehensive assessment of at-risk mental states. *Australian & New Zealand*  
432 *Journal of Psychiatry*. 2005;39:964–971.
- 433 5. Miller TJ, McGlashan TH, Rosen JL, Cadenhead K, Ventura J, McFarlane W, et al. Prodromal  
434 assessment with the structured interview for prodromal syndromes and the scale of prodromal  
435 symptoms: predictive validity, interrater reliability, and training to reliability. *Schizophrenia*  
436 *Bulletin*. 2003;29:703–715.
- 437 6. Fusar-Poli P, Borgwardt S, Bechdolf A, Addington J, Riecher-Rössler A, Schultze-Lutter F, et al.  
438 The Psychosis High-Risk State: A Comprehensive State-of-the-Art Review. *JAMA Psychiatry*.  
439 2013;70:107–120.
- 440 7. Salazar de Pablo G, Radua J, Pereira J, Bonoldi I, Arienti V, Besana F, et al. Probability of  
441 Transition to Psychosis in Individuals at Clinical High Risk: An Updated Meta-analysis. *JAMA*  
442 *Psychiatry*. 2021;78:970–978.
- 443 8. Fraguas D, Díaz-Caneja CM, Rodríguez-Quiroga A, Arango C. Oxidative stress and inflammation  
444 in early onset first episode psychosis: a systematic review and meta-analysis. *International*  
445 *Journal of Neuropsychopharmacology*. 2017;20:435–444.
- 446 9. Upthegrove R, Manzanares-Teson N, Barnes NM. Cytokine function in medication-naive first  
447 episode psychosis: a systematic review and meta-analysis. *Schizophrenia Research*.  
448 2014;155:101–108.
- 449 10. Perry BI, McIntosh G, Weich S, Singh S, Rees K. The association between first-episode psychosis  
450 and abnormal glycaemic control: systematic review and meta-analysis. *The Lancet Psychiatry*.  
451 2016;3:1049–1058.
- 452 11. Mongan D, Healy C, Jones HJ, Zammit S, Cannon M, Cotter DR. Plasma polyunsaturated fatty  
453 acids and mental disorders in adolescence and early adulthood: cross-sectional and  
454 longitudinal associations in a general population cohort. *Translational Psychiatry*. 2021;11:1–  
455 13.
- 456 12. Perry BI, Stochl J, Upthegrove R, Zammit S, Wareham N, Langenberg C, et al. Longitudinal  
457 Trends in Childhood Insulin Levels and Body Mass Index and Associations With Risks of  
458 Psychosis and Depression in Young Adults. *JAMA Psychiatry*. 2021;78:416–425.
- 459 13. Osimo EF, Baxter L, Stochl J, Perry BI, Metcalf SA, Kunutsor SK, et al. Longitudinal association  
460 between CRP levels and risk of psychosis: a meta-analysis of population-based cohort studies.  
461 *Npj Schizophrenia*. 2021;7:31.

Review: predicting psychosis using blood biomarkers

- 462 14. Firth J, Solmi M, Wootton RE, Vancampfort D, Schuch FB, Hoare E, et al. A meta-review of  
463 “lifestyle psychiatry”: the role of exercise, smoking, diet and sleep in the prevention and  
464 treatment of mental disorders. *World Psychiatry*. 2020;19:360–380.
- 465 15. Mongan D, Föcking M, Healy C, Susai SR, Heurich M, Wynne K, et al. Development of proteomic  
466 prediction models for transition to psychotic disorder in the clinical high-risk state and  
467 psychotic experiences in adolescence. *JAMA Psychiatry*. 2021;78:77–90.
- 468 16. Dickens AM, Sen P, Kempton MJ, Barrantes-Vidal N, Iyegbe C, Nordentoft M, et al.  
469 Dysregulated lipid metabolism precedes onset of psychosis. *Biological Psychiatry*.  
470 2021;89:288–297.
- 471 17. Koutsouleris N, Dwyer DB, Degenhardt F, Maj C, Urquijo-Castro MF, Sanfelici R, et al.  
472 Multimodal machine learning workflows for prediction of psychosis in patients with clinical  
473 high-risk syndromes and recent-onset depression. *JAMA Psychiatry*. 2021;78:195–209.
- 474 18. Li Z, Zhang T, Xu L, Wei Y, Cui H, Tang Y, et al. Plasma metabolic alterations and potential  
475 biomarkers in individuals at clinical high risk for psychosis. *Schizophrenia Research*.  
476 2022;239:19–28.
- 477 19. Song W, Xu L, Zhang T, Wang W, Fu Y, Xu Q, et al. Peripheral transcriptome of clinical high-risk  
478 psychosis reflects symptom alteration and helps prognosis prediction. *Psychiatry and Clinical  
479 Neurosciences*. 2022. 2022.
- 480 20. Meehan AJ, Lewis SJ, Fazel S, Fusar-Poli P, Steyerberg EW, Stahl D, et al. Clinical prediction  
481 models in psychiatry: a systematic review of two decades of progress and challenges.  
482 *Molecular Psychiatry*. 2022;27:2700–2708.
- 483 21. Studerus E, Rameyad A, Riecher-Rössler A. Prediction of transition to psychosis in patients with  
484 a clinical high risk for psychosis: a systematic review of methodology and reporting.  
485 *Psychological Medicine*. 2017;47:1163–1178.
- 486 22. Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis  
487 Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Medicine*.  
488 2013;10:e1001381.
- 489 23. Snell KIE, Levis B, Damen JAA, Dhiman P, Debray TPA, Hooft L, et al. Transparent reporting of  
490 multivariable prediction models for individual prognosis or diagnosis: checklist for systematic  
491 reviews and meta-analyses (TRIPOD-SRMA). *BMJ*. 2023;381:e073538.
- 492 24. Yung AR, Phillips LJ, McGorry PD, McFarlane CA, Francey S, Harrigan S, et al. Prediction of  
493 psychosis: a step towards indicated prevention of schizophrenia. *The British Journal of  
494 Psychiatry*. 1998;172:14–20.

Review: predicting psychosis using blood biomarkers

- 495 25. Geersing G-J, Bouwmeester W, Zuithoff P, Spijker R, Leeflang M, Moons K. Search filters for  
496 finding prognostic and diagnostic prediction studies in Medline to enhance systematic reviews.  
497 PLoS One. 2012;7:e32844.
- 498 26. Bramer WM, Giustini D, de Jonge GB, Holland L, Bekhuis T. De-duplication of database search  
499 results for systematic reviews in EndNote. Journal of the Medical Library Association: JMLA.  
500 2016;104:240.
- 501 27. Moons KG, de Groot JA, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical  
502 appraisal and data extraction for systematic reviews of prediction modelling studies: the  
503 CHARMS checklist. PLoS Medicine. 2014;11:e1001744.
- 504 28. Harrell Jr FE. Regression Modeling Strategies With Applications to Linear Models, Logistic and  
505 Ordinal Regression, and Survival Analysis. 2nd ed. Springer Cham; 2015.
- 506 29. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW, Bossuyt P, et al.  
507 Calibration: the Achilles heel of predictive analytics. BMC Medicine. 2019;17:230.
- 508 30. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A Tool  
509 to Assess the Risk of Bias and Applicability of Prediction Model Studies. Ann Intern Med.  
510 2019;170:51–58.
- 511 31. Moons KG, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to  
512 assess risk of bias and applicability of prediction model studies: explanation and elaboration.  
513 Annals of Internal Medicine. 2019;170:W1–W33.
- 514 32. McGuinness LA, Higgins JPT. Risk-of-bias VISualization (robvis): An R package and Shiny web  
515 app for visualizing risk-of-bias assessments. Research Synthesis Methods. 2021;12:55–61.
- 516 33. Pavlou M, Qu C, Omar RZ, Seaman SR, Steyerberg EW, White IR, et al. Estimation of required  
517 sample size for external validation of risk models for binary outcomes. Stat Methods Med Res.  
518 2021;30:2187–2206.
- 519 34. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the conduct  
520 of narrative synthesis in systematic reviews. A Product from the ESRC Methods Programme  
521 Version. 2006;1:b92.
- 522 35. Debray TP, Damen JA, Riley RD, Snell K, Reitsma JB, Hooft L, et al. A framework for meta-  
523 analysis of prediction model studies with binary and time-to-event outcomes. Statistical  
524 Methods in Medical Research. 2019;28:2768–2786.
- 525 36. Perkins DO, Jeffries CD, Addington J, Bearden CE, Cadenhead KS, Cannon TD, et al. Towards a  
526 psychosis risk blood diagnostic for persons experiencing high-risk symptoms: preliminary  
527 results from the NAPLS project. Schizophr Bull. 2015;41:419–428.

Review: predicting psychosis using blood biomarkers

- 528 37. Clark S, Baune B, Schubert K, Lavoie S, Smesny S, Rice S, et al. Prediction of transition from  
529 ultra-high risk to first-episode psychosis using a probabilistic model combining history, clinical  
530 assessment and fatty-acid biomarkers. *Translational Psychiatry*. 2016;6:e897–e897.
- 531 38. Jeffries C, Perkins D, Chandler S, Stark T, Yeo E, Addington J, et al. Insights into psychosis risk  
532 from leukocyte microRNA expression. *Translational Psychiatry*. 2016;6:e981–e981.
- 533 39. Lavoie S, Berger M, Schlögelhofer M, Schäfer M, Rice S, Kim S, et al. Erythrocyte glutathione  
534 levels as long-term predictor of transition to psychosis. *Translational Psychiatry*. 2017;7:e1064–  
535 e1064.
- 536 40. Bousman CA, Lee TY, Kim M, Lee J, Mostaid MS, Bang M, et al. Genetic variation in cytokine  
537 genes and risk for transition to psychosis among individuals at ultra-high risk. *Schizophrenia  
538 Research*. 2018;195:589–590.
- 539 41. Perkins DO, Olde Loohuis L, Barbee J, Ford J, Jeffries CD, Addington J, et al. Polygenic risk score  
540 contribution to psychosis prediction in a target population of persons at clinical high risk.  
541 *American Journal of Psychiatry*. 2020;177:155–163.
- 542 42. Mondelli V, Blackman G, Kempton MJ, Pollak TA, Iyegbe C, Valmaggia LR, et al. Serum immune  
543 markers and transition to psychosis in individuals at clinical high risk. *Brain, Behavior, and  
544 Immunity*. 2023;110:290–296.
- 545 43. Tavares V, Vassos E, Marquand A, Stone J, Valli I, Barker GJ, et al. Prediction of transition to  
546 psychosis from an at-risk mental state using structural neuroimaging, genetic, and  
547 environmental data. *Frontiers in Psychiatry*. 2022;13.
- 548 44. Zhang T, Zeng J, Wei Y, Ye J, Tang X, Xu L, et al. Changes in inflammatory balance correlates  
549 with conversion to psychosis among individuals at clinical high-risk: A prospective cohort study.  
550 *Psychiatry Research*. 2022;318:114938.
- 551 45. Zhang T, Wei Y, Zeng J, Ye J, Tang X, Xu L, et al. Interleukin-2/interleukin-6 imbalance correlates  
552 with conversion to psychosis from a clinical high-risk state. *Psychiatry and Clinical  
553 Neurosciences*. 2023;77:62–63.
- 554 46. Zhang T, Zeng J, Ye J, Gao Y, Hu Y, Xu L, et al. Serum complement proteins rather than  
555 inflammatory factors is effective in predicting psychosis in individuals at clinical high risk.  
556 *Translational Psychiatry*. 2023;13:9.
- 557 47. Chan MK, Krebs M, Cox D, Guest P, Yolken R, Rahmoune H, et al. Development of a blood-  
558 based molecular biomarker test for identification of schizophrenia before disease onset.  
559 *Translational Psychiatry*. 2015;5:e601–e601.
- 560 48. Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of  
561 a multivariable prognostic model: a resampling study. *Statistics in Medicine*. 2016;35:214–226.

Review: predicting psychosis using blood biomarkers

- 562 49. Riley RD, Debray TPA, Collins GS, Archer L, Ensor J, van Smeden M, et al. Minimum sample size  
563 for external validation of a clinical prediction model with a binary outcome. *Statistics in*  
564 *Medicine*. 2021;40:4230–4251.
- 565 50. Damen JAA, Moons KGM, van Smeden M, Hooft L. How to conduct a systematic review and  
566 meta-analysis of prognostic model studies. *Clinical Microbiology and Infection*.  
567 <https://doi.org/10.1016/j.cmi.2022.07.019>.
- 568 51. Moriarty AS, Meader N, Snell KI, Riley RD, Paton LW, Dawson S, et al. Predicting relapse or  
569 recurrence of depression: systematic review of prognostic models. *The British Journal of*  
570 *Psychiatry*. 2022:1–11.
- 571 52. Lee R, Leighton SP, Thomas L, Gkoutos GV, Wood SJ, Fenton S-JH, et al. Prediction models in  
572 first-episode psychosis: systematic review and critical appraisal. *The British Journal of*  
573 *Psychiatry*. 2022:1–13.
- 574 53. Wallace E, Smith SM, Perera-Salazar R, Vaucher P, McCowan C, Collins G, et al. Framework for  
575 the impact analysis and implementation of Clinical Prediction Rules (CPRs). *BMC Medical*  
576 *Informatics and Decision Making*. 2011;11:62.
- 577 54. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable  
578 prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Journal*  
579 *of British Surgery*. 2015;102:148–158.
- 580 55. Burger B, Vaudel M, Barsnes H. Importance of block randomization when designing proteomics  
581 experiments. *Journal of Proteome Research*. 2020;20:122–128.
- 582 56. Qin L-X, Zhou Q, Bogomolny F, Villafania L, Olvera N, Cavatore M, et al. Blocking and  
583 Randomization to Improve Molecular Biomarker Discovery Blocking and Randomization for  
584 Biomarker Discovery. *Clinical Cancer Research*. 2014;20:3371–3378.
- 585 57. Lång U, Yates K, Leacy FP, Clarke MC, McNicholas F, Cannon M, et al. Systematic Review and  
586 Meta-analysis: Psychosis Risk in Children and Adolescents With an At-Risk Mental State.  
587 *Journal of the American Academy of Child & Adolescent Psychiatry*. 2022;61:615–625.
- 588 58. Ajnakina O, Morgan C, Gayer-Anderson C, Oduola S, Bourque F, Bramley S, et al. Only a small  
589 proportion of patients with first episode psychosis come via prodromal services: a  
590 retrospective survey of a large UK mental health programme. *BMC Psychiatry*. 2017;17:1–9.
- 591 59. Burke T, Thompson A, Mifsud N, Yung AR, Nelson B, McGorry P, et al. Proportion and  
592 characteristics of young people in a first-episode psychosis clinic who first attended an at-risk  
593 mental state service or other specialist youth mental health service. *Schizophrenia Research*.  
594 2022;241:94–101.

Review: predicting psychosis using blood biomarkers

- 595 60. Bolhuis K, Lång U, Gyllenberg D, Kääriälä A, Veijola J, Gissler M, et al. Hospital presentation for  
596 self-harm in youth as a risk marker for later psychotic and bipolar disorders: a cohort study of  
597 59 476 Finns. *Schizophrenia Bulletin*. 2021;47:1685–1694.
- 598 61. Lång U, Ramsay H, Yates K, Veijola J, Gyllenberg D, Clarke MC, et al. Potential for prediction of  
599 psychosis and bipolar disorder in Child and Adolescent Mental Health Services: a longitudinal  
600 register study of all people born in Finland in 1987. *World Psychiatry*. 2022;21:436–443.
- 601 62. Cotter D, Healy C, Staines L, Mongan D, Cannon M. Broadening the parameters of clinical high  
602 risk for psychosis. *American Journal of Psychiatry*. 2022;179:593–595.
- 603 63. Wigman JT, van Nierop M, Vollebergh WA, Lieb R, Beesdo-Baum K, Wittchen H-U, et al.  
604 Evidence that psychotic symptoms are prevalent in disorders of anxiety and depression,  
605 impacting on illness onset, risk, and severity—implications for diagnosis and ultra-high risk  
606 research. *Schizophrenia Bulletin*. 2012;38:247–257.
- 607 64. Fusar-Poli P, Rutigliano G, Stahl D, Davies C, Bonoldi I, Reilly T, et al. Development and  
608 Validation of a Clinically Based Risk Calculator for the Transdiagnostic Prediction of Psychosis.  
609 *JAMA Psychiatry*. 2017;74:493–500.
- 610 65. Fusar-Poli P, Werbeloff N, Rutigliano G, Oliver D, Davies C, Stahl D, et al. Transdiagnostic Risk  
611 Calculator for the Automatic Detection of Individuals at Risk and the Prediction of Psychosis:  
612 Second Replication in an Independent National Health Service Trust. *Schizophrenia Bulletin*.  
613 2019;45:562–570.
- 614 66. Oliver D, Wong CMJ, Bøg M, Jönsson L, Kinon BJ, Wehnert A, et al. Transdiagnostic  
615 individualized clinically-based risk calculator for the automatic detection of individuals at-risk  
616 and the prediction of psychosis: external replication in 2,430,333 US patients. *Translational  
617 Psychiatry*. 2020;10:364.
- 618 67. Martin B, Wittmann M, Franck N, Cermolacce M, Berna F, Giersch A. Temporal structure of  
619 consciousness and minimal self in schizophrenia. *Frontiers in Psychology*. 2014;5.
- 620 68. Lunsford-Avery JR, Gonçalves B da SB, Brietzke E, Bressan RA, Gadelha A, Auerbach RP, et al.  
621 Adolescents at clinical-high risk for psychosis: Circadian rhythm disturbances predict worsened  
622 prognosis at 1-year follow-up. *Schizophrenia Research*. 2017;189:37–42.
- 623 69. Clementz BA, Parker DA, Trotti RL, McDowell JE, Keedy SK, Keshavan MS, et al. Psychosis  
624 Biotypes: Replication and Validation from the B-SNIP Consortium. *Schizophrenia Bulletin*.  
625 2022;48:56–68.
- 626 70. Sun G-W, Shook TL, Kay GL. Inappropriate use of bivariable analysis to screen risk factors for  
627 use in multivariable analysis. *Journal of Clinical Epidemiology*. 1996;49:907–916.

Review: predicting psychosis using blood biomarkers

- 628 71. Riley RD, Snell KIE, Martin GP, Whittle R, Archer L, Sperrin M, et al. Penalization and shrinkage  
629 methods produced unreliable clinical prediction models especially when sample size was small.  
630 *Journal of Clinical Epidemiology*. 2021;132:88–96.
- 631 72. Riley RD, Ensor J, Snell KI, Harrell FE, Martin GP, Reitsma JB, et al. Calculating the sample size  
632 required for developing a clinical prediction model. *Bmj*. 2020;368.
- 633 73. Riley RD, Snell KI, Ensor J, Burke DL, Harrell Jr FE, Moons KG, et al. Minimum sample size for  
634 developing a multivariable prediction model: PART II - binary and time-to-event outcomes.  
635 *Statistics in Medicine*. 2019;38:1276–1296.
- 636 74. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision  
637 curve analysis. *Diagnostic and Prognostic Research*. 2019;3:18.
- 638 75. Gardner ML, Freitas MA. Multiple Imputation Approaches Applied to the Missing Value  
639 Problem in Bottom-Up Proteomics. *International Journal of Molecular Sciences*. 2021;22.
- 640 76. van Os J, Guloksuz S. A critique of the “ultra-high risk” and “transition” paradigm. *World*  
641 *Psychiatry*. 2017;16:200–206.
- 642
- 643
- 644



Review: predicting psychosis using blood biomarkers

645 **Figure 1: PRISMA flow diagram of study selection.**

646

647 **Table 1: Study Characteristics**

648 Structured Interview for Psychosis Risk Syndromes (SIPS), Comprehensive Assessment of At-Risk  
649 Mental State (CAARMS), Social and Occupational Functioning Assessment Scale (SOFAS), micro-  
650 ribonucleic acid (miRNA), single-nucleotide polymorphisms (SNPs), least-absolute shrinkage and  
651 selection operator (LASSO), support vector machine (SVM), bayesian generalised linear model  
652 (bayesglm), logistic regression (LR), leave one site out cross-validation (LOSOCV).

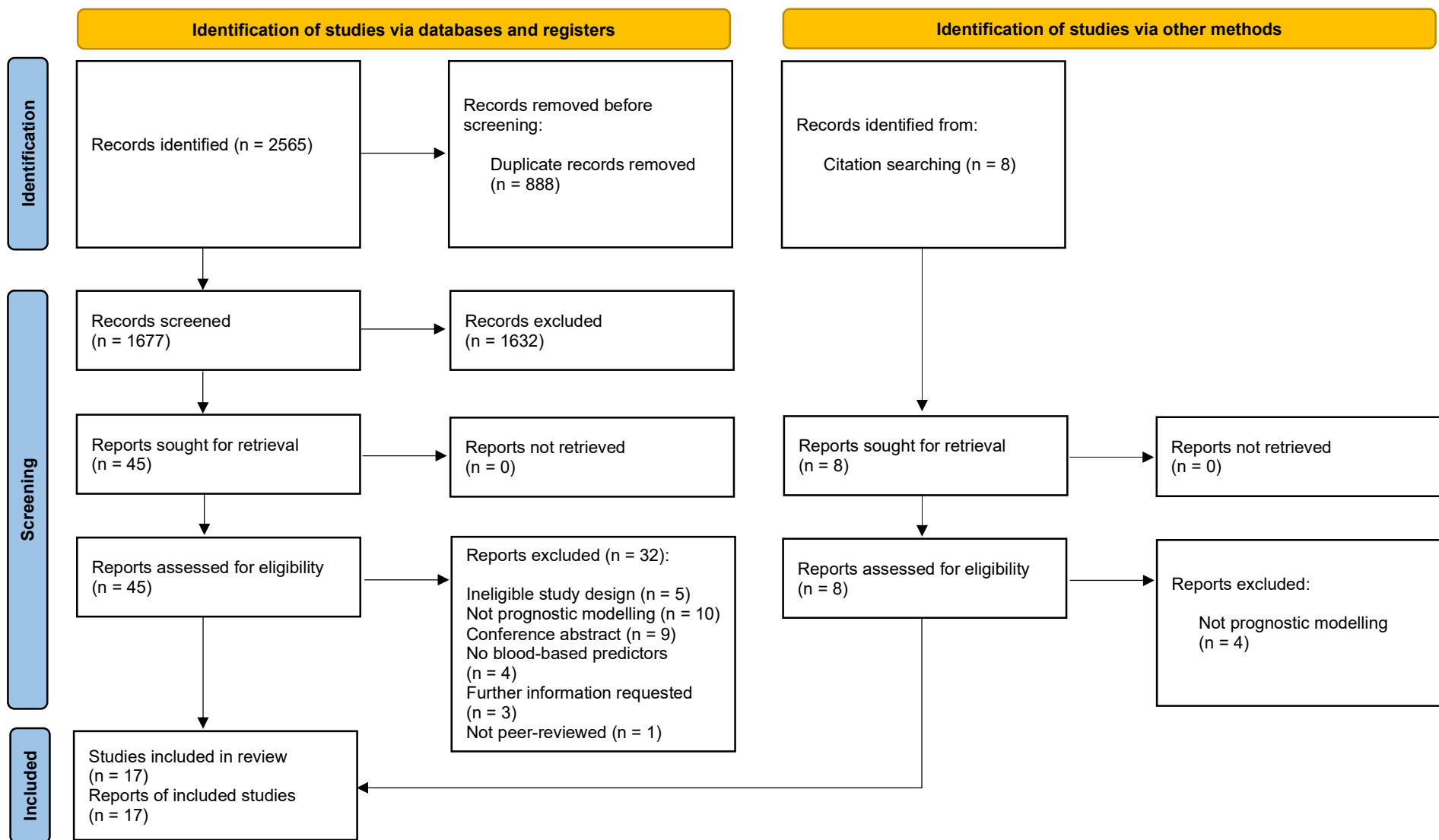
653

654 **Figure 2: Results of risk of bias assessments in each domain.**

655

656 **Figure 3: Summary of applicability concerns in each domain.**

Review: predicting psychosis using blood biomarkers



| Study                                     | Bousman et al., 2018         | Clark et al., 2016 | Dickens et al., 2021 | Chan et al., 2015                                     | Jeffries et al., 2016                | Koutsouleris et al., 2021               | Lavoie et al., 2017 | Li et al., 2022   | Mondelli et al 2023    | Mongan et al., 2021                     | Perkins et al., 2015                                  | Perkins et al., 2020 | Song et al., 2022 | Tavares et al., 2023   | Zhang et al., 2022  | Zhang et al., 2023a | Zhang et al., 2023b               |
|---|------------------------------|--------------------|----------------------|---|--------------------------------------|---|---------------------|-------------------|------------------------|---|---|----------------------|-------------------|------------------------|---------------------|---------------------|-----------------------------------|
| Type of study                             | Model Development            | Model Development  | Model Development    | Model Validation                                      | Model Development                    | Model Development                       | Model Development   | Model Development | Model Development      | Model Development                       | Model Development                                     | Model Development    | Model Development | Model Development      | Model Development   | Model Development   | Model Development                 |
| Study design                              | Cohort                       | RCT                | Cohort               | Cohort  | Cohort                               | Cohort                                  | RCT                 | Cohort            | Cohort                 | Cohort                                  | Cohort  | Cohort               | Cohort            | Cohort                 | Cohort              | Cohort              | Cohort                            |
| Study setting                             | Outpatient                   | Outpatient         | Outpatient           | Outpatient  | Outpatient                           | Outpatient                              | Outpatient          | Outpatient        | Outpatient             | Outpatient                              | Outpatient  | Outpatient           | Outpatient        | Outpatient             | Outpatient          | Outpatient          | Outpatient                        |
| Location                                  | Korea                        | Austria            | International        | France  | North America                        | Europe                                  | Austria             | China             | International          | International                           | North America   | North America        | China             | United Kingdom         | China               | China               | China                             |
| N =                                       | 134                          | 40                 | 263                  | 76  | 67                                   | 201                                     | 36                  | 90                | 269                    | 133                                     | 72  | 595                  | 56                | 75                     | 84                  | 37                  | 49                                |
| N transitions                             | 20                           | 11                 | 50                   | 18  | 30                                   | 25                                      | 15                  | 23                | 50                     | 49                                      | 32  | 84                   | NI                | 21                     | 16                  | 8                   | 25                                |
| N (transitions) for external validation   | N/A                          | N/A                | N/A                  | 76 (18)   | N/A                                  | N/A                                     | N/A                 | N/A               | N/A                    | N/A                                     | N/A   | N/A                  | N/A               | N/A                    | N/A                 | N/A                 | N/A                               |
| Age, mean (SD):                           | 21 (4)                       | 16 (1.7)           | 22.5 (4.8)           | 21.5 (3.9)  | 18.4 (3.7)                           | NI (full plus 18M cohort 24.6 (5.6))    | 15.8 (1.8)          | 18.1 (4.4)        | 22.6 (4.9)             | 22.8 (4.5)                              | 19.4 (4.2)  | 18.5 (4.3)           | 17.7 (4.2)        | 22.9 (4.4)             | 18.5 (5.2)          | 17.6 (3.7)          | 18.2 (4.1)                        |
| Sex, n (% female):                        | 85 (63.4%)                   | 27 (67.5%)         | 122 (46.4%)          | 32 (42.1%)  | 22 (32.8%)                           | NI (full plus 18M cohort 51.6%)         | 25 (69.4%)          | 38 (42.2%)        | 122 (45.4%)            | 65 (48.9%)                              | 25 (34.7%)  | 251 (42.2%)          | 25 (44.6%)        | 31 (41.3%)             | 35 (41.7%)          | 18 (48.6%)          | 14 (28.6%)                        |
| Method of risk assessment                 | SIPS                         | CAARMS-equivalent  | CAARMS               | SOFAS; N/A  | SIPS                                 | SIPS                                    | CAARMS-equivalent   | SIPS              | CAARMS                 | CAARMS                                  | SIPS  | SIPS                 | SIPS              | CAARMS                 | SIPS                | SIPS                | SIPS                              |
| Biomarkers measured:                      | SNPs of cytokine genes       | Fatty acids        | Lipidome             | Hormonal, inflammatory and metabolic-related analytes | Leukocyte miRNAs                     | Genome                                  | Glutathione         | Metabolome        | Cytokines              | Proteome                                | Hormonal, inflammatory and metabolic-related analytes | Genome               | Transcriptome     | Genome                 | Cytokines           | Cytokines           | Cytokines and complement proteins |
| No. of events per candidate predictor     | 0.87                         | 0.55               | 0.29                 | 0.78  | 0.22                                 | 1.25                                    | 15                  | 0.0012            | 1.85                   | 0.21                                    | 0.27  | 9.33                 | 0.00053           | 21                     | 16                  | 2.67                | 1.39                              |
| Modelling method                          | Crosstab/Logistic Regression | Bayes' Rule        | Ridge Regression     | LASSO Regression                                      | Coarse approximate linear classifier | SVM                                     | Cox regression      | Bayesglm          | Elastic Net Regression | SVM                                     | Coarse approximate linear classifier                  | Cox regression       | LASSO Regression  | Elastic Net Regression | Logistic regression | Logistic regression | Logistic regression               |
| Method used for testing model performance | Apparent                     | Apparent           | Internal             | External  | Internal                             | Internal-external (geographical)        | Apparent            | Apparent          | Internal               | Internal-external (geographical)        | Internal  | Internal             | Internal          | Internal               | Apparent            | Apparent            | Apparent                          |
| Type of internal validation if used       | N/A                          | N/A                | Bootstrap            | N/A   | Cross validation                     | Nested cross validation (outer: LOSOCV) | N/A                 | N/A               | Cross validation       | Nested cross validation (outer: LOSOCV) | Cross validation                                      | Bootstrap            | Cross validation  | Cross validation       | N/A                 | N/A                 | N/A                               |

Risk of bias domains

|                           | D1 | D2 | D3 | D4 | Overall |
|---------------------------|----|----|----|----|---------|
| Bousman et al., 2018      | -  | -  | +  | X  | X       |
| Chan et al., 2015         | -  | +  | -  | X  | X       |
| Clark et al., 2016        | -  | -  | -  | X  | X       |
| Dickens et al., 2021      | +  | -  | -  | X  | X       |
| Jeffries et al., 2016     | -  | -  | +  | X  | X       |
| Koutsouleris et al., 2021 | +  | -  | +  | X  | X       |
| Lavoie et al., 2017       | -  | +  | +  | X  | X       |
| Li et al., 2022           | -  | -  | +  | X  | X       |
| Mondelli et al., 2023     | +  | +  | -  | X  | X       |
| Mongan et al., 2020       | +  | +  | -  | X  | X       |
| Perkins et al., 2016      | -  | -  | +  | X  | X       |
| Perkins et al., 2017      | -  | +  | +  | X  | X       |
| Song et al., 2022         | -  | -  | -  | X  | X       |
| Tavares et al., 2023      | +  | -  | +  | X  | X       |
| Zhang et al., 2022        | +  | -  | +  | X  | X       |
| Zhang et al., 2023a       | +  | -  | +  | X  | X       |
| Zhang et al., 2023b       | -  | -  | +  | X  | X       |

Domains:

D1: Bias introduced by selection of participants.

D2: Bias introduced by predictors or their assessment.


D3: Bias introduced by the outcome or its determination.

D4: Bias introduced by the analysis.

Judgement

 High risk of bias

 Unclear

 Low risk of bias

Study

|                           | Applicability domains |    |    | Overall |
|---------------------------|-----------------------|----|----|---------|
|                           | D1                    | D2 | D3 |         |
| Bousman et al., 2018      | -                     | +  | +  | -       |
| Chan et al., 2015         | -                     | +  | -  | -       |
| Clark et al., 2016        | -                     | +  | +  | -       |
| Dickens et al., 2021      | +                     | +  | +  | +       |
| Jeffries et al., 2016     | -                     | -  | +  | -       |
| Koutsouleris et al., 2021 | +                     | +  | +  | +       |
| Lavoie et al., 2017       | -                     | +  | +  | -       |
| Li et al., 2022           | -                     | -  | +  | -       |
| Mondelli et al., 2023     | +                     | +  | +  | +       |
| Mongan et al., 2020       | +                     | -  | +  | -       |
| Perkins et al., 2016      | -                     | +  | +  | -       |
| Perkins et al., 2017      | -                     | +  | +  | -       |
| Song et al., 2022         | -                     | -  | +  | -       |
| Tavares et al., 2023      | +                     | +  | +  | +       |
| Zhang et al., 2022        | +                     | +  | +  | +       |
| Zhang et al., 2023a       | +                     | +  | +  | +       |
| Zhang et al., 2023b       | -                     | +  | +  | -       |

Domains:

D1: Applicability concern regarding the included participants and setting.

D2: Applicability concern regarding the definition, assessment or timing of predictors.

D3: Applicability concern regarding the outcome, its definition, timing or determination.

Judgement

- Unclear

+ Low concern

Study