Citation for published version (APA):
BIOMARKERS FOR PSYCHOSIS: ARE WE THERE YET?

UMBRELLA REVIEW OF 1478 BIOMARKERS

Paola Fuentes-Claramonte1,2†, Andrés Estradé3†, Aleix Solanes4,5, Valentina Ramella3, Maria Angeles Garcia-Leon1,2, Javier de Diego Adelino2,5,6, Conrad Molins7, Eric Fung1, Marc Valent2,4,8, Gerard Anmella2,4,8, Edith Pomarol-Clotet1,2, Dominic Oliver3,8-10, Eduard Vieta2,4,11, Joaquim Radua2,3,4,12+, Paolo Fusar-Poli3,13,+†

† Shared first authorship.
+ Shared senior authorship.

1FIDMAG Germanes Hospitalàries Research Foundation, Barcelona, Spain
2Biomedical Research Networking Centre Consortium on Mental Health (CIBERSAM), Barcelona, Spain
3Early Psychosis: Interventions & Clinical-detection (EPIC) Lab, Department of Psychosis Studies, Institute of Psychiatry Psychology and Neuroscience, King's College London, London, United Kingdom
4Institut d’Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain
5Barcelona Autonomous University (UAB), Barcelona, Spain
6Hospital de la Santa Creu i Sant Pau, Biomedical Research Institute Sant Pau (IIB Sant Pau), Barcelona, Spain
7Psychiatric Service, Hospital Universitari Santa Maria, Lleida, Catalonia, Spain
8Department of Psychiatry, University of Oxford, Oxford OX3 7JX, UK
9NIHR Oxford Health Biomedical Research Centre, Oxford OX3 7JX, UK
10OPEN early detection service, Oxford Health NHS Foundation Trust, Oxford OX3 7JX, UK†Bipolar and Depressive Disorders Unit, Institute of Neuroscience, Hospital Clinic de Barcelona, University of Barcelona, Barcelona, Spain
11Center for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
12OASIS service, South London and the Maudsley NHS Foundation Trust, London, United Kingdom

Corresponding author: Andrés Estradé, andres.estrade_vaz@kcl.ac.uk, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, 16 De Crespigny Park, London, SE5 8AF.

Running title: Umbrella review of 1478 psychosis biomarkers

Word count:

- Abstract: 248
- Body text (exc. Col and Funding): 4,464
- Figures: 4
- Tables: 1
**ABSTRACT**

**Background and Hypothesis:** This umbrella review aims to comprehensively synthesize the evidence of association between peripheral, electrophysiological, neuroimaging, neuropathological, and other biomarkers and diagnosis of psychotic disorders.

**Study Design:** We selected systematic reviews and meta-analyses of observational studies on diagnostic biomarkers for psychotic disorders, published until February 1st, 2018. Data extraction was conducted according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines. Evidence of association between biomarkers and psychotic disorders was classified as convincing, highly suggestive, suggestive, weak, or non-significant, using a standardized classification. Quality analyses used the Assessment of Multiple Systematic Reviews (AMSTAR) tool.

**Study Results:** The umbrella review included 110 meta-analyses or systematic reviews corresponding to 3,892 individual studies, 1,478 biomarkers, and 392,210 participants. No factor showed a convincing level of evidence. Highly suggestive evidence was observed for transglutaminase autoantibodies levels (odds ratio [OR]=7.32; 95% CI: 3.36, 15.94), mismatch negativity in auditory event-related potentials (standardised mean difference [SMD]=0.73; 95% CI: 0.5, 0.96), P300 component latency (SMD=-0.6; 95% CI: -0.83, -0.38), ventricle brain ratio (SMD=0.61; 95% CI: 0.5, 0.71), and minor physical anomalies (SMD=0.99; 95% CI: 0.64, 1.34). Suggestive evidence was observed for folate, malondialdehyde, brain-derived neurotrophic factor, homocysteine, P50 sensory gating (P50 S2/S1 ratio), frontal N-acetyl-aspartate, and high-frequency heart rate variability. Among the remaining biomarkers, weak evidence was found for 626 and non-significant association for 833 factors.

**Conclusions:** While several biomarkers present highly suggestive or suggestive evidence of association with psychotic disorders, methodological biases and underpowered studies call for future higher-quality research.

*Key words:* schizophrenia, psychotic disorders, peripheral biomarkers, electrophysiological biomarkers, neuroimaging biomarkers, neuropathological biomarkers.
INTRODUCTION

Schizophrenia spectrum and other psychotic disorders have an estimated mean lifetime prevalence of 9.57 per 1000 and usually begin during youth and early adulthood (meta-analytic peak age of onset at 20.5 years). The initial diagnosis of psychosis usually occurs at the time of the first episode of psychosis (FEP), occurring at a late stage in the neurodevelopmental trajectory of the disorder. Following a FEP, standard care with antipsychotics is primarily symptom-focused and limited for altering the course of the disorder. Psychotic disorders continue to be associated with poor clinical outcomes (one in seven recovery rate) and a 3-fold excess mortality rate when compared to general population (standardized mortality rate of 3.08) due to physical illness and increased suicide risk in the initial years following illness onset. Schizophrenia is a global leading cause of health-related disability, with associated global economic costs of up to 1.65% of the gross domestic product. Given the suboptimal results of current therapeutics, course altering preventive interventions during the early developmental stages, such as cannabidiol or oxytocin, constitute promising future avenues for next-stage interventions. Course-altering interventions require a refined understanding of the pathophysiology and neurobiological substrates of psychosis. After decades of research, however, the causes of psychotic disorders remain elusive. The etiological models that have received the strongest empirical support suggest a complex combination of direct and interactive effect of genetic, epigenetic and environmental factors across the developmental cycle that interfere with brain development and maturation.

Nonetheless, the past decades have witnessed an explosion of psychosis biomarker research. Biomarkers provide clues for understanding the pathophysiological basis of psychosis, and could become key tools in the real-world implementation of precision medicine and individualised prediction modelling. For example, biomarkers can aid in developing mechanism-driven preventive interventions, identifying illness subtypes via biological screening, or in the introduction of novel diagnostic frameworks based on pathophysiology of mental disorders. As such, biomarkers could be helpful for ascertaining the presence of a disorder or specific illness subtypes (‘diagnostic’ biomarkers), predicting therapeutic response (‘predictive’ biomarkers) and course of the disorder (‘prognostic’ biomarkers), and for monitoring illness progression (‘monitoring’ biomarkers). However, the progress achieved in biomarker research has not been fully translated into real-world clinical practice, hence the current ‘translational gap’. Over 100 years following Bleuler’s introduction of the term ‘schizophrenia’, diagnosis is still based on clinical examination in accordance with DSM (Diagnostic and Statistical Manual of Mental Disorders) or ICD (International Classification of Diseases) diagnostic criteria. General psychopathological
knowledge, and evidence-interventions, are implemented following a trial-and-error approach according to a general clinical profile.23,32

One obstacle for the translational potential of scientific findings is fragmentation of knowledge as primary biomedical studies proliferate and diversify.33 This also affects secondary research due to the accumulation of overlapping, often contradictory, systematic reviews (SRs) and meta-analyses (MAs).34,35 Umbrella reviews help overcome this challenge by providing a synthesis and critical appraisal of SRs and MAs. For example, umbrella reviews for schizophrenia and psychotic disorders have focused on sociodemographic, developmental and environmental risk factors, preventive treatments, or duration of untreated psychosis.39 Regarding biomarker research, umbrella reviews on schizophrenia and FEP have been limited to peripheral biomarkers, while also incorporating environmental exposures or other severe mental disorders. As such, to the best of our knowledge, no comprehensive umbrella review of biomarkers for psychosis has been published to date. Our study aims to close this gap by providing a state-of-the-art comprehensive evidence synthesis on the association between peripheral, electrophysiological, neuroimaging, neuropathological and other biomarkers and psychotic disorders. We focus on observational studies of diagnostic biomarkers comparing individuals with psychosis vs healthy controls, as the analysis of prognostic markers would require a different design and evidence synthesis method which accounts for time-dependency of outcomes. To increase sampling power, we use an extended sampling approach by including a diagnosis of any non-organic psychotic spectrum disorder, rather than limiting our selection to the chronic and more symptomatic cases, such as schizophrenia.

In addition, biomarker research is often affected by underpowered samples, methodological biases, and inconsistent reporting practices. To address this challenge, we provide a hierarchical classification of the robustness of the association for each factor. To achieve this, we conduct a systematic analysis of biases through a set of a priori criteria extensively validated in previous risk factors studies for physical, neurological, and mental disorders, as well as in clinical studies. This classification into hierarchical levels of evidence is essential for reducing the ambiguities and contradictions often found in SRs and MAs.

**METHODS**

We pre-registered the umbrella review protocol with the International Prospective Register of Systematic Reviews (PROSPERO; CRD42017084377).
Search strategy and eligibility criteria

Various researchers (PF-C and VR; plus JDA, EF, and CM for the neuroimaging part) systematically and independently searched PubMed, Web of Science, and Scopus through February 1st, 2018, using the search terms (“systematic review” OR “meta-analysis”) and (“psychosis” OR “schizophrenia”), to identify SRs and MAs of studies examining potential diagnostic biomarkers for psychotic disorders. Reference lists of the SRs and MAs reaching full-text review were also carefully reviewed. Eligibility criteria included: 1) a SR or MA of individual observational studies examining associations between biological markers and psychotic disorders; 2) studies considering only established DSM or ICD diagnoses of non-organic psychotic spectrum disorders (e.g., schizophrenia, schizoaffective, schizophreniform, affective psychosis [mania, depression or bipolar disorder with psychosis], drug-induced psychosis, delusional disorder, brief psychotic disorder/acute and transient psychotic disorder, psychosis not otherwise specified); 3) inclusion of a healthy control comparison group, and 4) studies reporting sufficient data to perform the analyses (or where data were retrievable from the authors). No language restrictions were applied.

When the biomarker dataset of an article was part of a larger dataset in another article, we only retained the latter. When two articles presented minimally overlapping datasets on the same biomarker, we used both SR or MA conjointly counting the overlapping primary studies only once. We excluded articles with an outcome other than established psychotic disorder, such as those related to relapse, remission, or treatment response of psychosis or symptom severity, and those investigating genetic markers for psychosis. We used the same inclusion/exclusion criteria for each study included in every eligible SR or MA.

Definition of biomarker

We used the following accepted definition of biomarker: "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes or pathogenic processes". We did not include potential genetic biomarkers because umbrella reviews of genetic variables require different analytical methods and criteria. Neither did we include potential biomarkers from whole-brain voxel-based neuroimaging studies (although we did include other types of neuroimaging data) because we would need to treat each voxel as a biomarker. Instead, we refer the reader to existing MAs of whole-brain imaging studies in psychosis.

We used the definition for each biomarker provided in the corresponding SR or MA. However, for reporting purposes, we classified biomarkers into the following categories: peripheral, electrophysiologic, neuropathological, neuroimaging, and other (e.g. minor physical...
anomalies, high frequency heart rate variability). The clustering of different biomarkers was pragmatically operationalised following the definition presented by each individual study. These groups hold only descriptive value as the actual analyses were performed at the single marker level. There was no assumption of biological mechanisms underlying these categories, most of which are yet to be fully elucidated.

Data extraction and selection

Data extraction was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA 2020) guidelines. Various investigators (PF-C and MAG-L; plus JDA, EF, CM, and AS for the neuroimaging part) conducted the following steps independently. First, we identified the potential biomarkers assessed in each selected SR or MA. Second, we confirmed that each article included in the SR or MA met our eligibility criteria for the umbrella review (i.e., ICD/DSM diagnosis, healthy control group, and sufficient data for analysis). Third, we extracted the following data (from the SR or MA or, otherwise, from the individual study): 1) first author and year of publication, 2) the number of cases and controls, 3) effect size (ES) measure (standardized mean difference [SMD] for continuous biomarkers, odds ratio [OR] for binary biomarkers) and corresponding 95% confidence interval (CI), 4) means and standard deviations for cases and controls for continuous biomarkers, and the number of cases and controls with and without the biomarker for binary biomarkers. An exception was neuroimaging biomarkers, for which we only relied on the information reported in the MAs unless this information showed that the biomarker's evidence could be stronger than class IV (see "Evidence stratification" below). An independent double extraction process was conducted for those biomarkers that presented unclear data. Fourth, for those biomarkers showing evidence of class I, II, or III, we rated the quality of the SRs or MAs that contributed studies for that biomarker using the Assessment of Multiple Systematic Reviews (AMSTAR) tool. Our quality ratings obtained a high interrater agreement (intraclass correlation = 0.924). For further information on the quality analysis, see the Supplementary Material (sTable 1).

Statistical analyses

We conducted all analyses with the package "metaumbrella" for R (https://metaumbrella.org/), that performs all the calculations necessary for stratifying evidence in umbrella reviews. Specifically, it calculated, for each biomarker, the ES (Hedges' g for continuous variables, OR for binary variables) and its confidence and prediction intervals, the between-study heterogeneity ($I^2$ statistic), the Egger test to detect
potential publication/reporting bias, and the excess significance bias test. Hedges' g values represent SMDs between the patient and control groups and are interpreted as indicative of a 'small' (g=0.2), 'medium' (g=0.5) or 'large' effect (g=0.80). OR provides a measure of the likelihood of presenting any specific biomarker in cases vs healthy controls, with OR>1 indicating increased likelihood and OR<1 decreased likelihood. The I² statistic represents the percentage of total variance resulting from heterogeneity (i.e. real differences in the studies' ES), rather than chance. Egger's test quantifies the relationship between sample size and ES, with significant results indicating risk of publication bias. Finally, the excess significance bias test evaluates the relative presence of studies with excessive significant findings, by comparing the observed vs the expected number of studies with significant results.

**Evidence stratification**

We classified the strength of the evidence according to previous criteria: class I (convincing) when the number of patients >1000, p<10⁻⁶, I²<50%, the 95% prediction interval excludes the null, and no publication/reporting or excess significance biases are detected; class II (highly suggestive) when the evidence is weaker than convincing but the number of patients >1000, p<10⁻⁶, and the largest study is statistically significant; class III (suggestive) when the evidence is weaker than highly suggestive but the number of patients >1000 and p<10⁻³; and class IV (weak) when the evidence is lower than suggestive but p<0.05.

**RESULTS**

We included 110 SRs and MAs (sFigure 1) covering 1,478 potential biomarkers with a cumulative sample size of 189,180 individuals with psychosis and 203,030 healthy controls. In Table 1 we include the biomarkers that achieved highly suggestive (class II) or suggestive (class III) level of evidence. For the full list of biomarkers, see the Supplementary Material (sTables 2-6).

**Peripheral biomarkers**

We included 284 peripheral biomarkers from 63,390 patients and 79,410 controls (cumulated sample sizes; sTable 2). Transglutaminase (tTG) autoantibody levels achieved highly suggestive (class II) evidence with an OR = 7.32 (CI: 3.36, 15.94) (Table 1; Figure 1); we could not assess heterogeneity and potential biases because only one study was available for
this analysis. Blood levels of folate, malondialdehyde (MDA), brain-derived neurotrophic factor (BDNF), and homocysteine (Hcy) showed class III (suggestive) evidence of association, with effect sizes of -1.44 (CI: -2.18, -0.71), 1.38 (CI: 0.82, 1.94), -0.69 (CI: -1.05, -0.33) and 0.60 (CI: 0.31, 0.89), respectively. These class III biomarkers showed very large heterogeneity and (except for folate) potential excess significance bias. One hundred sixty-three other peripheral biomarkers achieved class IV evidence.

Electrophysiologic biomarkers

We examined 79 electrophysiologic biomarkers in 18,151 patients vs 19,346 controls (cumulated sample sizes; sTable 3). Two biomarkers achieved class II evidence: mismatch negativity in auditory event-related potentials \( g = 0.73, \text{CI}: 0.50, 0.96 \) and P300 component latency \( g = -0.60, \text{CI}: -0.83, -0.38 \) (Table 1; Figure 2). However, they showed very large heterogeneity and potential excess significance bias. On the other hand, P50 sensory gating achieved class III evidence with an ES of \( g = 0.79 \) (CI: 0.43, 1.16), but showing very large heterogeneity. Forty-one other biomarkers achieved class IV evidence.

Neuroimaging biomarkers

We included 238 neuroimaging biomarkers in a cumulative sample size of 77,849 patients vs. 73,184 controls (sTable 4). Two of them showed class II or III evidence: ventricle-brain ratio (class II, \( g = 0.61, \text{CI}: 0.50, 0.71 \)) and frontal N-acetylaspartate (NAA) levels (class III, \( g = -0.34, \text{CI}: -0.47, -0.20 \)) (Table 1; Figure 3). The analyses suggested potential excess significance bias in both cases, and the heterogeneity was large (>50%). One hundred forty-nine other neuroimaging biomarkers achieved class IV evidence.

Neuropathological biomarkers

We studied a total of 406 neuropathological biomarkers in a cumulative sample size of 6,170 patients vs. 6,526 controls (sTable 5). One hundred thirteen biomarkers showed a statistically significant association with psychosis, all with class IV (weak) evidence.

Other biomarkers

The other biomarkers category enclosed 471 biomarkers in 23,620 patients vs. 24,564 controls (cumulated sample sizes; sTable 6). Minor physical anomalies achieved class II evidence \( g = 0.99, \text{CI}: 0.64, 1.34 \) (Table 1; Figure 4), although with very large heterogeneity
and potential excess significance bias. High-frequency heart rate variability achieved class III evidence (\(g = -0.99, \text{CI: } -1.41, -0.58\)), although, beyond very large heterogeneity, it also showed potential publication and excess significance biases. One hundred sixty other biomarkers achieved class IV evidence.

**DISCUSSION**

To the best of our knowledge, this is the first comprehensive umbrella review of diagnostic biomarkers for psychosis that incorporates peripheral, electrophysiologic, neuroimaging, neuropathological, and other biomarkers, in addition to a robust hierarchical classification of evidence. Overall, 110 SRs and MAs, a total of 3,892 individual studies and 392,210 participants, and 1,478 potential biomarkers were included. None of the evaluated biomarkers presented convincing evidence (class I) of association. Nonetheless, there was highly suggestive evidence (class II) for five biomarkers (0.3% of all evaluated factors), and suggestive evidence (class III) for seven biomarkers (0.5%). Among the remaining biomarkers, 626 (42.6%) presented significant but weak (class IV) evidence. The remaining 833 (56.6%) presented non-significant associations.

Associations supported by highly suggestive and suggestive evidence merit discussion. Regarding peripheral biomarkers, highly suggestive evidence was only found for elevated levels of tTG autoantibodies. This finding provides evidence, at for a subset of individuals, towards a link between psychosis and celiac disease\(^73,74\). The relationship between celiac disease or gluten sensitivity and neurologic and psychiatric disorders was initially reported almost 70 years ago\(^75,76\), with celiac disease having since been linked to an increased risk of various neurological\(^77\) and mental disorders\(^78\). In addition, it provides evidence for psychosis as a systemic disorder with potential autoimmune components\(^79,80\). While the specific nature of the association remains unknown, possible mechanisms include shared genetic susceptibility or immunological abnormalities\(^81\). However, the estimate is based on a single study. As a result, heterogeneity or publication bias could not be assessed. Nonetheless, the large sample of cases (n=2,301) grants this association a class II level of evidence. In addition, we found suggestive evidence for reduced blood levels of folate among individuals with psychosis, although this association was characterized by very large heterogeneity. This finding further indicates potential nutrition-related mechanisms in mental disorders\(^82-84\). For example, maternal exposure to nutritional deprivation during pregnancy\(^85\) could result in deficits in micronutrients involved in 1-carbon metabolism, particularly folate (vitamin B9) and vitamin B12, leading to increased risk due to epigenetic changes via disruptions in DNA methylation\(^86-88\). Findings of decreased levels of vitamin B12 have been inconsistent\(^89,90\) and
additional epidemiological and laboratory studies are required. We also acknowledge that there might be a bi-directional effect of psychosis on diet and therefore no causal assumption is made when interpreting diet-related biomarkers. The scarcity of prospective studies is problematic, as folate and vitamin B12 levels could be confounded by use of antipsychotics. We also observed suggestive evidence for increased plasma levels of Hcy, further indicating a potential role of 1-carbon metabolism. Hcy is a non-protein amino acid produced in 1-carbon methyl group-transfer metabolism with various functions in brain activity, and its regulation depends on dietary folate and other B vitamins. In schizophrenia patients, plasma Hcy levels were found to correlate negatively with folate and vitamin B12, and to be associated with symptom severity and a progressive course of illness. Elevated Hcy has been observed in FEP patients, although not in CHR-P individuals. However, research in the CHR-P population is still in an emerging state. Potential mechanisms involved in the association between Hcy and psychosis include aberrant DNA methylation, altered NMDA receptor and glutamatergic transmission, toxic effects on dopaminergic neurons, premature apoptosis, oxidative stress, and placental vascular damage and fetal hypoxia. Nonetheless, the question remains as to whether hyperhomocysteinemia is a contributor, a consequence or an epiphenomenon of psychosis. Adjunct folate therapy has been found to improve symptoms of depressive and bipolar disorders, but not for schizophrenia. However, emerging evidence suggests moderate effectiveness of pooled vitamin B (i.e. B6, B9, B12) supplementation on total schizophrenia symptoms. It is still unclear what specific symptoms and clinical groups (e.g. FEP, chronic) would benefit the most from nutritional interventions, and an opportunity for clinical stratification might exist. For example, folic acid appears to be more beneficial for the treatment of negative vs positive symptoms. Likewise, a recent RCT study suggested that B-vitamin supplementation could have neuroprotective effects in FEP patients with elevated Hcy. Also, it is unclear how nutritional deficiencies might interact with genetic variants. For example, genetic variants linked to folate-metabolism can be associated with clinical response to adjunct folate therapy. Overall, these findings call for further research into gene-environment interaction and potential avenues for personalized medicine.

Finally, regarding peripheral biomarkers, elevated levels of MDA and reduced levels of BDNF also presented suggestive evidence which, despite large heterogeneity, was based on large samples (n=1,795 and n=4,955, respectively). Alongside increased Hcy, Elevated MDA and reduced BDNF levels suggest a role for oxidative stress and inflammation in psychosis pathophysiology, possibly linked to immune dysregulation. Immune dysregulation in psychosis has been supported by genome-wide association and postmortem studies. Early-life adversity, including prenatal insults and childhood trauma, can also
promote a proinflammatory state later in life, potentially via epigenetic changes\textsuperscript{117,118}. We also observed weak evidence for other oxidative stress and inflammation biomarkers, such as docosahexaenoic acid (DHA), IL-6, TNF-alpha, and total antioxidant status. Overall, these findings are consistent with evidence suggestive of antioxidant status and pro-inflammatory imbalances among FEP patients\textsuperscript{101} and CHR-P individuals\textsuperscript{119}. However, these associations are often weakened by small samples, scarcity of prospective studies, and high heterogeneity potentially resulting from between-study sample differences (e.g. illness phase, developmental stage, antipsychotic status, and comorbid substance-user or other lifestyle factors)\textsuperscript{101,120}. Heterogeneity and small effect sizes might further indicate that immune system involvement varies between individuals and across illness phases\textsuperscript{121}. For example, anti-inflammatory agents have been found to be more effective for early-phase vs chronic schizophrenia\textsuperscript{122}. Also, anti-inflammatory biomarkers can be associated with antipsychotic treatment response\textsuperscript{123}. As such, an opportunity might also exist for personalized medicine approaches\textsuperscript{22} via inflammatory biomarker-informed patient stratification.

In terms of electrophysiologic biomarkers, mismatch negativity (MMN) in auditory event-related potentials (ERP) achieved highly suggestive evidence. Reduced MMN is a well-replicated feature in first-episode and chronic psychosis\textsuperscript{28,124–126}, and might reflect N-methyl-D-aspartate (NMDA) receptor function\textsuperscript{127,128}. MMN does not seem to always distinguish between early and chronic schizophrenia\textsuperscript{129,130}, and has been suggested to stabilize following deterioration during the first 1-2 years of diagnosis\textsuperscript{124}. In addition, abnormal auditory MMN has also been observed to precede illness onset among CHR-P individuals\textsuperscript{131–133}, although not consistently\textsuperscript{130}, and with other studies reporting similar patterns in CHR-P and FEP individuals\textsuperscript{134}. Overall, there is translational potential for MMN as a biomarker for monitoring illness progression and the identification of individuals at greater risk of transition\textsuperscript{124,135}. In addition, we observed highly suggestive evidence for P300 component latency. Significant evidence, albeit weak, was also observed for P300 amplitude disturbances. Both P300 component latency and amplitude, a more direct biomarker of cognition than MMN\textsuperscript{136–138}, have been widely replicated in first-episode\textsuperscript{39–141} and chronic psychosis patients\textsuperscript{142,143}, and first-degree relatives\textsuperscript{144}. P300 anomalies have also been reported to respond to antipsychotic treatment in schizophrenia patients\textsuperscript{145}, and might be linked to grey matter volume reduction in both CHR-P\textsuperscript{146,147} and psychosis samples\textsuperscript{141}. As such, ERP deficits have potential as biomarkers for transition risk in clinical high-risk individuals\textsuperscript{140,148–150} in whom both neurocognitive decline\textsuperscript{151} and reduction in grey matter volume\textsuperscript{152} precede illness onset. Finally, P50 sensory gating (P50 S2/S1 ratio) presented suggestive evidence, further supporting the existence of ERP deficits in psychosis\textsuperscript{28}. All class II and class III electrophysiologic biomarkers were affected by large heterogeneity.
Among neuroimaging biomarkers, we observed highly suggestive evidence of increased ventricle-brain ratio among individuals with psychosis. Since the pioneer neuroimaging studies of the 1970-1980’s\textsuperscript{153-156}, ventricular enlargement has remained one of the most replicated neuroanatomical correlated of schizophrenia, as reflected in the large number of participants in our summary (n=6,099). Ventricular enlargement has also been observed in unaffected first-degree relatives\textsuperscript{157} and in transitioning CHR-P samples\textsuperscript{158}. In addition, decreased frontal NAA achieved suggestive level of evidence. Weak evidence for reduced NAA was also present for other brain regions (e.g. temporal, cerebellum, thalamus), further supporting an association between psychosis and loss of neuronal integrity\textsuperscript{3}. Deterioration in neuronal health might precede illness onset, as suggested by evidence for CHR-P individuals\textsuperscript{159,160}, although findings in this group have been inconsistent\textsuperscript{161} and better-powered prospective studies are required. Overall, neuroimaging evidence is indicative of aberrant brain trajectories that seems to precede illness onset\textsuperscript{162,163}, suggesting a neurodevelopmental component to the disorder\textsuperscript{16}, and that have potential as markers for transition risk in high-risk individuals\textsuperscript{164}. Following illness onset, comorbid substance use, antipsychotic medication, stress and lifestyle factors can contribute to progressive ventricular volume increases\textsuperscript{165,166}, in addition to underlying neurobiological factors. As such, high between-study heterogeneity observed in neuroimaging biomarkers might reflect methodological differences affecting selection of controls, diagnostic criteria, and clinical characteristics of included cases\textsuperscript{167}.

For neuropathological biomarkers, none of the 406 assessed factors obtained highly suggestive or suggestive evidence. While 113 biomarkers did present significant associations, the level of evidence was weak, with most (95.6%) estimates being based on a single study, and all of them employing small samples (range of n=9 to 176).

Regarding other biomarkers, minor physical abnormalities presented highly suggestive evidence, although very large heterogeneity and potential excess significance bias were observed. This finding is consistent with extensive evidence of early-life neurodevelopmental deviance in psychosis, including an excess of prenatal and perinatal insults\textsuperscript{85}, childhood motor\textsuperscript{168} and cognitive\textsuperscript{169} abnormalities, and structural brain alterations\textsuperscript{163,170,171}. Originally, these observations led to the neurodevelopmental hypothesis of schizophrenia\textsuperscript{172,173}, later expanded to incorporate environmental adversity occurring during childhood and adolescence\textsuperscript{16,174}. Finally, lower high-frequency heart rate variability achieved suggestive evidence. However, this association is affected by large heterogeneity and potential publication and excess significance bias. Nonetheless, this finding supports the hypothesis of reduced vagal activity among individuals with psychotic disorders as a potential endophenotype associated with executive function\textsuperscript{175}, emotional regulation\textsuperscript{176,177}, and threat inhibition\textsuperscript{178}. This is also consistent with the evidence of increased risk of cardiovascular
disease among individuals with schizophrenia\textsuperscript{179}. While it is unclear how autonomic function alterations manifest in CHR-P individuals\textsuperscript{180,181}, an increase in high-frequency heart rate variability has been observed in CHR-P men following intranasal oxytocin\textsuperscript{14}, suggesting possible novel disease-engagement psychopharmacological targets.

Overall, the evidence summarized in this study presents some limitations. First, we provide estimates of association and do not assume a cause-effect relationship between psychotic disorders and the reported biomarkers. Reverse causality\textsuperscript{182} is an important component in biomarker research, as the temporality of the exposure cannot always be ascertained. Furthermore, we found common methodological limitations in psychosis biomarker research. On one hand, large statistical heterogeneity ($I^2 >50\%$) was observed in 12.1\% of all 1,478 factors, and in 91.7\% of factors with class II or III evidence. In addition, between-study heterogeneity could not be evaluated for 74.5\% of biomarkers, and for one factor with class II evidence. This could indicate between-study methodological differences related to data collection and reporting, or sampling criteria, which can limit the reproducibility and real-world translation of findings. High heterogeneity can also be indicative of real difference between clinical or biological subtypes, meriting further exploration\textsuperscript{24,183}. In addition to high heterogeneity, excess significance bias (affecting 75\% of class II and III biomarkers) and limited number of cases of <1,000 (affecting 92.4\% of all 1,478 factors) were frequently observed. Inflated effect sizes have been suggested to particularly affect newly discovered associations due to suboptimal power\textsuperscript{184}. In our sample, 38.8\% of all factors based on a single study reported statistically significant associations. Among 865 estimates based on a single study (affecting 58.5\% of all factors), all but one had a sample of cases of <1,000 participants. High heterogeneity, excess significance bias and small samples are methodological limitations not restrictive of psychosis research, and have also been observed in biomarker research for other mental disorders\textsuperscript{41,67,185,186}. Furthermore, reporting bias via Egger test could not be assessed for 87.5\% of all factors due to insufficient data or due to only one study being included, and was significant for 1.1\% and non-significant for 11.4\% of biomarkers. Overall, these results highlight the need for better-powered replication studies with representative samples that will allow us to assess whether these biomarkers are generalisable to other settings with different sociodemographics and service configurations. Without this understanding, the prospects of implementing these biomarkers in real-world clinical care and benefitting patients is severely limited\textsuperscript{187}. Moreover, implementing a biomarker that is only useful in certain (usually majority) sub-populations or settings can lead to harm and perpetuation of existing health inequities\textsuperscript{188,189}. Ongoing large multisite research networks are being conducted to address these limitations\textsuperscript{190,191}. In addition, compliance with common guidelines and practices for evaluating and reporting biomarkers\textsuperscript{192} is needed to increase the
replicability of finding and minimize missing data. Another limitation is that the current literature search is not fully updated. This is due to the high complexity of data extraction, data quality check procedures and synthesis that are needed for this type of analyses. However, our robust method could be leveraged by subsequent evidence syntheses in this field to mainstream updated summaries. As a final limitation of our report, we did not assess the quality of individual primary studies included in the evidence synthesis as this was outside the scope of our review. In the same lines, In the same lines, the scope of this review was limited to diagnostic biomarkers and we did not assess prognostic biomarkers or those useful to forecast drug response\textsuperscript{193}.

**CONCLUSION**

Several biomarkers presented a highly suggestive or suggestive evidence of association with psychotic disorders, indicating various translational opportunities for personalized medicine. However, the widespread presence of methodological biases and underpowered studies in biomarker research highlight the need for future higher-quality research.

**CONFLICT OF INTEREST**

GA has received CME-related honoraria, or consulting fees from Janssen-Cilag, Lundbeck, Lundbeck/Otsuka, and Angelini, with no financial or other relationship relevant to the subject of this article. EV has received grants and served as consultant, advisor or CME speaker for the following entities: AB-Biotics, AbbVie, Angelini, Biogen, Boehringer-Ingelheim, Celon Pharma, Dainippon Sumitomo Pharma, Ferrer, Gedeon Richter, GH Research, Glaxo-Smith Kline, Janssen, Lundbeck, Novartis, Orion Corporation, Organon, Otsuka, Sage, Sanofi-Aventis, Sunovion, Takeda, and Viatris, outside the submitted work. PFP has received research funds or personal fees from Lundbeck, Angelini, Menarini, Sunovion, Boehringer Ingelheim, Mindstrong, Proxymm Science, outside the current study.

**FUNDING**

GA is supported by a Rio Hortega 2021 grant (CM21/00017) financed by the Instituto de Salud Carlos III (ISCIII) and cofinanced by the Fondo Social Europeo Plus (FSE+).
REFERENCES


Fleischhacker WW. A meta view on meta-analyses. *JAMA Psychiatry.*


100. Applebaum J, Shimon H, Sela BA, Belmaker RH, Levine J. Homocysteine levels in


112. Landek-salgado MA, Faust TE, Sawa A. Molecular substrates of schizophrenia:


207. de Wilde OM, Bour LJ, Dingemans PM, Koelman JHTM, Linszen DH. A meta-analysis of P50 studies in patients with schizophrenia and relatives: differences in methodology between research groups. *Schizophr Res*. 2007;97(1-3):137-151.


### Table 1. Biomarkers with highly suggestive or suggestive evidence of association with psychotic disorders

<table>
<thead>
<tr>
<th>Factor</th>
<th>k</th>
<th>Diagnosis</th>
<th>ES (95% CI)</th>
<th>N</th>
<th>Cases</th>
<th>Controls</th>
<th>p</th>
<th>I²</th>
<th>PI 95% CI</th>
<th>Egger</th>
<th>ESB</th>
<th>LS</th>
<th>eOR</th>
<th>CE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Peripheral biomarkers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tTG autoantibodies(^{194})</td>
<td>1</td>
<td>SZ</td>
<td>OR, 7.32 (3.36, 15.94)</td>
<td>2301</td>
<td>1401</td>
<td>900</td>
<td>1 x 10^{-6}</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>No</td>
<td>Yes</td>
<td>7.32</td>
<td>II</td>
</tr>
<tr>
<td>Folate(^{195})</td>
<td>16</td>
<td>SZ</td>
<td>SMD, -1.44 (-2.18, -0.71)</td>
<td>2599</td>
<td>1260</td>
<td>1339</td>
<td>1.20 x 10^{-4}</td>
<td>98%</td>
<td>-4.71, 1.83</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>0.07</td>
<td>III</td>
</tr>
<tr>
<td>Malondialdehyde(^{120,196,197})</td>
<td>26</td>
<td>SZ</td>
<td>SMD, 1.38 (0.82, 1.94)</td>
<td>1795</td>
<td>1077</td>
<td>718</td>
<td>1 x 10^{-6}</td>
<td>94%</td>
<td>-1.56, 4.32</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>12.22</td>
<td>III</td>
</tr>
<tr>
<td>BDNF(^{198–201})</td>
<td>47</td>
<td>FEP, SZ</td>
<td>SMD, -0.69 (-1.05, -0.33)</td>
<td>4955</td>
<td>2756</td>
<td>2199</td>
<td>2.00 x 10^{-4}</td>
<td>91%</td>
<td>-3.2, 1.82</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>0.29</td>
<td>III</td>
</tr>
<tr>
<td>Homocysteine(^{94})</td>
<td>19</td>
<td>SZ</td>
<td>SMD, 0.6 (0.31, 0.89)</td>
<td>3320</td>
<td>1303</td>
<td>2017</td>
<td>5.1 x 10^{-5}</td>
<td>89%</td>
<td>-0.7, 1.91</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>2.98</td>
<td>III</td>
</tr>
<tr>
<td><strong>Electrophysiologic biomarkers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mismatch negativity in auditory event-related potentials(^{125,126,202,203})</td>
<td>47</td>
<td>SZ</td>
<td>SMD, 0.73 (0.5, 0.96)</td>
<td>5649</td>
<td>2871</td>
<td>2778</td>
<td>&lt;1 x 10^{-6}</td>
<td>86%</td>
<td>-0.81, 2.28</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>3.77</td>
<td>II</td>
</tr>
<tr>
<td>P300 component latency(^{143,204})</td>
<td>56</td>
<td>SZ</td>
<td>SMD, -0.6 (-0.83, -0.38)</td>
<td>3502</td>
<td>1735</td>
<td>1767</td>
<td>&lt;1 x 10^{-6}</td>
<td>90%</td>
<td>-2.22, 1.01</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>0.33</td>
<td>II</td>
</tr>
<tr>
<td>P50 sensory gating (P50 S2/S1 ratio)(^{143,205–208})</td>
<td>80</td>
<td>AP, SZ</td>
<td>SMD, 0.79 (0.43, 1.16)</td>
<td>4999</td>
<td>2107</td>
<td>2892</td>
<td>2.0 x 10^{-5}</td>
<td>93%</td>
<td>-2.43, 4.02</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>4.22</td>
<td>III</td>
</tr>
<tr>
<td><strong>Neuroimaging biomarkers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ventricle brain ratio(^{209})</td>
<td>72</td>
<td>SZ</td>
<td>SMD, 0.61 (0.5, 0.71)</td>
<td>6099</td>
<td>3463</td>
<td>2636</td>
<td>&lt;1 x 10^{-6}</td>
<td>68%</td>
<td>-0.14, 1.35</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>3.00</td>
<td>II</td>
</tr>
<tr>
<td>Frontal NAA(^{159})</td>
<td>68</td>
<td>SZ</td>
<td>SMD, -0.34 (-0.47, -0.2)</td>
<td>2868</td>
<td>1444</td>
<td>1424</td>
<td>1 x 10^{-6}</td>
<td>63%</td>
<td>-1.22, 0.55</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>0.54</td>
<td>III</td>
</tr>
<tr>
<td><strong>Other biomarkers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minor physical anomalies(^{210})</td>
<td>14</td>
<td>SZ</td>
<td>SMD, 0.99 (0.64, 1.34)</td>
<td>2160</td>
<td>1153</td>
<td>1007</td>
<td>&lt;1 x 10^{-6}</td>
<td>93%</td>
<td>-0.45, 2.42</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>5.99</td>
<td>II</td>
</tr>
<tr>
<td>High frequency heart rate variability(^{211})</td>
<td>28</td>
<td>SZ</td>
<td>SMD, -0.99 (-1.41, -0.58)</td>
<td>3008</td>
<td>1330</td>
<td>1678</td>
<td>2 x 10^{-6}</td>
<td>98%</td>
<td>-3.24, 1.25</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>0.16</td>
<td>III</td>
</tr>
</tbody>
</table>

Figure 1. Standardized mean differences for peripheral biomarkers and psychotic disorders

[add Figure 1 here]

OR for tTG autoantibodies have been converted to equivalent Hedges’ g (eG) to allow for a comparison across risk factors. BDNF - brain-derived neurotrophic factor, CI - confidence interval, eG – equivalent Hedges’ g, SMD – standardized mean difference, tTG - transglutaminase.
Figure 2. Standardized mean differences for electrophysiological biomarkers and psychotic disorders

[add Figure 2 here]

CI - confidence interval, SMD – standardized mean difference.
**Figure 3.** Standardized mean differences for neuroimaging biomarkers and psychotic disorders

[add Figure 3 here]

CI - confidence interval, NAA - N-acetylaspartate, SMD – standardized mean difference.
Figure 4. Standardized mean differences for other biomarkers and psychotic disorders

[add Figure 4 here]

CI - confidence interval, SMD – standardized mean difference.