Received: 11 April 2024 | Revised: 12 August 2024 | Accepted: 17 August 2024

DOI: 10.1111/acps.13752

ORIGINAL ARTICLE

Acta Psychiatrica Scandinavica WILEY

Psychosis metabolic risk calculator (PsyMetRiC) in early psychosis: External validation study in Finland

Jaakko Keinänen[1,2] | **Saana Eskelinen**[1,2] | **Tiina From**[3] | **Heikki Laurikainen**[3] | **Jarmo Hietala**[3] | **Graham K. Murray**[4,5] | **Jaana Suvisaari**[1,6] | **Benjamin I. Perry**[4,5]

[1]Mental Health Team, Department of Public Health and Welfare, Finnish Institute for Health and Welfare, Helsinki, Finland

[2]Department of Psychiatry, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

[3]Department of Psychiatry, University of Turku and Turku University Hospital, Turku, Finland

[4]Department of Psychiatry, University of Cambridge, Cambridge, UK

[5]Cambridgeshire and Peterborough NHS Foundation Trust, Fulbourn, UK

[6]The Finnish Medical Society Duodecim, Helsinki, Finland

Correspondence
Jaakko Keinänen, Department of Psychiatry, Helsinki University Hospital, Välskärinkatu 12, 00260, Helsinki, Finland.
Email: jaakko.keinanen@hus.fi

Funding information
Helsingin ja Uudenmaan Sairaanhoitopiiri; Turun Yliopistollinen Keskussairaala; Academy of Finland; Sigrid Juséliuksen Säätiö; Suomen Kulttuurirahasto

Abstract

Introduction: Accurate detection of cardiometabolic risk in early psychosis is crucial to reducing somatic morbidity and mortality in people with psychotic disorders. We conducted an external validation of the psychosis metabolic risk calculator (PsyMetRiC), a cardiometabolic risk prediction tool developed in the UK and tailored for young people with psychosis. We compared the predictive accuracy and clinical usefulness of PsyMetRiC and a general population-based risk prediction tool for type 2 diabetes, the Finnish Diabetes Risk Score (FINDRISC).

Methods: We included first-episode psychosis and ultra-high-risk for psychosis patients without metabolic syndrome aged 18–35 years from the Helsinki Early Psychosis and Turku Early Psychosis Study cohorts. We tested two versions of PsyMetRiC: the full model including age, sex, ethnicity, body-mass index, smoking status, prescription of metabolically-active antipsychotic medication, high-density lipoprotein, and triglyceride concentrations, and the partial-model excluding biochemical predictors, and the simplified FINDRISC including BMI, sex, systolic blood pressure, and fasting glucose. Discrimination, calibration, and decision curve analyses were used to assess the predictive performance and clinical usefulness of both PsyMetRiC and FINDRISC. We performed a site-specific re-calibration of PsyMetRiC (PsyMetRiC-Fi).

Results: The study sample consisted of 278 individuals (all White European ethnicity, 58.6% male, mean age 24.8 years, 37.8% smoking, mean BMI 23.5). Discrimination was marginally better in the PsyMetRiC full model ($C = 0.72$,

95% CI, 0.59–0.82) compared with partial model ($C = 0.70$, 95% CI 0.59–0.80) or FINDRISC ($C = 0.63$, 95% CI 0.54–0.71). Calibration plots displayed evidence of minor miscalibration for PsyMetRiC, which corrected following recalibration. Miscalibration was more pronounced for FINDRISC. Decision curve analysis showed that PsyMetRiC offers likely clinical usefulness in improving cardiometabolic risk management in early psychosis compared with giving everyone or no one an intervention.

Conclusion: PsyMetRiC has utility in predicting cardiometabolic risk in Finnish patients with early psychosis. It has better discriminatory accuracy and offers more accurate risk prediction compared to other available strategies.

KEYWORDS

early intervention, metabolic syndrome, psychosis, risk prediction algorithm

1 | INTRODUCTION

Patients with psychotic disorders are known to have excessive somatic morbidity.[1] Cardiovascular diseases (CVD), type 2 diabetes (T2D) and obesity are up to five-times more prevalent in this group compared to the general population.[2] The life expectancy of people with psychotic disorders is shortened by 10–15 years on average, with general medical conditions including CVD and metabolic disorders an important contributor.[3] Mortality rates of 15% are observed in patients with psychotic disorders already by around age 40.[4] In addition to the years of life lost, comorbid psychotic and cardiometabolic disorders are a substantial economic burden on society. Thus, the optimization of preventive strategies of CVD should be a priority for psychiatric clinicians from the onset of psychosis.[5]

Metabolic syndrome (MetS) comprises abdominal obesity, hyperglycaemia, dyslipidaemia and elevated blood pressure.[6] It is an established risk factor for CVD, T2D, and premature mortality.[7] While the prevalence of MetS is not higher at baseline in first-episode psychosis (FEP), meta-analytic evidence suggests that early markers of developing cardiometabolic morbidity, e.g. insulin resistance[8] and dyslipidemia[9] are commonly present. Weight gain during the first year after a diagnosis of a psychotic disorder, accelerated by antipsychotic medication, predisposes to metabolically harmful abdominal obesity, and deterioration of glucose-insulin homeostasis.[10]

Cardiometabolic risk prediction algorithms are routinely used by clinicians to reduce the risk of future adverse cardiometabolic events in the general population. The Finnish Diabetes Risk Score (FINDRISC) was developed to estimate the risk of developing type 2 diabetes in the Finnish population, and it has also been validated for screening[11–13] and predicting[14] MetS.

Significant outcomes

- Psychosis metabolic risk calculator (PsyMetRiC) provides likely clinical benefit in predicting metabolic risk in early psychosis in Finland.
- The predictive performance of PsyMetRiC in this sample was comparable to other external validations of PsyMetRiC conducted in other European samples.
- PsyMetRiC is likely a more useful tool for metabolic risk assessment in early psychosis compared to other risk prediction tools developed for the general population.

Limitations

- The sample size in the study is limited and for estimation of wider generalizability, external validations with larger samples are needed.
- There was a relatively large amount of missing data at follow-up. We did not detect patterns of not-random missingness, which allowed for the use of imputation in the statistical analyses.

Existing risk-prediction algorithms, including FINDRISC, were developed for populations of middle-to-older adults due to the highest incidences of cardiometabolic disturbances in these age groups.[15] Because these disturbances emerge much earlier in patients with psychotic disorders, the algorithms severely underpredict cardiometabolic risk in this high-risk group.[16,17] Therefore, recently the Psychosis Metabolic Risk Calculator (PsyMetRiC) was developed in the UK to predict the

6-year risk for MetS among FEP patients aged 18–35 years.[18] PsyMetRiC has been externally validated in the United Kingdom (UK), Spain and Swizerland.[19] PsyMetRiC has potential for clinical usefulness in treating young adult patients with psychotic disorders, and could help clinicians to personalize counseling, pharmacological and behavioral interventions in diminishing CVD risk.

PsyMetRiC has not yet been tested in Nordic countries, and so its generalizability in those regions cannot be assumed. Even large-scale general population-based cardiometabolic risk prediction algorithms show variability in predictive performance when tested across borders.[20–22] There are a multitude of cultural, social, health behavioral or environmental confounders that could limit the performance of PsyMetRiC outside previously validated populations.

We studied the predictive accuracy of PsyMetRiC across two clinical samples in Finland by performing detailed external validation analysis to assess whether PsyMetRiC may be generalizable for potential future use in Finland. Then, where we identified evidence of differing predictive accuracy across borders, we performed population-specific recalibration to develop a revised version of the PsyMetRiC algorithm tailored for the Finnish population (PsyMetRiC-Fi). Finally, we studied the accuracy and likely clinical usefulness of PsyMetRiC-Fi compared with the FINDRISC algorithm in three early psychosis samples in Finland. This study is reported according to TRIPOD guidelines (see Table S1).[23]

The aim of this study was to validate and recalibrate PsyMetRiC in a Finnish early psychosis sample and compare its clinical usefulness to another readily available risk prediction calculator aimed for the general population. Ultimately, accurate metabolic risk prediction would inform both the clinician and the patient about individual-level risk and help clinicians target interventions to those who are most likely to benefit from them.

2 | METHODS

2.1 | Data sources

2.1.1 | Helsinki Early Psychosis Study

The patients in the Helsinki Early Psychosis Study (HEPS) study were recruited from the catchment areas of Helsinki University Hospital and Helsinki Psychiatric Services from December 2010 through June 2016. At the time of recruitment, participants were receiving either inpatient or outpatient care. Inclusion criteria were: (1) receiving a minimum score of 4 in the Brief Psychiatric Rating Scale—Expanded (BPRS-E)[24] Unusual thought content or Hallucinations and (2) fluency in Finnish. Patients with previous psychotic episodes, substance induced psychosis and psychosis due to a general medical condition were excluded. The study was carried out in accordance with the International Code of Medical Ethics of the World Medical Association (the Declaration of Helsinki). The study protocol was approved by the Ethics Committee of the hospital district of Helsinki and Uusimaa (257/12/03/03/2009; 226/13/03/03/2013) and the institutional review boards of the Finnish Institute for Health and Welfare, Helsinki, Finland, and the University of Helsinki. All participants gave a written informed consent.

At baseline, 61% (59/97) of patients were hospitalized. Weight, waist circumference and blood pressure were measured, and blood samples collected at baseline and at the follow-up measurement. Participants' physical activity, diet and smoking were also assessed. Mean time between baseline and the follow-up measurements was 391 days (range 356–545 days). Psychiatric diagnostic interviews were done at the follow-ups of two and 12 months. The diagnostic assessments were based on the Research Version of the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I)[25] and available medical records. Detailed methods of the HEPS study have previously been described in Keinänen et al. 2018.[26]

2.1.2 | Turku Early Psychosis Study

First-episode psychosis patients and clinical-high-risk for psychosis patients (CHR) (ages 18–50 years) with at least adequate skills in Finnish were recruited from the psychiatric inpatient and outpatient units of Turku Health Services and the Hospital Districts of Southwest Finland and Satakunta. The time period of recruitment was from October 2011 to December 2017. The study design and protocols were approved by the ethics committee of the Turku University Hospital (Approval numbers Turku Early Psychosis Study (TEPS) ETMK 64/180/2011, PRONIA ETMK 99/180/2013 and METSY ETMK 98/180/2013). All participants gave a written informed consent.

FEP was defined as fulfilling the criteria for DSM-IV Axis I psychosis diagnosis, including schizophrenia, delusional and bipolar psychoses, acute transient psychoses and other psychoses, and having received treatment for the psychosis for less than 2 years. Clinical high risk for psychosis was determined according to the Structured Interview for Prodromal Symptoms.[27] Exclusion criteria for study patients were a previous psychotic disorder and IQ < 70.

'Intent to study'-CHR (Clinical High Risk) group consisted originally of patients that had a high risk for psychosis according to a clinical estimate by the treating clinical team in psychiatric health care services. These

TABLE 1 Sociodemographic characteristics of the original PsyMetRiC development sample and included external validation sample.

Characteristic	Original PsyMetRiC development sample (UK)	Full analytic sample (Finland)	Between-group differences[a]
Sample before inclusion/exclusion criteria applied[b], *N*	1504	334	–
Included sample size[b], *N* (%)	651 (43.28)	278 (83.23)	–
Age in years, mean (SD)	24.52 (4.91)	24.8 (4.60)	$t = 0.21$, $p = 0.835$
White European/NR ethnicity, *N* (%)	360 (55.3)	100 (100)	$\chi = 141.62$, $p < 0.001$
Male sex, *N* (%)	440 (67.59)	163 (58.60)	$\chi = 5.14$, $p = 0.023$
HDL at baseline (mmol/L) mean (SD)	1.88 (0.57)	1.46 (0.38)	$t = -10.02$, $p < 0.001$
Triglycerides at baseline (mmol/L) mean (SD)	1.39 (1.06)	1.03 (0.53)	$t = -4.58$, $p < 0.001$
BMI at baseline (kg/m^2), mean (SD)	23.63 (5.43)	23.50 (4.32)	$t = 0.17$, $p = 0.866$
FPG at baseline (mmol/L), mean (SD)	5.19 (1.28)	4.91 (1.15)	$t = -4.00$, $p < 0.001$
Systolic BP at baseline (mmHg), mean (SD)	120.65 (11.68)	122.00 (13.40)	$t = 1.42$, $p = 0.156$
Prescribed a more-metabolically active antipsychotic[c], *N* (%)	455 (69.89)	203 (73.0)	$\chi = 15.24$, $p < 0.001$
Smoking at baseline, *N* (%)	315 (48.39)	105 (37.80)	$\chi = 5.58$, $p = 0.018$
MetSyndrome at baseline, *N* (%)[d]	49 (6.58)	27 (8.90)	$\chi = 5.15$, $p = 0.023$
MetSyndrome at follow-up, *N* (%)	109 (16.74)	53 (19.11)	$\chi = 0.76$, $p = 0.384$

Abbreviations: BMI, body mass index; BP, blood pressure; FPG, fasting plasma glucose; HDL, high-density lipoprotein; MetS, metabolic syndrome.
[a]Analysis of means was conducted using unpaired t-tests. Analysis of proportions was conducted using the chi-square equality of proportions test.
[b]See Figure 1 for a flow-chart of included participants in the study.
[c]Definitions of Metabolically-active antipsychotics per Perry et al (2021)[16] reported in Table S3.
[d]Corresponds to percentage of total sample before participants with MetS were excluded.

patients were subsequently examined with Structured Interview for Prodromal Symptoms (SIPS), including Global Assessment of Functioning (GAF). Following high-risk criteria were used: Attenuated Psychotic Symptoms (APS), Brief Limited Psychotic Symptoms (BLIPS), and Genetic risk and reduction of function (GRD). Some of the clinician-defined psychosis risk patients did not quite fulfill the SIPS criteria. However, as observed in a previous publication, both CHR groups (SIPS positive and SIPS negative) were clinically similar[28] Thus, a combined CHR group was used in this study.

At baseline, 82% (107/130) of patients were hospitalized. Weight, waist circumference and blood pressure were measured, and blood samples collected at baseline and at follow-up points.

The mean time between the baseline and the follow-up measurements was 366 days (range 138–501 days).

2.1.3 | Comparisons between study populations

While there are similarities between Finland and the UK, where PsyMetRiC was originally developed, there are also notable differences. Table 1 shows the sociodemographic characteristics of the study samples. See Table S2 for a detailed comparison of key sociodemographic, economic, and healthcare-related metrics between the overall UK and Finnish populations.

Regarding the study regions specifically, the Greater Helsinki Region has a population of 1,733,033 inhabitants, of whom 664,028 live in the capital city, Helsinki.[29] The wellbeing services county of Southwest Finland (Turku region) has a population of 485,567 inhabitants, of whom 197,900 live in Turku.[29]

2.2 | Inclusion and exclusion criteria

Following the methodology of the original PsyMetRiC study,[18] in both samples we excluded participants who: were aged <16 years or >35 years at the time of baseline assessment; met the outcome criteria at baseline; or had missing data on all predictor or outcome constituent variables. See Figure 1 for a flow-chart of included participants in the study from both samples.

2.3 | Outcome

As per the original PsyMetRiC study,[18] we used the harmonized definition[30] of MetS as a binary outcome:

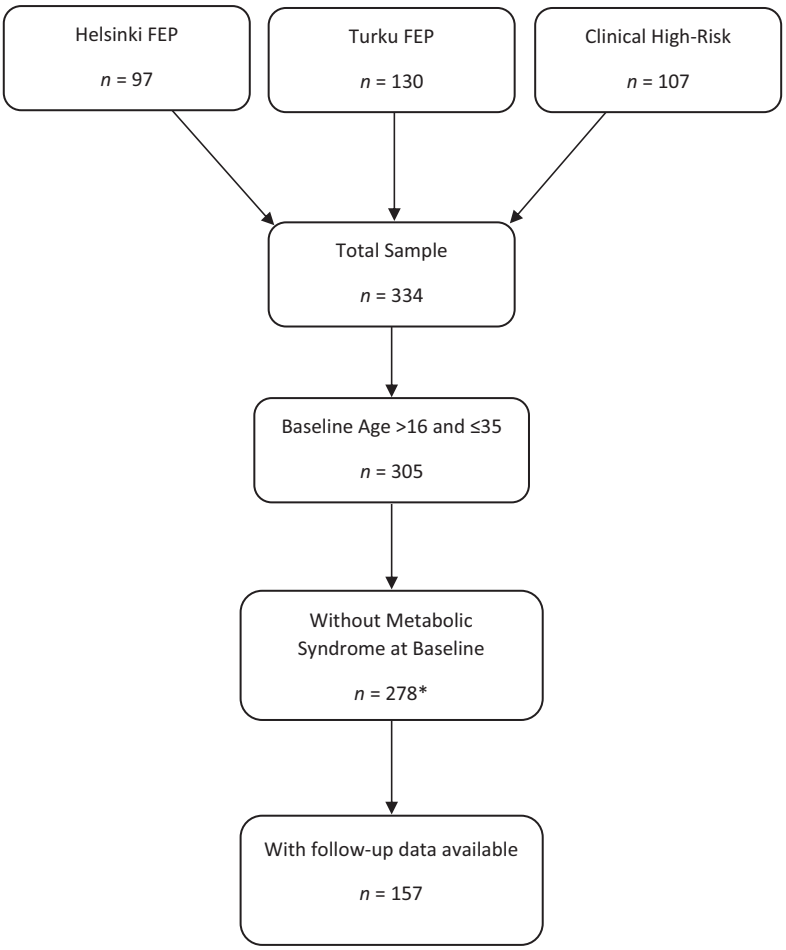


FIGURE 1 Flow-chart of participants included in the study. *Number of individuals included in the final analyses ($n = 278$).

ethnicity-specific waist circumference ≥94 cm in males and ≥80 cm in females for Caucasians; ≥90 cm in males and ≥80 cm in females for other ethnic groups, or body mass index (BMI) >29.9; alongside two of: triglycerides ε1.70 mmol/L; high-density lipoprotein (HDL) <1.03 mmol/L (males) or <1.29 mmol/L (females); systolic blood pressure >130 mmHg; fasting plasma glucose (FPG) >5.60 mmol/L. In each sample, where multiple follow-ups were available for each participant, we used the latest follow-up available between 1 and 6 years after baseline with the least amount of missing data.

2.4 | Algorithms tested for predictive performance

2.4.1 | Psychosis Metabolic Risk Calculator

Psychosis Metabolic Risk Calculator (PsyMetRiC), developed in the UK,[18] consists of two forced-entry multivariable penalized logistic regression equations: the full-model and the partial-model. The full-model includes age, sex, ethnic background, current smoking status, psychotropic medication at baseline, BMI, HDL and triglycerides. The partial-model is the same minus the biochemical predictors (HDL and triglycerides). Briefly, predictors were selected during the original PsyMetRiC model development study[18] based on clinical knowledge, prior research, and likely clinical usefulness/patient acceptability. Antipsychotic medications were categorized as more or less metabolically-active (Table S3). The PsyMetRiC algorithm coefficients are presented in Table S4. See the original PsyMetRiC study[18] for further details.

2.4.2 | Finnish Diabetes Risk Score

The Finnish Diabetes Risk Score (FINDRISC)[31] consists of a forced-entry multivariable logistic regression equation developed using data from the Finnish national Population Register in 1987 and another independent sample in 1992 (the FINRISK studies). Predictors, including age,

BMI, waist circumference, hypertension history, fasting blood glucose, diet and physical activity levels were included on a balance of clinical knowledge, prior research, and based on their likely convenience for use in primary care. We performed analysis of the concise FINDRISC version (all predictors except hypertension history, diet and physical activity). As reported by the authors of FINDRISC, diet and physical activity were added to the FINDRISC full model to emphasize lifestyle factors in the prevention of diabetes, although they did not meaningfully improve the predictive performance of the model.[31] FINDRISC was originally developed to predict type 2 diabetes but has also been independently validated to predict MetS.[11–13]

2.5 | Statistical analysis

2.5.1 | Sample preparation and estimation of analytic precision

Where necessary, the units of biochemical values were converted to mmol/L. Recently developed criteria[32] to estimate analytic precision given the fixed sample sizes (Supplementary Methods) were applied. Briefly, the expected SE for the C-statistic was 0.04. The expected SEs for the calibration slope and calibration-in-the-large were 0.22 and 0.20 respectively. Multiple imputation using chained equations was considered for missing data (Supplementary Methods). Rubin's rules were used to pool the estimates for numerical-based analyses. For plot-based analyses, plots generated in each imputed dataset were checked for similarity. In the main manuscript, one randomly selected plot per analysis is shown and the remaining plots are presented in the Supplementary Results.

2.5.2 | Primary external validation analysis

After applying the algorithms to the analytic sample, the distribution of predicted outcome probabilities was inspected using histograms. Primary assessment of algorithm performance was done with measures of discrimination (concordance (C-) statistic), and calibration (calibration plots) (Supplementary Methods). We also recorded the Nagelkerke-Cox-Snell-Maddala-Magee r^2 index, the calibration intercept (ideally close to 0), calibration slope (ideally close to 1), and the Brier score (ideally close to 0, with scores >0.25 indicating poor performance).

2.5.3 | Recalibration and generation of Finland-specific PsyMetRiC version

Miscalibration was estimated by visual inspection of calibration plots and where miscalibration was identified (between the observed proportion and predicted probability), logistic calibration approach was considered. By completing this step, we obtained a Finland-specific version of PsyMetRiC (PsyMetRiC-Fi) (Supplementary Methods). Performance estimates are presented accompanied by 95% CIs for all results in our analysis. We did not perform recalibration analysis for FINDRISC.

2.5.4 | Clinical usefulness

Clinical usefulness was assessed using decision curve analysis.[33] Net benefit was estimated across a range of feasible thresholds (i.e., the risk score at which an intervention would be deemed necessary) (Supplementary Methods). We considered a risk threshold upper-bound of 0.30 as in previous PsyMetRiC studies.[18,19] Using net benefit instead of related measures such as sensitivity and specificity was preferred, as net benefit incorporates the consequences of the decisions made on the basis of an algorithm.[34] We reported the net benefit and standardized net benefit (net benefit / outcome prevalence that is, the additional percentage of cases that could be intervened on with use of PsyMetRiC with no increase in false-positives) across a range of reasonable risk thresholds. Decision curve plot was drawn to visualize and compare the net benefit of PsyMetRiC, PsyMetRiC-Fi and FINDRISC, compared with intervening in all or intervening in none. Classical decision theory proposes that at a chosen risk-threshold, the choice with the greatest net-benefit should be preferred.[34]

2.5.5 | Sensitivity analyses—Predictive accuracy in young people at risk of developing psychosis

The original PsyMetRiC study included sensitivity analysis in young people at risk for developing psychosis as measured using the Psychosis-Like Symptom Interview and showed that PsyMetRiC may be generalizable for use in these individuals.[18] To externally validate this finding in an international sample for the first time, we conducted analysis in 107 individuals from the TEPS study with high-risk for psychosis. Analytic methods followed those described above.

TABLE 2 Predictive performance statistics of PsyMetRiC and FINDRISC in analytic sample.

Measure of predictive performance	Primary analysis, estimate (95% CI)		After logistic calibration, estimate (95% CI)	
PsyMetRiC				
	Full-model	**Partial-model**	**Full-model**	**Partial-model**
C-statistic	0.72 (0.59, 0.82)	0.70 (0.59, 0.80)	0.72 (0.59, 0.82)	0.70 (0.59, 0.80)
r^2	0.17 (0.15, 0.20)	0.14 (0.11, 0.17)	0.19 (0.16, 0.22)	0.16 (0.13, 0.19)
Calibration intercept	0.16 (−0.04, 0.37)	−0.14 (−0.25, −0.02)	0.01 (0.00, 0.01)	0.01 (0.00, 0.01)
Calibration slope	1.03 (0.86, 1.21)	0.94 (0.81, 1.15)	1.00 (0.99, 1.01)	1.01 (0.99. 1.02)
Brier score	0.05 (0.01, 0.08)	0.06 (0.01, 0.12)	0.04 (0.01, 0.07)	0.05 (0.01, 0.08)
FINDRISC				
C-statistic	0.63 (0.54, 0.71)			
r^2	0.06 (0.02, 0.10)			
Calibration intercept	2.28 (1.45, 3.10)			
Calibration slope	0.58 (0.23, 0.93)			
Brier score	0.15 (0.11, 0.20)			

Note: The C-statistic is a measure of discrimination and estimates the probability that a randomly selected 'case' will have a higher predicted probability than a randomly selected non-case. Scores of 1.0 indicate perfect discrimination; scores of >0.70 are generally considered acceptable. The calibration intercept (ideally close to 0) and calibration slope (ideally close to 1) are estimates of model calibration (i.e., the agreement between the observed proportion and predicted risk). The Brier score (ideally close to 0, with scores >0.25 indicating poor performance) is an overall measure of algorithm performance. We did not perform logistic calibration for FINDRISC.

2.5.6 | Data visualization

An online data visualization website for PsyMetRiC was created to accompany the original study (https://psymetric.shinyapps.io/psymetric). The website was updated with the Finland-specific PsyMetRiC version obtained through recalibration analysis. All statistical analyses were done locally by researchers in Finland with R version 4.3.2.[35]

3 | RESULTS

3.1 | Samples

After applying inclusion criteria, we included 278 participants (Table 1, Figure 1). The Finnish sample differed from the UK PsyMetRiC development sample on most sociodemographic, lifestyle and biochemical characteristics (Table 1).

3.2 | Primary external validation analysis

The distribution of predicted probabilities for PsyMetRiC were similar to the original PsyMetRiC study (Figure S1). The distribution of predicted probabilities for FINDRISC showed that the vast majority of included participants scored <10% predicted risk. Predictive performance statistics are reported in Table 2. In summary, there was acceptable discriminative performance of PsyMetRiC (full-model $C = 0.72$, 95% C.I., 0.59, 0.82; partial-model $C = 0.70$, 95% CI, 0.59–0.80), but poor discriminative performance of FINDRISC ($C = 0.63$, 95% CI, 0.54–0.71). Calibration plots (Figure 2, Figure S3) show varying degrees of miscalibration for PsyMetRiC across imputed datasets, with predominantly risk underprediction in the full-model, and overprediction in the partial-model. For FINDRISC there was consistent under-prediction of risk across all imputed datasets.

3.3 | Recalibration and generation of site-specific PsyMetRiC version (PsyMetRiC-Fi)

After logistic calibration of PsyMetRiC (Table S5), the shape of the distributions of predicted probabilities was similar to the primary analysis (Figure S1). Recalibrated performance statistics for PsyMetRiC-Fi are reported in Table 2. Discriminative performance was unchanged, and calibration plots for both PsyMetRiC-Fi versions were similar across imputed datasets (Figure 1; Figure S4) and showed substantially improved agreement between observed risk proportions and predicted risk.

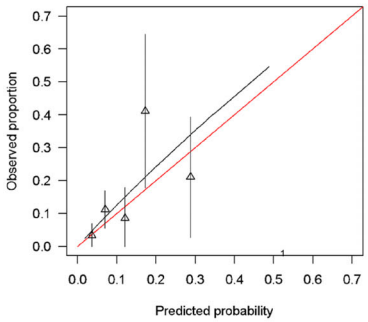
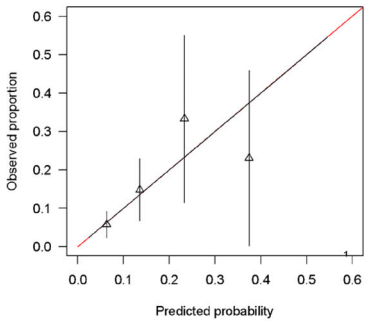
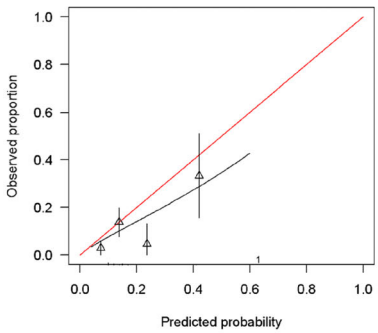
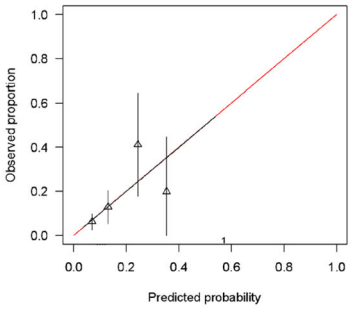
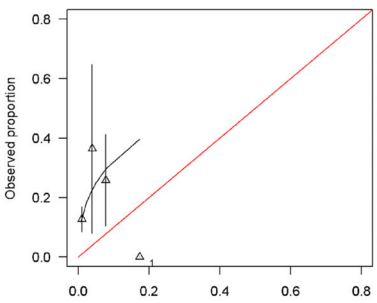
	Primary Analysis	After Logistic Calibration[a]
PsyMetRiC Full-Model		
PsyMetRiC Partial-Model		
FINDRISC		N/A

FIGURE 2 Calibration plots in analytic sample. Calibration plots illustrate agreement between the observed (*y* axis) and predicted risk (*x* axis). Perfect agreement would trace the red line. Algorithm calibration is illustrated by the black line. Triangles denote grouped observations for participants at deciles of predicted risk, with 95% CI's indicated by the vertical black lines. The letter a in superscript denotes logistic calibration takes into account differences in baseline risk that may exist between populations by re-estimating the intercept term, and also re-estimates the slope term thus assuming similar *relative* effects of the predictors but allowing for a larger or smaller *absolute* effect of the predictors. We did not perform recalibration for FINDRISC. See Methods.

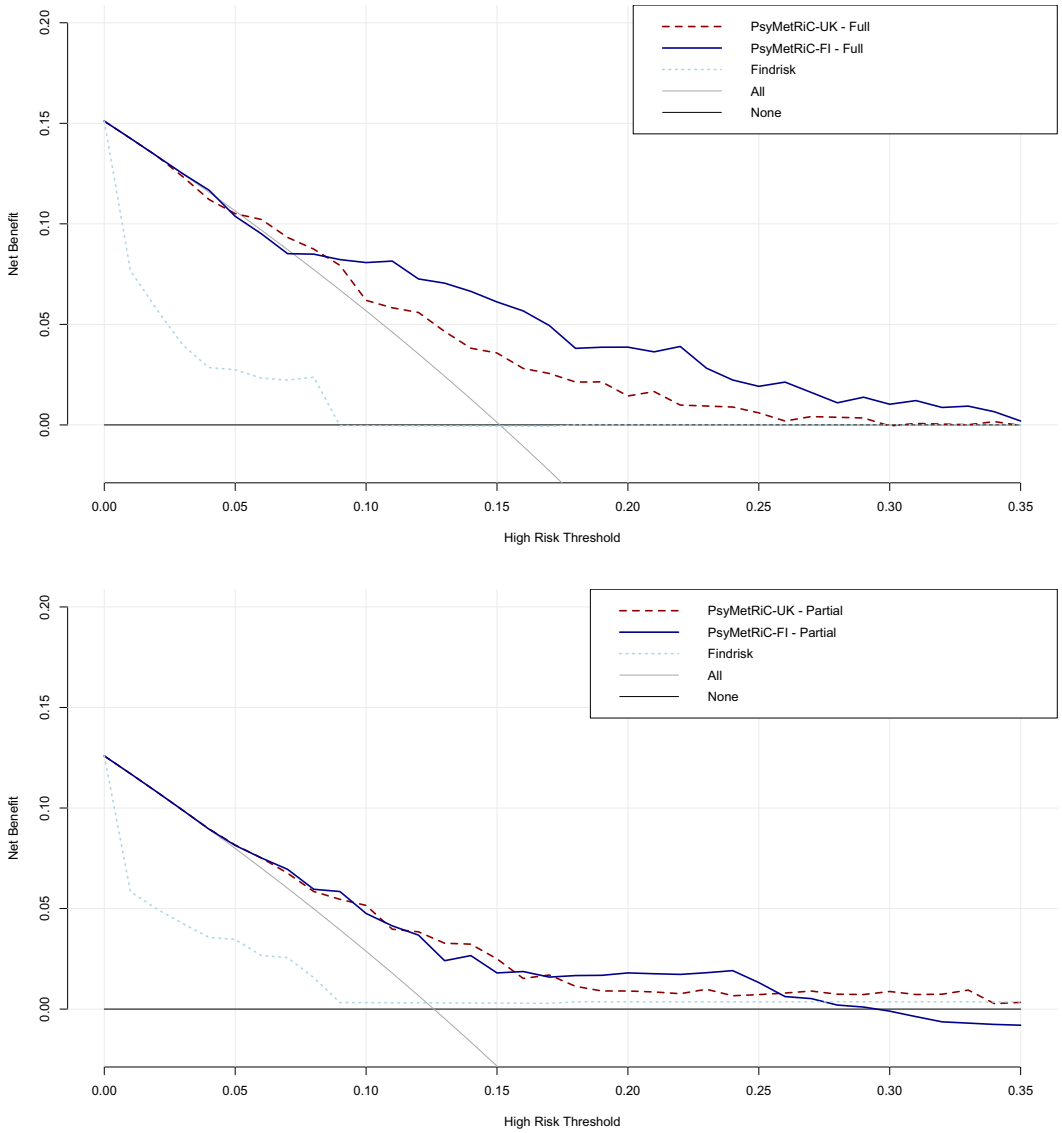


FIGURE 3 Clinical usefulness of PsyMetRiC and FINDRISC in the analytic sample. The plot reports net benefit (y axis) of the original PsyMetRiC and recalibrated PsyMetRiC-Fi full- and partial models, and FINDRISC across a range of risk thresholds (x axis) compared with intervening in all (gray line) or intervening in none (black line). In decision curve analysis, it is customary to consider only the range of risk-thresholds that may reasonably be considered in clinical practice. Our upper bound of 0.35 represents over a one-in-three chance of developing MetS should nothing change, and it is unlikely that risk thresholds greater would be tolerated. Net harm (i.e., more false positives than true positives exposed to an intervention at a selected risk threshold) is indicated when the decision curve line is plotted at $y < 0$.

3.4 | Clinical usefulness

Full decision curve analysis results are presented in Figure 3, Figure S5 and Tables S6–S8. The decision curve analysis plots indicate the net benefit provided by the original (un-calibrated) PsyMetRiC, the recalibrated version (PsyMetRiC-Fi), and FINDRISC in comparison to intervening in all and intervening in none. In summary, results were broadly similar across imputed datasets and indicate that both PsyMetRiC versions are likely to convey clinical usefulness across the range of studied risk thresholds compared with competing strategies of using FINDRISC, intervening in all or intervening in none. Likely clinical usefulness for both PsyMetRiC versions improved further on recalibration. Across all imputed datasets FINDRISC appeared to be less clinically useful than intervening in all (Figure 3, Figure S5) for most studied risk thresholds. For example, at a risk threshold of 15%, the recalibrated full and partial PsyMetRiC-Fi versions provided net benefits of 0.13 (95% CI, 0.08–0.19) and 0.13 (95% CI, 0.07–0.20) respectively, and FINDRISC provided a net benefit of 0.01 (95% CI, 0.00, 0.03). This

equates to an additional 52% of metabolic syndrome cases that could be potentially prevented with the full model, an additional 50% with the partial-model, and an additional 4% with FINDRISC, without an increase in false positives (Tables S6–S8).

3.5 | Sensitivity analysis

In the CHR group, the C-statistic for PsyMetRiC full model was 0.71 (95% CI 0.58–0.85) and partial model 0.66 (95% CI 0.51–0.80). The calibration plots shown in the Figure S7 indicate similar performance of PsyMetRiC in the CHR group as in the main analysis with mainly risk underprediction in the full model and overprediction in the partial model.

4 | DISCUSSION

In this study, we evaluated the performance of the PsyMetRiC cardiometabolic risk prediction algorithm in a sample of Finnish individuals with early psychosis, comparing its performance with FINDRISC, a type 2 diabetes risk prediction tool developed locally for the general population. Our results indicate that PsyMetRiC-Fi has similar utility in this sample as in the previous external validations conducted internationally, and that it is more appropriate for use in the Finnish early psychosis population compared with equivalent general population-based tools, for example, FINDRISC.

As in previous validations of PsyMetRiC, we tested both versions of the prediction model. Each model includes age, sex, BMI, waist circumference, smoking, and antipsychotic medication as predictors. In addition, the full model includes the blood levels of triglycerides and HDL cholesterol. Discrimination accuracy, measured by the C-statistic, reports the proportion of individuals who developed MetS who received a higher risk score than individuals who did not develop MetS. Our results show stable discriminative accuracy for both PsyMetRiC versions compared with validations done in FEP cohorts in Spain and Switzerland, each showing slightly lower discriminative accuracy compared with the original external validation in the UK. As in the previous PsyMetRiC validations, the full model had marginally better discriminative accuracy, supporting the utility of blood tests in evaluating the risk of MetS in individuals with early psychosis. However, validation in larger samples will be required to confirm whether the PsyMetRiC-Fi full model outperforms the partial model. Both PsyMetRiC versions showed superior discriminative accuracy in this early psychosis population compared with using FINDRISC.

Appropriate assessments of algorithm calibration are essential to assess the agreement of predicted risk estimates compared with observed risk proportions. A poorly calibrated algorithm (i.e., where predicted risk estimates under- or over-predict risk) may lead to individuals being inappropriately withheld from potentially impactful, necessary interventions, or vice versa—being offered interventions they do not need. In the calibration analysis we detected that the PsyMetRiC full model tended to underpredict risk, and that the partial model overpredicted it. One potential explanation for this pattern of miscalibration is the difference in sociodemographic and clinical characteristics of the Finnish sample compared with original UK samples. Most notably, the ethnicity of the individuals in the Finnish sample was exclusively White European, while in the UK sample roughly half of the participants were of White European ethnicity.

Following recalibration of PsyMetRiC-Fi to account for population differences, calibration of both models improved substantially with no detriment to discriminative accuracy. The advantage of recalibration is that it does not change the predictor coefficients. This means the algorithm benefits from the large sample used to develop and externally validate it in the first place, and each additional successful validation study adds further confidence in the overall generalisability of the algorithm, with a larger total sample size involved across all validations. These are notable benefits that are relatively rare in risk prediction research across biomedicine.

Calibration analysis of FINDRISC demonstrated a consistent underprediction of risk in this early psychosis sample. This result is consistent with other studies showing that cardiometabolic risk prediction algorithms developed for older adults from the general population are unsuitable for the early psychosis population because they substantially underpredict risk in this group.

We used decision curve analysis to estimate the likely clinical usefulness of PsyMetRiC, compared to competing strategies of using FINDRISC, intervening in all, or intervening in none. Both PsyMetRiC versions displayed superior net benefit compared with competing strategies across a range of reasonable risk thresholds, and the likely clinical usefulness improved even further following recalibration.

In sensitivity analysis, we found that the predictive accuracy of PsyMetRiC was stable in the psychosis risk group as compared to the FEP group, the full model being again more accurate in predicting risk than the partial model. Individuals with high risk for psychosis have a high prevalence of metabolic risk factors similar to those with FEP, including insulin resistance and dyslipidemia.[36]

When considering the potential generalisability of our results to the whole Finnish population with early psychosis, the relatively large geographical variation in the prevalence of cardiovascular disease and metabolic risk factors across Finland should be taken into consideration.[37] The sample in our study was recruited from two areas in the South-Western part of Finland and it would also therefore be important to validate PsyMetRiC-FI in other Finnish early psychosis samples.

There are a few other key limitations of our study. First, the sample size, while relatively large for a longitudinal study of early psychosis, limited the precision with which we could estimate algorithmic accuracy. The testing of PsyMetRiC in larger international samples in future will help to confirm its accuracy across borders and permit the extension of the model with additional predictors of local importance. Second, the relatively large proportion of missing data at follow-up is another limitation. However, we did not detect any significant patterns of not-random missingness and thus were able to use imputation to account for the missing observations. Third, we did not use the complete FINDRISC algorithm, leaving out family history of diabetes, use of antihypertensive medication, and data on diet and physical activity. We do not expect this to significantly influence the risk assessment, as the diet and physical activity variables were originally added to the FINDRISC model to highlight the importance of modifiable risk factors of T2D and did not contribute meaningfully to the predictive accuracy of FINDRISC.[31] Fourth, while FINDRISC has been shown to be accurate in MetS prediction in general population samples, it was developed to predict risk of T2DM rather than MetS, meaning our comparison to PsyMetRiC should be considered in this context.

5 | CONCLUSION

We report the first validation study of a cardiometabolic risk prediction algorithm tailored specifically for the early psychosis population in the Nordic region. PsyMetRiC-Fi appears promising for future routine use by health professionals caring for young people with or at risk of developing psychosis in Finland. Further validation studies in larger samples will be the critical next step toward clinical implementation of the model.

ACKNOWLEDGMENTS

The authors would like to thank all the study participants. We thank the HEPS and TEPS study groups for the data collection and interviews and Marjut Grainger for the HEPS study data management.

FUNDING INFORMATION

Helsinki Early Psychosis Study was supported by Academy of Finland (#278171 to JS), the Sigrid Juselius Foundation (JS), the Finnish Cultural Foundation grant (JS), and by the Helsinki University Central Hospital (Reference Numbers: TYH2013332) Turku Early Psychosis Study was supported by the Academy of Finland (#26080794) and Turku University Hospital grant P3848. The funding organizations had no role in the design and conduct of the study; collection, management, analysis, or interpretation of the data; or preparation, review, or approval of the manuscript.

CONFLICT OF INTEREST STATEMENT

GM has consulted for ieso Digital Health. The remaining authors have no conflicts of interest to declare.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/acps.13752.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Jaakko Keinänen https://orcid.org/0000-0003-3202-7382

REFERENCES

1. O'Connor RC, Worthman CM, Abanga M, et al. Gone too soon: priorities for action to prevent premature mortality associated with mental illness and mental distress. *Lancet Psychiatry*. 2023;10(6):452-464. doi:10.1016/S2215-0366(23)00058-5
2. De Hert M, Correll CU, Bobes J, et al. Physical illness in patients with severe mental disorders. I. Prevalence, impact of medications and disparities in health care. *World Psychiatry*. 2011;10(1):52-77. doi:10.1002/j.2051-5545.2011.tb00014.x
3. Plana-Ripoll O, Pedersen CB, Agerbo E, et al. A comprehensive analysis of mortality-related health metrics associated with mental disorders: a nationwide, register-based cohort study. *Lancet*. 2019;394(10211):1827-1835. doi:10.1016/S0140-6736(19)32316-5
4. Hansen HG, Starzer M, Nilsson SF, Hjorthøj C, Albert N, Nordentoft M. Clinical recovery and long-term Association of Specialized Early Intervention Services vs treatment as usual among individuals with first-episode schizophrenia Spectrum disorder: 20-year follow-up of the OPUS trial. *JAMA Psychiatry*. 2023;80(4):371-379. doi:10.1001/jamapsychiatry.2022.5164
5. Perry B, Holt R, Chew-Graham C, et al. Update (with acknowledgement to the late Helen Lester for her contribution to the original 2012 version) positive cardiometabolic health resource:

an intervention framework for people experiencing psychosis and schizophrenia. 2023 update. Published online 2023. 2023.
6. Alberti K, Zimmet P, Shaw J. The metabolic syndrome—a new worldwide definition. *Lancet*. 2005;366(9491):1059-1062. www.thelancet.com
7. Malik S, Wong ND, Franklin SS, et al. Impact of the metabolic syndrome on mortality from coronary heart disease, cardiovascular disease, and all causes in United States adults. *Circulation*. 2004;110(10):1245-1250. doi:10.1161/01.CIR.0000140677.20606.0E
8. Perry BI, McIntosh G, Weich S, Singh S, Rees K. The association between first-episode psychosis and abnormal glycaemic control: systematic review and meta-analysis. *Lancet Psychiatry*. 2016;3(11):1049-1058. doi:10.1016/S2215-0366(16)30262-0
9. Pillinger T, Beck K, Stubbs B, Howes OD. Cholesterol and triglyceride levels in first-episode psychosis: systematic review and meta-analysis. *Br J Psychiatry*. 2017;211(6):339-349. doi:10.1192/bjp.bp.117.200907
10. Vázquez-Bourgon J, Gómez-Revuelta M, Mayoral-van Son J, et al. Pattern of long-term weight and metabolic changes after a first episode of psychosis: results from a 10-year prospective follow-up of the PAFIP program for early intervention in psychosis cohort. *Eur Psychiatry*. 2022;65(1):e48. doi:10.1192/j.eurpsy.2022.2308
11. Abdallah M, Sharbaji S, Sharbaji M, et al. Diagnostic accuracy of the Finnish diabetes risk score for the prediction of undiagnosed type 2 diabetes, prediabetes, and metabolic syndrome in the Lebanese University. *Diabetol Metab Syndr*. 2020;12(1):84. doi:10.1186/s13098-020-00590-8
12. Makrilakis K, Liatis S, Grammatikou S, et al. Validation du questionnaire finlandais calculant un score de risque de diabète (FINDRISC) pour le dépistage du diabète de type 2, des anomalies de la glycorégulation et du syndrome métabolique en Grèce. *Diabetes Metab*. 2011;37(2):144-151. doi:10.1016/j.diabet.2010.09.006
13. Saaristo T, Peltonen M, Lindström J, et al. Cross-sectional evaluation of the Finnish diabetes risk score: a tool to identify undetected type 2 diabetes, abnormal glucose tolerance and metabolic syndrome. *Diab Vasc Dis Res*. 2005;2(2):67-72.
14. Janghorbani M, Adineh H, Amini M. Evaluation of the Finnish diabetes risk score (FINDRISC) as a screening tool for the metabolic syndrome. *Rev Diabet Stud*. 2013;10(4):283-292. doi:10.1900/RDS.2013.10.283
15. Tabák AG, Jokela M, Akbaraly TN, Brunner EJ, Kivimäki M, Witte DR. Trajectories of glycaemia, insulin sensitivity, and insulin secretion before diagnosis of type 2 diabetes: an analysis from the Whitehall II study. *Lancet*. 2009;373(9682):2215-2221.
16. Perry BI, Upthegrove R, Crawford O, et al. Cardiometabolic risk prediction algorithms for young people with psychosis: a systematic review and exploratory analysis. *Acta Psychiatr Scand*. 2020;142(3):215-232. doi:10.1111/acps.13212
17. Quadackers D, Lieburg E, Bos F, et al. Cardiovascular risk assessment methods yield unequal risk predictions: a large cross-sectional study in psychiatric secondary care outpatients. *BMC Psychiatry*. 2023;23(1):536. doi:10.1186/s12888-023-05022-1
18. Perry BI, Osimo E, Upthegrove R, et al. Development and external validation of the psychosis metabolic risk calculator (PsyMetRiC): a cardiometabolic risk prediction algorithm for young people with psychosis. *Lancet Psychiatry*. 2021;8(7):589-598. doi:10.1016/S2215-0366(21)00114-0
19. Perry B, Vandenberghe F, Garrido-Torres N, et al. The psychosis metabolic risk calculator (PsyMetRiC) for young people with psychosis: international external validation and site-specific recalibration in two independent European samples. *Lancet Regional Health - Eur*. 2022;22:22. doi:10.1016/j.lanepe.2022.100498
20. Read SH, Van Diepen M, Colhoun HM, et al. Performance of cardiovascular disease risk scores in people diagnosed with type 2 diabetes: external validation using data from the National Scottish diabetes register. *Diabetes Care*. Vol 41. American Diabetes Association Inc; 2018:2010-2018. doi:10.2337/dc18-0578
21. Damen JA, Pajouheshnia R, Heus P, et al. Performance of the Framingham risk models and pooled cohort equations for predicting 10-year risk of cardiovascular disease: a systematic review and meta-analysis. *BMC Med*. 2019;17(1):109. doi:10.1186/s12916-019-1340-7
22. Pike MM, Decker PA, Larson NB, et al. Improvement in cardiovascular risk prediction with electronic health records. *J Cardiovasc Transl Res*. 2016;9(3):214-222. doi:10.1007/s12265-016-9687-z
23. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMC Med*. 2015;13(1):1. doi:10.1186/s12916-014-0241-z
24. Ventura J, Lukoff D, Nuechterlein K, Liberman R, Green M, Shaner A. Brief psychiatric rating scale (BPRS), expanded version (4.0): scales, anchor points, and administration manual. *Int J Methods Psychiatr Res*. 1993;3:227-243.
25. First MB, Spitzer RL, Gibbon M, Williams JBW. *Structured clinical interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. (SCID-I/P). Biometrics Research*. New York State Psychiatric Institute; 2002.
26. Keinänen J, Suvisaari J, Reinikainen J, et al. Low-grade inflammation in first-episode psychosis is determined by waist circumference increase. *Psychiatry Res*. 2018;12(270):547-553. doi:10.1093/schbul/sby017.697
27. Miller TJ, McGlashan TH, Lifshey Rosen J, et al. Prospective diagnosis of the initial prodrome for schizophrenia based on the Structured Interview for Prodromal Syndromes: preliminary evidence of interrater reliability and predictive validity. *Am J Psychiatry*. 2002;159(5):863-865. doi:10.1176/appi.ajp.159.5.863
28. Salokangas RKR, From T, Ilonen T, et al. Short-term functional outcome in psychotic patients: results of the Turku early psychosis study (TEPS). *BMC Psychiatry*. 2021;21(1):602. doi:10.1186/s12888-021-03516-4
29. National Institute for Health and Welfare. Sotkanet Statistics and Indicator Bank. Population at year end (id:127). 2024.
30. Alberti KGMM, Eckel RH, Grundy SM, et al. Harmonizing the metabolic syndrome: a joint interim statement of the international diabetes federation task force on epidemiology and prevention. *Circulation*. 2009;120(16):1640-1645. doi:10.1161/CIRCULATIONAHA.109.192644
31. Lindstrom J, Tuomilehto J. The diabetes risk score a practical tool to predict type 2 diabetes risk. *Diabetes Care*. 2003;26(3):725-731.

32. Pavlou M, Qu C, Omar RZ, et al. Estimation of required sample size for external validation of risk models for binary outcomes. *Stat Methods Med Res*. 2021;30(10):2187-2206. doi:10.1177/09622802211007522
33. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Mak*. 2006;26(6):565-574. doi:10.1177/0272989X06295361
34. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. *Diagn Progn Res*. 2019;3(1):18. doi:10.1186/s41512-019-0064-7
35. R Core Team. R: a language and environment for statistical computing. 2022. Accessed March 26, 2024. https://www.R-project.org/
36. Cadenhead KS, Minichino A, Kelsven S, et al. Metabolic abnormalities and low dietary omega 3 are associated with symptom severity and worse functioning prior to the onset of psychosis: findings from the north American Prodrome longitudinal studies consortium. *Schizophr Res*. 2019;204:96-103. doi:10.1016/j.schres.2018.09.022
37. Parikka S, Koskela T, Pietilä A, Koponen P, Koskinen S. THL's Morbidity Index 2019: Large regional differences in morbidity. Statistical report. 2022.

SUPPORTING INFORMATION

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

How to cite this article: Keinänen J, Eskelinen S, From T, et al. Psychosis metabolic risk calculator (PsyMetRiC) in early psychosis: External validation study in Finland. *Acta Psychiatr Scand*. 2024;1-13. doi:10.1111/acps.13752