Assessing the generalisability of the psychosis metabolic risk calculator (PsyMetRiC) for young people with first-episode psychosis with validation in a Hong Kong Chinese Han population: a 4-year follow-up study

Wing Tse,[a] Golam M. Khandaker,[b] Huiquan Zhou,[a,c] Hao Luo,[c] Wai Ching Yan,[d] Man Wah Siu,[d] Lap Tak Poon,[e] Edwin Ho Ming Lee,[a] Qingpeng Zhang,[f,g] Rachel Upthegrove,[h] Emanuele F. Osimo,[i,j] Benjamin I. Perry,[i,j,l,]* and Sherry Kit Wa Chan[a,k,l,]**

[a]Department of Psychiatry, School of Clinical Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, China
[b]Bristol Medical School, University of Bristol, Bristol, England
[c]Department of Social Work and Social Administration, Faculty of Social Sciences, The University of Hong Kong, Hong Kong Special Administrative Region, China
[d]Department of Psychiatry, Kowloon Hospital, Hong Kong Special Administrative Region, China
[e]Department of Psychiatry, United Christian Hospital, Hong Kong Special Administrative Region, China
[f]Musketeers Foundation Institute of Data Science, The University of Hong Kong, Hong Kong Special Administrative Region, China
[g]Department of Pharmacology and Pharmacy, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, China
[h]Institute for Mental Health, University of Birmingham, England
[i]Department of Psychiatry, University of Cambridge School of Clinical Medicine, Cambridge, England
[j]Cambridgeshire and Peterborough NHS Foundation Trust, Cambridge, England
[k]The State Key Laboratory of Brain and Cognitive Sciences, The University of Hong Kong, Hong Kong Special Administrative Region, China

Summary

Background Metabolic syndrome (MetS) is common following first-episode psychosis (FEP), contributing to substantial morbidity and mortality. The Psychosis Metabolic Risk Calculator (PsyMetRiC), a risk prediction algorithm for MetS following a FEP diagnosis, was developed in the United Kingdom and has been validated in other European populations. However, the predictive accuracy of PsyMetRiC in Chinese populations is unknown.

Methods FEP patients aged 15–35 y, first presented to the Early Assessment Service for Young People with Early Psychosis (EASY) Programme in Hong Kong (HK) between 2012 and 2021 were included. A binary MetS outcome was determined based on the latest available follow-up clinical information between 1 and 12 years after baseline assessment. The PsyMetRiC Full and Partial algorithms were assessed for discrimination, calibration and clinical utility in the HK sample, and logistic calibration was conducted to account for population differences. Sensitivity analysis was performed in patients aged >35 years and using Chinese MetS criteria.

Findings The main analysis included 416 FEP patients (mean age = 23.8 y, male sex = 40.4%, 22.4% MetS prevalence at follow-up). PsyMetRiC showed adequate discriminative performance (full-model C = 0.76, 95% C.I. = 0.69–0.81; partial-model: C = 0.73, 95% C.I. = 0.65–0.8). Systematic risk underestimation in both models was corrected using logistic calibration to refine PsyMetRiC for HK Chinese FEP population (PsyMetRiC-HK). PsyMetRiC-HK provided a greater net benefit than competing strategies. Results remained robust with a Chinese MetS definition, but worse for the older age group.

Interpretation With good predictive performance for incident MetS, PsyMetRiC-HK presents a step forward for personalized preventative strategies of cardiometabolic morbidity and mortality in young Hong Kong Chinese FEP patients.

---

*Corresponding author. Department of Psychiatry, University of Cambridge, Cambridge, CB2 0SZ, UK.
**Corresponding author. Room 219, New Clinical Building, Department of Psychiatry, The University of Hong Kong, Queen Mary Hospital, 102 Pokfulam Road, Hong Kong Special Administrative Region, China.
    E-mail addresses: bip20@medschl.cam.ac.uk (B.I. Perry), kwsherry@hku.hk (S.K.W. Chan).
[l]These authors contributed equally.

Research in context

Evidence before this study
The pooled prevalence of metabolic syndrome (MetS) in people with severe mental illness including psychosis was 32.6% which contributed significantly to the excess mortality rate of this population. Even at the onset of psychosis, the rate of MetS was 13% and continued to increase over time. There are only two prognostic risk prediction algorithms of MetS for general population and people with general SMI. There is only one metabolic risk calculator for younger people with first-episode psychosis (FEP), the PsyMetRiC, developed in the UK population, and validated in two independent European samples (Switzerland and Spain). PsyMetRic performed consistently in both populations regarding clinical usefulness and reliability of predicting the risk of developing MetS in young individuals with psychosis. However, MetS risk varies by ethnicity, which warrants further testing of this risk prediction algorithm in non-European populations.

Added value of this study
This study reports the first validation of PsyMetRiC in a non-European sample. Detailed sensitivity analysis also includes

the first assessment of PsyMetRiC's predictive performance for the Chinese MetS criteria, and the first assessment of PsyMetRiC's predictive performance in adults with first episode psychosis aged >35 yrs. Results suggested that PsyMetRiC showed stability in discriminative ability in the Hong Kong sample but led to a systematic under-prediction of metabolic risk. Site-specific recalibration of PsyMetRiC greatly improved calibration performance and likely clinical usefulness with no detriment to discrimination.

Implications of all the available evidence
The PsyMetRic is generalizable in the Hong Kong Chinese Han population with satisfactory predictive accuracy and the PsyMetRiC-HK is a thus a reliable prognostic risk prediction algorithm of MetS in the younger FEP in Hong Kong. This is the first step towards personalized preventive strategies of cardiometabolic morbidity and mortality in young Chinese FEP patients. Further validation studies in the Chinese or Asian population will help to further probe the generalizability of PsyMetRiC.

Introduction

People with severe mental illness (SMI) including schizophrenia-spectrum, uni- and bipolar affective disorders have two to three times higher mortality rates than the general population.[1,2] It is estimated that 60% of this excess mortality could be attributed to physical morbidities, predominantly diabetes, obesity and cardiovascular diseases (CVD).[2,3] A combination of cardiometabolic risk factors including central obesity, high blood pressure, low high-density lipoprotein (HDL) cholesterol, elevated triglycerides and hyperglycaemia are termed the Metabolic syndrome (MetS), and MetS is a key risk factor of longer-term cardiometabolic morbidity[4] and mortality.[5] Clear definitions of MetS have been established by different organizations,[6–9] with slight variations over time and across international borders, aiming to facilitate the identification of a high-risk group so as to prevent the development of major morbidity and reduce mortality.

Meta-analyses found that the prevalence of MetS in patients with schizophrenia was 32.5%[10] and the pooled prevalence of MetS of patients with SMI was 32.6%.[2] Compared with the general population, people with SMI have a significantly higher risk of MetS.[2] Even

among younger individuals at psychosis onset, the rate of MetS was found to be 13%[10] and significantly increased over time.[11] Furthermore, about twice as many patients with a first-episode psychosis (FEP) as age and sex matched healthy controls had at least one altered MetS component.[12] FEP are patients who experienced signs or symptoms of psychosis, such as hallucinations, delusions, with the level of severity and period that warrant a psychiatric diagnosis, for the first time in their life. Therefore, identification of people at psychosis onset who are at higher risk of developing MetS would allow the personalised tailoring of treatments to reduce the development of CVD-related morbidity and mortality. Prognostic risk prediction algorithms are routinely used in the general population to identify individuals at higher cardiometabolic risk. However, psychosis has a peak incidence around the early twenties, and existing cardiometabolic prognostic algorithms substantially under-estimate cardiometabolic risk in this group.[13]

Therefore, a Psychosis Metabolic Risk Calculator (PsyMetRiC) was developed in the UK to predict the risk of developing MetS up to six-years after FEP for patients aged 16–35 years, and external validation in the UK demonstrated the potential for generalisability in the UK

population (C-statistics of full-model ranged from 0.75 to 0.80).[14] PsyMetRiC has recently been further validated in two independent Western European FEP samples of the same age group (Spain and Switzerland) and retained good discriminative accuracy (C-statistics ranged from 0.72 to 0.73).[15] The results suggest the generalizability of PsyMetRiC in European populations. Differences in MetS prevalence between different ethnicities and regions have been observed including the Western and Chinese population. A comparative study found the prevalence of MetS doubled in the Western population compared with the Hong Kong Chinese population.[16] Some of these differences may be due to genetics, however many could be attributed to the differences in lifestyle, diet, health systems, and/or urbanicity.[17] Indeed, it has been suggested that about 3% of the heterogeneity between the prevalence of MetS in FEP could be attributed to ethnicity in general.[18] Therefore, PsyMetRiC requires validation in the Chinese population to ensure it is generalisable.

The current study aimed to validate and recalibrate PsyMetRiC in a Hong Kong Chinese Han population (PsyMetRiC-HK) of people with a FEP, aged 16–35 years, who were receiving care from psychosis early intervention services (EIS) in Hong Kong. In 2001, the Early Assessment Service for Young People with Early Psychosis (EASY) Programme was established by the Hospital Authority as a region-wide service to provide two-year phase-specific multi-disciplinary interventions to patients with FEP of age 15–25.[19,20] The service was further expanded to a three-year service covering FEP with ages 15–64 in 2011.[21] In order to examine the possible generalizability of the algorithms for the whole population served by the early intervention service locally and across different MetS criteria, the prediction performances of PsyMetRiC-HK using Chinese Diagnostic Criteria for MetS proposed by the Chinese Diabetes Society (CDS) or among those age over 35 were also examined as sensitivity analyses in the current study.

Methods
Study sample and data sources
We included patients with a FEP who first presented to the Early Assessment Service for Young People with Early Psychosis (EASY) Programme of Queen Mary Hospital, Hong Kong, between 2012 and 2021. The EASY service of Queen Mary Hospital is one of the seven centers of the EASY programme in Hong Kong, serving a population of 0.55 million approximately. Inclusion and exclusion criteria follow that of the previous PsyMetRiC study.[14] In detail, patients with FEP, Han Chinese, aged 16–35 years at first service contact, and who had ≥1 year of follow-up data available were included. Patients who had missing data on all predictor or outcome variables were excluded, as were individuals who had MetS at baseline (first mental health service contact). Information of subjects aged 36–64 years were also obtained for sensitivity analysis. Patients with a diagnosis of a non-organic psychosis-spectrum disorder at baseline (ICD-10 codes F06.0–2, F20–F31, F32.3, F33.3, F53.1 as defined in the original PsyMetRiC study) were included. As the EASY service excludes people with co-morbid organic brain conditions, drug-induced psychosis or moderate to severe learning disabilities, individuals with these diagnoses were also excluded. Supplementary Table S1 shows the diagnostic classification of included patients, and a recruitment flow-chart is shown in Supplementary Figure S1. Clinical information was obtained from the written clinical records and the Clinical Management System (CMS), which is a central electronic health record system of the public health system in Hong Kong, including psychiatric services. All clinical information including investigation results, inpatient and outpatient records, medication prescribed, and clinical diagnosis are included in the CMS.[22] Institutional ethical approval was obtained from the Institutional Review Board of The University of Hong Kong/Hospital Authority Hong Kong West Cluster (IRB reference number: UW 21–626). All information was obtained anonymously, and written consent was waived. Data was collected from 6 December 2022 to 12 April 2023.

Baseline information and outcome
Baseline information included age at presentation, biological sex, waist circumference (cm), weight (kg), height (cm), body mass index (BMI), current smoking status, prescription of more metabolically-active antipsychotic medications (Supplementary Table S2), triglycerides (mmol/L), high-density lipoprotein (HDL; mmol/L), systolic and diastolic blood pressure (mmHg), and fasting blood glucose (FBG; mmol/L). As per the original PsyMetRiC study,[15] we used the harmonized definition of MetS as a binary outcome[8]: ethnicity-specific waist circumference ≥90 cm in males and ≥80 cm in females for our sample, or BMI >29.9; alongside two of: triglycerides ≥1.70 mmol/L; HDL <1.03 mmol/L (males) or <1.29 mmol/L (females); systolic blood pressure >130 mmHg; FBG >5.60 mmol/L. When multiple follow-ups were available for each participant, we used the latest follow-up available between 1 and 12 years after baseline with the most available data.

Study population characteristics
The Hong Kong population is mostly Han Chinese (98.3%), the sociodemographic, economic, and healthcare-related information were recorded in Supplementary Table S3. Hong Kong has a population of 7.498 million (mid-2023) inhabitants, of whom 0.55 million is served by the Queen Mary Hospital, covering the Central, Western and Southern districts of the Hong Kong Island. According to the Census and Statistics

Department (C&SD)–The Government of the Hong Kong Special Administrative Region (HKSAR), of the employed persons in 2022, 87.5% work in services, 11.7% work in industry, and 0.8% work in the primary sector, with unemployment rate to be 3.7% in Quarter 4 (September–November, 2022), and the gross domestic product per capita in 2022 was HK$383,611.

The PsyMetRiC algorithms

Two forced-entry multivariable penalized logistic regression equations are included in the PsyMetRiC: the full-model and the partial-model. Predictors were included on a balance of clinical knowledge, prior research, and likely clinical usefulness/patient acceptability. The partial-model was developed to cover eventualities where biochemical results may not be available. The PsyMetRiC algorithm coefficients are presented in Supplementary Table S4. See the original PsyMetRiC[14] study for further details.

Statistical analysis

Sample preparation and estimation of analytic precision
Biochemical values of triglycerides, HDL, and FBG, were converted to mmol/L where necessary. Multiple imputation with chained equations was used to address missing data (Supplementary Methods). For numerical-based analyses, Rubin's rules were used to pool the estimates. For plot-based analyses, plots were generated in each imputed dataset and checked for similarity, with one randomly selected plot per analysis presented in the main manuscript and Supplementary Figures S5–S7. Recently developed criteria[23] to estimate analytic precision given the fixed sample sizes were applied (Supplementary Methods). Briefly, the expected SEs for the C-statistic were 0.032. The expected standard errors (SEs) for the calibration slope and calibration-in-the-large were 0.16 & 0.15 respectively. Comparisons between the original PsyMetRiC development sample and EASY samples for key sociodemographic, lifestyle and biochemical characteristics were performed using ANOVA (for means) and the chi-square equality of proportions test (for proportions). All analyses presented herein were performed independently in Hong Kong, conducted using R version 4.2.3.

Primary external validation analysis
Histograms were used to check the distribution of the predicted probabilities of outcomes. Algorithm performance was primarily assessed 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). TRIPOD reporting guideline was followed[24] (Supplementary Table S5).

Recalibration and generation of site-specific PsyMetRiC versions
Due to the differences of data from international sources, variations in the calibration performance compared to the original PsyMetRiC was anticipated. Visual examination of the calibration plots was conducted to detect any miscalibration (i.e., disagreement between the observed proportions and predicted probability). Logistic calibration was conducted to adjust for variances in baseline risk that may exist between populations by re-estimating the intercept and the slope. Consequently, logistic calibration assumes that the relative effects of the predictors are similar while accommodating the possibility of larger or smaller absolute effects. By completing this step, we obtained a site-specific version of PsyMetRiC (PsyMetRiC-HK) (Supplementary Methods for the recalibration with logistic calibration). For all results, we present performance estimates accompanied by 95% CIs. The relative importance of predictors in the original PsyMetRiC algorithm were checked and the same predictors of the PsyMetRic was used to fit a new regression model in the EASY data. The relative importance of the predictors in this regression model was checked and compared with the those in the original PsyMetRiC.

Clinical usefulness
To evaluate the clinical usefulness of the algorithm, we employed decision curve analysis, which assessed the net benefit over a threshold. The threshold represents the risk score at which an intervention would be considered necessary (see Supplementary Methods). In our study, we set a risk threshold upper bound of 0.30, indicating a roughly one-third chance of developing MetS if no intervention takes place, as risk thresholds exceeding this value may likely require intervention. Net benefit considers the consequences of decisions made using the algorithm and is therefore a more favourable measure. We presented the net benefit and standardized net benefit (net benefit divided by outcome prevalence, representing the additional percentage of cases that could be addressed using PsyMetRiC without increasing false positives) across a range of reasonable risk thresholds. To visualize and compare the net benefit of the original PsyMetRiC and PsyMetRiC-HK, a decision curve plot was created. According to Classical decision theory, the option with the highest net benefit should be preferred when a risk threshold is chosen.

Data visualisation and sensitivity analyses
An online data visualization website for PsyMetRiC was created to accompany the original study (https://psymetric.shinyapps.io/psymetric/). The website was updated with site-specific PsyMetRiC-HK versions obtained through recalibration analysis.

Sensitivity analysis was carried out in patients >35 years old to examine whether PsyMeRiC is generalisable

to older age groups. In addition, we also conducted sensitivity analysis using MetS criteria developed by the Chinese Diabetes Society (CDS). The definition of MetS for Chinese as per the CDS MetS criteria in the 2019 revision[25] was adapted, with the presence of at least three of the following: waist circumference ≥90 cm for male and ≥85 cm for female; triglycerides ≥1.70 mmol/L; HDL-C <1.04 mmol/L; systolic blood pressure ≥130 mmHg; diastolic blood pressure ≥85 mmHg; FBG ≥5.60 mmol/L. For FBG, we adapted the cut-off value according to the revision recommended by a nationwide study in China.[9]

Results
Samples
In total, our sample frame included 608 patients aged 15–64 years. After applying inclusion and exclusion criteria (we excluded $n = 13$ [3.03%] due to the presence of the outcome, MetS, at baseline; n = 179 were excluded for the main analysis based on being >35 years old; no patients were excluded due to missing data on all predictor or outcome variables), our main analysis included 416 patients (Table 1). Supplementary Table S6 shows the comparison of the characteristics of included and excluded samples. The included sample has a significantly younger age, higher in HDL, and lower in triglycerides, FBG, BMI, systolic BP, number of smokers, follow-up time, and metabolic syndrome at baseline and follow-up. Ninety-three included subjects (22.4%) were identified to have MetS at follow-up. The

Hong Kong samples differed from the original UK PsyMetRiC development sample on most sociodemographic, lifestyle and biochemical characteristics (Table 1). The Hong Kong sample has a significantly smaller proportion of males, fewer smokers, metabolic syndrome at baseline, lower in BMI, and biochemical values at baseline including triglycerides, HDL, and FBG. However, the Hong Kong sample has a significantly higher number of patients prescribed with a more-metabolically-active antipsychotic, and having a longer follow-up time (mean = 4.13 years, SD = 3.16 years). The pattern of missing values was assessed (Supplementary Figure S2), 25% of missing data was observed among the predictor variables. The matrix of missing values per pair of predictor variables (Supplementary Figure S3) and proportion of missing data per variable (Supplementary Table S7) were supplemented.

Primary external validation analysis in EASY, Hong Kong
The shape of the distribution of predicted probabilities was similar to the original PsyMetRiC study (Supplementary Figure S4). Predictive performance statistics before and after recalibration are reported in Table 2. Assessment of discrimination showed full-model C = 0.76, 95% C.I., 0.69–0.81; partial-model: C = 0.73, 95% C.I., 0.65–0.8. Calibration plots for the full-model and partial-model were similar across imputed datasets (Fig. 1, Supplementary Table S8, Supplementary Figure S5).

Characteristic	Original PsyMetRiC development sample (UK)	EASY validation sample (Hong Kong)	Between-group differences[d]
Sample before Inclusion/Exclusion Criteria Applied[a], N.	1504	608	/
Included sample size[a], N. (%)	651 (43.28)	416 (68.42)	/
Age in Years, mean (SD)	24.52 (4.91)	23.81 (5.69)	t = 4.68, p = 0.031
White European/NR Ethnicity, N. (%)	360 (55.3)	/	/
Black/African-Caribbean Ethnicity, N. (%)	109 (16.74)	/	/
Asian/Other Ethnicity, N. (%)	181 (27.80)	416 (100)	/
Male Sex, N. (%)	440 (67.59)	168 (40.38)	χ = 75.52, p < 0.001
HDL at baseline, mmol/L, mean (SD)	1.88 (0.57)	1.44 (0.38)	t = 193.03, p < 0.001
Triglycerides at baseline, mmol/L, mean (SD)	1.39 (1.06)	1.03 (0.59)	t = 40.05, p < 0.001
BMI at baseline, kg/m², mean (SD)	23.63 (5.43)	22.77 (5.28)	t = 6.50, p = 0.011
FBG at baseline (mmol/L), mean (SD)	5.19 (1.28)	4.65 (0.66)	t = 63.27, p < 0.001
Systolic BP at baseline (mmHg), mean (SD)	120.65 (11.68)	119.27 (16.59)	t = 2.54, p = 0.112
Prescribed a More-Metabolically-Active Antipsychotic[b], N. (%)	455 (69.89)	359 (86.30)	χ = 36.86, p < 0.001
Smoking at baseline, N. (%)	315 (48.39)	57 (13.70)	χ = 132.94, p < 0.001
Follow-up time, years, mean (SD)	1.86 (1.32)	4.13 (3.16)	t = 263.97, p < 0.001
Antipsychotic Naïve at baseline, N. (%)	NR	NR	/
Metabolic Syndrome at baseline, N. (%)[c]	49 (6.58)	13 (3.03)	χ = 7.33, p = 0.007
Metabolic Syndrome at Follow-up, N. (%)	109 (16.74)	93 (22.36)	χ = 4.85, p = 0.028

HDL, high-density lipoprotein; BMI, body mass index; FPG, fasting plasma glucose; BP, blood pressure; NR, Not recorded. [a]See Supplementary Figure S1 for a flow-chart of included participants in the study. [b]Definitions of Metabolically-active antipsychotics are listed in Supplementary Table S2. [c]Corresponds to the percentage of sample before those participants were excluded. [d]Analysis of means was conducted using t-tests. Analysis of proportions was conducted using the chi-square equality of proportions test.

Table 1: Sociodemographic characteristics and metabolic measurements of the original PsyMetRiC development sample and PsyMetRiC-HK.

Measure of predictive performance	Primary analysis, estimate (95% C.I.)		After logistic recalibration, estimate (95% C.I.)	
	Full-Model	Partial-Model	Full-Model	Partial-Model
PsyMetRiC-HK				
C-Statistic	0.76 (0.69, 0.81)	0.73 (0.65, 0.8)	0.76 (0.69, 0.81)	0.73 (0.65, 0.8)
r^2	0.14 (0.10, 0.18)	0.12 (0.09, 0.15)	0.17 (0.14, 0.21)	0.14 (0.11, 0.17)
Calibration intercept	0.78 (0.20, 1.37)	0.79 (0.42, 1.16)	0.04 (0.01, 0.06)	0.08 (−0.01, 0.17)
Calibration slope	1.30 (0.89, 1.72)	1.15 (0.73, 1.58)	1.05 (0.97, 1.13)	1.07 (0.94, 1.13)
Brier score	−0.04 (−0.12, 0.04)	0.04 (−0.04, 0.11)	0.03 (0.01, 0.05)	0.04 (0.01, 0.08)
PsyMetRiC–external validation in the UK				
C-Statistic	0.75 (0.69, 0.80)	0.74 (0.67, 0.79)		
r^2	0.21 (0.18, 0.25)	0.17 (0.14, 0.20)		
Calibration intercept	−0.05 (−0.08, −0.02)	−0.07 (−0.11, −0.03)		
Brier score	0.07 (0.04, 0.10)	0.08 (0.05, 0.11)		

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. For comparison, results from the original PsyMetRiC external validation in the UK are shown in the table, see the original PsyMetRiC manuscript for further details.[14]

Table 2: **Predictive performance statistics of the PsyMetRiC full- and partial models before and after logistic calibration in PsyMetRiC-HK.**

Algorithm recalibration and generation of site-specific PsyMetRiC versions (PsyMetRiC-HK)

After logistic calibration (Fig. 1, Supplementary Table S8), the shape of the distributions of predicted probabilities changed slightly (Supplementary Figure S4). Recalibrated performance statistics are reported in Table 2. Recalibration plots for both PsyMetRiC versions were similar across imputed datasets (Fig. 1; Supplementary Figure S6) and showed improved calibration performance. We assessed for the presence of predictor multicollinearity by measuring the variance inflation factor (Supplementary Table S9). See Supplementary Table S10 for the associations of individual PsyMetRiC predictors with MetS in the EASY sample. The relative importance of the predictors in the original UK PsyMetRiC and in the recalibrated PsyMetRiC-HK after the new regression model was fitted are shown in Supplementary Table S11 and Fig. 2.
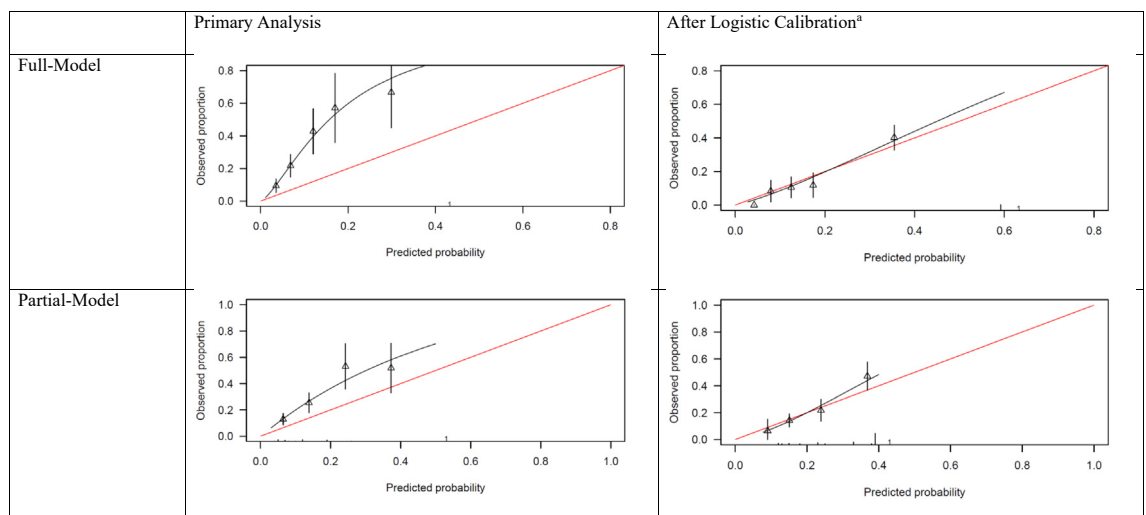
Fig. 1: **Calibration plots of PsyMetRiC-HK**. Calibration plots illustrate agreement between the observed (y axis) and predicted risk (x axis). Perfect agreement would trace the red line. Algorithm calibration is illustrated by the black line. Triangles denote grouped observations for participants at deciles of predicted risk, with 95% C.I.'s indicated by the vertical black lines. [a]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. See Methods.
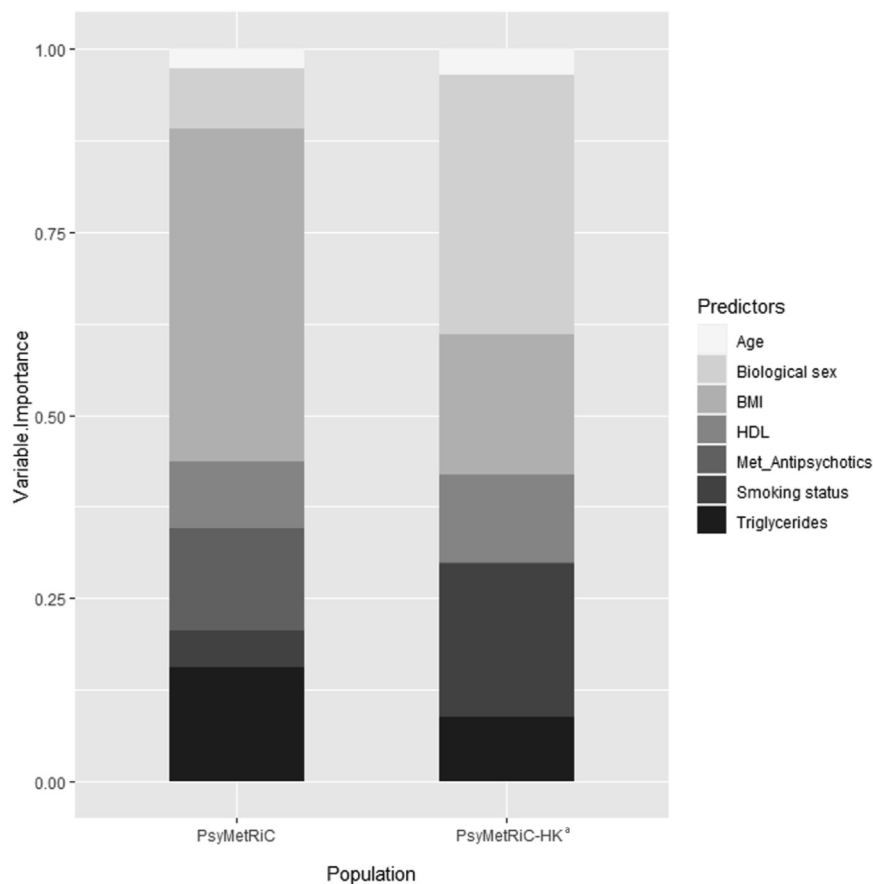
Fig. 2: Comparison of variable importance between the original PsyMetRiC regression model and a logistic regression model with the same predictors fit in EASY sample. BMI, body mass index; HDL, high-density lipoprotein cholesterol; Met_Antipsychotics, metabolically-active antipsychotics; FPG, fasting plasma glucose; BP, blood pressure. [a]Corresponds to the regression model fit in the EASY sample. *Ethnicity is a constant variable in PsyMetRiC-HK, as there was no ethnicity diversity in the EASY sample. Therefore, it is not reported in the variable importance.

Ethnicity is a constant variable in recalibrated PsyMetRiC-HK, as only Han Chinese was included in the EASY sample. Therefore, ethnicity was not reported in the variable importance. Only small differences in the relative importance of different predictors in the EASY data compared with the original PsyMetRiC model were seen (Fig. 2), which explained the differences in algorithm performances, but the relative variable importance is generally consistent.
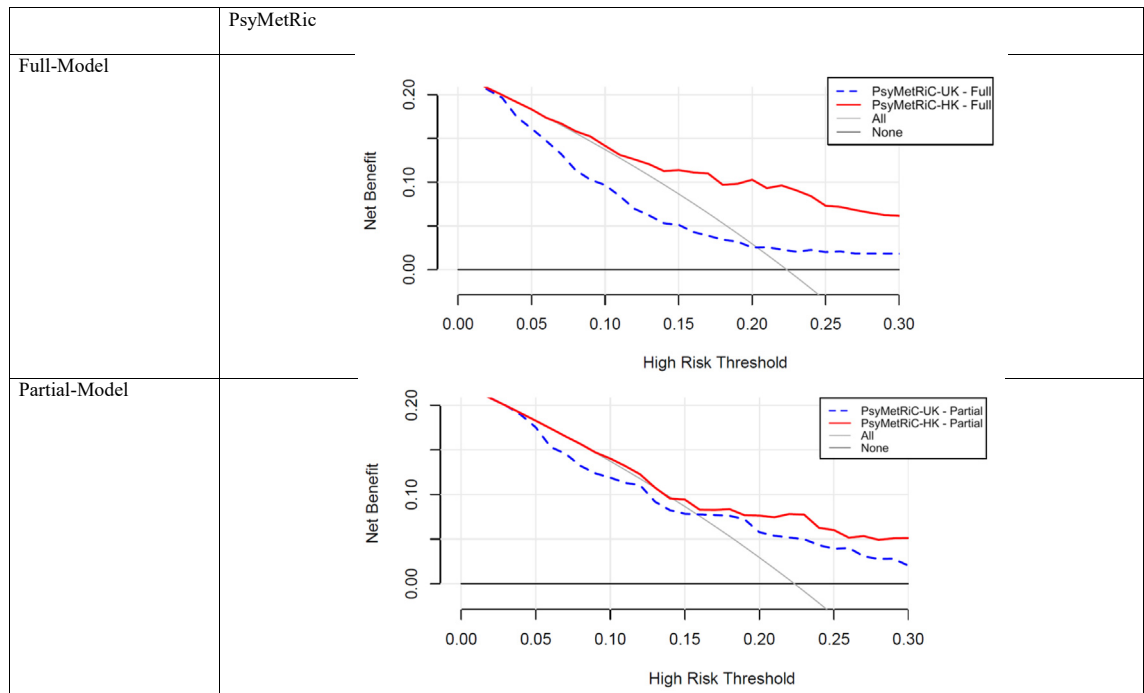
Clinical usefulness

Before recalibration, decision curve analysis (Fig. 3; Supplementary Figure S7) showed that in the EASY sample, both PsyMetRiC versions only displayed evidence of net benefit compared with competing strategies at relatively high risk-thresholds (>0.20 for the full-model; >0.16 for the partial-model). After recalibration, clinical utility vastly improved and both PsyMetRiC versions showed universally greater net benefit than competing strategies, net benefit was 0.073 at threshold

0.30, which means a 7.3% improvement in detecting MetS compared with competing strategies, and 6.4% improved compared with the original PsyMetRiC version before calibration. After recalibration, net benefit was about equal between PsyMetRiC versions. For example, if an intervention was considered for patients scoring higher than 0.15, the recalibrated full- and partial-models provided net benefits of 0.13 (95% C.I., 0.09–0.18) and 0.13 (95% C.I., 0.09–0.19) respectively, meaning that an additional 52% of metabolic syndrome cases could be prevented with both PsyMetRiC versions, with no increase in false positives (Fig. 3; Supplementary Figure S7; Supplementary Tables S12 and S13).

Sensitivity analysis

The harmonized definition of MetS used in the original PsyMetRiC study was replaced with the revised CDS MetS criteria[9] as a sensitivity analysis. After re-applying exclusion criteria, eight (1.86%) patients with MetS at

Fig. 3: Clinical usefulness of PsyMetRiC before and after logistic calibration. The plot reports net benefit (y axis) of PsyMetRiC Full- and Partial-Models (blue line = original PsyMetRiC algorithm applied to the sample; red line = recalibrated site-specific version PsyMetRiC-HK) across a range of risk thresholds (x axis) compared with intervening in all (grey 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 0.30 represents around 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.

baseline were excluded. Supplementary Table S14 shows the comparisons of demographics and baseline information between the included and excluded sample using the CDS MetS criteria. The predictive performance statistics were full-model C = 0.73, 95% C.I., 0.65–0.8; partial-model: C = 0.71, 95% C.I., 0.61–0.79 (Supplementary Figure S9).

Further sensitivity analysis was carried out in older population aged 36–64 (see sample characteristics in Supplementary Table S13). Results showed predictive performance of full-model C = 0.66 (95% C.I., 0.49–0.81) and partial-model: C = 0.62 (95% C.I., 0.5–0.73) (Supplementary Figure S8).

Discussion

The results of the current study suggest that PsyMetRiC, a MetS risk prediction algorithm for patients with FEP aged 16–35 developed in the UK, generalises well to the Chinese population in Hong Kong as the value of C-statistics and the confidence intervals of PsyMetRiC of UK and Hong Kong (HK) were similar. We found that PsyMetRiC in the HK sample discriminates cases of MetS adequately and with stability compared to the original UK study (C-statistics >0.70 for both PsyMetRiC versions in this study, which is generally considered acceptable/good, indicates over 70% of individuals who developed MetS received a higher risk score than individuals who did not). However, we found in calibration analysis that both versions of PsyMetRiC systematically underpredicted risk in the Hong Kong sample. This miscalibration was improved through logistic recalibration, enabling the development of a Hong Kong specific PsyMetRiC version (PsyMetRiC-HK), with no negative impact on discrimination performance. Results were similar in a sensitivity analysis including a Chinese MetS definition. However, there was evidence of worse predictive performance in older age-groups, which is unsurprising as PsyMetRiC was specifically developed to overcome the issue of models developed in non-SMI older populations.

The HK sample has a longer follow up period than that of the original UK sample, which might explain the slightly higher proportion of patients developing MetS in the HK sample but generally similar as that was reported in previous studies.[10,11] Demographically, the current sample obtained from one HK region has a similar pattern of demographic characteristics as the

previous territory-wide studies in HK,[19,26] but differed significantly from that of the UK sample, specifically being younger, with fewer smokers and a higher proportion of females. The HK population also differed significantly from the UK sample in many of the metabolic measurements at baseline, including triglycerides, HDL, and FBG. These differences in the baseline predictors as well as the slightly higher MetS rate are likely to have contributed to the less satisfactory calibration performance. Indeed, differences in the relative importance of the baseline variables in the model between the two samples were seen. Nevertheless, logistic calibration improved the performance of the algorithm by enhancing the agreement between the observed proportion and predicted risk. With a different MetS criteria specific for the Chinese population, the performance of PsyMetRiC-HK remained stable. However, the performance of PsyMetRiC-HK drops in the older population with confidence intervals of both full and partial models overlap with chance performance. Therefore, PsyMetRiC-HK of current form is not recommended for use in FEP aged 36 or over in Hong Kong. The small sample size could be one possible reason. On the other hand, age is a more important cardiometabolic risk factor in older age groups,[27] and the coefficient of age in PsyMetRiC was trained between 16 and 25. Although age might seem to be a relatively weak predictor in PsyMetRiC (Fig. 2), as it was only trained in the younger populations, it is possible to vary non-linearly with advancing age and may contribute to worse performance of PsyMetRiC in the older age group. Despite the positive findings of the validity of PsyMetRiC in this Hong Kong sample, further validation of the PsyMetRiC-HK with a bigger sample of different EASY service centers will still be required to ensure the generalizability. Larger sample sizes will permit further improvements to the predictive performance of PsyMetRiC-HK through the addition of other important baseline predictors such as low-density lipoprotein (LDL), family history of MetS and education level.[28,29] Though there are age-appropriate algorithms developed for older adults from the general population (e.g., QRISK)[30] and older adults with SMI (e.g., PRIMROSE),[27] none have yet been validated in the Chinese population. Therefore, developing or validating the existing tools in predicting the MetS in older FEP patients from the Chinese population is required. Apart from the age of population that PsyMetRiC was developed that is distinct from the other existing models including QRISK and PRIMROSE, PsyMetRiC used a different definition of metabolically-active antipsychotics, developed specifically in FEP population and with partial model which allows for the use of the tool without blood test results which might not be available in some settings and for some group of people. Details of distinction of PsyMetRiC compared with other existing tools were provided in previous literature[14]

These features may allow PsyMetRiC to be more readily usable in the clinical settings.

In the current study an upper risk threshold bound of 0.30 was set for the decision curve analysis. After logistic calibration, we found at risk threshold close to 0.07, the full model of PsyMetRiC-HK has higher net benefit than intervention for all, and at-risk threshold close to 0.10, the partial model has higher net benefit than intervention for all. This suggests the likely clinical usefulness of the PsyMetRiC-HK for a multitude of different potential interventions. However, an acceptable cut-off of risk threshold for the decision curve analysis should be decided and agreed by all key stakeholders including both patients and clinicians. Furthermore, types of interventions provided to the individuals scoring above the cut-off risk to prevent the development of MetS will require further study.

A population-based approach in monitoring the risk MetS in patients with psychosis should certainly be part of the routine care, however, targeted interventions for high-risk population would be crucial for the development of cost-effective service model.[31] People with psychotic disorders died 20 years sooner than the rest of the population and the mortality gaps is getting bigger over time.[32] Furthermore, over half of the total cost of treating psychotic disorders are account for by costs of co-morbidity.[33] In fact, improvement of physical health with precision medicine approach has been specifically stated as a priority in a recent patient-led consensus statement.[34] Indeed, the high-risk approach has been adopted routinely for decades to facilitate cardiometabolic risk prediction in the general population (e.g., QRISK in the UK which is in NICE guidelines). Yet, existing tools are inaccurate for young people with psychosis and will lead to further health inequalities in this population. Therefore, PsyMetRiC could be one of the useful clinical tools for the development of targeted approach to reduce the MetS and thus other related physical comorbidities in young people with FEP.

One of the key limitations of the study is sample size, which prevents more complex approaches to improve predictive performance (e.g., the addition of new predictors). In addition, included patients were from one clinical center in Hong Kong which may limit generalizability. In future, including other EASY service centers in analysis may help to address this. Secondly, some of the follow-up variables had a relatively high degree of missingness. This likely reflects the lack of awareness of measuring or documenting the key MetS measurements for younger patients with FEP. While we carefully used multiple imputation, a data-driven approach, to address this limitation, the impact of missingness may still impact the validity of our results. Furthermore, some important predictors of MetS were not included in the study which will also limit the predictive power of the tool. In addition, there are differential risk factors of MetS between different Asian populations, such as

adipose tissue distribution.[35] PsyMetRiC-HK only included Hong Kong Chinese Han population, thus cannot be generalizable for other Asian populations and further validation studies for other Asian population are needed.

Conclusions

This validation and recalibration study of a risk prediction model for MetS follows best-practice methods (including TRIPOD guidelines[24]), and uses a well characterized "real life" FEP sample of all suitable patients entering the early intervention service for psychosis in one clinical center in Hong Kong over 10 years. Therefore, the predictive performance of PsyMetRiC-HK is likely to match real clinical practice. Further studies are needed to determine a clinically acceptable risk threshold for MetS and types of interventions provided. Nonetheless, the establishment of this HK-specific version of PsyMetRiC is a clear step towards development of personalized preventive strategies of MetS in young FEP patients in Hong Kong.

References

1 Osborn DPJ, Levy G, Nazareth I, Petersen I, Islam A, King MB. Relative risk of cardiovascular and cancer mortality in people with severe mental illness from the United Kingdom's general practice rsearch database. *Arch Gen Psychiatry*. 2007;64:242–249.

2 Vancampfort D, Stubbs B, Mitchell AJ, et al. Risk of metabolic syndrome and its components in people with schizophrenia and related psychotic disorders, bipolar disorder and major depressive disorder: a systematic review and meta-analysis. *World Psychiatry*. 2015;14:339–347.

3 Hoang U, Goldacre MJ, Stewart R. Avoidable mortality in people with schizophrenia or bipolar disorder in England. *Acta Psychiatr Scand*. 2013;127:195–201.

4 Wilson PWF, D'Agostino RB, Parise H, Sullivan L, Meigs JB. Metabolic syndrome as a precursor of cardiovascular disease and type 2 diabetes mellitus. *Circulation*. 2005;112:3066–3072.

5 Hildrum B, Mykletun A, Dahl AA, Midthjell K. Metabolic syndrome and risk of mortality in middle-aged versus elderly individuals: the Nord-Trøndelag health study (HUNT). *Diabetologia*. 2009;52:583–590.

6 Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in adults. Executive summary of the third report of the national cholesterol education program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel III). *JAMA*. 2001;285:2486–2497.

7 Grundy SM, Cleeman JI, Daniels SR, et al. Diagnosis and management of the metabolic syndrome: an American heart association/national heart, lung, and blood institute scientific statement. *Circulation*. 2005;112:2735–2752.

8 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; national heart, lung, and blood institute; American heart association; world heart federation; international atherosclerosis society; and international association for the study of obesity. *Circulation*. 2009;120:1640–1645.

9 Xing Y, Xu S, Jia A, et al. Recommendations for revision of Chinese diagnostic criteria for metabolic syndrome: a nationwide study. *J Diabetes*. 2018;10:232–239.

10 Mitchell AJ, Vancampfort D, Sweers K, van Winkel R, Yu W, De Hert M. Prevalence of metabolic syndrome and metabolic abnormalities in schizophrenia and related disorders—a systematic review and meta-analysis. *Schizophr Bull*. 2013;39:306–318.

11 Coentre R, Levy P, Góis C, Figueira ML. Metabolic syndrome following a first episode of psychosis: results of a 1-year longitudinal study conducted in metropolitan Lisbon, Portugal. *J Int Med Res*. 2022;50:3000605221106703.

12 Garrido-Torres N, Ruiz-Veguilla M, Alameda L, et al. Prevalence of metabolic syndrome and related factors in a large sample of antipsychotic naïve patients with first-episode psychosis: baseline results from the PAFIP cohort. *Schizophr Res*. 2022;246:277–285.

13 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;142:215–232.

14 Perry BI, Osimo EF, 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;8:589–598.

15 Perry BI, 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 Reg Health Eur*. 2022;22:100493.

16 Patel A, Huang KC, Janus ED, et al. Is a single definition of the metabolic syndrome appropriate?—a comparative study of the USA and Asia. *Atherosclerosis*. 2006;184(1):225–232.

17 Saklayen MG. The global epidemic of the metabolic syndrome. *Curr Hypertens Rep*. 2018;20:12.

18 Garrido-Torres N, Rocha-Gonzalez I, Alameda L, et al. Metabolic syndrome in antipsychotic-naïve patients with first-episode psychosis: a systematic review and meta-analysis. *Psychol Med*. 2021;51:2307–2320.

19 Chan SKW, So HC, Hui CLM, et al. 10-year outcome study of an early intervention program for psychosis compared with standard care service. *Psychol Med*. 2015;45:1181–1193.

20 Tang JYM, Wong GHY, Hui CLM, et al. Early intervention for psychosis in Hong Kong–the EASY programme. *Early Interv Psychiatry*. 2010;4:214–219.

21 Lau KW, Chan SKW, Hui CLM, et al. Rates and predictors of disengagement of patients with first-episode psychosis from the early intervention service for sychosis service (EASY) covering 15 to 64 years of age in Hong Kong. *Early Interv Psychiatry*. 2019;13:398–404.

22 Chan SKW, Chan SWY, Pang HH, et al. Association of an early intervention service for psychosis with suicide rate among patients with first-episode schizophrenia-spectrum disorders. *JAMA Psychiatry*. 2018;75:458–464.

23 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;30:2187–2206.

24 Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Br J Surg*. 2015;102:148–158.

25 Jia W, Weng J, Zhu D, et al. Standards of medical care for type 2 diabetes in China 2019. *Diabetes Metab Res Rev*. 2019;35:e3158.

26 Wong TY, Chan SKW, Cheung C, et al. Dynamic patterns of symptoms and functioning in predicting deliberate self-harm in patients with first-episode schizophrenia-spectrum disorders over 3 years. *Schizophr Bull*. 2022;48:1043–1052.

27 Osborn DPJ, Hardoon S, Omar RZ, et al. Cardiovascular risk prediction models for people with severe mental illness: results from the prediction and management of cardiovascular risk in people with severe mental illnesses (PRIMROSE) research program. *JAMA Psychiatry*. 2015;72:143–151.

28 Ng S-M, Su X. Prevalence and correlates of metabolic syndrome in Hong Kong Chinese adults-a random community sample study. *Psychol Health Med*. 2018;23:485–495.

29 Feng Y, Yang X, Li Y, et al. Cohort study evaluation of new Chinese diabetes risk score: a new non-invasive indicator for predicting metabolic syndrome. *Prim Care Diabetes*. 2021;15:825–831.

30 Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M, Brindle P. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *BMJ*. 2007;335:136.

31 Rose G. Sick individuals and sick populations. *Int J Epidemiol*. 2001;30:427–432.

32 Younis A, Younis A, Tzur B, et al. Metabolic syndrome is independently associated with increased 20-year mortality in patients with stable coronary artery disease. *Cardiovasc Diabetol*. 2016;15(1):149.

33 Ride J, Kasteridis P, Gutacker N, Aragon Aragon MJ, Jacobs R. Healthcare costs for people with serious mental illness in England: an analysis of costs across primary care, hospital care, and specialist mental healthcare. *Appl Health Econ Health Policy*. 2020;18(2):177–188.

34 Tobias DK, Merino J, Ahmad A, et al. Second international consensus report on gaps and opportunities for the clinical translation of precision diabetes medicine. *Nat Med*. 2023;29(10):2438–2457.

35 Boon MR, Bakker LE, van der Linden RA, et al. High prevalence of cardiovascular disease in South Asians: central role for brown adipose tissue? *Crit Rev Clin Lab Sci*. 2015;52:150–157.