*DOI:*
10.1016/j.schres.2021.01.005

*Document Version*
Peer reviewed version

# THIRD EXTERNAL REPLICATON OF AN INDIVIDUALISED TRANSDIAGNOSTIC PREDICTION MODEL FOR THE AUTOMATIC DETECTION OF INDIVIDUALS AT RISK OF PSYCHOSIS USING ELECTRONIC HEALTH RECORDS

Stephen Puntis[1], Dominic Oliver[2], Paolo Fusar-Poli[2, 3, 4, 5]

[1] Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford, United Kingdom

[2]Early Psychosis: Interventions and Clinical-detection (EPIC) Lab, Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom

[3]National Institute for Health Research, Maudsley Biomedical Research Centre, South London and Maudsley National Health Service (NHS) Foundation Trust, London, United Kingdom;

[4]OASIS Service, South London and Maudsley National Health Service (NHS) Foundation Trust, London, United Kingdom

[5]Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy

Abstract: 249 words

Word count: 4183

Tables, Figures: 3

Corresponding author with email and address: Stephen.puntis@psych.ox.ac.uk

## Abstract

**Background:** Primary indicated prevention is a key target for reducing the incidence and burden of schizophrenia and related psychotic disorders. An individualised, clinically-based transdiagnostic model for the detection of individuals at risk of psychosis has been developed and validated in two large, urban healthcare providers. We tested its external validity in a geographically and demographically different non-urban population.

**Method:** Retrospective EHR cohort study. All individuals accessing secondary healthcare provided by Oxford Health NHS Foundation Trust between 1st January 2011 and 30th November 2019 and receiving a primary index diagnosis of a non-psychotic or non-organic mental disorder were considered eligible. The previously developed model was applied to this database and its external prognostic accuracy was measured with Harrell's C.

**Findings:** The study included n=33,710 eligible individuals, with an average age of 27.7 years (SD=19.8), mostly white (92.0%) and female (57.3%). The mean follow-up was 1863.9 days (SD=948.9), with 868 transitions to psychosis and a cumulative incidence of psychosis at 6 years of 2.9% (95%CI: 2.7-3.1). Compared to the urban development database, Oxford Health was characterised by a relevant case mix, lower incidence of psychosis, different distribution of baseline predictors, higher proportion of white females, and a lack of specialised clinical services for at risk individuals. Despite these differences the model retained an adequate prognostic performance (Harrell's C=0.79, 95%CI: 0.78-0.81), with no major miscalibration.

**Interpretation** The transdiagnostic, individualised, clinically-based risk calculator is transportable outside urban healthcare providers. Further research should test transportability of this risk prediction model in an international setting.

Keywords: Psychosis, Schizophrenia, Prevention, Detection, Electronic Health Records

# 1. INTRODUCTION

Schizophrenia and related psychotic disorders are large contributors to the global burden of disease (Mangalore and Knapp, 2007), and the impact on the individuals' quality of life, health, social functioning and education and employment can be severe, debilitating, and life-limiting (Marshall and Rathbone, 2011; Wiersma et al., 1998). Recent evidence suggests that the long-term prognosis for recovery of those with first episode of psychosis is better than previously thought, at an estimated 38% (Lally et al., 2017), but there has been little change in that rate over the last 70 years despite vast changes in the provision of mental health care and the delivery of treatment (Jääskeläinen et al., 2013). Recovery rates may not increase over time following a first episode of psychosis (FEP), with recovery rates the same at two years post-FEP as they are at six years post-FEP (Lally et al., 2017). This stasis suggests some utility in a focus on primary prevention of the disorder and early intervention in young individuals with subtle signs or symptoms of the disorder (termed as primary indicated prevention) (Fusar-Poli et al., 2019a). Identification of those at risk of psychosis could provide early and comprehensive treatment in order to prevent a psychotic illness or at least limit its progression (Fusar-Poli et al., 2017; McGorry and Mei, 2018; Millan et al., 2016; Nelson and McGorry, 2020).

A typical strategy for primary indicated prevention of psychosis is the creation of treatment pathways for those at Clinical High Risk for Psychosis (CHR-P) (McGorry et al., 1990; McGorry et al., 1996). A problem with only targeting CHR-P cases is that their identification is limited to those who are detected and referred to specialised CHR-P mental health services. Specifically, their detection is dependent on an accumulation of risk factors for psychosis that produce functional impairment and trigger help seeking behaviours (Falkenberg et al., 2015; Radua et al., 2018). The detection and recruitment of these samples is therefore non-systematic and selective (Fusar-Poli et al., 2016b; Fusar-Poli et al., 2019c). This is illustrated by the small number of CHR-P cases who transition to psychosis as a proportion of the total number of first episode psychosis cases treated in clinical services, with estimates as low as 5% (Fusar-Poli et al., 2017b) , and only as high as 20% (McGorry et al., 2020). Although the proportion of first episode psychoses detected through the CHR-P paradigm is currently modest, CHR-P features are present in the majority of those developing a first episode of psychosis. A first episode of psychosis occurring after a CHR-P stage is also similar to that observed in typical first episode psychosis patients (Fusar-Poli et al., 2016a; Sykes et al., 2020). While there have been significant efforts to better predict psychosis onset within CHR-P individuals (Cannon et al., 2016; Carrión et al., 2016; Osborne and Mittal, 2019; Zhang et al., 2018), these are dependent on prior adequate detection by CHR-P services. Therefore, there is a window of missed opportunity to improve the detection of individuals at risk for psychosis.

A previous study developed (n=33,820) and externally validated (n=54,716) a pragmatic individualised, clinically-based risk prediction tool for the detection of individuals at risk of psychosis in South London and Maudsley (SLaM) NHS Foundation Trust (Fusar-Poli et al., 2017b). This study demonstrated the clinical utility of expanding the risk detection beyond those with CHR-P to the entire population of individuals who present to mental health services (Fusar-Poli, 2019; Fusar-Poli et al., 2019). By using routinely collected health care data and routinely recorded predictors (age, gender, age*gender interaction, ethnicity and ICD-10 diagnosis) that are available in Electronic Health Records (EHRs), the prediction model showed adequate performance when externally validated (Harrell's C=0.79, 95%CI = 0.79-0.81) in detecting individuals at risk of psychosis over the following six years (Fusar-Poli et al., 2017b). Further to this, decision curve analysis showed that testing on the basis of

the transdiagnostic risk calculator may provide a net benefit compared to no test at all (Fusar-Poli et al., 2017b). This model has been externally validated a second time in a similar urban inner-city North London NHS healthcare provider (Camden and Islington, n=13,702), that was characterised by different sociodemographics and service configuration. The model retained its adequate prognostic performance (Harrell's C=0.73), suggesting the utility of using this prediction model within clinical services (Fusar-Poli et al., 2019d). Building on this evidence, the model has been prospectively implemented within SLaM where it was developed in order to pilot real-word usability (Oliver et al., 2020) . It is the only implementation study of risk prediction models in this field to date (Salazar de Pablo et al., 2020).

As the two previous validations were restricted to urban areas, the model's ability to maintain its performance in a different non-urban setting is unknown. Urbanicity has emerged as one of the strongest risk factors for psychosis in a recent umbrella review (Radua et al., 2018). The incidence of psychosis in urban and non-urban areas in developed countries can differ, with higher incidence of psychosis in urban settings which is often attributed to different accumulation of risk factors for the disorders, including demographic, social, and economic individual differences (Kirkbride et al., 2017; March et al., 2008; Richardson et al., 2018).

In this study we aimed to externally validate the individualised transdiagnostic risk calculator, investigating whether the model retains performance in a non-urban area which is geographically and demographically different to previously urban validation datasets.

# 2. METHODS

### 2.1 Design

We conducted a retrospective cohort study using routine EHRs. This study is in accordance with the REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement (Supplementary Table 1) (Benchimol et al., 2015).

### 2.2 Data source

We used the UK Clinical Record Interactive Search Tool (UK-CRIS) to access the Oxford Health NHS Foundation Trust (hereafter Oxford Health) electronic health records clinical register. UK-CRIS is a platform that provides a technological and governance model to allow researchers to access pseudonymised clinical records held in mental health NHS Trusts. The use of UK-CRIS for anonymised secondary data retrieval has been approved by the National Health Service Health Research Authority (HRA) and does not require individual study ethical approval. All UK-CRIS projects in Oxford Health are submitted to an independent CRIS Oversight Group for approval.

Oxford Health is the primary healthcare provider of both inpatient and outpatient mental health care in the counties of Oxfordshire and Buckinghamshire, England, serving a population of 1.2 million. The counties have both rural and urban areas with mostly lower deprivation than the national average, although Oxford city has pockets of very high deprivation (amongst the 20% most deprived in England). There is lower incidence of psychosis in Oxfordshire and Buckinghamshire in comparison to London (estimated with Psymaptic [http://www.psymaptic.org/] (Kirkbride et al., 2013), with crude incidence rates between 37–50 per 100,000 person-years in SLaM compared to 16–31 per 100,000 person-years in Oxfordshire and Buckinghamshire. Oxford Health also differs from SLaM in that it does not provide services for CHR-P individuals or substance misuse services.

### 2.3 Study population

Our eligible sample was all individuals who accessed secondary mental health care in Oxford Health between 1st January 2011 and 30th November 2019 and who received an index diagnosis of a non-organic, non-psychotic disorder (defined below).

### 2.4 Follow-up

Follow-up started at the time of the ICD-10 index diagnosis and ended when a transition to psychosis (see below) was recorded or at the end of study date.

### 2.5 Model Specification

Details of the population and data source for model development have been described previously (Fusar-Poli et al., 2019b; Fusar-Poli et al., 2017b; Fusar-Poli et al., 2019d). In brief, the transdiagnostic psychosis prediction risk calculator was developed and internally validated using retrospective routine secondary mental health care data in the SLaM. The model was developed initially in the SLaM boroughs of Lambeth and Southwark (n=33,820), firstly validated in the SLaM boroughs of Croydon and Lewisham (n=54,716) and secondly validated in Camden & Islington NHS Trust in London (n=13,702), England (Fusar-Poli et al., 2017b; Fusar-Poli et al., 2019d).

All individuals who received an index diagnosis of a non-organic or non-psychotic illness were eligible. A Cox survival model was used to predict the hazard ratio of developing a non-organic psychotic disorder over time (defined in Supplementary Table 2). It included the predictors of age (at the time of index diagnosis), gender, age by gender, ethnicity (categorised as in Supplementary Table 3), and cluster index diagnosis (defined in Supplementary Table 4). The latter included the ICD-10 clusters: acute and transient psychotic disorders (ATPD), bipolar mood disorders, non-bipolar mood disorders, anxiety disorders, personality disorders, developmental disorders, childhood/adolescence onset disorders, physiological syndromes, and mental retardation.

The first model was originally developed on a retrospective cohort and excluded cases with an onset of psychosis within the first three months to minimise the short-term diagnostic instability of baseline ICD-10 index diagnoses. In a subsequent implementation study (Oliver et al., 2020) an updated version of the model was adapted for prospective use (i.e. that did not exclude transitions occurring in the first three months). This model demonstrated similar prognostic performance, which can be seen in Supplementary Table 5. The revised model was developed in the original derivation dataset (SLaM boroughs Lambeth and Southwark, with the respective beta coefficients seen in Supplementary Table 6. Our validation study seeks to replicate the performance of this revised model.

### 2.6 Analysis

We conducted this study in accordance with the guidelines of Royston and Altman (Royston et al., 2010), Steyerberg and colleagues (Steyerberg et al., 2010), and the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) guidelines (Supplementary Table 7) (Collins et al., 2015).

All data collection and analyses were performed by a researcher independent of the model development team (SP). The development team had no access to the original data. The development team provided model coefficients (Supplementary Methods 1) and shared their analysis script, however all analyses were independently coded by SP, with the development team analysis script used only to compare coding for accuracy following completion of the analysis. The development team also shared summary sociodemographic data, the

Prognostic Index (PI: weighted sum of covariates with the model weights from the Cox model), and cumulative incidence data from the SLaM dataset for comparability purposes.

In order to interpret the performance of a risk model in the context of external validation, we first quantified the similarities between development and validation samples (Debray et al., 2015). We investigated external model transportability by the extent to which the SLaM and Oxford Health datasets comprised patients with sets of prognostically relevant predictors in common, comparable time-to-event outcomes with roughly similar follow-up times, and the same clinical condition observed in similar settings (Royston et al., 2010). As a first step, we described the Oxford Health patient population including the configuration of clinical services and compared it with the SLaM dataset. Baseline clinical and sociodemographic characteristics of the sample (including missing data) were described by means and standard deviations (SDs) for continuous variables, and absolute and relative frequencies for categorical variables. Differences between baseline continuous variables in Oxford Health and SLaM were assessed using independent sample t-tests; differences between categorical variables were assessed using Fisher's exact tests. As per the development cohort, participants with missing data were excluded prior to analysis.

We then visually compared the Kaplan–Meier failure functions of the Oxford Health and SLaM datasets. The overall cumulative risk of psychosis onset in Oxford Health was visualised with the Kaplan–Meier failure function (1—survival) and 95% confidence intervals (CIs). Curves that vary noticeably may indicate systematic differences within the study populations (Royston et al., 2010).

We calculated and compared the spread (SD) and the mean of the PI in the two datasets. An increased (or decreased) variability of the PI would indicate more (or less) heterogeneity of case mix between the two datasets, and therefore, of their overarching target populations (Debray et al., 2015). Differences in the mean PI would indicate differences in the overall (predicted) outcome frequency, reflecting case-mix severity between the two (Debray et al., 2015).

We then performed the formal external validation. We calculated the predicted probabilities for each participant in the Oxford Health dataset from the regression coefficients obtained from the model developed in the SLaM dataset through the application of the PI (see Supplementary Table 6). We determined discrimination of the external model (the ability to discriminate between those with and those without the outcome) using the Harrell's C-index (Harrell Jr, 2015), which is the probability that for a random pair of "case" and "control," the predicted risk of an event (PI) is higher for the "case". In interpreting Harrell's C statistic, guidance suggests values of 0.9–1.0 are considered outstanding, 0.8–0.9 excellent and 0.7–0.8 acceptable. We estimated overall model performance using the Brier score (the average mean squared difference between predicted probabilities and actual outcomes, which also captures calibration and discrimination aspects), the estimates of which range between 0 (most accurate) and 1 (least accurate) (Steyerberg et al., 2010). Calibration (the agreement between observed outcomes and predictions) was assessed using the regression slope of the PI (Royston et al., 2010). We then updated the model using the regression slope of the PI as a shrinkage factor for recalibration, in line with the Royston and colleagues guidelines (Royston et al., 2010).

In a sensitivity analysis, we validated the original development model (not adapted for prospective use, i.e. excluding individuals who developed a psychotic illness within three months of their index diagnosis).

We used R version 3.5.0 for all cleaning and analysis of data (R Core Team, 2013). Significance was set to $P$ <.05.

# 3. RESULTS

## 3.1 Sample characteristics of the Oxford Health cohort

We identified 65,278 individuals accessing Oxford Health NHS Foundation Trust between 1st January 2011 and 30th November 2019 who received an index ICD-10 diagnosis. After excluding those with an index diagnosis of a psychotic disorder or an organic psychiatric disorder there were 41,167 eligible participants (flow diagram in Figure 1). After excluding individuals with missing data, the final sample included in the study was n = 33,710. Patients included in the current study had an average age of 27.70 years (SD=19.76; median = 17.7, IQR 14.2; 39.0), 57.3% were female, and white ethnicity was particularly common (92.0%). The most frequent index diagnoses were non-bipolar mood disorders (25.9%). Table 1 outlines the demographic and clinical characteristics of the sample.

## 3.2 Differences between Oxford Health and SLaM databases

### *3.2.1 Sociodemographic and service configuration differences*

Individuals in Oxford Health were younger than their SLaM counterparts (a mean age of 27.70 years vs 34.43 years, p<0.001), more were female (57.3% vs 48.8%, p<0.001), and there were more people of white ethnicity (92.0% vs 59.0%, Table 1). As expected, given the sociodemographic and service differences across the two databases the incidence of ATPD, substance use disorders, and anxiety disorders all appeared lower in Oxford Health in comparison to SLaM. However, there seem to be higher proportions of bipolar mood disorders, nonbipolar mood disorders, personality disorders, developmental disorders, childhood/adolescent disorders and physiological syndromes in Oxford Health compared to SLaM (Table 1). Oxford Health also had more missing ethnicity data (18.1% vs 7.1%). The most important difference is that while SLaM included CHR-P services, these were not available in Oxford Health. Similarly SLaM provides substance use disorder treatment while Oxford Health does not. Furthermore, the incidence of psychosis in the general population covered by Oxford Health is lower than in SLaM (the urban area of South London is characterised by one of the highest incidences of psychosis worldwide) (Kirkbride et al., 2013).

### *3.2.2 Cumulative Risk of Psychosis*

The mean follow up time in the Oxford Health database was 1863.85 days with a standard deviation of 948.99 (median = 2006.0, IQR 1071.0; 2663.0) compared to 1580.64 days (SD=927.72) in SLaM. There were 868 transitions to psychosis in the Oxford Health database compared to 1,273 in SLaM. The mean number of days from index diagnosis to transition to psychosis in the Oxford Health database was 741.20 days (SD=722.87; median = 725, IQR 335; 1296), compared to 664.03 days (SD=621.04) in SLaM. The 2-year cumulative risk of psychosis in the Oxford Health database was 1.63 (95%CI: 1.49-1.77), with the last transition being observed at 3282 days), compared to 2.57 (95%CI: 2.40-2.75), with the last transition being observed at 3246 days in SLaM. The cumulative incidences curves (1-Kaplan–Meier) plotted alongside the cumulative risk in the SLaM database is depicted in Figure 2. The mean PI in Oxford Health was -1.66 (SD=0.96) in comparison to −1.18 (SD=0.94) in SLaM (p<0.001).

## 3.3 External validation in Oxford Health

The model was predicting significantly better than chance in Oxford Health with a Harrell's C of 0.79 (95%CI: 0.78-0.81; Harrell's C in SLaM=0.79). The 2-year Brier score was 0.010 and 6-year Brier score was 0.019 (2-year Brier score in SLaM=0.012; 6-year Brier score in

SLaM=0.027). The model did not show major calibration issues, with a regression slope close to 1: 0.996 (95%CI: 0.945-1.048 p<0.001).

Recalibrating the model, using the regression slope of the PI as an offset term, resulted in no change in model performance, with the recalibrated PI mean of -1.65 (SD=0.96), and Harrell's C of 0.79 (95%CI: 0.78-0.81).

In the sensitivity analysis, which excluded those who transitioned to psychosis within 90 days of their index diagnosis (as per the original development model, final sample size = 33,543), there were 701 transitions to psychosis. The mean PI was -1.79 (SD=0.91), Harrell's C 0.78 (95%CI: 0.76-0.80), the 2-year Brier score was 0.006 and 6-year Brier score was 0.014, and the model calibration slope = 0.926 (95%CI: 0.870-0.982, P<0.001).

# 4. DISCUSSION

This is the third external validation of a transdiagnostic, clinically-based, individualized risk calculator for psychosis and the first study to test the potential transportability of this prediction model outside an urban area. Compared to the original development urban database, Oxford Health was characterised by a relevant case mix; but with differences in a lower proportion of transitions to psychosis, a lower proportion of ATPD, substance use disorders (there were no specialised substance misuse services), and anxiety disorders and higher proportions of bipolar mood disorders, nonbipolar mood disorders, personality disorders, developmental disorders, childhood/adolescent disorders and physiological syndromes. It had a higher proportion of females and those of white ethnicity in comparison to SLaM. Despite these differences, the prognostic model retained an adequate prognostic performance (Harrell's C = 0.79; 95%CI: 0.78-0.81), with no major miscalibration issues.

This study provides the third independent external replication of a prediction model for the detection of individuals at risk of psychosis in secondary car; secondary care represents the source of 75% of all referrals to FEP services (Fusar-Poli et al., In press). Furthermore, the external prognostic accuracy of this model aligns with several other models developed in this field, several of which are more complex and include domains difficult to access at a wider scale (e.g. neuroimaging or peripheral biomarkers)(Sanfelici et al., 2020; Worthington et al., 2019). This step represents an important accomplishment in the field of prediction modelling in psychiatry. Clinical psychiatry is currently affected by an important replication crisis (Szucs and Ioannidis, 2017), to the point that external replications become as equally, or more, important than original and new discoveries (Ioannidis, 2006). In line with these findings, a systematic review and meta-analysis of clinical prediction models for predicting the onset of psychosis in individuals at risk found no studies across the 91 reviewed that performed a true external validation of an existing model (Studerus et al., 2017).

Oxford Health NHS Foundation Trust boundaries cover both urban and rural space and differs greatly in in density in comparison to London's highly urban geography. There is large variation in psychosis risk between urban and rural environments with excess rates in urban cities (Kirkbride et al., 2017) that potentially contribute to the lower incidence of psychosis in Oxford Health (2.9% risk of psychosis at 6 years in Oxford Health vs 4.79% at 6 years in SLaM). Differences in incidence of psychosis were confirmed by a lower mean of PI in Oxford Health vs SLaM. Oxford Health also provides treatment to a very different population, with less ethnic diversity, higher socioeconomic status, and a higher proportion of females than males attending mental health services, all factors that may reduce the incidence of psychosis. The development model in SLaM was trained on a sample of individuals from an urban city area with a far higher proportion of black males while the Oxford Health validation dataset was largely white and majority female. The higher incidence of psychosis in black

Caribbean and African migrants and their descendants in higher income countries has been widely published, an increased risk which is independent of demographic and socioeconomic differences but still poorly understood (Jones and Jongsma, 2020; Morgan et al., 2019; Tortelli et al., 2015). There are also robust findings of a higher incidence of schizophrenia and psychoses in males (Van der Werf et al., 2014) that also likely contribute to the differences we found, although this risk changes over time with the risk in females higher than males in adolescence and early years (Dalsgaard et al., 2019). One would expect poorer performance of the model in Oxford Health due to these differences, but the prediction model retained adequate performance, comparable to that of the development model. This suggests that the model is not wholly dependent on gender and ethnicity parameters.

Similarly, there were profound diagnostic differences between samples. Oxford Health does not provide a CHR-P service and therefore had no index diagnoses of CHR-P, additionally having fewer index cases of ATPD, which is a strong predictor of later development of persistent psychotic disorders (Fusar-Poli et al., 2017a; Minichino et al., 2019; Rutigliano et al., 2018) and this did not seem to affect model performance which suggests that the prediction model is not driven by those psychosis-adjacent diagnoses. There was also substantially fewer index diagnoses of substance misuse, likely due to the fact that Oxford Health does not provide substance misuse services. Rates of substance misuse (another key factor which substantially increases psychosis risk (Moore et al., 2007) in inner-city, socially deprived areas are higher than those in more rural, more affluent areas like Oxfordshire and Buckinghamshire (Department of Health, 2018). At the same time there were more diagnoses of affective disorders (bipolar and non-bipolar) and neurodevelopmental disorders. Given the transdiagnostic nature of this model, the high frequency of these diagnoses may have played a significant role in increasing the risk of psychosis in this cohort. Furthermore, the relative higher frequency of affective and neurodevelopmental disorders versus ATPD or substance misuse disorders may explain the longer time to transition observed in Oxford Health compared to SLaM (Murrie et al., 2019). Overall, the differences in case-mix and incidence of psychosis led to a similar variance in predictions (the SD of the PI was similar across Oxford Health and SLaM) and the model retained its performance. The differences in case-mix, highlights the transportability of the transdiagnostic risk calculator to healthcare providers with differences in service configuration, such as a lack of CHR-P services. This adds further evidence to replications in Camden & Islington NHS Foundation Trust (Fusar-Poli et al., 2019d) and an international commercial insurance dataset (Oliver et al., 2020). Together, this suggests that the transdiagnostic risk calculator could be a valuable tool in improving detection of individuals at risk for psychosis in areas where CHR-P service provision is not available. The simplicity of the model has allowed for ease of implementation, with its increasing automation leading to greater cost-effectiveness, which is essential within under-funded and under-resourced mental health services, which have negatively impacted perceptions of early intervention (McGorry and Mei, 2020; Woods et al., 2020).

The limitations of using a retrospective cohort from routine clinical data have been discussed previously (Fusar-Poli et al., 2017b). The lack of validated structured psychometric interviews for the diagnosis means that while our diagnostic criteria have high ecological validity (i.e. they represent real world practice), they have not been subjected to formal validation. Second, both development and validation models excluded participants with missing data. The missing data in the Oxford Health dataset was substantial (18%) and unlikely to be missing completely at random, which may have produced biased estimates. Finally, the ‘gold standard’ validation of a prediction model requires a completely independent study team in order to reduce bias, as non-independent validations often result

in overall more optimistic results (Collins et al., 2014). We attempted to minimize this by ensuring that data collection and analysis were performed independently from the original development study team, and not allowing the development study team access to the original data.

### 4.1 Conclusions

This validation study has demonstrated the reproducibility and transportability of an individualised, clinically-based transdiagnostic model for the automatic screening of EHRs and the detection of individuals at risk of psychosis. Despite large differences in case-mix and urban and rural geography, the model retained its performance and was well calibrated, suggesting the risk calculator would be appropriate for use throughout the United Kingdom in the detection of at-risk cases in secondary mental health care. Further testing of model transportability to mental health care providers internationally is required.

## 5. Funding

S.P. is funded by a National Institute for Health Research Post Doctoral Fellowship award (grant number PDF-2017-10-029). This study was supported by the UK Clinical Record Interactive Search (UK-CRIS) system using data and systems of the NIHR Oxford Health Biomedical Research Centre (BRC-1215-20005). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. This study was also supported by the King's College London Confidence in Concept award from the Medical Research Council (MRC) (MC_PC_16048) to PF-P. DO is supported by the UK Medical Research Council (MR/N013700/1) and King’s College London member of the MRC Doctoral Training Partnership in Biomedical Sciences.

## 6. Acknowledgements

PF-P has received research fees from Lundbeck and advisory consultancy fees from Lundbeck, Menarini and Angelini outside of this work.

We would like to acknowledge the work and support of the Oxford CRIS Team, Tanya Smith, Adam Pill, Suzanne Fisher, and Lulu Kane.

## Tables and Figures

Figure 1: Flow chart of the study population

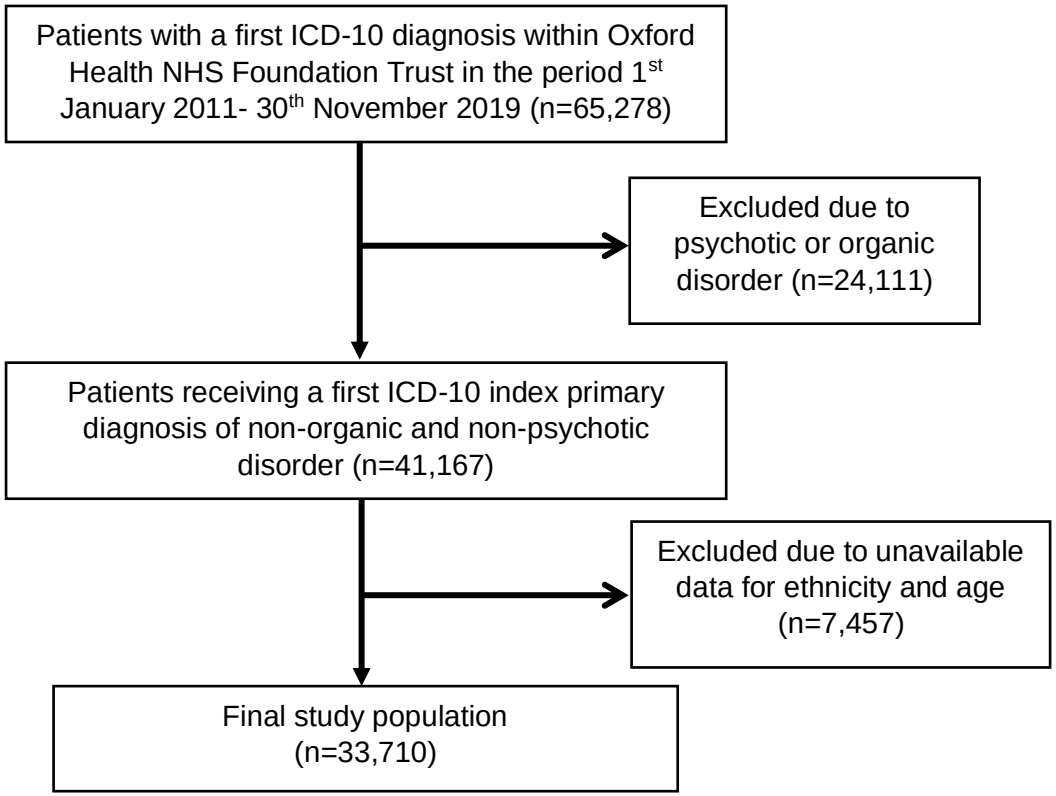

**Table 1. Demographics and index diagnoses comparison between South London and Maudsley (SLaM) and Oxford Health datasets**

| Variable | SLaM (n=34,209)[a] | Oxford Health (n=33,710) | p-value |
|---|---|---|---|
| | **Mean (SD)** | **Mean (SD)** | |
| **Age**, years | 34.43 (18.89) | 27.7 (19.8) | <0.001 |
| | **No. (%)** | **No. (%)** | |
| **Sex** | - | - | <0.001 |
| Male | 17,511 (51.20) | 14,397 (42.7) | |
| Female | 16,688 (48.80) | 19,313 (57.3) | |
| **Ethnicity** | - | - | <0.001 |
| Black | 7,055 (22.19) | 341 (1.0) | |
| White | 18,768 (59.03) | 31,015 (92.0) | |
| Asian | 1,149 (3.61) | 925 (2.7) | |
| Mixed | 1,319 (4.15) | 1,107 (3.3) | |
| Other | 3,502 (11.02) | 322 (1.0) | |
| **Index diagnosis** | - | - | <0.001 |
| CHR-P | 314 (0.92) | - | |
| ATPD[b] | 747 (2.18) | 357 (1.1) | |
| Substance use disorders | 7,187 (21.01) | 734 (1.78) | |
| Bipolar mood disorders | 980 (2.86) | 1,816 (5.4) | |
| Nonbipolar mood disorders | 6,364 (18.60) | 8,719 (25.9) | |
| Anxiety disorders | 8,279 (24.20) | 7,311 (21.7) | |
| Personality disorders | 1,297 (3.79) | 1,873 (5.6) | |
| Developmental disorders | 1,413 (4.13) | 3,747 (11.1) | |
| Childhood/adolescence onset disorders | 4,201 (12.28) | 4,947 (14.7) | |
| Physiological syndromes | 2,560 (7.48) | 3,601 (10.7) | |
| Mental retardation | 867 (2.53) | 853 (2.5) | |

[a]South London and Maudsley NHS Foundation Trust
[b]Acute and Transient Psychotic Disorders

**Figure 2: Cumulative incidence for the risk of development of psychotic disorders in the Oxford Health (left) and SLaM (right) databases**

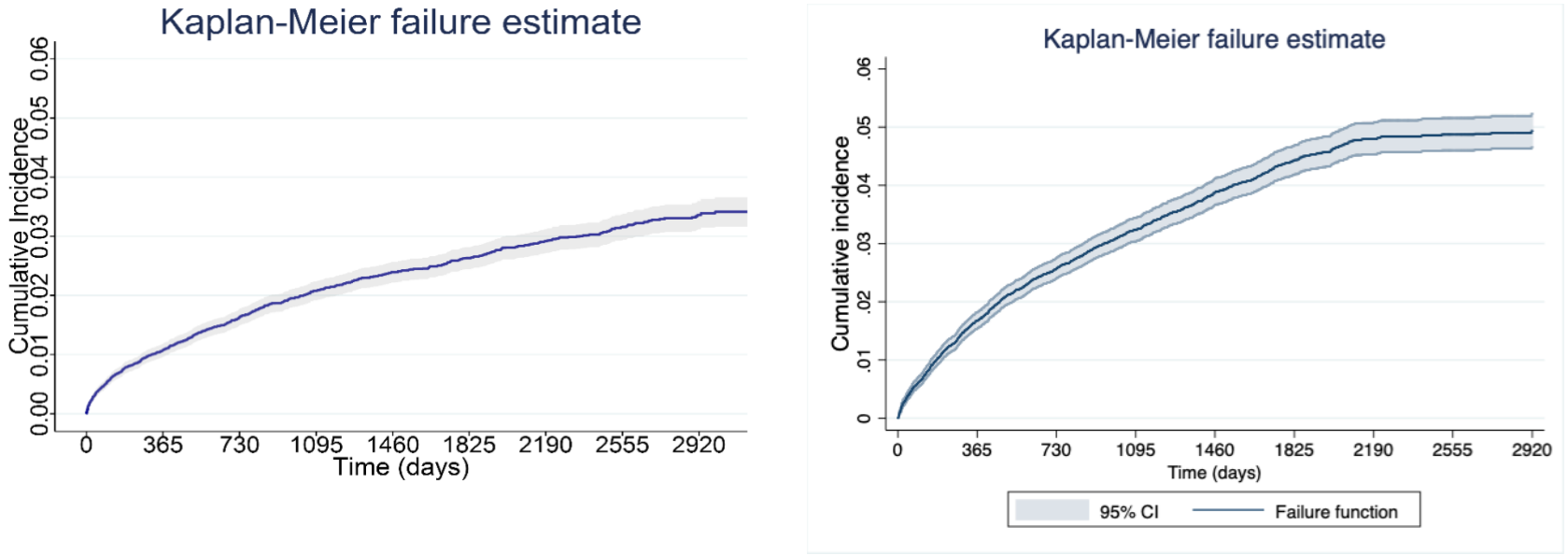


**Figure 2** Left figure: cumulative incidence (Kaplan–Meier failure function) for risk of development of psychotic disorders in the Oxford Health Database. There were a total of 868 events (transition to psychosis): 356 in the first 365 days, 168 in the interval 366–730 days, 121 in the interval 731–1095 days, 74 in the interval 1095–1460 days, 50 in the interval 1460–1825 days, 50 in the interval 1825–2190 days, 30 in the interval 2190–2555 days, 16 in the interval 2555-2920, 3 in the interval 2920-2997 days (end of follow-up). The last event was observed at 2997 days, when 4873 individuals were still at risk. The cumulative incidence of psychosis was: 1.1 (95%CI: 1.0-1.2, 31233 individuals at risk) at 1 year , 1.6 (95%CI: 1.5-1.8, 28178 individuals at risk) at 2 years, 2.1 (95%CI: 1.9-2.2, 24931 individuals at risk) at 3 years, 2.4 (95%CI: 2.2-2.6, 21688 individuals at risk) at 4 years, 2.6 (95%CI: 2.4-2.8, 18759) at 5 years, 2.9 (95%CI: 2.7-3.1, 14914) at 6 years, 3.2 (95%CI: 2.9-3.4, 10041 individuals at risk) at 7 years, 3.4 (95%CI: 3.1-3.6, 5551 individuals at risk) at 8 years . Right figure: cumulative incidence in the SLaM derivation database, truncated at 2920 days for visual comparability. Cumulative incidence of psychosis: 1.67 (95%CI: 1.61-1.89, 30102 individuals still at risk) at 1 year, 2.57 (95%CI: 2.40-2.75, 26337 individuals still at risk) at 2 years, 3.88%(95%CI 3.66-4.12, 18285 individuals still at risk) at 3 years, 4.42% (95%CI 4.18-4.68, 14091 individuals still at risk) at 5 years, 4.79% (95%CI 4.53-5.07, 9590 individuals still at risk) at 6 years, 4.87%(95%CI 4.60-5.16, 6626 individuals at risk) at 7 years, 4.93 (95%CI 4.65-5.23, 9590 individuals still at risk) at 8 years.

# 7. References


Benchimol, E.I., Smeeth, L., Guttmann, A., Harron, K., Moher, D., Petersen, I., Sørensen, H.T., von Elm, E., Langan, S.M., Committee, R.W., 2015. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Med 12(10), e1001885.

Cannon, T.D., Yu, C., Addington, J., Bearden, C.E., Cadenhead, K.S., Cornblatt, B.A., Heinssen, R., Jeffries, C.D., Mathalon, D.H., McGlashan, T.H., 2016. An individualized risk calculator for research in prodromal psychosis. American Journal of Psychiatry 173(10), 980-988.

Carrión, R.E., Cornblatt, B.A., Burton, C.Z., Tso, I.F., Auther, A.M., Adelsheim, S., Calkins, R., Carter, C.S., Niendam, T., Sale, T.G., 2016. Personalized prediction of psychosis: external validation of the NAPLS-2 psychosis risk calculator with the EDIPPP project. American Journal of Psychiatry 173(10), 989-996.

Collins, G.S., de Groot, J.A., Dutton, S., Omar, O., Shanyinde, M., Tajar, A., Voysey, M., Wharton, R., Yu, L.-M., Moons, K.G., 2014. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. BMC medical research methodology 14(1), 40.

Collins, G.S., Reitsma, J.B., Altman, D.G., Moons, K.G., 2015. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. British Journal of Surgery 102(3), 148-158.

Dalsgaard, S., Thorsteinsson, E., Trabjerg, B.B., Schullehner, J., Plana-Ripoll, O., Brikell, I., Wimberley, T., Thygesen, M., Madsen, K.B., Timmerman, A., 2019. Incidence Rates and Cumulative Incidences of the Full Spectrum of Diagnosed Mental Disorders in Childhood and Adolescence. JAMA psychiatry.

Debray, T.P., Vergouwe, Y., Koffijberg, H., Nieboer, D., Steyerberg, E.W., Moons, K.G., 2015. A new framework to enhance the interpretation of external validation studies of clinical prediction models. Journal of clinical epidemiology 68(3), 279-289.

Department of Health, 2018. Statistics on Drugs Misuse England, 2018. Dandy Booksellers Limited.

Falkenberg, I., Valmaggia, L., Byrnes, M., Frascarelli, M., Jones, C., Rocchetti, M., Straube, B., Badger, S., McGuire, P., Fusar-Poli, P., 2015. Why are help-seeking subjects at ultra-high risk for psychosis help-seeking? Psychiatry research 228(3), 808-815.

Fusar-Poli, P., Bauer, M., Borgwardt, S., Bechdolf, A., Correll, C.U., Do, K.Q., Domschke, K., Galderisi, S., Kessing, L.V., Koutsouleris, N., 2019a. European college of neuropsychopharmacology network on the prevention of mental disorders and mental health promotion (ECNP PMD-MHP). European Neuropsychopharmacology 29(12), 1301-1311.

Fusar-Poli, P., Cappucciati, M., De Micheli, A., Rutigliano, G., Bonoldi, I., Tognin, S., Ramella-Cravaro, V., Castagnini, A., McGuire, P., 2017a. Diagnostic and prognostic significance of brief limited intermittent psychotic symptoms (BLIPS) in individuals at ultra high risk. Schizophrenia bulletin 43(1), 48-56.

Fusar-Poli, P., Lai, S., Di Forti, M., Iacopani, E., Thornicroft, G., McGuire, P., Jauhar, S., In press. Early intervention services for first episode of psychosis in south london and the maudsley: twenty years of care and research for young people. Frontiers in Psychiatry.

Fusar-Poli, P., Oliver, D., Spada, G., Patel, R., Stewart, R., Dobson, R., McGuire, P., 2019b. Real-world Implementation of a Transdiagnostic Risk Calculator for the Automatic Detection of Individuals at Risk of Psychosis in Clinical Routine: Study Protocol. Frontiers in Psychiatry 10, 109.

Fusar-Poli, P., Rutigliano, G., Stahl, D., Davies, C., Bonoldi, I., Reilly, T., McGuire, P., 2017b. Development and validation of a clinically based risk calculator for the transdiagnostic prediction of psychosis. JAMA psychiatry 74(5), 493-500.

Fusar-Poli, P., Rutigliano, G., Stahl, D., Schmidt, A., Ramella-Cravaro, V., Hitesh, S., McGuire, P., 2016a. Deconstructing pretest risk enrichment to optimize prediction of psychosis in individuals at clinical high risk. JAMA psychiatry 73(12), 1260-1267.
Fusar-Poli, P., Schultze-Lutter, F., Cappucciati, M., Rutigliano, G., Bonoldi, I., Stahl, D., Borgwardt, S., Riecher-Rössler, A., Addington, J., Perkins, D.O., 2016b. The dark side of the moon: meta-analytical impact of recruitment strategies on risk enrichment in the clinical high risk state for psychosis. Schizophrenia bulletin 42(3), 732-743.
Fusar-Poli, P., Sullivan, S., Shah, J., Uhlhaas, P., 2019c. Improving the detection of individuals at clinical risk for psychosis in the community, primary and secondary care: an integrated evidence-based approach. Frontiers in psychiatry 10, 774.
Fusar-Poli, P., Werbeloff, N., Rutigliano, G., Oliver, D., Davies, C., Stahl, D., McGuire, P., Osborn, D., 2019d. Transdiagnostic risk calculator for the automatic detection of individuals at risk and the prediction of psychosis: second replication in an independent national health service trust. Schizophrenia bulletin 45(3), 562-570.
Fusar-Poli, P., 2019. TRANSD recommendations: improving transdiagnostic research in psychiatry. World Psychiatry 18(3), 361.
Fusar-Poli, P., McGorry, P.D., Kane, J.M., 2017. Improving outcomes of first-episode psychosis: an overview. World psychiatry 16(3), 251-265.
Fusar-Poli, P., Solmi, M., Brondino, N., Davies, C., Chae, C., Politi, P., Borgwardt, S., Lawrie, S.M., Parnas, J., McGuire, P., 2019. Transdiagnostic psychiatry: a systematic review. World Psychiatry 18(2), 192-207.
Harrell Jr, F.E., 2015. Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis. Springer.
Ioannidis, J.P., 2006. Evolution and translation of research findings: from bench to where? PLoS clinical trials 1(7).
Jääskeläinen, E., Juola, P., Hirvonen, N., McGrath, J.J., Saha, S., Isohanni, M., Veijola, J., Miettunen, J., 2013. A systematic review and meta-analysis of recovery in schizophrenia. Schizophrenia bulletin 39(6), 1296-1306.
Jones, P., Jongsma, H., 2020. Social disadvantage, linguistic distance, ethnic minority status and first episode psychosis: results from the EU-GEI case-control study.
Kirkbride, J.B., Hameed, Y., Ankireddypalli, G., Ioannidis, K., Crane, C.M., Nasir, M., Kabacs, N., Metastasio, A., Jenkins, O., Espandian, A., 2017. The epidemiology of first-episode psychosis in early intervention in psychosis services: findings from the social epidemiology of psychoses in East Anglia [SEPEA] study. American Journal of Psychiatry 174(2), 143-153.
Kirkbride, J.B., Jackson, D., Perez, J., Fowler, D., Winton, F., Coid, J.W., Murray, R.M., Jones, P.B., 2013. A population-level prediction tool for the incidence of first-episode psychosis: translational epidemiology based on cross-sectional data. BMJ open 3(2), e001998.
Lally, J., Ajnakina, O., Stubbs, B., Cullinane, M., Murphy, K.C., Gaughran, F., Murray, R.M., 2017. Remission and recovery from first-episode psychosis in adults: systematic review and meta-analysis of long-term outcome studies. The British Journal of Psychiatry 211(6), 350-358.
Mangalore, R., Knapp, M., 2007. Cost of schizophrenia in England. The journal of mental health policy and economics 10(1), 23-41.
March, D., Hatch, S.L., Morgan, C., Kirkbride, J.B., Bresnahan, M., Fearon, P., Susser, E., 2008. Psychosis and place. Epidemiologic reviews 30(1), 84-100.
Marshall, M., Rathbone, J., 2011. Early intervention for psychosis. Cochrane Database of Systematic Reviews(6).
McGorry, P.D., Copolov, D.L., Singh, B.S., 1990. Current concepts in functional psychosis: the case for a loosening of associations. Schizophrenia research 3(4), 221-234.
McGorry, P.D., Edwards, J., Mihalopoulos, C., Harrigan, S.M., Jackson, H.J., 1996. EPPIC: an evolving system of early detection and optimal management. Schizophrenia bulletin 22(2), 305-326.

McGorry, P.D., Mei, C., 2018. Ultra-high-risk paradigm: lessons learnt and new directions. Evidence-Based Mental Health 21(4), 131-133.
McGorry, P.D., Mei, C., 2020. Why do psychiatrists doubt the value of early intervention? The power of illusion. Australasian Psychiatry, 1039856220924323.
McGorry, P.D., Nelson, B., Wood, S.J., Shah, J.L., Malla, A., Yung, A., 2020. Transcending false dichotomies and diagnostic silos to reduce disease burden in mental disorders. Social psychiatry and psychiatric epidemiology, 1-9.
Millan, M.J., Andrieux, A., Bartzokis, G., Cadenhead, K., Dazzan, P., Fusar-Poli, P., Gallinat, J., Giedd, J., Grayson, D.R., Heinrichs, M., 2016. Altering the course of schizophrenia: progress and perspectives. Nature Reviews Drug Discovery 15(7), 485.
Minichino, A., Rutigliano, G., Merlino, S., Davies, C., Oliver, D., De Micheli, A., Patel, R., McGuire, P., Fusar-Poli, P., 2019. Unmet needs in patients with brief psychotic disorders: Too ill for clinical high risk services and not ill enough for first episode services. European Psychiatry 57, 26-32.
Moore, T.H., Zammit, S., Lingford-Hughes, A., Barnes, T.R., Jones, P.B., Burke, M., Lewis, G., 2007. Cannabis use and risk of psychotic or affective mental health outcomes: a systematic review. The Lancet 370(9584), 319-328.
Morgan, C., Knowles, G., Hutchinson, G., 2019. Migration, ethnicity and psychoses: evidence, models and future directions. World Psychiatry 18(3), 247-258.
Murrie, B., Lappin, J., Large, M., Sara, G., 2019. Transition of Substance-Induced, Brief, and Atypical Psychoses to Schizophrenia: A Systematic Review and Meta-analysis. Schizophrenia Bulletin.
Nelson, B., McGorry, P., 2020. The Prodrome of Psychotic Disorders: Identification, Prediction, and Preventive Treatment. Child and Adolescent Psychiatric Clinics 29(1), 57-69.
Oliver, D., Spada, G., Colling, C., Broadbent, M., Baldwin, H., Patel, R., Stewart, R., Stahl, D., Dobson, R., McGuire, P., 2020. Real-world implementation of precision psychiatry: transdiagnostic risk calculator for the automatic detection of individuals at-risk of psychosis. Schizophrenia research. DOI: 10.1016/j.schres.2020.05.007
Osborne, K.J., Mittal, V.A., 2019. External validation and extension of the NAPLS-2 and SIPS-RC personalized risk calculators in an independent clinical high-risk sample. Psychiatry research 279, 9-14.
R Core Team, 2013. R: A language and environment for statistical computing.
Radua, J., Ramella-Cravaro, V., Ioannidis, J.P., Reichenberg, A., Phiphopthatsanee, N., Amir, T., Yenn Thoo, H., Oliver, D., Davies, C., Morgan, C., 2018. What causes psychosis? An umbrella review of risk and protective factors. World Psychiatry 17(1), 49-66.
Richardson, L., Hameed, Y., Perez, J., Jones, P.B., Kirkbride, J.B., 2018. Association of environment with the risk of developing psychotic disorders in rural populations: findings from the social epidemiology of psychoses in East Anglia study. JAMA psychiatry 75(1), 75-83.
Royston, P., Parmar, M.K., Altman, D.G., 2010. External validation and updating of a prognostic survival model. Hub for Trials Methodology Research, MRC Clinical Trials Unit and University College London 222.
Rutigliano, G., Merlino, S., Minichino, A., Patel, R., Davies, C., Oliver, D., De Micheli, A., McGuire, P., Fusar-Poli, P., 2018. Long term outcomes of acute and transient psychotic disorders: The missed opportunity of preventive interventions. European Psychiatry 52, 126-133.
Salazar de Pablo, G., Studerus, E., Vaquerizo-Serrano, J., Irving, J., Catalan, A., Oliver, D., Baldwin, H., Danese, A., Fazel, S., Steyerberg, E.W., 2020. Implementing Precision Psychiatry: A Systematic Review of Individualized Prediction Models for Clinical Practice. Schizophrenia Bulletin.
Sanfelici, R., Dwyer, D.B., Antonucci, L.A., Koutsouleris, N., 2020. Individualized diagnostic and prognostic models for patients with psychosis risk syndromes: A meta-analytic view on the state-of-the-art. Biological Psychiatry.
Steyerberg, E.W., Vickers, A.J., Cook, N.R., Gerds, T., Gonen, M., Obuchowski, N., Pencina, M.J., Kattan, M.W., 2010. Assessing the performance of prediction models: a

framework for some traditional and novel measures. Epidemiology (Cambridge, Mass.) 21(1), 128.
Studerus, E., Ramyead, A., Riecher-Rössler, A., 2017. Prediction of transition to psychosis in patients with a clinical high risk for psychosis: a systematic review of methodology and reporting. Psychological medicine 47(7), 1163-1178.
Sykes, L.A.Y., Ferrara, M., Addington, J., Bearden, C.E., Cadenhead, K.S., Cannon, T.D., Cornblatt, B.A., Perkins, D.O., Mathalon, D.H., Seidman, L.J., 2020. Predictive validity of conversion from the clinical high risk syndrome to frank psychosis. Schizophrenia research 216, 184-191.
Szucs, D., Ioannidis, J.P., 2017. Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature. PLoS biology 15(3).
Tortelli, A., Errazuriz, A., Croudace, T., Morgan, C., Murray, R., Jones, P., Szoke, A., Kirkbride, J., 2015. Schizophrenia and other psychotic disorders in Caribbean-born migrants and their descendants in England: systematic review and meta-analysis of incidence rates, 1950–2013. Social psychiatry and psychiatric epidemiology 50(7), 1039-1055.
Van der Werf, M., Hanssen, M., Köhler, S., Verkaaik, M., Verhey, F., van Winkel, R., Van Os, J., Allardyce, J., Investigators, R., 2014. Systematic review and collaborative recalculation of 133 693 incident cases of schizophrenia. Psychological medicine 44(1), 9-16.
Wiersma, D., Nienhuis, F.J., Slooff, C.J., Giel, R., 1998. Natural course of schizophrenic disorders: a 15-year followup of a Dutch incidence cohort. Schizophrenia bulletin 24(1), 75-85.
Woods, S.W., Bearden, C.E., Sabb, F.W., Stone, W.S., Torous, J., Cornblatt, B.A., Perkins, D.O., Cadenhead, K.S., Addington, J., Powers III, A.R., 2020. Counterpoint. Early intervention for psychosis risk syndromes: Minimizing risk and maximizing benefit. Schizophrenia research.
Worthington, M.A., Cao, H., Cannon, T.D., 2019. Discovery and Validation of Prediction Algorithms for Psychosis in Youths at Clinical High Risk. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging.
Zhang, T., Li, H., Tang, Y., Niznikiewicz, M.A., Shenton, M.E., Keshavan, M.S., Stone, W.S., McCarley, R.W., Seidman, L.J., Wang, J., 2018. Validating the predictive accuracy of the NAPLS-2 psychosis risk calculator in a clinical high-risk sample from the SHARP (Shanghai At Risk for Psychosis) program. American Journal of Psychiatry 175(9), 906-908.