Improving prediction of psychosis risk

Oliver, Dom

Awarding institution:  
King’s College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT

Unless another licence is stated on the immediately following page this work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence. https://creativecommons.org/licenses/by-nc-nd/4.0/

You are free to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Improving prediction of psychosis risk

Dominic Oliver

Thesis submitted for the degree of
Doctor of Philosophy

Institute of Psychiatry, Psychology & Neuroscience
King’s College London

2020
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>5</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>7</td>
</tr>
<tr>
<td>THESIS OUTLINE</td>
<td>9</td>
</tr>
<tr>
<td>STATEMENT OF PERSONAL CONTRIBUTION</td>
<td>10</td>
</tr>
<tr>
<td>PREFACE</td>
<td>11</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>13</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>15</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>16</td>
</tr>
<tr>
<td>PART A: IMPROVING DETECTION OF INDIVIDUALS AT RISK OF DEVELOPING PSYCHOSIS</td>
<td>17</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>18</td>
</tr>
<tr>
<td>1.1. PRIMARY INDICATED PREVENTION AND THE CLINICAL HIGH RISK STATE</td>
<td>18</td>
</tr>
<tr>
<td>1.2. DETECTING INDIVIDUALS AT RISK</td>
<td>19</td>
</tr>
<tr>
<td>1.3. DEVELOPMENT AND VALIDATION OF A TRANSDIAGNOSTIC RISK CALCULATOR FOR IMPROVING DETECTION OF PSYCHOSIS</td>
<td>21</td>
</tr>
<tr>
<td>1.4. REPLICATION AND IMPLEMENTATION IN PSYCHIATRY</td>
<td>23</td>
</tr>
<tr>
<td>1.5. AIMS AND OBJECTIVES</td>
<td>26</td>
</tr>
<tr>
<td>2. TRANSDIAGNOSTIC INDIVIDUALIZED CLINICALLY-BASED RISK CALCULATOR FOR THE AUTOMATIC DETECTION OF INDIVIDUALS AT-RISK AND THE PREDICTION OF PSYCHOSIS: EXTERNAL REPLICATION IN 2,430,333 US PATIENTS</td>
<td>30</td>
</tr>
<tr>
<td>2.1. ABSTRACT</td>
<td>31</td>
</tr>
<tr>
<td>2.2. INTRODUCTION</td>
<td>32</td>
</tr>
<tr>
<td>2.3. MATERIALS AND METHODS</td>
<td>33</td>
</tr>
<tr>
<td>2.4. RESULTS</td>
<td>39</td>
</tr>
<tr>
<td>2.5. DISCUSSION</td>
<td>42</td>
</tr>
<tr>
<td>2.6. REFERENCES</td>
<td>50</td>
</tr>
<tr>
<td>2.7. SUPPLEMENTARY DATA</td>
<td>53</td>
</tr>
<tr>
<td>2.8. SUPPLEMENTARY REFERENCES</td>
<td>67</td>
</tr>
<tr>
<td>3. REAL-WORLD IMPLEMENTATION OF PRECISION PSYCHIATRY: TRANSDIAGNOSTIC RISK CALCULATOR FOR THE AUTOMATIC DETECTION OF INDIVIDUALS AT-RISK OF PSYCHOSIS</td>
<td>68</td>
</tr>
<tr>
<td>3.1. ABSTRACT</td>
<td>69</td>
</tr>
<tr>
<td>3.2. INTRODUCTION</td>
<td>70</td>
</tr>
</tbody>
</table>
PART C: GENERAL DISCUSSION

6. HIGH RISK? A META-ANALYSIS OF RISK AND PROTECTIVE FACTORS IN PSYCHOSIS

5. WHAT CAUSES THE ONSET OF PSYCHOSIS IN INDIVIDUALS AT CLINICAL HIGH RISK? A META-ANALYSIS OF RISK AND PROTECTIVE FACTORS

4. INTRODUCTION

3.3. METHODS

3.4. RESULTS

3.5. DISCUSSION

3.6. REFERENCES

3.7. SUPPLEMENTARY DATA

3.8. SUPPLEMENTARY REFERENCES

PART B: IMPROVING PROGNOSTICATION OF OUTCOMES IN CHR-P INDIVIDUALS

4. INTRODUCTION

4.1. IMPROVING PROGNOSTICATION OF OUTCOMES

4.2. GENETICS IN PSYCHOSIS

4.3. NON-GENETIC RISK AND PROTECTIVE FACTORS IN PSYCHOSIS

4.4. PSYCHOSIS POLYRISK SCORE (PPS)

4.5. SUMMARY, AIMS AND OBJECTIVES

4.6. REFERENCES

5. WHAT CAUSES THE ONSET OF PSYCHOSIS IN INDIVIDUALS AT CLINICAL HIGH RISK? A META-ANALYSIS OF RISK AND PROTECTIVE FACTORS

5.1. ABSTRACT

5.2. INTRODUCTION

5.3. METHODS

5.4. RESULTS

5.5. DISCUSSION

5.6. REFERENCES

5.7. SUPPLEMENTARY DATA

6. REAL-WORLD DIGITAL IMPLEMENTATION OF THE PSYCHOSIS POLYRISK SCORE (PPS): A PILOT FEASIBILITY STUDY

6.1. ABSTRACT

6.2. INTRODUCTION

6.3. MATERIALS AND METHODS

6.4. RESULTS

6.5. DISCUSSION

6.6. REFERENCES

6.7. SUPPLEMENTARY DATA

6.8. SUPPLEMENTARY REFERENCES

PART C: GENERAL DISCUSSION
7. GENERAL DISCUSSION........................................................................................................... 169
7.1. SUMMARY OF FINDINGS.............................................................................................. 169
7.2. IMPACT OF FINDINGS................................................................................................. 171
7.3. LIMITATIONS................................................................................................................ 174
7.4. FUTURE RESEARCH.................................................................................................... 177
7.5. CONCLUSIONS............................................................................................................. 183
7.6. REFERENCES ............................................................................................................... 184
8. APPENDIX: WORK PUBLISHED DURING THESIS ....................................................... 187
ACKNOWLEDGEMENTS

I owe tremendous thanks to my supervisors, Paolo Fusar-Poli and Philip McGuire. Paolo, thank you for the huge amount of support you have given me to achieve my goals, for pushing me to produce high quality work and presenting me with opportunities that have aided my development as an academic (and occasionally helping expand my travel plans!). I have really enjoyed working with you and I am very thankful that we will continue working together for the foreseeable future. Philip, you have helped me recognise the importance of considering the field as a whole in my work and expanding my horizons. Your advice and direction, as well as the opportunities you have given me, have been invaluable to my development. I hope that I can repay both of you for your guidance and support through my continued contributions to the field and to the department.

Mum and Dad, thank you for always nurturing my curiosity and instilling an interest in science, medicine and the brain from when I was young. That set me on this path and through your support, I have managed to go even further down it than I ever thought possible. To Grandad, who fought so hard to see this day, your passion and encouragement for learning and education pushed me forward to where I am today. I remember when I wanted to be an author as a child, Seeya told me to “be a doctor first and you can be an author when you retire”. Instead, it looks like I’ve done a bit of both at the same time by doing a doctorate and writing this thesis, I think he would have been proud.

Simi, thank you for your unwavering love and support through all the stress, work and evenings/weekends of writing that have gone into this PhD. You continue to inspire me every day with your ambition, drive and dedication. Writing this thesis in your company has definitely made the whole process far less overwhelming and I cannot be more grateful for that.

Cathy, since my first day at the IoPPN you have made me feel at home (and adequately caffeinated) in the department and I am so grateful for your friendship throughout this PhD. I hope that in reading this you do not decide to deplatform me on methodological grounds. Thank you for being one of the few people to read this thesis cover-to-cover, I owe you a coffee when the Quad is finally reconvened.
Amir, working with you has been a highlight of the past few years and it’s a shame that so little of that work has ended up in the thesis! I appreciate the trust and responsibility you gave me in the day-to-day running of eCBD, which helped me grow so much. I also value the environment of togetherness and inclusiveness you have consistently fostered with the team. It has made work feel a lot more enjoyable. Thanks to you, I’ve given talks, given interviews for articles and the radio, and can be glimpsed in a documentary. Although, I will never forgive Horizon for leaving all my insightful lines on the cutting room floor.

Ed, it’s been an absolute joy to work with you, both on eCBD and all our various side projects, which have, for the most part, worked out well (don’t ever talk to me about cost-effectiveness again). The various late night phone calls about inclusion criteria for our psychiatric presentations associated with coronaviruses meta-analysis was definitely a highlight of lockdown. Looking forward to many more years of trying to convince you that the clinical high risk state is more encouraging than certain papers will have you believe.

To Andrea, Aisling, Jack, Lucy, Giulia, Uzma, Sophie, Alexis, Umberto, Jessica, Helen, Gonzalo, George, Graham, Tom, Rob, Jonathan, Sarah, Anai, Kate, Candace, Mary Jane and everyone else at the IoPPN, BRC Nucleus and Clinical Research Facility who have made the past few years so fulfilling. I hope we can celebrate this together when the IoPPN opens up again. Regardless of whether you are in Ireland or Australia, I expect you there without excuse.

Dan, James, Katie and Simon, you’ve been my best friends for a ridiculous number of years at this point. You’ve seen this journey from start to finish and I hope we can all celebrate together soon. Adding to this: Dan, thank you for making the most beautiful figure in the thesis (Figure 5-2) and your commitment to napping for science. James, to this day I am still using that poster template you made for me four years ago and intend to continue using it until there is a new poster trend. Please consider that in any future invoices. Thank you both for helping me make my science more aesthetically pleasing.

Finally, thank you to all the participants who entrusted me with their time and data. Without you, none of this would have been possible.
ABSTRACT

Individuals at clinical high risk for psychosis (CHR-P) experience attenuated positive psychotic symptoms and impaired level of overall functioning. Currently, methods available for detecting individuals at risk are sub-optimal with early detection services for CHR-P individuals identifying only a small minority (5-12%) of first episode psychosis (FEP) cases prior to illness onset. Moreover, once CHR-P subjects have been identified, available methods of predicting their clinical outcomes have limited accuracy. This thesis aims to improve the detection of individuals at high risk of developing psychosis (Part A), and the estimation of psychosis risk in this population (Part B). It will also review gaps in current knowledge and provide directions for future research (Part C).

Part A of this thesis aimed to improve the detection of individuals at risk of developing psychosis. In Chapter 2, I performed the largest replication study of a risk prediction model in psychiatry. As part of this, I validated the discrimination performance of an individualised transdiagnostic psychosis risk calculator that leverages electronic health record (EHR) data, previously developed in our group, in an international US dataset comprised of 2.4 million patients. This was the first replication of the transdiagnostic risk calculator outside of the UK. This validation indicated that the transdiagnostic risk calculator retained significantly better discrimination performance than chance (Harrell’s C = 0.68), highlighting its clinical transportability and potential for automated screening for individuals at risk for psychosis at scale in international settings. In Chapter 3, I assessed the feasibility of implementing the transdiagnostic risk calculator in a clinical EHR system. This represents one of the first clinical applications of a risk prediction model in a mental health setting. The primary barrier to successful use of prediction models in clinical care is clinician endorsement. As such, I investigated the real-world feasibility of screening all individuals accessing secondary mental healthcare at the South London and Maudsley NHS Foundation Trust (SLaM) with clinician adherence being the primary outcome. Clinician adherence to the calculator was high with 78% responding to recommendations of the transdiagnostic risk calculator, emphasising the feasibility of implementing the risk calculator to improve detection of individuals at risk of developing psychosis. Through these studies, I have advanced knowledge and the potential of real-world, automated, systematic detection of individuals at risk for psychosis.
Part B aimed to improve the prognostication of outcomes in CHR-P individuals. Chapter 5 presents the identification of robust, non-genetic factors that modulate transition risk in CHR-P individuals through the synthesis of the available evidence. The results showed that no factors met the criteria for the highest classification of evidence (class I, convincing). However, attenuated positive psychotic symptoms and global functioning were associated with highly suggestive evidence (class II) and negative psychotic symptoms with suggestive evidence (class III). The remaining factors were associated with either weak (class IV) evidence or were non-significant. Chapter 6 outlines the development, digital implementation and piloting of a novel multivariate assessment for non-genetic risk and protective factors for psychosis, termed the Psychosis Polyrisk Score (PPS). This assessment can be completed online, in less than 15 minutes, to provide an individualised estimate of exposure to non-genetic risk and protective factors for psychosis “en masse”. A simulated general population dataset was also used to show the range and distribution of scores. Pilot data found that individuals referred for a CHR-P assessment had higher PPS scores compared to healthy controls, highlighting the feasibility of its use in real world settings and potential for refining prognostication of outcomes, complementing the CHR-P assessment. Together, these studies advance the potential for refining prognostic estimates of clinical outcomes for CHR-P individuals.

Finally, in Part C, I discussed the collective implications of the findings from Part A and Part B, reviewed their limitations and considered how to overcome these in future work. I then discussed ways to take work from this thesis forward in order to further improve the identification of people at high risk and the prognostication of their outcomes.
THESIS OUTLINE

PART A: Improving detection of individuals at risk
- Chapter 1: Introduction to CHR-P detection methods
- Chapter 2: International replication of transdiagnostic risk calculator for psychosis
- Chapter 3: Implementation of transdiagnostic risk calculator for psychosis

PART B: Improving prognostication of outcomes
- Chapter 4: Introduction to CHR-P prognostication methods
- Chapter 5: Meta-analysis of risk and protective factors for transition in CHR-P individuals
- Chapter 6: Psychosis Polyrisk Score (PPS) pilot feasibility study

PART C: General discussion
- Summary of findings
- Limitations
- Future work
- Conclusions
STATEMENT OF PERSONAL CONTRIBUTION

For the international replication of transdiagnostic risk calculator for psychosis (Chapter 2), I wrote the original protocol, performed all the analyses and wrote the first draft of the manuscript. For the implementation of transdiagnostic risk calculator for psychosis (Chapter 3), I wrote the original protocol and obtained ethical approval for the study and subsequent amendments, I conducted recruitment and data collection alongside study clinicians, independently conducted data management and cleaning, performed all data analyses and wrote the first draft of the manuscript. For the meta-analysis of risk and protective factors for transition in CHR-P individuals (Chapter 5), I designed the original study protocol, conducted the literature search and data extraction, led the team of researchers, performed all the analyses in the meta-analysis and wrote the first draft of the manuscript. For the Psychosis Polyrisk Score (PPS) pilot feasibility study (Chapter 6), I designed the study, conducted recruitment alongside study clinicians, conducted data collection, independently conducted data management and cleaning, performed all data analyses and wrote the first draft of the manuscript. Finally, I wrote this thesis in its entirety, with the following exception: the published papers, after being drafted by me, were circulated to co-authors and underwent peer review prior to acceptance, leading to final editing of the manuscripts.
PREFACE

This thesis is a “thesis incorporating publications”. This refers to the fact that four chapters are composed of published journal articles of which I am the first author.

Publications relating to the work presented in this thesis – four chapters are composed of the following journal articles which are reproduced in full:


In addition, the introduction to Part B (Chapter 4) and general discussion (Part C) contains text from a conceptual review of which I am first author:
LIST OF TABLES

Table 1-1 Core characteristics of the transdiagnostic psychosis risk calculator .................24
Table 1-2 Statistics for individual predictor variables in the transdiagnostic risk calculator in the derivation data set .....................................................................................................................25
Table 2-1 Sociodemographic Characteristics of the Commercial dataset compared with the SLaM dataset ........................................................................................................................38
Table 2-2 Individualized clinical prediction models that have been externally validated for early psychosis..........................................................................................................................44
Supplementary Table 2-3 The REporting of studies Conducted using Observational Routinely-collected Data (RECORD) statement – checklist of items, extended from the STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) statement, that should be reported in observational studies using routinely collected health data ........................................................................................................54
Supplementary Table 2-4 Predictor definitions: Primary index diagnoses of non-organic and non-psychotic mental disorder formulated at baseline (time of the first contact with the NHS Trust) .................................................................................................................................60
Supplementary Table 2-5 Predictor definitions: self-assigned ethnicity. Individual-level ethnicity data was not available for the MarketScan database so was imputed using Integrated Public Use Microdata Series (IPUMS) as discussed in the methods section ................62
Supplementary Table 2-6 Model performance of the Transdiagnostic Individualised Clinically-based Risk Calculator for the Automatic Detection of Individuals at Risk and the Prediction of Psychosis (revised version) in the original South London and Maudsley NHS Foundation Trust (SLaM) derivation and validation datasets ...........................................................................64
Supplementary Table 2-7 Transdiagnostic Individualised Clinically-based Risk Calculator for the Automatic Detection of Individuals at Risk and the Prediction of Psychosis (revised version), original derivation dataset (SLaM boroughs Lambeth & Southwark). This model has been used for external validation in the current study .......................................................................65
Supplementary Table 2-8 Model performance of the Transdiagnostic Individualised Clinically-based Risk Calculator for the Automatic Detection of Individuals at Risk and the Prediction of Psychosis (revised version) in the original SLaM derivation dataset with removed ethnicity predictor and with aggregate ethnicity coefficients .................................................66
Table 3-1 Sociodemographic characteristics of the study population, both all patients automatically screened in addition to those patients automatically detected ......................................75
Table 3-2 Sociodemographic characteristics of the study population, both all patients automatically screened in addition to those patients automatically detected. ..........................81

Supplementary Table 3-3 Performance of the Transdiagnostic Individualised Clinically-based Risk Calculator for the Automatic Detection of Individuals at-Risk and the Prediction of Psychosis (revised version) in the original SLaM derivation and validation datasets, refined without a three month lag after the index diagnosis. .................................................................92

Table 5-1 Level of evidence for the association of sociodemographic/parental factors and psychotic disorders........................................................................................................125

Table 5-2 Level of evidence for the association of later factors and psychotic disorders ....126

Table 5-3 Level of evidence for the association of antecedents and psychotic disorders.....127

Table 5-4 Level of evidence for the association of symptom scores/clinical factors and psychotic disorders........................................................................................................128

Table 6-1 Scoring system for the Psychosis Polyrisk Score (PPS).................................149

Table 6-2 Sociodemographics of study participants: individuals referred for a CHR-P assessment and healthy controls. .................................................................................................151

Supplementary Table 6-3 List of included factors, along with their definitions and the tentative cut-offs for defining each respective Risk Ratio.................................................................165
LIST OF FIGURES

Figure 2-1 Flow chart of the study population.................................................................37
Figure 2-2 Cumulative incidence (Kaplan–Meier failure function) for risk of development of psychotic disorders in the Commercial Database and SLaM derivation database. .......41
Figure 3-1 Overview of in-vitro and in-vivo phases.........................................................72
Figure 3-2 CFIR Implementation barriers addressed during the in-vitro phase. .............73
Figure 3-3 Outline of study design. ..................................................................................80
Supplementary Figure 3-4 Cumulative incidence (Kaplan–Meier failure function) for risk of development of psychotic disorders in individuals screened and detected by the transdiagnostic risk calculator and followed-up over a period of six-months........97
Supplementary Figure 3-5 Log-rank test comparing the cumulative incidence (Kaplan–Meier failure function) for risk of development of psychotic disorders in 52 individuals detected and referred vs 49 individuals detected but not referred. ...........................98
Figure 5-1 PRISMA flowchart outlining study selection process ..................................121
Figure 5-2 Graphical summary of risk/protective factors for psychosis onset in the CHR-P state. No factors met criteria for a convincing level of evidence (class I), two factors for a highly suggestive level of evidence (class II), one factor for a suggestive level of evidence (class III) and eight for a weak level of evidence (class IV). .............................122
Figure 6-1 Distribution of PPS scores in the simulated general population. Blue bars indicate the proportion of individuals receiving each PPS score based on the prevalence of risk factors and 10,000,000 permutations ..................................................................................153
Figure 6-2 Violin plot comparing PPS scores distribution in individuals referred for a CHR-P assessment, controls and the (simulated) general population ........................................154
Figure 7-1 Putative PPS assessment for the detection of at-risk individuals and the prediction of psychosis..............................................................................................................181
LIST OF ABBREVIATIONS

APS: Attenuated psychotic syndrome
APSS: Adolescent Psychotic-like Symptom Screener
ATPD: Acute and transient psychotic disorder
BAC: Balanced accuracy
BIPS/BLIPS: Brief (limited) intermittent psychotic symptoms
BPRS: Brief Psychiatric Rating Scale
BSIP: Basel Screening Instrument for Psychosis
CAARMS: Comprehensive Assessment for At Risk Mental States
CFIR: Consolidated Framework for Implementation Research
CHR-P: Clinical high risk for psychosis
C&I: Camden & Islington NHS Foundation Trust
CRIS: Clinical Record Interactive Search
DSM: Diagnostic and Statistical Manual of Mental Disorders
DUP: Duration of untreated psychosis
EHR: Electronic health record
FEP: First episode psychosis
GAF: Global Assessment of Functioning
GRD: Genetic Risk and Deterioration
GWAS: Genome-wide association studies
HR: Hazard Ratio
ICD: International Classification of Diseases
IPUMS: Integrated Public Use Microdata Series
MSA: Metropolitan State Area
NAPLS: North American Prodrome Longitudinal Study
NHS: National Health Service
NICE: National Institute for Health and Care Excellence
OASIS: Outreach and Support in South London
OR: Odds Ratio
PACE: Personal Assessment and Crisis Evaluation
PANSS: Positive and Negative Syndrome Scale
PI: Prognostic Index
PPS: Psychosis Polyrisk Score
PPV: Positive predictive validity
PRS: Polygenic risk score
PQ-16: 16-item Prodromal Questionnaire
RR: Relative Risk/Risk Ratio
SCID: Structured Clinical Interview for DSM-IV
SIPS: Structured Instrument for Prodromal Syndromes
SlaM: South London and Maudsley NHS Foundation Trust
SMD: Standardised Mean Difference
SOPS: Scale of Prodromal Symptoms
YouR-Study: Youth-Mental Risk and Resilience study
PART A: IMPROVING DETECTION OF INDIVIDUALS AT RISK OF DEVELOPING PSYCHOSIS
1. INTRODUCTION

1.1. PRIMARY INDICATED PREVENTION AND THE CLINICAL HIGH RISK STATE

Over the past two decades, there has been increasing interest in the possibility of primary indicated prevention (i.e. targeting individuals identified as having signs or symptoms that indicate increased risk of developing a disorder) through the clinical high risk for psychosis state (CHR-P).\(^1\) CHR-P individuals are help-seeking, experiencing attenuated positive psychotic symptoms (unusual thought content, non-bizarre thinking, perceptual abnormalities and disorganised speech) accompanied by functional impairment.\(^1\) Approximately 50% of first episode psychosis (FEP) patients retrospectively report experiencing a prodromal phase prior to the onset of their first episode,\(^2\)–\(^5\) highlighting a clear clinical opportunity for indicated prevention in this population.\(^6\) The CHR-P construct is now well-developed, with the National Institute for Health and Care Excellence (NICE) regarding assessment, treatment and monitoring of CHR-P individuals as an essential component of early intervention for psychosis.\(^7\) Guidelines are similar internationally leading to the CHR-P construct being well implemented with 47 specialist CHR-P services providing care to over 20,000 CHR-P individuals worldwide.\(^8\)

Preventing transition to a psychotic disorder is a key aim of CHR-P clinical services and the main clinical outcome that has been examined in CHR-P research.\(^9\) The CHR-P state provides a unique opportunity to alter the trajectory of psychotic disorders before symptoms or pathophysiology become too severe and enduring.\(^10\) Meeting CHR-P criteria is associated with a greatly increased risk of developing psychosis, with 22% of individuals transitioning to FEP within three years of presenting to services.\(^11\) In addition to the long term aim of preventing transition, in the shorter term, CHR-P services also aim to reduce the severity of the presenting attenuated positive and negative symptoms, and to improve quality of life and impairments in social and vocational functioning. If an individual with CHR-P develops psychosis, early detection services can quickly refer them to FEP services and start antipsychotic treatment. The greater the length of time between psychosis onset and the provision of adequate antipsychotic treatment (the duration of untreated psychosis; DUP),\(^12\) the worse the clinical,\(^13\)–\(^17\) cognitive\(^16,18,19\) and social outcomes in FEP.\(^16,18,20\) CHR-P services can thus provide a very effective way of reducing DUP, thereby improving outcomes after psychosis onset.\(^21\)
Overall, the real-world impact of the CHR-P construct is dependent on three concatenated factors:

i) Efficient detection of individuals at risk of developing psychosis
ii) Accurate prognostication of clinical outcomes
iii) Effective preventative interventions

This thesis presents the results of several studies designed to improve the detection of psychosis risk and to refine the accurate prognostication of outcomes. This would expand the benefit of the CHR-P construct to a greater number of young people who may be at risk of developing serious mental disorders. The limitations of the current state of these two factors, their resultant impact on psychosis prevention and the rationale for improving them are analysed in the introductions to Part A and Part B of this thesis, in addition to the introductory sections of later chapters. The extent to which these limitations are addressed are explored in Part C, as well as the discussion sections of later chapters.

In this introduction to Part A, I will detail current strategies for detecting individuals at risk for psychosis, the limitations of those strategies and introduce the transdiagnostic risk calculator for psychosis that seeks to address these limitations. Finally, I will explore the rationale for the replication and implementation studies that comprise Chapters 2 and 3 of this PhD thesis.

1.2. CHALLENGES OF DETECTING INDIVIDUALS AT RISK
Efficient detection of individuals at risk of developing psychosis is the first key rate limiting step in permitting effective preventative intervention. The impact of preventative treatment will be modest if it can only be provided to a small proportion of those at risk. Unfortunately, current detection methods are suboptimal. In this section, I will outline the challenges relating to detection strategies, which my PhD seeks to address.

The first challenge associated with identifying individuals at risk is that the CHR-P state is not necessarily the prototypical pre-psychotic stage. While the majority of FEP patients have some form of CHR-P features, approximately one third of FEP patients do not experience a CHR-P stage.22,23 This can often be attributed to a short-lived psychotic episode lasting a
few weeks. Therefore, if identification is entirely contingent on recognising the CHR-P state, then it is unlikely for all future FEP cases to be detected prior to psychosis onset. It is evident that to expand our approach to potentially identifying all FEP patients prior to psychosis onset, then CHR-P assessments will need to be supplemented with information from other sources.

The second challenge is that the current detection strategies for CHR-P services could be more efficient. Detection approaches vary widely between clinical services, with some relying solely on clinical referrals and others running intensive outreach campaigns. There have been various specific interventions within early intervention programmes to improve detection of FEP patients and CHR-P individuals. These were defined by the LYRIKS study: screening assessments and recruitment (outpatient and satellite clinics, armed forces, private hospitals, government organisations, internet gaming shops, and youth hubs); workshops involving various community partners such as counsellors and mental healthcare professionals; roadshows; student internships; print media (brochures and posters, articles and advertorials, newsletter); and social media (Facebook, Twitter, blogs, websites). Regardless of the characteristics of the recruitment strategy employed, at present services are only detecting between 5% (UK) and 12% (Australia) of future FEP cases prior to psychosis onset. Thus, the vast majority of individuals who develop a first episode of psychosis will not have had access to help and support that might have either reduced their risk of transition or improved their prognosis after psychosis onset.

This in turn has a knock-on effect in terms of the utility of the CHR-P assessment. The probability of an individual developing psychosis after the result of an assessment is known (post-test risk) is dependent on the characteristics of the assessments (e.g. Comprehensive Assessment for At Risk Mental States [CAARMS], Structured Interview for Prodromal Syndromes [SIPS]) themselves, in particular their sensitivity and specificity. However, on the basis of Bayes’ theorem, the post-test probability is also dependent on the probability of the individual developing psychosis before the test result is known (pre-test risk). While the average pre-test risk in those referred to CHR-P services is 15% over 38 months, significantly higher than the 0.1% seen in the general population over the same time period, the heterogeneity of pre-test risk is high (95%CI: 9%-24%). This heterogeneity can largely be explained by variation in strategies of recruitment. Pre-test risk is high when outreach campaigns are designed to recruit samples that are enriched for risk and are directed towards
mental healthcare services. Among these samples, self-referrals are relatively uncommon. However, pre-test risk is diluted when outreach efforts are extended to the general public (e.g. through social media), resulting in high numbers of self-referrals. This reduction in pre-test risk (i.e. a reduced prevalence in the sample) has a negative impact on the positive predictive value (PPV) of CHR-P assessments, meaning that a larger number of individuals will meet at-risk criteria, be treated by services, but will not develop a psychotic disorder.

Overall, these findings indicate that any intervention to improve detection of CHR-P individuals prior to the onset of their first episode of psychosis should be systematic and encompass strategies that do not dilute pre-test risk.

1.3. DEVELOPMENT AND VALIDATION OF A TRANSDIAGNOSTIC RISK CALCULATOR FOR IMPROVING DETECTION OF PSYCHOSIS

Precision medicine tailors clinical recommendations and/or treatment to the individual, informed by data from sources that include sociodemographics, diagnoses and environmental risk/protective factors. Current methods of detection are unstructured and idiosyncratic, with a resultant dilution of pre-test risk enrichment. My PhD employs precision medicine methods to systematically screen databases at scale, with a focus on secondary care to improve detection of individuals at risk of psychosis while preserving risk enrichment. This is particularly prescient in South London and Maudsley NHS Foundation Trust (SLaM), where 95% of individuals developing a FEP have not been detected during their potential CHR-P stage. This is a clear missed clinical opportunity for preventing psychosis in individuals who are already under the care of mental health services. To rectify these limitations in detection while maintaining enrichment of pre-test psychosis risk, a transdiagnostic risk calculator for psychosis was developed and externally validated by Paolo Fusar-Poli and colleagues in SLaM. I was not involved in the initial development and validation of the risk calculator. SLaM is a mental health trust that provides secondary mental healthcare for 1.3 million people in four discrete London boroughs (Croydon, Lambeth, Lewisham and Southwark). The trust is paper-free with all clinical records available and maintained digitally. Patients’ records are continually updated throughout their care, regardless of referral to other services or discharge from SLaM care. A Clinical Record Interactive Search (CRIS) function was implemented in SLaM to allow for searching and analysing real-world, real-time, anonymised routine clinical information from mental healthcare for research purposes. This also enhanced implementation potential as
the transdiagnostic risk calculator could easily transition from being applied to retrospective CRIS data to prospective use in clinical routine, systematically screening the local electronic health record (EHR) system.

The core characteristics of the risk calculator emphasise the strengths of this approach (Table 1-1). Being transdiagnostic, the approach allows for the use of this calculator outside of CHR-P samples, across different psychopathological domains. This approach can be applied to any patient receiving an ICD-10 index diagnosis of a non-psychotic mental disorder, thus overcoming the above limitations. The tool is also clinically-focused, using predictors that are widely available and rarely missing in EHRs, thereby reducing additional burden on clinicians and patients. Furthermore, the model is lifespan-inclusive, meaning it can work across all ages (including the age range of peak of risk for psychosis). As the model leverages information from EHRs, screening can be automated. This automation has a number of benefits: it allows for screening on a large scale, it allows for standardised screening and reduces costs associated with use. This transdiagnostic risk calculator aimed to use clinical and sociodemographic variables routinely collected as part of clinical care, and therefore available in CRIS, to enhance implementation potential. Predictor selection was supported by a priori clinical knowledge and meta-analytical evidence, as advised by model building guidelines. Age, gender, age*gender interaction, ethnicity and ICD-10 diagnosis were used in the model. The model is presented in Table 1-2 below.

A Cox proportional hazards model was used to assess the effects of these prespecified predictors on the transdiagnostic development of non-organic psychotic disorders and time to psychosis onset. The overall SLaM sample were split between development (Lambeth and Southwark; n = 33,820) and external validation (Lewisham and Croydon; n = 54,716) datasets in a non-random fashion to mitigate over-fitting. Non-random splitting was preferred to preserve the significant differences in sociodemographics between SLaM boroughs, ensuring distinct development and validation datasets that would promote future transportability of the model. Significant sociodemographic differences between SLaM boroughs meant that validation in this fashion would allow for a more generalisable model. Importantly, borough was defined according to the address of each patient’s registered GP practice at the time of index diagnosis so no individual could belong to more than one borough in the dataset. Individuals with missing GP or outside SLaM boroughs were included in the external validation dataset. Model performance was good in the development
dataset (Harrell’s C = 0.80) and fair-to-good in the external validation dataset (Harrell’s C = 0.79).32

1.4. REPLICATION AND IMPLEMENTATION IN PSYCHIATRY

In Chapters 2 and 3 of this thesis, I will be presenting the first replication of the transdiagnostic risk calculator outside of the UK and the first implementation attempt of a prediction model in psychiatry. In this section, I will detail the importance of replication and implementation of prediction models in psychiatry.

While development of prediction models is extremely important, there is a large discrepancy between the number of published models and the subset that go on to be implemented in clinical care, both in psychiatry and in physical health.42,43 The transdiagnostic risk calculator was developed with clinical implementation in mind; it is therefore important to consider the barriers that impede other models. Successful translation of a model depends not only on its prognostic accuracy but also on its independent replication, and then an implementation process.

Replication is integral in psychiatric research for two key reasons. Firstly, published science has a statistical power issue at its core that results in a likely high false report probability,44,45 therefore a replicated model is a more reliable model. Moreover, replication can provide evidence of clinical transportability46 through demonstrating evidence of prognostic accuracy in different populations with different case-mixes in settings with different care configurations. Despite these benefits, replication in early psychosis research is rare,47 and this dearth of replication restricts the possibility of clinical translation.48 The transdiagnostic psychosis risk calculator has been replicated in another UK setting, which had a key difference in service configuration as there were no CHR-P services (Harrell’s C = 0.73).49 While the transdiagnostic risk calculator was developed with maximising generalisability in mind, its prognostic performance has yet to be tested in samples from outside the UK. A successful replication result in a US dataset would be encouraging for expanding the clinical utility of the transdiagnostic risk calculator to enable other countries with EHR infrastructure to improve detection of individuals at risk for psychosis through cheap, automated screening at scale.
Table 1-1 Core characteristics of the transdiagnostic psychosis risk calculator

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lifespan-inclusive</td>
<td>Works with any age</td>
</tr>
<tr>
<td>Clinically-based</td>
<td>Predictors selected through <em>a priori</em> clinical knowledge</td>
</tr>
<tr>
<td>Transdiagnostic</td>
<td>Works across all ICD-10 diagnostic spectra</td>
</tr>
<tr>
<td>Individualised</td>
<td>Individual subject-level risk estimates</td>
</tr>
<tr>
<td>Cheap</td>
<td>Predictors routinely collected by clinicians</td>
</tr>
<tr>
<td>Automated</td>
<td>Electronic health records as well as manual entry of predictors</td>
</tr>
<tr>
<td>e-Health</td>
<td>Implemented online</td>
</tr>
<tr>
<td>Scalable</td>
<td>Screens electronic health records at scale</td>
</tr>
<tr>
<td>Optimisable</td>
<td>Further refined by the inclusion of other predictors</td>
</tr>
<tr>
<td>Sequential testing</td>
<td>Can be used as part of a staged assessment framework</td>
</tr>
<tr>
<td>Implementable</td>
<td>Can be integrated with existing structures for use in clinical settings</td>
</tr>
<tr>
<td>Predictor</td>
<td>HR</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Age (years)</td>
<td>1.011</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1.764</td>
</tr>
<tr>
<td>Female</td>
<td>1</td>
</tr>
<tr>
<td>Age*gender (male)</td>
<td>0.988</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>1</td>
</tr>
<tr>
<td>Black</td>
<td>2.823</td>
</tr>
<tr>
<td>Asian</td>
<td>1.671</td>
</tr>
<tr>
<td>Mixed</td>
<td>1.839</td>
</tr>
<tr>
<td>Other</td>
<td>1.504</td>
</tr>
<tr>
<td>ICD-10 index diagnosis</td>
<td>CHRP</td>
</tr>
<tr>
<td>Acute and transient psychotic disorders</td>
<td>2.682</td>
</tr>
<tr>
<td>Substance use disorders</td>
<td>0.146</td>
</tr>
<tr>
<td>Bipolar mood disorders</td>
<td>0.839</td>
</tr>
<tr>
<td>Nonbipolar mood disorders</td>
<td>0.152</td>
</tr>
<tr>
<td>Anxiety disorders</td>
<td>0.107</td>
</tr>
<tr>
<td>Personality disorders</td>
<td>0.213</td>
</tr>
<tr>
<td>Developmental disorders</td>
<td>0.031</td>
</tr>
<tr>
<td>Childhood/adolescence onset disorders</td>
<td>0.039</td>
</tr>
<tr>
<td>Physiological syndromes</td>
<td>0.085</td>
</tr>
<tr>
<td>Mental retardation</td>
<td>0.086</td>
</tr>
</tbody>
</table>

Abbreviations: CHRP, clinical high risk for psychosis; HR, hazard ratio
Similarly, the gap between publication and clinical use of prediction models highlights the clear importance of implementation research: the scientific study of methods translating research findings into practical, useful outcomes. Implementation research seeks to understand and work pragmatically within real-world conditions, rather than trying to control for them.\textsuperscript{50–52} Many methods of predicting psychosis risk are unlikely to be implemented into clinical care in the near future, partially due to pragmatic concerns relating to high costs (e.g. neuroimaging modalities), labour (e.g. cognitive tasks) or applicability (e.g. genetics). Implementation research aims at solving a wide range of practical problems relating to the real-world usability of precision medicine and digital health in clinical practice. For example, prediction models are unlikely to impact clinical pathways unless they are used by clinicians in day-to-day practice;\textsuperscript{53} therefore clinicians’ compliance with the recommendations made by a prediction model represents the first key barrier to implementation.\textsuperscript{54,55} Showcasing clinician adherence to the recommendations of the transdiagnostic psychosis risk calculator is an imperative step for it to be used effectively in clinical routine.

1.5. AIMS AND OBJECTIVES

The overarching aim of this work was to validate and implement a pragmatic, clinically-focused, automated transdiagnostic risk calculator for psychosis to improve detection of individuals at risk of developing psychosis, and then assess the feasibility of its application in a clinical setting. More specifically, I aimed to achieve the following:

**Chapter 2:** Validate a transdiagnostic psychosis risk calculator in a large, independent international database, which involved different referral pathways and a different case-mix compared to the original development study and previous external validations

**Chapter 3:** Implement the transdiagnostic psychosis risk calculator in a clinical setting and test the feasibility of its use in clinical practice.
1.6. References


2. TRANSDIAGNOSTIC INDIVIDUALIZED CLINICALLY-BASED RISK CALCULATOR FOR THE AUTOMATIC DETECTION OF INDIVIDUALS AT-RISK AND THE PREDICTION OF PSYCHOSIS: EXTERNAL REPLICATION IN 2,430,333 US PATIENTS

Published in Translational Psychiatry 2020
doi: https://doi.org/10.1038/s41398-020-01032-9

Dominic Oliver1*, Chiew Meng Johnny Wong2*, Martin Bøg3, Linus Jönsson3,4, Bruce J. Kinon5, Allan Wehnert3, Kristian Tore Jørgensen3, Jessica Irving1, Daniel Stahl6, Philip McGuire7,8, Lars Lau Raket3,9, Paolo Fusar-Poli1,7,10

*Both authors contributed equally to this manuscript
1 Early Psychosis: Interventions & Clinical detection (EPIC) Lab, Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, London, SE5 8AF, UK,
2 Lundbeck Singapore Pte, Ltd., Singapore 307591,
3 H. Lundbeck A/S, Valby, Denmark,
4 Karolinska Institutet, Stockholm, Sweden,
5 Lundbeck Pharmaceuticals LLC, Deerfield, IL, USA,
6 Department of Biostatistics, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, London, SE5 8AF, UK,
7 OASIS Service, South London and the Maudsley NHS National Health Service Foundation Trust, London, UK,
8 Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, London, SE5 8AF, UK,
9 Clinical Memory Research Unit, Lund University, Lund, Sweden,
10 Department of Brain and Behavioural Sciences, University of Pavia, Pavia, Italy.

Permissions
Copyright Oliver et al., 2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0). The license can be viewed at: https://creativecommons.org/licenses/by/4.0/legalcode. No changes were made in the reproduction of this article.
2.1. ABSTRACT

The real-world impact of psychosis prevention is reliant on effective strategies for identifying individuals at risk. A transdiagnostic, individualized, clinically-based risk calculator to improve this has been developed and externally validated twice in two different UK healthcare trusts with convincing results. The prognostic performance of this risk calculator outside the UK is unknown. All individuals who accessed primary or secondary health care services belonging to the IBM® MarketScan® Commercial Database between January 2015 and December 2017, and received a first ICD-10 index diagnosis of nonorganic/nonpsychotic mental disorder, were included. According to the risk calculator, age, gender, ethnicity, age-by-gender, and ICD-10 cluster diagnosis at index date were used to predict development of any ICD-10 nonorganic psychotic disorder. Because patient-level ethnicity data were not available city-level ethnicity proportions were used as proxy. The study included 2,430,333 patients with a mean follow-up of 15.36 months and cumulative incidence of psychosis at two years of 1.43%. There were profound differences compared to the original development UK database in terms of case-mix, psychosis incidence, distribution of baseline predictors (ICD-10 cluster diagnoses), availability of patient-level ethnicity data, follow-up time and availability of specialized clinical services for at-risk individuals. Despite these important differences, the model retained accuracy significantly above chance (Harrell’s C=0.676, 95% CI:0.672-0.679). To date, this is the largest international external replication of an individualized prognostic model in the field of psychiatry. This risk calculator is transportable on an international scale to improve the automatic detection of individuals at risk of psychosis.
2.2. INTRODUCTION

Under standard care, clinical outcomes in psychosis are suboptimal; prevention and early intervention are essential to improve outcomes of this disorder.\(^1\) Primary indicated prevention of psychosis revolves around the ability to detect, assess and care for individuals at risk of psychosis. The Clinical High Risk state for Psychosis (CHR-P)\(^2\) includes individuals who present with attenuated psychotic symptoms, impaired functioning\(^3\) and help-seeking behaviour. 20% of these individuals develop a psychotic disorder within two years.\(^4\) Primary indicated prevention of psychosis through specialized CHR-P clinical services\(^5\) is uniquely positioned to alter the course of the disorder and improve outcomes.\(^1\)

The impact of the CHR-P approach is contingent on effective identification of individuals at risk of developing psychosis. Because of complex interactions between help-seeking behaviours, recruitment strategies and referral pathways,\(^6\) detection of at-risk individuals is currently inefficient: only 5\% - 12\%\(^7\) - 8\%\(^8\) of first-episode cases are identified by specialized or youth mental health CHR-P services. Moreover, these services are only available to a limited number of individuals, with only 48 services mapped worldwide.\(^9\) To overcome these problems, a transdiagnostic, individualized, clinically-based risk calculator has been developed in the South London and Maudsley (SLaM) NHS Trust boroughs of Lambeth and Southwark \(n=33,820\).\(^7\) This prognostic model uses core predictors that were selected on \textit{a priori} meta-analytical knowledge\(^10\) (age, gender, ethnicity, primary index diagnosis and age*gender interaction), that are routinely collected in clinical care, to forecast individual level of psychosis risk up to six years. This model leverages electronic health record (EHR) data, therefore allowing for the automatic detection of at-risk individuals. This prognostic model has shown adequate performance in a first external validation in the SLaM boroughs of Lewisham and Croydon \(n=54,716\), Harrell’s C=0.79\(^7\) and in a second external validation in the Camden and Islington NHS Foundation Trust (C&I; \(n=13,702\), Harrell’s C=0.73),\(^11\) with Harrell’s C demonstrating the probability that a randomly selected patient who experienced an event will have a higher score than a patient who did not. This prognostic model is also currently being piloted for real-world implementation in clinical routine in the UK.\(^12\)

Despite these promising results, it is not yet clear whether this prognostic model is transportable to international healthcare settings. External validation studies are scarce in psychiatry, undermining the translational impact of research discoveries. The current study
aims to investigate the international external validity of the original transdiagnostic, clinically-based, individualized risk calculator using large scale EHRs from the US.

2.3. MATERIALS AND METHODS

2.3.1. DESIGN
Retrospective cohort study using Electronic Health Records (EHRs) conducted according to the *Reporting of studies Conducted using Observational Routinely-collected health Data (RECORD)* statement\(^1\) (see checklist reported in Supplementary Table 2-3).

2.3.2. DATA SOURCE
The IBM\(^\circ\) MarketScan\(^\circ\) Commercial Database (hereafter Commercial) contains data from approximately 65 million people from multiple geographically dispersed US states, who are covered by employer-sponsored health insurance plans. This data includes all medical and pharmaceutical claims for these individuals and their dependents (Supplementary Methods 2-1). It provides contemporaneous and ‘real-world’ data on both routine primary and secondary mental healthcare.

2.3.3. STUDY POPULATION
All patients accessing primary or secondary healthcare between 1 January 2015 and 31 December 2017 who received an ICD-10 primary index diagnosis of a nonorganic and nonpsychotic mental disorder (Supplementary Methods 2-2). To ensure correct diagnosis classification, a lookback period of six months was applied to each patient (Supplementary Methods 2-3).

2.3.4. FOLLOW-UP
Follow-up started at the time of the ICD-10 index diagnosis and ended when a transition to psychosis was recorded, or when the patient dropped out of the EHR (as documented by the last entry on Commercial).

2.3.5. MODEL SPECIFICATIONS
The original transdiagnostic, clinically-based, individualized risk calculator was developed using a retrospective cohort study leveraging EHRs of the SLaM boroughs of Lambeth and Southwark, firstly validated in the SLaM boroughs of Croydon and Lewisham\(^7\) and secondly
validated in C&I\textsuperscript{11} in the UK. In summary, a Cox model was used to predict the hazard ratio of developing any psychotic disorder over time (see Supplementary Methods 2-2 for definition) as primary outcome of interest. The predictors included age (at the time of the index diagnosis), gender, age*gender, self-assigned ethnicity, and cluster index diagnosis (ICD-10 diagnostic spectra: acute and transient psychotic disorders (ATPD), bipolar mood disorders, nonbipolar mood disorders, anxiety disorders, personality disorders, developmental disorders, childhood/adolescence onset disorders, physiological syndromes, mental retardation). Self-assigned ethnicity and index diagnoses were operationalized as indicated in Supplementary Tables 2-4 & 2-5. A weighted sum of covariates with the model weights from the Cox model resulted in the Prognostic Index (PI). From this, the risk of the individual developing a psychotic disorder within a time period (between one and six years) could be calculated.\textsuperscript{14}

Since this model was originally developed on a retrospective cohort,\textsuperscript{7} it excluded cases with an onset of psychosis within the first three months to minimize the short-term diagnostic instability of baseline ICD-10 index diagnoses. However, during the subsequent implementation study\textsuperscript{12,15} an updated version of the model was adapted for prospective use (i.e. not excluding transitions occurring in the first three months), demonstrating similar prognostic performance (Supplementary Table 2-6). Furthermore, a lookback period was additionally used in the current study (Supplementary Methods 3), to minimize the risk of misclassification of index diagnosis date. The specifications of the present model are fully detailed in Supplementary Table 2-7.

A main difference compared to the SLaM dataset was that there were no patient-level ethnicity data in Commercial. To mitigate this issue, aggregate ethnicity coefficients were generated for patients who had Metropolitan Statistical Area (MSA) and state-level ethnicity data using Integrated Public Use Microdata Series (IPUMS) census data (www.ipums.org). The geographical information from IPUMS were matched with the geographical data available for each patient in the study population from Commercial, assigning each patient with a vector of ethnic weights for each level of the ethnicity predictor. For example, if a patient were matched for New York (NY) state and Ithaca, NY MSA and was diagnosed in 2016, the proportions of White individuals in the MSA in the year of index diagnosis was 0.82, Black individuals was 0.03, Asian individuals was 0.10, Mixed individuals was 0.03 and Other was 0.01. For comparability purposes we also reported the performance of the
original model\textsuperscript{7} (i) without ethnicity as a predictor and computing (ii) aggregate ethnicity using census data\textsuperscript{16} (Supplementary Table 2-8).

2.3.6. \textit{STATISTICAL ANALYSIS}

Model external validation followed the guidelines of Royston and Altman,\textsuperscript{17} Steyerberg et al.,\textsuperscript{18} and the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD).\textsuperscript{19} The study protocol is uploaded in the Research Registry database (www.researchregistry.com, researchregistry5130).

To interpret the performance of a risk model in the context of external validation, it is essential to first quantify the similarities between development and validation samples.\textsuperscript{20} External validity only assesses model transportability if validation samples have a different case-mix, the greater the difference in the case-mixes, the greater the possibility of generalising to other populations. Thus, we investigated the extent to which the SLaM and Commercial 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.\textsuperscript{21}

As a first step, we described the Commercial patient population, including the configuration of clinical services and compared with SLaM. Baseline clinical and sociodemographic characteristics of the sample (including missing data) were described by means and SDs for continuous variables, and absolute and relative frequencies for categorical variables.\textsuperscript{21} In a second step, we visually compared the two Kaplan–Meier failure functions, showing the number of patients developing a psychotic disorder as well as those still at risk over time. The overall cumulative risk of psychosis onset in Commercial was visualized with the Kaplan–Meier failure function (1—survival)\textsuperscript{22} and Greenwood 95\% confidence intervals (CIs).\textsuperscript{23} Curves that vary noticeably may indicate systematic differences within the study populations\textsuperscript{21}.

In a third step, we reported the spread (SD) and 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.\textsuperscript{20} Differences in the mean PI indicate differences in overall (predicted) outcome frequency,
reflecting case-mix severity between the two datasets (and revealing the model’s calibration-in-the-large in the Commercial database). Continuous variables were tested with independent t-tests.

We then performed the formal external validation, assessing the prognostic accuracy of the model in the Commercial database. Accordingly, the regression coefficients obtained from our model developed in SLaM (see Supplementary Table 2-8) were applied to each case in the external Commercial database, to generate the PI in the Commercial database. In the case of ethnicity, the ethnic weights were multiplied by their respective regression coefficients to provide an aggregate coefficient for that patient. The sum of an individual’s regression coefficients resulted in an individualised PI. The greater the PI, the higher the risk of the individual developing a psychotic disorder.

Since we were interested in discrimination, the primary outcome measure for this study was the external model performance (accurate predictions discriminate between those with and those without the outcome), defined with the Harrell’s C-index. Harrell’s C is a recommended measure for external validation of Cox models according to established guidelines. Harrell’s C 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 addition, we estimated the overall model performance using the Brier score (average mean squared difference between predicted probabilities and actual outcomes, which also captures calibration and discrimination aspects). Calibration (agreement between observed outcomes and predictions) was assessed using the regression slope of the PI.

As a further exploratory step, we updated the model using the regression slope on the PI as a shrinkage factor for recalibration, in line with the Royston et al. guidelines.

All analyses were conducted in R version 3.3.2 using the survival package, and significance was set to $P < .05$. 
Figure 2-1 Flow chart of the study population

Patients receiving a first ICD-10 index primary diagnosis of nonorganic psychotic disorder within MarketScan in the period 2015-2017 (n=3,828,791)

Excluded due to unavailable data for ethnicity imputation (n=1,398,458)

Final study population (n=2,430,333)

Other missing predictors (age, gender): (n=0)

Number of events i.e. individuals developing an ICD-10 diagnosis of non-organic psychotic disorder (n=24,941); individuals not developing psychotic disorders (n=2,405,392)
<table>
<thead>
<tr>
<th></th>
<th>Commercial (external validation database)</th>
<th>SLaM (original development database)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 2,430,333)</td>
<td>(n = 34,209)</td>
</tr>
<tr>
<td>Age, y</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td></td>
<td>34.2 (16.88)</td>
<td>34.43 (18.89)</td>
</tr>
<tr>
<td>Ethnicity(a)</td>
<td>No. (%)</td>
<td>No. (%)</td>
</tr>
<tr>
<td>Black</td>
<td>0.12 (0.10)</td>
<td>7,055 (22.19)</td>
</tr>
<tr>
<td>White</td>
<td>0.79 (0.11)</td>
<td>18,768 (59.03)</td>
</tr>
<tr>
<td>Asian</td>
<td>0.04 (0.04)</td>
<td>1,149 (3.61)</td>
</tr>
<tr>
<td>Mixed</td>
<td>0.03 (0.01)</td>
<td>1,319 (4.15)</td>
</tr>
<tr>
<td>Other</td>
<td>0.02 (0.03)</td>
<td>3,502 (11.02)</td>
</tr>
<tr>
<td>Sex</td>
<td>No. (%)</td>
<td>No. (%)</td>
</tr>
<tr>
<td>Male</td>
<td>995,262 (40.95)</td>
<td>17,511 (51.20)</td>
</tr>
<tr>
<td>Female</td>
<td>1,435,071 (59.05)</td>
<td>16,688 (48.80)</td>
</tr>
<tr>
<td>Index diagnosis</td>
<td>No. (%)</td>
<td>No. (%)</td>
</tr>
<tr>
<td>CHR-P</td>
<td>-</td>
<td>314 (0.92)</td>
</tr>
<tr>
<td>Acute and transient psychotic disorders</td>
<td>1,316 (0.05)</td>
<td>747 (2.18)</td>
</tr>
<tr>
<td>Substance use disorders</td>
<td>153,401 (6.31)</td>
<td>7,187 (21.01)</td>
</tr>
<tr>
<td>Bipolar mood disorders</td>
<td>64,623 (2.66)</td>
<td>980 (2.86)</td>
</tr>
<tr>
<td>Nonbipolar mood disorders</td>
<td>543,854 (22.38)</td>
<td>6,364 (18.60)</td>
</tr>
<tr>
<td>Anxiety disorders</td>
<td>1,092,893 (44.97)</td>
<td>8,279 (24.20)</td>
</tr>
<tr>
<td>Personality disorders</td>
<td>11,572 (0.48)</td>
<td>1,297 (3.79)</td>
</tr>
<tr>
<td>Developmental disorders</td>
<td>74,072 (3.05)</td>
<td>1,413 (4.13)</td>
</tr>
<tr>
<td>Childhood/adolescence onset disorders</td>
<td>418,316 (17.21)</td>
<td>4,201 (12.28)</td>
</tr>
<tr>
<td>Physiological syndromes</td>
<td>68,476 (2.82)</td>
<td>2,560 (7.48)</td>
</tr>
<tr>
<td>Mental retardation</td>
<td>1,810 (0.07)</td>
<td>867 (2.53)</td>
</tr>
</tbody>
</table>

a) Ethnicity data in Commercial were imputed so they are not directly comparable with SLaM. The means and SDs presented here represent the average proportion of ethnicities in patients’ Metropolitan Statistical Area (MSA)
2.4. RESULTS

2.4.1. COMMERCIAL SAMPLE CHARACTERISTICS

A total of 3,828,791 patients accessing primary or secondary healthcare between January 2015 and December 2017 received an ICD-10 primary index diagnosis of a nonorganic and nonpsychotic mental disorder. 2,430,333 (63.5%) of these individuals could be matched with ethnicity data, and were included in the analysis, as detailed in the study flow-diagram (Figure 2-1). Patients accessing Commercial and included in the current study had an average age of 34.2 years (95%CI: 34.19-34.23), 59% were female, and White ethnicity was particularly common in patients’ MSAs (79%). The most frequent index diagnosis was anxiety disorders (45%). Full sociodemographic information is provided in Table 2-1.

2.4.2. DIFFERENCES BETWEEN THE COMMERCIAL AND SLAM DATABASES

Sociodemographic and service configuration differences

The most important difference is that while the SLaM database contains data on individuals accessing publicly funded secondary mental healthcare, Commercial is limited to individuals covered by employer-sponsored health insurance plans. Compared to the full population, incidence of psychosis may be rarer in those covered by private insurance such as in the Commercial dataset. Similar to the C&I Trust that was the basis of the second external replication study, Commercial did not include CHR-P services; therefore, there were no CHR-P diagnoses. Additional differences are that Commercial data incorporates both primary and secondary healthcare, compared to solely secondary healthcare in SLaM and C&I, as well as the aggregation of ethnicity data as discussed in Methods. The average patient’s age in the Commercial was 0.2 years lower than in SLaM (p=0.03). Compared with SLaM, there was a lower incidence of ATPD, substance use disorders, bipolar mood disorders, personality disorders, developmental disorders, physiological syndromes and mental retardation in the
Commercial dataset. Conversely, there were higher rates of nonbipolar mood disorders, anxiety disorders and childhood/adolescence onset disorders. Finally, there were fewer males in Commercial than in SLaM (Table 2-1).

**Cumulative Risk of Psychosis in Commercial compared with the SLaM derivation dataset**

The average follow-up time in Commercial was 460.89 days (SD=280.04) compared with 1580.64 days (SD=927.72) in SLaM. There were 24,941 (1.03% of the sample size) events (transition to psychosis) in Commercial compared with 1,273 (3.72% of the sample size) in SLaM. The average time to transition to psychosis in those who transitioned was 199.77 days (SD=204.48) in Commercial compared to 664.03 days (SD=621.04) in SLaM. The two-year cumulative risk of psychosis in the Commercial was 1.43% (95%CI: 1.41%-1.45%, with the last transition being observed at 819 days), compared to 2.57% (95%CI: 2.40%-2.75%, with the last transition being observed at 3,246 days) in SLaM. The cumulative incidences curves (Kaplan–Meier) from the Commercial and SLaM datasets are compared in Figure 2. Mean values of the PI within the Commercial and SLaM databases were −1.51 and −1.18, respectively ($P<.001$). SD of the PI in the Commercial and SLaM databases were 0.70 and 0.94, respectively ($P<.001$).
**Figure 2-2** Cumulative incidence (Kaplan–Meier failure function) for risk of development of psychotic disorders in the Commercial Database and SLaM derivation database.

Upper part of the figure: cumulative incidence (Kaplan–Meier failure function) for risk of development of psychotic disorders in the Commercial Database. There were a total of 24,941 events (transition to psychosis): 19,687 in the first 365 days, 4,851 in the interval 366–730 days, 403 in the interval 731–819 days. The last event was observed at 819 days, when 360,396 individuals were still at risk. The cumulative incidence of psychosis was: 0.94 (95%CI: 0.93-0.95) at one year and 1.43 (95%CI: 1.41-1.45) at two years. Lower part of the figure: cumulative incidence (Kaplan–Meier failure function) for risk of development of psychotic disorders in the SLaM derivation database, truncated at 1,460 days for visual comparability. Cumulative incidence of psychosis: 1.67 (95%CI: 1.61-1.89, 30,102 individuals still at risk) at one year, 2.57 (95%CI: 2.40-2.75, 26,337 individuals still at risk) at two years.
2.4.3. EXTERNAL VALIDATION IN THE COMMERCIAL DATABASE

The comparative model performance in the SLaM dataset using aggregate ethnicity data was 0.761 (Table S5). In the Commercial dataset, the model predicted significantly better than chance, with a Harrell’s C of 0.676 (95%CI: 0.672-0.679, Harrell’s C in SLaM=0.79). The two-year Brier score was 0.013 (two-year Brier score in SLaM=0.012). The model did not show major calibration issues, with a regression slope close to 1: 0.93, 95% CI: 0.91–0.94 ($P<.001$).

Updating the model optimized calibration (regression slope=1) but conferred no substantial improvement in model performance (full model specifications are appended in Supplementary Table 2-8).

2.5. DISCUSSION

This is the largest ever replication study of a risk prediction model in psychiatry. The study demonstrates that the transdiagnostic, individualized risk calculator was able to detect individuals at risk of psychosis in an international setting with a prognostic discriminative performance that was significantly above chance.

To our knowledge, this is the largest ever external replication study of a risk calculator not only in early psychosis but also in clinical psychiatry. Importantly, this study included 24,941 events (transitions to psychosis) which are over one hundred times the minimum recommended amount of 100 events required to produce accurate estimates of external prognostic accuracy.\textsuperscript{26,27} The previous largest external validation study of this kind was our first external replication of this calculator conducted in SLaM (n=33,820),\textsuperscript{7} followed by a validation study of a calculator that predicts major depressive disorder (n=29,621)\textsuperscript{28} and by another calculator that predicts risk of violent crime in patients with severe mental illness (n=16,387),\textsuperscript{29} all smaller
than our sample size of 2,40,333. This is a substantial achievement given that prognostic modelling in psychiatry is affected by a severe scarcity of replication efforts,\textsuperscript{30} to the point that replication has become equally as—or even more—important than discovery.\textsuperscript{31} A systematic review and meta-analysis of clinical prediction models for predicting the onset of psychosis in CHR-P people uncovered 91 studies, none of which performed a true external validation of an existing model.\textsuperscript{32} This is the only transdiagnostic clinical prediction model to be externally validated in three different populations (Lewisham & Croydon SLaM NHS Trust, C&I and now Commercial); another risk prediction model for use in CHR-P patients has also received three independent validations.\textsuperscript{33-35} A full list of individualized risk prediction models that have been externally replicated in the field of early psychosis is detailed in Table 2-2.
Table 2-2 Individualized clinical prediction models that have been externally validated for early psychosis.

This table presents key features of the target populations, discrimination/prognostic performance and type of data used in externally validated individualized clinical prediction models for early psychosis. Abbreviations: Population: CHR-P clinical high risk for psychosis, FEP first-episode psychosis; Performance: AUC area under the curve, BAC balanced accuracy, NR not reported; Data: CLIN clinical data, NPSY neuropsychological data, Y yes.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Targets</th>
<th>Population</th>
<th>Derivation sample size (Location)</th>
<th>Performance</th>
<th>Validation sample size (Location)</th>
<th>Performance</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fusar-Poli¹</td>
<td>2016</td>
<td>Detection</td>
<td>CHR-P</td>
<td>321 (UK)</td>
<td>Harrell’s C = 0.66</td>
<td>389 (UK)</td>
<td>Harrell’s C = 0.66</td>
<td>Y</td>
</tr>
<tr>
<td>Fusar-Poli²</td>
<td>2017</td>
<td>Detection</td>
<td>CHR-P</td>
<td>33,820 (UK)</td>
<td>Harrell’s C = 0.80</td>
<td>54,716 (UK), 13,702 (UK)¹</td>
<td>Harrell’s C = 0.79</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Harrell’s C = 0.86</td>
<td>2,430,333 (USA)</td>
<td>Harrell’s C = 0.68</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Harrell’s C = 0.86</td>
<td>63,854 (UK)</td>
<td>Harrell’s C = 0.85</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Refined: Natural language processing⁴</td>
<td>28,297 (UK)</td>
<td></td>
<td>32,430,333 (USA)</td>
<td>Harrell’s C = 0.68</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Refined: Non-linear modelling of age⁵</td>
<td>33,820 (UK)</td>
<td></td>
<td>54,716 (UK)</td>
<td>Harrell’s C = 0.81</td>
<td>Y</td>
</tr>
<tr>
<td>Cannon⁶</td>
<td>2016</td>
<td>Prognosis (Transition)</td>
<td>CHR-P</td>
<td>596 (USA)</td>
<td>Harrell’s C = 0.71</td>
<td>176 (USA)¹, 199 (China)⁹</td>
<td>AUC = 0.79</td>
<td>Y</td>
</tr>
<tr>
<td>Zhang⁷⁸</td>
<td>2019</td>
<td>Prognosis (Transition)</td>
<td>CHR-P</td>
<td>349 (China)</td>
<td>AUC = 0.74</td>
<td>100 (China), 68 (USA)⁹</td>
<td>AUC = 0.80</td>
<td>Y</td>
</tr>
<tr>
<td>Koutsouleris¹⁰</td>
<td>2016</td>
<td>Prognosis (Functioning)</td>
<td>FEP</td>
<td>334 (Europe, Israel)</td>
<td>BAC = 0.75</td>
<td>108 (Europe, Israel)</td>
<td>BAC = 0.72</td>
<td>Y</td>
</tr>
<tr>
<td>Leighton¹²</td>
<td>2019</td>
<td>Prognosis (Functioning)</td>
<td>FEP</td>
<td>83 (UK)</td>
<td>NR</td>
<td>79 (UK)</td>
<td>AUC = 0.88</td>
<td></td>
</tr>
<tr>
<td>Leighton¹³</td>
<td>2019</td>
<td>Prognosis (Remission, Recovery, Quality of life)</td>
<td>FEP</td>
<td>Remission: 673 (UK)</td>
<td>Remission: AUC = 0.70</td>
<td>Social recovery: 829 (UK)</td>
<td>Remission: 131 (UK)</td>
<td>Remission: 338 (Denmark)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Social recovery: 829 (UK)</td>
<td>Remission: AUC = 0.70</td>
<td>Vocational recovery: 807 (UK)</td>
<td>Vocational recovery: AUC = 0.73</td>
<td>Remission: AUC = 0.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Vocational recovery: 807 (UK)</td>
<td>Remission: AUC = 0.70</td>
<td>Quality of life: 729 (UK)</td>
<td>Vocational recovery: AUC = 0.73</td>
<td>Social recovery: AUC = 0.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Quality of life: 729 (UK)</td>
<td>Remission: AUC = 0.70</td>
<td>Remission: 142 (UK)</td>
<td>Remission: Vocational recovery: AUC = 0.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Quality of life: 47 (UK)</td>
<td>Quality of life: AUC = 0.68</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

44
The additional strength of this study is that it provides further empirical support for the use of EHRs in the context of precision psychiatry. Transporting risk prediction models across different EHRs representing heterogeneous clinical settings is complex because they reflect underlying differences in the patient population. A first empirical challenge is the availability of predictors and outcomes. The vast majority of predictors were available in the Commercial database, with the exception of ethnicity; patient-level ethnicity variables were computed to compensate for this. There was also a shorter follow-up time in Commercial compared to SLaM, as ICD-10 was only integrated into United States healthcare on 1 October 2015. Use of ICD-9 diagnoses was considered to extend follow-up but converting diagnostic clusters to ICD-9 proved inexact and therefore inappropriate. A second challenge is to quantify the differences between development and validation databases to interpret the performance of a risk model in the context of external validation. For example, compared with SLaM, where the model was developed, there were apparent differences in sociodemographic characteristics in Commercial (fewer males and fewer patients of Black ethnicity and different frequency of ICD diagnoses, reflected by smaller spread of the PI) and time to event (shorter). Furthermore, similar to our second replication in C&I, there were no CHR-P services in Commercial and, therefore, no CHR-P designations. However, as ATPD diagnoses are typically not made in CHR-P or early intervention services, the number of ATPD diagnoses in Commercial are unlikely to be affected by this difference in service configuration. Because of this case-mix, the incidence of psychosis was about half in Commercial (1.43/2.57 at two years, reflected by a lower mean value of the PI). The most important difference is that, while previous replications were performed in data collected from publicly funded secondary mental healthcare alone, the Commercial database was composed of both primary and secondary healthcare data composed of commercially insured patients. Given such relevant differences, it was expected that the risk calculator could not be easily transported to the
Commercial setting and that it would achieve a lower prognostic performance and calibration than that observed in the first two external validations.

Despite these differences in clinical setting and populations, the overall prognostic accuracy of the transdiagnostic, clinically-based risk calculator remained significantly above chance. As expected, the level of prognostic performance (Harrell’s C=0.68) was suboptimal and lower than our previous external validation (Harrell’s C=0.73). Yet, this level of accuracy is comparable to that of structural neuroimaging methods (i.e. grey matter volume) to detect a first-episode of psychosis at the individual level, with accuracies ranging from 0.5 to 0.63. A recent machine-learning study externally validated a risk calculator to predict treatment outcome in depression in 151 patients. The study reported a one year prognostic accuracy of 0.59 and concluded that, if implemented at scale, performance even only significantly above chance can be considered to be clinically useful. Given that our risk calculator has been developed on real-world EHR data, it offers the potential for automatically screening large mental health populations. Psychiatry is undergoing a digital revolution, and there is an ongoing expansion of EHR adoption worldwide. More to this point, this risk calculator was evidently developed with a clear vision of future implementation as decision support in clinical routine and is currently being piloted in this capacity. For example, it uses simple predictors that can easily be understood by clinicians, as compared to complex black-box machine-learning-derived algorithms. Furthermore, harnessing data from EHRs is cheaper than other methods such as patient recruitment, because most of the predictors are available as part of clinical routine. There are no competing algorithms (CHR-P instruments are not usable for screening purposes) to screen the at-risk population at scale. Other risk prediction tools in early psychosis have shown promise, however they predominantly rely on clinical symptom scores, which means they are more financially and labour intensive than this tool; potential for automation is therefore limited. Moreover, these tools are focused on identifying transition to psychosis and
are reliant on prior identification of CHR-P, whereas our tool is able to predict psychosis risk transdiagnostically outside of this designation. Thus, there is potential benefit in utilising this risk calculator to screen for psychosis risk in large numbers.

There is scope for optimisation of the current risk calculator through stepped risk stratification and model refinement. As a first step, this risk calculator could be deployed in a screening pathway where an individual’s risk is calculated upon entry into secondary mental health services. Individuals flagged by our risk calculator as being at risk for psychosis would progress to a more thorough clinical CHR-P assessment in the context of a staged sequential risk assessment.48,49 A potential further step would be using additional information (environmental, genetic or biomarkers) to improve prognostic accuracy further,48,49 refine estimates of individuals’ risk and stratify them accordingly. This is in keeping with the current clinical staging model of early psychosis, which aims to improve preventative care and reduce the duration of untreated psychosis to improve outcomes.1 In addition to its clinical utility, this risk calculator could improve CHR-P research by aiding recruitment for much needed large-scale international collaborations in the vein of the HARMONY project, incorporating NAPLS (https://campuspress.yale.edu/napls/), PRONIA (https://www.pronia.eu/) and PSYSCAN (http://psyscan.eu), and the proposed 26-site ProNET cohort study. Furthermore, this prognostic model can be refined. In companion studies, we have tested whether using machine-learning methods and expanding the range of,50 or redefining,51 predictors might improve the prognostic accuracy of this risk calculator.

The limitations of this study are largely inherited from the original study. We did not employ structured psychometric interviews to ascertain the type of emerging psychotic diagnoses at follow-up. However, we predicted psychotic disorders rather than specific ICD-10 diagnoses, a category which has good prognostic stability.52 Therefore, while the psychotic diagnoses in our analyses are high in ecological validity (i.e. they represent real-world clinical practice), they have
not been subjected to formal validation with research-based criteria. However, the use of structured diagnostic interviews can lead to selection biases, decreasing the transportability of models.\textsuperscript{53} There is also meta-analytical evidence indicating that within psychotic disorders, administrative data recorded in clinical registers are generally predictive of true validated diagnoses.\textsuperscript{54}

Other limitations were inherent in the Commercial database, mostly due to the lack of patient-level ethnicity data and a short follow-up time. These two issues reduced the prognostic performance of the model \textit{a priori}, in particular considering that risk for psychosis may well extend beyond two years.\textsuperscript{55} It is therefore possible that prognostic performance of this model in the longer term may actually be better than the performance reported here. A further limitation is that the study team for this replication are not completely independent from the team who completed the original study,\textsuperscript{56} which is particularly relevant given the support of a pharmaceutical company. As this study involved a large commercial dataset and a refined version of the model, it was logistically impossible to conduct this research independently from the original team. To mitigate against this overlap, we adhered to the Royston,\textsuperscript{21} RECORD,\textsuperscript{13} and TRIPOD\textsuperscript{19} guidelines to ensure transparency. Finally, although we welcome further external validation studies, it must be noted that even strong replication does not automatically imply the potential for successful adoption in clinical or public health practice. Ideally, randomised clinical trials or economic modelling are needed to assess whether our risk calculator effectively improves patient outcomes.

\textbf{Conclusion}

The largest international external replication of an individualized prognostic model in psychiatry confirms that precision medicine in this discipline is feasible even at large scale. The transdiagnostic, individualized, clinically-based risk calculator is potentially transportable on an international scale to improve the automatic detection of individuals at risk of psychosis. Further
research should refine the model and test the benefit of implementing this risk prediction model in clinical routine.

**Author Contributions**

P.F-P. developed the original model, validated it and conceived the current study. D.O., M.B., L.J., K.T.J. and P.F-P. developed the protocol. D.O. and C.M.J.W. wrote all analysis scripts and led the analyses. D.O. drafted the first version of the current manuscript. M.B., L.J., B.K., A.W., K.T.J., J.I., D.S. and L.L.R. advised on data organisation, cleaning and statistical analysis. D.O., P.F-P. and P.M. interpreted the results of the analyses. All authors approved the final manuscript.

**Acknowledgements**

We declare no conflict of interest in relation to the current manuscript. D.O. 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. P.F-P. is supported by a research grant from H. Lundbeck A/S. These funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.
2.6. REFERENCES


Baker JT, Germinne LT, Ressler KJ, Rauch SL, Carlezon WA. Digital devices and continuous telemetry:


2.7. SUPPLEMENTARY DATA


Supplementary Table 2.3 REporting of studies Conducted using Observational Routinely-collected Data (RECORD) statement

Supplementary Methods 2.1 Description of IBM® MarketScan® Commercial Database

Supplementary Methods 2.2 Definition of emerging ICD-10 primary diagnosis of non-organic psychotic disorder

Supplementary Table 2.4 Predictor definitions: Primary index diagnoses of non-organic and non-psychotic mental disorder formulated at baseline (time of the first contact with the NHS Trust).

Supplementary Table 2.5 Predictor definitions: self-assigned ethnicity

Supplementary Methods 2.3 Definition of lookback period

Supplementary Table 2.6 Model performance of the Transdiagnostic Individualised Clinically-based Risk Calculator for the Automatic Detection of Individuals at Risk and the Prediction of Psychosis (revised version) in the original South London and Maudsley NHS Foundation Trust (SLaM) derivation and validation datasets.

Supplementary Table 2.7 Transdiagnostic Individualised Clinically-based Risk Calculator for the Automatic Detection of Individuals at Risk and the Prediction of Psychosis (revised version), original derivation dataset (SLaM boroughs Lambeth & Southwark). This model has been used for external validation in the current study.

Supplementary Table 2.8 Model performance of the Transdiagnostic Individualised Clinically-based Risk Calculator for the Automatic Detection of Individuals at Risk and the Prediction of Psychosis (revised version) in the original SLaM derivation dataset with removed ethnicity predictor and with aggregate ethnicity coefficients.

Supplementary References
### Supplementary Table 2-3 The REporting of studies Conducted using Observational Routinely-collected Data (RECORD) statement – checklist of items, extended from the STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) statement, that should be reported in observational studies using routinely collected health data.

<table>
<thead>
<tr>
<th>Item no.</th>
<th>STROBE items</th>
<th>Location in manuscript where items are reported</th>
<th>RECORD items</th>
<th>Location in manuscript where items are reported</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title and abstract</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>(a) Indicate the study’s design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found.</td>
<td>Abstract</td>
<td>RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.</td>
<td>Abstract</td>
</tr>
<tr>
<td><strong>Introduction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Explain the scientific background and rationale for the investigation being reported.</td>
<td>Introduction</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Objectives</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>State specific objectives, including any prespecified hypotheses.</td>
<td>Abstract, Introduction, Materials and methods</td>
<td>RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided. RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided.</td>
<td>Materials and methods, referenced</td>
</tr>
<tr>
<td><strong>Methods</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Present key elements of study design early in the paper.</td>
<td>Abstract</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection.</td>
<td>Abstract, Materials and methods</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>(a) Cohort study - Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up Case-control study - Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls</td>
<td>Abstract, Materials and methods</td>
<td>RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided. RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided.</td>
<td>Materials and methods, referenced</td>
</tr>
<tr>
<td>Variables</td>
<td>7</td>
<td>Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.</td>
<td>Materials and methods, Table 2-1</td>
<td>RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Data sources/measurement</td>
<td>8</td>
<td>For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group.</td>
<td>Materials and methods</td>
<td></td>
</tr>
<tr>
<td>Bias</td>
<td>9</td>
<td>Describe any efforts to address potential sources of bias.</td>
<td>Materials and methods</td>
<td></td>
</tr>
<tr>
<td>Study size</td>
<td>10</td>
<td>Explain how the study size was arrived at.</td>
<td>Materials and methods, Figure 2-1</td>
<td></td>
</tr>
<tr>
<td>Quantitative variables</td>
<td>11</td>
<td>Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why.</td>
<td>Statistical analysis</td>
<td></td>
</tr>
<tr>
<td>Statistical methods</td>
<td>12</td>
<td>(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) Cohort study - If applicable, explain how loss to follow-up was addressed (e) Case-control study - If applicable, explain how matching of cases and controls was addressed (f) Cross-sectional study - If applicable, describe analytical methods taking account of sampling strategy (g) Describe any sensitivity analyses</td>
<td>Statistical analysis</td>
<td></td>
</tr>
</tbody>
</table>
### Data access and cleaning methods

**RECORD 12.1:** Authors should describe the extent to which the investigators had access to the database population used to create the study population.

**RECORD 12.2:** Authors should provide information on the data cleaning methods used in the study.

### Linkage

**RECORD 12.3:** State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.

### Results

#### Participants

| 13 | (a) Report the numbers of individuals at each stage of the study (e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed) | Results | RECORD 13.1: Describe in detail the selection of the persons included in the study (i.e., study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram. |
| 13 | (b) Give reasons for non-participation at each stage. | NA | |
| 13 | (c) Consider use of a flow diagram. | Figure 1 |

#### Descriptive data

| 14 | (a) Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders | Table 1 |
| 14 | (b) Indicate the number of participants with missing data for each variable of interest | Figure 2-1 |
| 14 | (c) Cohort study - summarise follow-up time (e.g., average and total amount). | Results |

#### Outcome data

| 15 | **Cohort study** - Report numbers of outcome events or summary measures over time | Results and Figure 2-2 |
| 15 | **Case-control study** - Report numbers in each exposure category, or summary measures of exposure | NA |
| 15 | **Cross-sectional study** - Report numbers of outcome events or summary measures | NA |

#### Main results

<p>| 16 | (a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included | Results |
| 16 | (b) Report category boundaries when continuous variables were categorized | NA |
| 16 | (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period. | Figure 2-2 |</p>
<table>
<thead>
<tr>
<th>Other analyses</th>
<th>17</th>
<th>Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses.</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Discussion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Key results</td>
<td>18</td>
<td>Summarise key results with reference to study objectives.</td>
<td>First paragraph of discussion</td>
</tr>
<tr>
<td>Limitations</td>
<td>19</td>
<td>Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.</td>
<td>Discussion (strengths and weaknesses of the study section) RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.</td>
</tr>
<tr>
<td>Interpretation</td>
<td>20</td>
<td>Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.</td>
<td>Discussion (interpretation of findings and strengths and weaknesses of the study sections)</td>
</tr>
<tr>
<td>Generalisability</td>
<td>21</td>
<td>Discuss the generalisability (external validity) of the study results.</td>
<td>Discussion</td>
</tr>
<tr>
<td><strong>Other information</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Funding</td>
<td>22</td>
<td>Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based.</td>
<td>Financial support statement in the abstract</td>
</tr>
<tr>
<td>Accessibility of protocol, raw data, and programming code</td>
<td></td>
<td></td>
<td>RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code. Statistical analysis, Methods 2</td>
</tr>
</tbody>
</table>
Supplementary Methods 2-1 Description of IBM® MarketScan® Commercial Database

This database is based on data from Truven MarketScan Commercial Claims and Encounters Database ("Commercial database" in short), which contains information from active employees, early retirees, Consolidated Omnibus Budget Reconciliation Act continues, and dependents insured by employer-sponsored plans (i.e. individuals not eligible for Medicare). It incorporates data from hundreds of payers, including commercial insurance companies, Blue Cross Blue Shield plans, and third-party administrators. It incorporates data from hundreds of payers, including commercial insurance companies and third-party administrators.
**Supplementary Methods 2. Definition of emerging ICD-10 primary diagnosis of non-organic psychotic disorder**

This was defined as the emergence of the first ICD-10 primary diagnosis of non-organic psychotic disorder after the index diagnosis as recorded in the local electronic medical records: schizophrenia spectrum psychoses (schizophrenia [F20.x, except F20.4/F20.5], schizoaffective disorder [F25.x], delusional disorders [F22.x, F24], Acute and Transient Psychotic Disorders [ATPD, F23.x]), unspecified nonorganic psychosis (F28/F29), psychotic disorders due to psychoactive substance use ([F10-F19].5), and affective psychoses (mania with psychotic symptoms [F30.2], bipolar affective disorder with psychotic symptoms [F31.2, F31.5], and depression with psychotic symptoms [F32.3/F33.3]). Accordingly, baseline ICD-10 psychotic disorders were excluded, with the exception of ATPD (F23.x), which are, by definition, clinically remitting and non-psychotic within three months (short-lived). The rationale for including the ATPD is due to the fact that this group is prognostically similar to the Brief Limited Intermittent Psychotic Symptom (BLIPS) or Brief Intermittent Psychotic Symptoms (BIPS) subgroups of the CHR-P construct (for details on these competing operationalisation see previous publications on the diagnostic and prognostic significance of BLIPS\(^1,2\)). On a diagnostic level, about two thirds (68%) of BLIPS meet ATPD criteria\(^1\).
**Supplementary Table 2-4** Predictor definitions: Primary index diagnoses of non-organic and non-psychotic mental disorder formulated at baseline (time of the first contact with the NHS Trust).

<table>
<thead>
<tr>
<th>Primary index diagnosis</th>
<th>ICD-10 code</th>
<th>ICD-10 diagnosis name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute and transient psychotic disorders</td>
<td>F23.x</td>
<td>Acute and transient psychotic disorders</td>
</tr>
<tr>
<td>Bipolar mood disorders</td>
<td>F31.x (excluding F31.2 and F31.5) F34.0 F30.x (excluding F30.2)</td>
<td>Non-psychotic bipolar disorder Cyclothymia Non-psychotic mania or hypomania</td>
</tr>
<tr>
<td>Non-bipolar mood disorders</td>
<td>[F32-F33].x (excluding F32.3 and F33.3) F34.1 F34.8, F34.9, F38.x, F39</td>
<td>Non-psychotic depressive disorder Dysthymia Unspecified mood disorders</td>
</tr>
<tr>
<td>Anxiety disorders</td>
<td>F40.x F41.0 F41.1 F41.2-F41.9 F42.x F43.x F44.x F45.x F48.x</td>
<td>Phobic anxiety disorders Panic disorder Generalised anxiety disorder Other anxiety disorders Obsessive compulsive disorders Reaction to severe stress, and adjustment disorders Dissociative [conversion] disorders Somatoform disorders Other neurotic disorders</td>
</tr>
<tr>
<td>Personality disorders</td>
<td>F60.0 F60.1 F60.2 F60.3 F60.4 F60.5 F60.6</td>
<td>Paranoid personality disorder Schizoid personality disorder Dissocial personality disorder Emotionally unstable personality disorder Histrionic personality disorder Anankastic personality disorder Anxious [avoidant] personality disorder</td>
</tr>
<tr>
<td>Primary index diagnosis</td>
<td>ICD-10 code</td>
<td>ICD-10 diagnosis name</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>F60.7</td>
<td>Dependent personality disorder</td>
<td></td>
</tr>
<tr>
<td>F60.8-F60.9, F61, F62.x, F68.x, F69</td>
<td>Other personality disorders</td>
<td></td>
</tr>
<tr>
<td>F21</td>
<td>Schizotypal Disorder</td>
<td></td>
</tr>
<tr>
<td>F63.x</td>
<td>Habit and impulse disorders</td>
<td></td>
</tr>
<tr>
<td>F64.x, F65.x, F66.x</td>
<td>Sexual disorders</td>
<td></td>
</tr>
<tr>
<td>F80.x</td>
<td>Specific developmental disorders of speech and language</td>
<td></td>
</tr>
<tr>
<td>F81.x, F82, F83</td>
<td>Other specific developmental disorders</td>
<td></td>
</tr>
<tr>
<td>F84.x</td>
<td>Pervasive developmental disorders</td>
<td></td>
</tr>
<tr>
<td>F88, F89</td>
<td>Other and unspecified disorders of psychological development</td>
<td></td>
</tr>
<tr>
<td>F90.x</td>
<td>Hyperkinetic disorders</td>
<td></td>
</tr>
<tr>
<td>F91.x</td>
<td>Conduct disorders</td>
<td></td>
</tr>
<tr>
<td>F92.x, F93.x, F94.x, F98.x, F95.x</td>
<td>Other emotional and behavioural disorders with childhood or adolescence onset</td>
<td></td>
</tr>
<tr>
<td>F50.x</td>
<td>Eating disorders</td>
<td></td>
</tr>
<tr>
<td>F51.x</td>
<td>Nonorganic sleep disorders</td>
<td></td>
</tr>
<tr>
<td>F52.x</td>
<td>Sexual dysfunction, not caused by organic disorder or disease</td>
<td></td>
</tr>
<tr>
<td>F53.x (excluding F53.1)</td>
<td>Non-psychotic Mental and behavioural disorders associated with the puerperium, not elsewhere classified</td>
<td></td>
</tr>
<tr>
<td>F54.x, F55, F59</td>
<td>Other physiological syndromes</td>
<td></td>
</tr>
<tr>
<td>F70.x</td>
<td>Mild mental retardation</td>
<td></td>
</tr>
<tr>
<td>F71.x</td>
<td>Moderate mental retardation</td>
<td></td>
</tr>
<tr>
<td>F72.x</td>
<td>Severe mental retardation</td>
<td></td>
</tr>
<tr>
<td>F73.x</td>
<td>Profound mental retardation</td>
<td></td>
</tr>
<tr>
<td>F78.x</td>
<td>Other mental retardation</td>
<td></td>
</tr>
<tr>
<td>F79.x</td>
<td>Unspecified mental retardation</td>
<td></td>
</tr>
</tbody>
</table>

*F00-F09 organic mental disorders and all psychotic disorders other than F23.x were excluded*
Supplementary Table 2-5  Predictor definitions: self-assigned ethnicity. Individual-level ethnicity data was not available for the MarketScan database so was imputed using Integrated Public Use Microdata Series (IPUMS) as discussed in the methods section.

<table>
<thead>
<tr>
<th>Ethnic group</th>
<th>Ethnicity as recorded in patient electronic health records</th>
<th>Ethnicity as recorded in IPUMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>Black or Black British - African</td>
<td>prop_Black_African_American</td>
</tr>
<tr>
<td></td>
<td>Black or Black British - Caribbean</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Black or Black British - Any other</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Black background</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>White - British</td>
<td>prop_White</td>
</tr>
<tr>
<td></td>
<td>White - Irish</td>
<td></td>
</tr>
<tr>
<td></td>
<td>White - Any other White background</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>Asian or Asian British - Bangladeshi</td>
<td>prop_Japanese</td>
</tr>
<tr>
<td></td>
<td>Asian or Asian British - Indian</td>
<td>prop_Chinese</td>
</tr>
<tr>
<td></td>
<td>Asian or Asian British - Pakistani</td>
<td>prop_Other_Asian_or_Pacific_Isla</td>
</tr>
<tr>
<td></td>
<td>Asian or Asian British - Any other</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asian background</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other Ethnic Groups - Chinese</td>
<td></td>
</tr>
<tr>
<td>Mixed</td>
<td>Mixed - White and Asian</td>
<td>prop_Two_major_races</td>
</tr>
<tr>
<td></td>
<td>Mixed - White and Black African</td>
<td>prop_Three_or_more_major_races</td>
</tr>
<tr>
<td></td>
<td>Mixed - White and Black Caribbean</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mixed - Any other mixed background</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>Other Ethnic Groups - Any other ethnic group</td>
<td>prop_American_Indian_or_Alaska_N</td>
</tr>
<tr>
<td></td>
<td>Not Known</td>
<td>prop_Other_race__nec</td>
</tr>
<tr>
<td>Missing</td>
<td>Not Recorded</td>
<td></td>
</tr>
</tbody>
</table>
Supplementary Methods 2-3 Definition of lookback period
A look back period is the length of an individual’s medical history prior to index diagnosis. This means that additional diagnoses, potentially unreported at index, are more likely to be detected. In particular, this limits discharge coding bias, where less serious, chronic conditions may be unrecorded within Electronic Health Records (EHRs) when a patient is being treated for an acute episode, and, likewise, if acute episodes are missed due to ongoing chronic symptomology. As such, a look back period of six months was applied to each patient to ensure stability of diagnoses.
**Supplementary Table 2-6** Model performance of the Transdiagnostic Individualised Clinically-based Risk Calculator for the Automatic Detection of Individuals at Risk and the Prediction of Psychosis (revised version) in the original South London and Maudsley NHS Foundation Trust (SLaM) derivation and validation datasets.

<table>
<thead>
<tr>
<th>Performance measure</th>
<th>Derivation (n = 34,209) (^{(a)})</th>
<th>Validation (n = 54,716) (^{(b)})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean 95%CI</td>
<td>Mean 95%CI</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brier ((^{(c)}))</td>
<td>0.028</td>
<td>0.021</td>
</tr>
<tr>
<td>R(^{2})</td>
<td>0.746 0.704-0.785</td>
<td>0.719 0.673-0.761</td>
</tr>
<tr>
<td><strong>Discrimination</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harrell's C</td>
<td>0.809 0.795-0.822</td>
<td>0.790 0.775-0.806</td>
</tr>
<tr>
<td><strong>Calibration</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calibration slope</td>
<td>1</td>
<td>0.968 0.929-1.015</td>
</tr>
</tbody>
</table>

\(^{(a)}\) The sample size is larger than the original 33,820 because individuals with psychosis onset within the first 3 months since baseline were not excluded.

\(^{(b)}\) The sample size of the external validation matches that used in the original study to facilitate comparability.

\(^{(c)}\) at 10-years.
**Supplementary Table 2-7** Transdiagnostic Individualised Clinically-based Risk Calculator for the Automatic Detection of Individuals at Risk and the Prediction of Psychosis (revised version), original derivation dataset (SLaM boroughs Lambeth & Southwark). This model has been used for external validation in the current study.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Multivariable model</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta coefficient</td>
<td>95% CI</td>
<td></td>
<td>P</td>
</tr>
<tr>
<td>Age (years)</td>
<td>0.010</td>
<td>0.005</td>
<td>0.143</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.457</td>
<td>0.185</td>
<td>0.730</td>
<td>0.001</td>
</tr>
<tr>
<td>Female (R)</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age by gender (male)</td>
<td>-0.009</td>
<td>-0.015</td>
<td>-0.002</td>
<td>0.009</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White (R)</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0.995</td>
<td>0.865</td>
<td>1.125</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Asian</td>
<td>0.487</td>
<td>0.207</td>
<td>0.767</td>
<td>0.001</td>
</tr>
<tr>
<td>Mixed</td>
<td>0.686</td>
<td>0.372</td>
<td>1.000</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Other</td>
<td>0.340</td>
<td>0.143</td>
<td>0.537</td>
<td>0.001</td>
</tr>
<tr>
<td>ICD-10 diagnostic spectra</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHR-P (R)</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute and transient psychotic</td>
<td>1.169</td>
<td>0.876</td>
<td>1.464</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>disorders</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Substance use disorders</td>
<td>-1.748</td>
<td>-2.065</td>
<td>-1.431</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Bipolar mood disorders</td>
<td>0.003</td>
<td>-0.327</td>
<td>0.333</td>
<td>0.986</td>
</tr>
<tr>
<td>Non-bipolar mood disorders</td>
<td>-1.560</td>
<td>-1.876</td>
<td>-1.245</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Anxiety disorders</td>
<td>-2.006</td>
<td>-2.320</td>
<td>-1.691</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Personality disorders</td>
<td>-1.363</td>
<td>-1.754</td>
<td>-0.971</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Developmental disorders</td>
<td>-3.337</td>
<td>-4.018</td>
<td>-2.656</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Childhood/adolescence onset</td>
<td>-3.200</td>
<td>-3.641</td>
<td>-2.753</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>disorders</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physiological syndromes</td>
<td>-2.310</td>
<td>-2.764</td>
<td>-1.856</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mental retardation</td>
<td>-2.326</td>
<td>-2.864</td>
<td>-1.788</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

(R) Reference category
**Supplementary Table 2-8** Model performance of the Transdiagnostic Individualised Clinically-based Risk Calculator for the Automatic Detection of Individuals at Risk and the Prediction of Psychosis (revised version) in the original SLaM derivation dataset with removed ethnicity predictor and with aggregate ethnicity coefficients.

<table>
<thead>
<tr>
<th>Performance measure</th>
<th>Mean</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No ethnicity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harrell's C (mean, 95%CI)</td>
<td>0.790</td>
<td>0.776-0.803</td>
</tr>
<tr>
<td><strong>Aggregate ethnicity coefficients</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harrell's C (mean, 95%CI)</td>
<td>0.761</td>
<td>0.745-0.776</td>
</tr>
</tbody>
</table>
2.8. SUPPLEMENTARY REFERENCES


3. REAL-WORLD IMPLEMENTATION OF PRECISION PSYCHIATRY: TRANSDIAGNOSTIC RISK CALCULATOR FOR THE AUTOMATIC DETECTION OF INDIVIDUALS AT-RISK OF PSYCHOSIS

Published in Schizophrenia Research 2020

doi: https://doi.org/10.1016/j.schres.2020.05.007

Dominic Oliver¹, Giulia Spada¹, Craig Colling², Matthew Broadbent², Helen Baldwin¹,², Rashmi Patel³, Robert Stewart⁴,⁶, Daniel Stahl⁸, Richard Dobson²,⁴,⁵,⁷, Philip McGuire³,⁹, Paolo Fusar-Poli¹,²,⁹,¹⁰*

¹ Early Psychosis: Interventions and Clinical-detection (EPIC) Lab, Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom
² National Institute for Health Research, Maudsley Biomedical Research Centre, South London and Maudsley National Health Service (NHS) Foundation Trust, London, United Kingdom;
³ Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom
⁴ Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom
⁵ Institute of Health Informatics Research, University College London, London, United Kingdom
⁶ South London and Maudsley Foundation Trust, London, United Kingdom
⁷ Health Data Research UK London, University College London, London, United Kingdom
⁸ Department of Biostatistics, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, United Kingdom
⁹ OASIS Service, South London and Maudsley National Health Service (NHS) Foundation Trust, London, United Kingdom
¹⁰ Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy

Permissions
Copyright Oliver et al., 2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0). The license can be viewed at: https://creativecommons.org/licenses/by/4.0/legalcode. No changes were made in the reproduction of this article.
3.1. ABSTRACT

Background
Risk estimation models integrated into Electronic Health Records (EHRs) can deliver innovative approaches in psychiatry, but clinicians’ endorsement and their real-world usability are unknown. This study aimed to investigate the real-world feasibility of implementing a transdiagnostic, clinically-based, lifespan-inclusive, individualised risk calculator to automatically screen EHRs and detect individuals at-risk for psychosis.

Methods

Feasibility implementation study encompassing an in-vitro phase (March 2018 to May 2018) and in-vivo phase (May 2018 to April 2019). The in-vitro phase addressed implementation barriers and embedded the risk calculator (predictors: age, gender, ethnicity, index cluster diagnosis, age*gender) into the local EHR. The in-vivo phase investigated the real-world feasibility of screening all individuals accessing secondary mental healthcare at the South London and Maudsley NHS Trust (SLaM). The primary outcome was adherence of clinicians to automatic EHR screening, defined by the proportion of clinicians who responded to alerts from the risk calculator, over those contacted.

Results

in-vitro phase: implementation barriers were identified/overcome with clinician and service user engagement, and the calculator was successfully integrated into the local EHR. in-vivo phase: 3,722 individuals were automatically screened and 115 were detected. Clinician adherence was 74% without outreach and 85% with outreach. One-third of clinicians responded to the first email (37.1%) or phone calls (33.7%). Among those detected, cumulative risk of developing psychosis was 0.12 at six-month follow-up.

Conclusion

This is the first implementation study suggesting that combining precision psychiatry and EHR methods to improve detection of individuals with emerging psychosis is feasible. Future psychiatric implementation research is urgently needed.
3.2. INTRODUCTION

Precision medicine and digital health are two pillars of contemporary clinical research in medicine and psychiatry.\(^1\)\(^-\)\(^7\) Precision medicine involves the development and validation of individualised risk estimation models to estimate several clinical outcomes of interest.\(^8\) Digital health approaches can involve Electronic Health Records (EHRs),\(^9\) which represent real-world clinical information (e.g. diagnoses, treatment plans, prescriptions) and are increasingly adopted across healthcare systems.\(^10\),\(^11\) Despite their potential, the use of precision medicine in EHRs has not yet entered clinical practice in psychiatry,\(^12\),\(^13\) highlighting a clear implementation challenge.

Implementation research is the scientific study of methods translating research findings into practical, useful outcomes; it seeks to understand and work within real-world conditions, rather than trying to control for them.\(^14\)\(^-\)\(^16\) Implementation research aims at solving a wide range of practical problems relating to the real-world usability of precision medicine and digital health in clinical practice. For example, risk estimation systems are unlikely to impact clinical pathways unless they are used by clinicians in day-to-day practice;\(^17\) clinicians’ compliance with the recommendations made by a risk calculator represents the first key barrier to implementation.\(^18\),\(^19\) Implementation research of precision medicine in EHRs—not only in psychiatry—is still in its infancy, and very few examples are available.\(^20\),\(^21\)

The current study addresses this gap of knowledge by focusing on prevention of psychosis in individuals at clinical high risk (CHR-P),\(^22\),\(^23\) which is one of the most established preventive approaches in psychiatry. These individuals accumulate risk factors for the disorder\(^24\),\(^25\) that lead to subtle symptoms,\(^26\) help-seeking behaviours,\(^27\) functional impairment\(^28\) and ultimately a heightened risk of developing psychosis.\(^29\) Prevention of psychosis in CHR-P individuals is limited by an insufficient ability to detect them, with only 5-12% identified before the first episode.\(^22\),\(^30\),\(^31\) We have previously developed a pragmatic, clinically-based, lifespan-inclusive, individualised transdiagnostic risk calculator to improve the detection of individuals at risk for psychosis in secondary mental healthcare at scale.\(^32\) This calculator uses five routinely collected variables in EHRs selected a priori based on meta-analytical evidence (age, gender, ethnicity, ICD-10 diagnosis and age*gender interaction) to estimate individualised, annualised risk of developing a psychotic disorder within six years of index diagnosis. This calculator has already
demonstrated adequate external prognostic accuracy in two different independent EHRs (n=54,716, Harrell’s C=0.79; n=13,702, Harrell’s C=0.73). Given the replication crisis in science, these findings represent a relevant result towards implementation. Because this calculator uses clinical information from EHRs, it is capable of improving real-world detection of individuals at-risk of psychosis by automatically screening large populations from multiple mental healthcare providers. Building on these results, we present the feasibility implementation study of this calculator. The primary aim of this study was to test the feasibility of the risk calculator in real-world clinical practice.

3.3. METHODS
This study was approved by the East of England - Cambridgeshire and Hertfordshire Research Ethics Committee (Reference number: 18/EE/0066) and by the SLaM Caldicott Guardian. The protocol for this study was published and is reported according to STROBE guidelines (Supplementary Table 3-4). We developed an innovative two-phase methodology for implementing risk calculators in EHRs (Figure 3-1).

3.3.1. IN-VITRO PHASE
Since this phase was conducted using data from the local EHR and without contacting clinicians or patients, it was termed “in-vitro”. This phase had two manifold aims: (i) to address implementation barriers according to the Consolidated Framework for Implementation Research (CFIR) (Figure 3-2), and (ii) to integrate the transdiagnostic risk calculator into the local EHR.
Figure 3-1 Overview of *in-vitro* and *in-vivo* phases

1. Addressed implementation barriers

2. Implemented the risk calculator in local EHR

1. Automatic estimate of predicted risk of psychosis

   Is patient over predicted risk threshold?

   No → Not recommended for further assessment for psychosis risk

   Yes → For any patient accessing SLaM services for the first time and receiving any ICD-10 diagnosis of non-psychotic mental disorder

2. Responsible clinicians received alerts informing them of individual's risk recommending further assessment

   Is referral initiated?

   No → Patient not referred. Treatment as usual

   Yes

3. Patients contacted and invited to undergo a standard assessment for psychosis risk (CHR-P)

4. Upon completion of the assessment, researchers inform the responsible clinicians

**in-vitro phase**
March 2018 - May 2018

**in-vivo phase**
May 2018 - May 2019
Figure 3-2 CFIR Implementation barriers addressed during the in-vitro phase.

The CFIR frameworks identifies five core implementation domains: characteristics of the intervention (the ‘core components’—that is, the essential elements of the intervention—and the ‘adaptable periphery’—that is, the adaptable elements in which the intervention occurs); outer setting (the ‘economic, political, and social context within which an organisation resides’); inner setting (the ‘structural, political, and cultural context through which the intervention proceeds’ and the relationship between these elements); individuals (the individuals responsible for carrying out the intervention or otherwise related to the intervention, their agency, and their relationships to each other and the intervention) and process (the active process through which the desired changes are achieved).³⁸

<table>
<thead>
<tr>
<th>Core Components</th>
<th>Adaptable Periphery</th>
<th>Inner Setting</th>
<th>Process</th>
<th>Individuals</th>
<th>Outer Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Lifespan-inclusive</td>
<td>• Adapted calculator for prospective use</td>
<td>• Use of predictors widely available in clinical routine and EHR</td>
<td>• Developed a digital EHR interface to better communicate psychosis risk to clinicians</td>
<td>• Collected patients’ feedback</td>
<td>• Replicated in other NHS Trusts</td>
</tr>
<tr>
<td>• Individualised</td>
<td></td>
<td>• Consent for Contact (C4C)</td>
<td></td>
<td>• Conducted workshops with local clinicians</td>
<td></td>
</tr>
<tr>
<td>• Automatic data acquisition and risk estimation</td>
<td></td>
<td>• Checked data quality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Clinically-based</td>
<td></td>
<td>• Cutting edge EHR infrastructure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Transdiagnostic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Cheap</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Usability in a sequential assessment framework</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

³⁸
Crucially, several aspects were carefully planned at the stage of model development to facilitate its subsequent implementation. The “core components” of the CFIR characteristics of the transdiagnostic risk calculator (i.e. lifespan-inclusive, individualised, clinically-based, transdiagnostic, automatic data acquisition and risk estimation, cheap, use of predictors widely available in clinical routine, usability in a sequential assessment framework) were selected *a priori* to match the CFIR “inner” and “outer” settings (Figure 3-2). The “adaptable periphery” of the CFIR characteristics of the intervention (Figure 3-2), required to adapt the transdiagnostic risk calculator (developed using retrospective data) to prospective use, after checking for data quality and missingness (which may differ in real-world prospective vs retrospective EHR datasets). In the retrospective version of the risk calculator, individuals who developed psychosis within three months following their index diagnosis were excluded to mitigate for ICD-10 diagnostic instability. However, prospective implementation of this diagnostic lag was inefficient. Subsequent analyses confirmed that a refined version of the risk calculator without this lag period had similar external prognostic accuracy (Harrell’s C=0.79). We therefore used the refined model, optimised for prospective usability (for full details see Table 3-1 & Supplementary Table 3-3).
Table 3-1  Sociodemographic characteristics of the study population, both all patients automatically screened in addition to those patients automatically detected.

The patients automatically detected here are all patients above the threshold of 5% risk of developing a psychotic disorder in two years, minus those excluded (n = 27: 16 patients moved out of SLaM, 6 organic psychiatric condition, 3 lack of good English, 1 psychotic diagnosis emerged from collateral clinical information, and 1 patient declined participation).

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Beta coefficient</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>0.010</td>
<td>0.005</td>
<td>0.143</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male (R)</td>
<td>0.457</td>
<td>0.185</td>
<td>0.730</td>
</tr>
<tr>
<td>Female</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age by gender (male)</td>
<td>-0.009</td>
<td>-0.015</td>
<td>-0.002</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White (R)</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0.995</td>
<td>0.865</td>
<td>1.125</td>
</tr>
<tr>
<td>Asian</td>
<td>0.487</td>
<td>0.207</td>
<td>0.767</td>
</tr>
<tr>
<td>Mixed</td>
<td>0.686</td>
<td>0.372</td>
<td>1.000</td>
</tr>
<tr>
<td>Other</td>
<td>0.340</td>
<td>0.143</td>
<td>0.537</td>
</tr>
<tr>
<td>Index diagnosis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHR-P (R)</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute and transient psychotic disorders</td>
<td>1.169</td>
<td>0.876</td>
<td>1.464</td>
</tr>
<tr>
<td>Substance use disorders</td>
<td>-1.748</td>
<td>-2.065</td>
<td>-1.431</td>
</tr>
<tr>
<td>Bipolar mood disorders</td>
<td>0.003</td>
<td>-0.327</td>
<td>0.333</td>
</tr>
<tr>
<td>Non-bipolar mood disorders</td>
<td>-1.560</td>
<td>-1.876</td>
<td>-1.245</td>
</tr>
<tr>
<td>Anxiety disorders</td>
<td>-2.006</td>
<td>-2.320</td>
<td>-1.691</td>
</tr>
<tr>
<td>Personality disorders</td>
<td>-1.363</td>
<td>-1.754</td>
<td>-0.971</td>
</tr>
<tr>
<td>Developmental disorders</td>
<td>-3.337</td>
<td>-4.018</td>
<td>-2.656</td>
</tr>
<tr>
<td>Childhood/adolescence onset disorders</td>
<td>-3.200</td>
<td>-3.641</td>
<td>-2.753</td>
</tr>
<tr>
<td>Physiological syndromes</td>
<td>-2.310</td>
<td>-2.764</td>
<td>-1.856</td>
</tr>
<tr>
<td>Mental retardation</td>
<td>-2.326</td>
<td>-2.864</td>
<td>-1.788</td>
</tr>
</tbody>
</table>

(R) Reference category
The CFIR “inner setting” (Figure 3-2) was characterised by cutting-edge digital EHR infrastructures (South London and the Maudsley, SLaM was awarded Global Digital Exemplar status by NHS England in 2017). SLaM is one of Europe’s largest secondary mental healthcare providers.40 Its main catchment area of 1.36 million individuals covers four socioeconomically diverse South London boroughs: Croydon, Lambeth, Lewisham and Southwark, alongside tertiary referrals from the rest of London and the United Kingdom. SLaM has one of the highest rates of psychosis in the world.41 SLaM is paper-free, and the local EHR comprehensively includes all clinical information recorded throughout mental healthcare episodes, including demographic and contact information, dates and other details of referrals and transfers, detailed clinical assessments, care plans, medication and any clinical activity. Deidentified information from the EHR is rendered available for research use by the Clinical Record Interactive Search (CRIS) platform, developed by the National Institute of Health Research Maudsley Biomedical Research Centre (NIHR Maudsley BRC). Upon SLaM access, each patient is asked to register for (or decline) Consent for Contact (C4C), which indicates their willingness to be contacted for research, without affecting quality of care. The CRIS platform—including its governance framework—has been fully described elsewhere10 and similar resources have been implemented across 12 NHS Trusts in the UK, harnessing over 2 million deidentified patient records (https://crisnetwork.co). The CFIR “outer setting” was addressed by the previous external replications of the risk calculator in other EHRs.32,33

The CFIR “process” domain (Figure 2) was addressed by collaborating with the Centre for Translational Informatics (CTI) to develop a digital system embedded in the local EHR to automatically run the risk calculator in CRIS/EHR. This digital interface was based on CogStack, an information retrieval and extraction platform for EHRs and has been fully detailed in an associated publication.42

Most of the in-vitro phase focused on fine-tuning the use of the transdiagnostic risk calculator to the CFIR “individuals’/users’ domain (Figure 3-2). We consulted the Outreach and Support in South London (OASIS) CHR-P service users group43 as well as the national Young Persons Mental Health Advisory Group (https://ypmhag.org) to collect patients’ feedback on practical and ethical issues relating to the real-world use of this risk calculator. We also conducted two group meetings with SLaM clinicians, to appraise barriers to clinicians’ adherence and optimise sharing of recommendations made by the risk calculator.
3.3.2. **IN-VIVO PHASE**

*Participants and study design*

Once the transdiagnostic risk calculator was embedded in the EHR (via CRIS), we started the *in-vivo* phase. During the study period (May 14\(^{th}\) 2018 to April 29\(^{th}\) 2019), all individuals (i) older than 14 years (ii) who were accessing any SLaM service (iii) receiving a first ICD-10 index primary diagnosis of any non-organic, non-psychotic mental disorder (Supplementary Methods 3-1), or a CHR-P designation and with (iv) existing contact details were deemed eligible. Clinicians were not required to enter any data; all predictors were recorded as part of clinical routine. During the study period, outreach was conducted with clinicians in Lambeth only, informing them about the study and the risk calculator. Clinicians in other boroughs were only aware of the study through alerts.

Every week, all new individuals accessing SLaM who met eligibility criteria were automatically screened. If predictor data was missing, the calculator rechecked their availability each subsequent week, until the end of the study period. Although the original transdiagnostic risk prediction model can provide individualised estimates of psychosis risk up to a period of six years—with no predetermined thresholds and associated sensitivity or specificity—the primary aim of the *in-vivo* phase of the study was to test the feasibility of use in clinical routine and not its effectiveness. Consequently, a tentative threshold of ≥5% risk of psychosis at two years was used to detect at-risk cases, following discussions in group meetings with SLaM clinicians of what would be considered clinically useful. If the patient's predicted risk was above the threshold, the contact details of the responsible clinician were automatically extracted, and the clinician was approached following the procedure established during *in-vitro* engagement work with SLaM clinicians. In the first step, the research team sent an email to the responsible clinician, making a recommendation to respond to the research team via email or phone to discuss the action to be taken. This included the clinician informing the patient that a face-to-face assessment was available in the local CHR-P clinic (OASIS) or at King’s College London. If contact details of the responsible clinicians were incorrect, a third email was sent to the last clinician registered on EHR (e.g. their care coordinator) or to their GP. In a second step, if there was still no response following a further week, phone calls were initiated to the SLaM clinician or GP practice. Following discussion with the research team, the clinician then decided whether to formally initiate the referral—asking the patient if they consented to contact details being
shared with the research team—or not. If the patient consented, they were contacted by the research team, and informed consent for face-to-face research was formally sought. Additional inclusion criteria applied at this stage were: (i) sufficient understanding of English language, (ii) absence of ICD-10 organic psychiatric disorders, substance-induced psychotic disorders (Supplementary Methods 3-2) or psychotic disorder confirmed by collateral clinical information and (iii) patient remained within the geographical catchment served by SLaM. If an individual did not reach the threshold, the research team recommended no further assessment. Since the risk calculator is transdiagnostic, it was expected to detect individuals meeting CHR-P criteria as well as at-risk individuals outside the CHR-P state, i.e. not experiencing psychotic symptoms.

3.3.3. OUTCOMES
The primary outcome was the adherence of clinicians to the use of the automatic EHR screening by the transdiagnostic risk calculator. This was operationalised as the proportion of clinicians who responded to recommendations of the calculator over those who were contacted by the research team. Secondary outcomes included: impact of different alerts on clinicians' adherence, the raw number of referrals initiated from secondary mental healthcare clinicians for an assessment of psychosis risk, and proportion of new ICD-10 diagnoses of psychotic disorders (Supplementary Methods 3-3) by six-month follow-up detected before their onset by the calculator across those screened.

3.3.4. STATISTICAL ANALYSIS
Baseline clinical and sociodemographic characteristics of the sample were described by means and standard deviations for continuous variables, and absolute and relative frequencies for categorical variables. Differences between continuous variables in patients screened and detected by the risk calculator (for being over the predicted risk threshold) were assessed using independent sample two-tailed t-tests; differences between categorical variables were assessed using two-tailed Fisher’s exact test. For categorical variables with only two categories, an Odd’s Ratio (OR) was calculated using Fisher’s exact test. The cumulative incidence of psychosis was measured with Kaplan-Meier curves and 95% Greenwood confidence intervals, and log-rank test. Statistical analyses were performed in R Version 3.2, and the threshold for statistical significance was 0.05.
3.4. RESULTS

3.4.1. STUDY POPULATION

3,722 patients presenting to SLaM clinical services during the study period and with data available in the EHR were eligible and automatically screened (Figure 3-3). 117 patients were detected for being at-risk by the transdiagnostic risk calculator (see Table 3-2, for missing data see Supplementary Results 3-1). Patients screened were aged 37.5 years on average (SD=18.4), 37.9% were male and mostly (60.4%) of White ethnicity; the most frequent index diagnosis was non-bipolar mood disorders (28.9%). Patients detected were on average aged 39.1 years (SD=18.3); 37.5% were male and mostly (39.9%) still of White ethnicity. The most frequent index diagnosis was bipolar mood disorders (70.5%) (see also Supplementary Results 3-2).

3.4.2. PRIMARY OUTCOME: CLINICIAN ADHERENCE TO THE RECOMMENDATIONS MADE BY THE TRANSDIAGNOSTIC RISK CALCULATOR

For two patients, no clinician contact details were available on the EHR; 115 prompts were therefore sent to clinicians. Of these, 89 clinicians (77.4%) responded to prompts sent on the recommendation of the transdiagnostic risk calculator.

3.4.3. SECONDARY OUTCOMES: IMPACT OF DIFFERENT ALERTS, NUMBER OF REFERRALS, AND PROPORTION OF FIRST-EPIsODE CASES DETECTED

33 clinicians (37.1%) responded to the first email, 20 (22.5%) to the second, six (6.7%) to the third and 30 (33.7%) responded to phone calls. Including patient names in SLaM emails instead of citing Trust IDs (REC approval was given for using de-anonymised patient data) raised the response rate from 37.5% (15/40) to 58.7% (44/75) (OR=2.35, 95%CI: 1.00, 5.64, p=0.03). Clinicians’ response to the prompts increased from 74.1% (60/81 in Croydon, Southwark, Lewisham) to 85.3% (29/34 in Lambeth) when outreach was deployed, but this difference was non-significant (OR=2.02, 95%CI: 0.65, 7.55, p=0.23). In terms of referral, among the 89 patients for whom clinicians responded, 18 (20.2%) patients were excluded (3 (3.4%) lack of good English, four (4.5%) organic psychiatric condition, nine (10.1%) patients moved out of SLaM, one (1.1%) psychotic diagnosis emerged from collateral clinical information, and one (1.1%) patient declined participation). Of the remaining 71 patients, 39 (54.9%) were referred by their responsible clinician to OASIS for a face-to-face assessment. The predominant reason for non-referral was patients experiencing acute phases
Figure 3-3 Outline of study design.

Step 1: Patients receiving first non-organic, non-psychotic diagnoses in SLaM are considered for the calculator. Step 2: Patients were automatically screened for psychosis risk with the transdiagnostic risk calculator if all predictors were entered. Step 3: The results of this automated screening with those above the threshold of 5% psychosis risk within two years. Step 4: If patients had contact information, clinicians received alerts informing them that their patient was at-risk for psychosis and recommending that they refer them for a psychosis risk assessment. Step 5: Clinicians responded to the alert, deciding whether to initiate a referral or not. Step 6: Patients were referred to OASIS for a psychosis risk assessment.
Table 3-2 Sociodemographic characteristics of the study population, both all patients automatically screened in addition to those patients automatically detected.

The patients automatically detected here are all patients above the threshold of 5% risk of developing a psychotic disorder in two years, minus those excluded (n=27: 3 lack of good English, 6 organic psychiatric condition, 16 patients moved out of SLaM, 1 psychotic diagnosis emerged from collateral clinical information, and 1 patient declined participation)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Patients automatically screened (n = 3,722), No. (%)</th>
<th>Patients automatically detected (n = 88), No. (%)</th>
<th>Screened vs Detected Test</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD), y</td>
<td>37.51 (18.44)</td>
<td>39.05 (18.27)</td>
<td>t = -0.78</td>
<td>0.437</td>
</tr>
<tr>
<td>Sex (% Male)</td>
<td>1412 (37.93%)</td>
<td>33 (37.5%)</td>
<td>OR = 1.04</td>
<td>0.912</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(95%CI: 0.66, 1.66)</td>
<td></td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>2249 (60.42%)</td>
<td>35 (39.77%)</td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Black</td>
<td>660 (17.73%)</td>
<td>21 (23.86%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>249 (6.69%)</td>
<td>8 (9.09%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixed</td>
<td>166 (4.46%)</td>
<td>5 (5.68%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>398 (10.69%)</td>
<td>18 (20.45%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index diagnosis</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Acute and transient psychotic disorders</td>
<td>46 (1.24%)</td>
<td>22 (25.00%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bipolar mood disorders</td>
<td>99 (2.66%)</td>
<td>62 (70.45%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-bipolar mood disorders</td>
<td>1076 (28.91%)</td>
<td>3 (3.41%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personality disorders</td>
<td>181 (4.86%)</td>
<td>1 (1.14%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Developmental disorders</td>
<td>57 (1.53%)</td>
<td>0 (0%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Childhood/adolescence onset disorders</td>
<td>240 (6.45%)</td>
<td>0 (0%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physiological syndromes</td>
<td>237 (6.37%)</td>
<td>0 (0%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental retardation</td>
<td>43 (1.16%)</td>
<td>0 (0%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Substance use disorders</td>
<td>545 (1.64%)</td>
<td>0 (0%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety disorders</td>
<td>1198 (32.19%)</td>
<td>0 (0%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
of high psychiatric symptom severity, either in inpatient units or the community, other reasons for non-referral are presented in Supplementary Results 3-3. Among those screened (n=3,722), 3,640 (97.79%) were followed-up through the EHR, and 38 (1.04%) developed a psychotic disorder by six-month follow-up. The cumulative incidence of psychosis in those screened was 0.016 (95%CI: 0.010-0.022, when 1,302 individuals were still at-risk) at six-months (Supplementary Figure 3-4).

Among those detected (n=115), 101 (87.82%) were followed-up through the EHR and nine (8.9%) developed a psychotic disorder by six-months. The cumulative incidence of psychosis in those detected was 0.12 (95%CI: 0.04-0.19, when 56 individuals were still at risk) at six-months (Supplementary Figure 3-4), which was significantly higher than in those screened (log-rank test: p<0.001).

Among the 49 patients detected but not referred (either through non-response or non-initiated referral) and with a six-month follow-up in the EHR, three (6.1%) developed a psychotic disorder. The cumulative incidence of psychosis in those detected but not referred was 0.147 (95%CI: 0.030-0.249, when 32 individuals were still at risk) at six-months (Supplementary Figure 3-5) and comparable to that observed in those detected and referred (0.094, 95%CI: 0-0.191, p=0.40, Supplementary Figure 3-5).

3.5. DISCUSSION

To our best knowledge, this is the first study reporting on the implementation of an individualised risk calculator in EHRs. This study demonstrates that it is feasible to combine precision medicine and digital health to embed a transdiagnostic, clinically-based, individualised psychosis risk calculator in EHR and potentially inform clinical practice.

This study advances knowledge in the field of precision psychiatry and digital health for early psychosis in several ways. This feasibility study was built on the pragmatic assumption that risk estimation models provide little value unless used by clinicians in day-to-day practice. While the risk calculator tested in this study has previously been shown to have modest-to-good prognostic performance in multiple case settings, its real-world usability was untested. The main result of this study is that 77% of clinicians responded to prompts issued by the risk calculator (85% if outreach was conducted, see below), indicating good adherence. This positive result has been accomplished through strong consideration into optimising the risk calculator for implementation already during the early phases of
model building, development and testing. First, this model is clinically-based, relying on predictors that are collected routinely in clinical practice to reduce clinician burden. Second, these predictors were extracted by the local EHR to allow scalability and automation of screening procedures. Third, the risk calculator does not require labour- and time-intensive assessments, which further facilitates implementation. Fourth, it is deliberately transdiagnostic, including at-risk patients meeting CHR-P criteria as well as those who might develop psychosis outside the CHR-P state (about one one-third of first-episode psychosis cases). Fifth, it is lifespan-inclusive (although in this study only patients aged 14 or above were recruited). These characteristics configure a risk estimation model with potential broader screening potential in the real-world. Sixth, this risk calculator makes individualised predictions, contrasting to the current CHR-P strategy and the majority of group level (i.e. at-risk vs not at-risk) prognostic models in this field. While other individualised risk calculators are available (e.g. NAPLS), these are only applicable to the relatively small group of individuals who have already met criteria for CHR-P, thereby these models cannot improve detection of at-risk individuals. Our transdiagnostic risk calculator is the only available model to extend detection and therefore primary indicated prevention of psychosis, with potential benefits to patients, carers and society as a whole.

This study also advances methodological knowledge in the field of implementation research for precision psychiatry. Implementation science, although much needed, is contested and complex. For example, we sought to follow the CFIR, but this framework is rather theoretical and does not offer specific pragmatic guidance to precision psychiatry. A recent systematic review concluded that only 6% of studies acknowledging the CFIR actually used it meaningfully. To our best knowledge, this is the first study to have addressed specific implementation barriers and have developed empirical methods to overcome them. The first innovation was to adopt a pragmatic approach to carefully pre-empt most implementation challenges, considering the CFIR domains during early model building. Our approach, encompassing two subsequent phases (in-vitro and in-vivo), may represent a viable method to overcome the implementation gap, which has led some to question the utility of precision psychiatry. For example, since most risk prediction models are developed on “artificial” retrospective datasets, the in-vitro phase seems particularly suited to address barriers relating to accessibility of predictors, outcome data and model refinements that are needed to use risk calculators in real-world EHRs prospectively. The in-vivo phase can be used to address core ethical, legal and societal barriers for implementing non-stigmatising
precision medicine through an encrypted network of EHR servers and databases.\textsuperscript{11} For example, while pilot research showed that patients generally welcomed knowing their risk for severe mental disorders,\textsuperscript{59} producing profound benefits by enabling more accurate, effective therapeutic decision-making, individualised outcome prediction of risk may also influence stigmatisation. To this tension, our \textit{in-vitro} phase included qualitative work with local clinicians and service users. Further to this, we demonstrated that in feasibility studies (as opposed to effectiveness studies) clinicians may be primarily approached to discuss the need of further optional referral with the research team to maintain continuity of care at their current mental healthcare service. This strategy aligned with the transdiagnostic nature of the risk calculator whereby different concurrent comorbid ICD-10 disorders, in addition to psychosis risk are allowed.\textsuperscript{43,60–62} Another methodologically-relevant finding is that conducting outreach can increase the clinicians’ adherence to the use of risk calculators in clinical routine. Furthermore, the sequential use of combined prompts can improve clinicians’ adherence: our automatic screening platform (CogStack) has since been developed to streamline this process.\textsuperscript{62} Finally, our method demonstrated that sequential risk assessment frameworks\textsuperscript{39} could have higher implementability into clinical routine. Failure to implement most risk calculators may be partially attributed to the complexity of models which involve high cost (e.g. neuroimaging modalities) or labour (e.g. cognitive tasks) to produce their predictions in a single run. Because the current risk calculator is simple, it can be used to screen large populations, with more complex (e.g. the recently developed Psychosis Polyrisk Score, PPS)\textsuperscript{25,63,64} or costly prognostic models reserved to subsequently refine risk estimates in individuals with uncertain prognostic estimates.\textsuperscript{65}

This study also opens several lines of future research. Firstly, this risk calculator can be improved, refining the current predictors (such as better modelling the higher psychosis risk of late adolescence and early adulthood through non-linear methods)\textsuperscript{66} or adding new predictors leveraging advanced data mining methods for EHRs (e.g. Natural Language Processing, NLP).\textsuperscript{67} Secondly, this study identified new implementation barriers that are unaddressed, such as the deployment of well-established governance frameworks and guidance to implement precision psychiatry into EHRs. Thirdly, in this study incidence of psychosis was 12\% within six-months in the individuals detected, comparable to the level of risk observed in the CHR-P paradigm (10\% at six-months).\textsuperscript{68} Interestingly, the incidence of psychosis at six-months was 14.7\% among those not referred by clinicians for face-to-face assessment, and comparable to that in those referred. In previous studies, clinicians’
predictions have typically been shown to be overoptimistic.\textsuperscript{65,69,70} It is thus evident that effective implementation of risk calculators in EHRs requires not only intensive outreach but an adequate provision of training and teaching for future clinicians.\textsuperscript{8,19}

The main limitation of this feasibility study is that it is only addressing pragmatic implementation barriers; as such it is clearly not sufficient either in terms of sample size or follow-up time to demonstrate effectiveness in real-world care, as has been observed in inadequate uptake of the human papillomavirus vaccine\textsuperscript{71} and smoking cessation initiatives.\textsuperscript{72} These aspects will need to be tested in a subsequent large-scale effectiveness study, which is currently being planned, in addition to using organisation-level collaboration, such as the 26-site ProNET and HARMONY, which incorporates NAPLS,\textsuperscript{73} PRONIA (https://www.pronia.eu/) and PSYSCAN.\textsuperscript{74} Data missingness is common issue within EHRs\textsuperscript{75} and was prominent here, with 35\% of individuals unable to be screened, restricting the potential impact of this calculator. Imputation of missing data through Bayesian methods may be one way to reduce the impact of this issue\textsuperscript{76} but more work needs to be done to establish utility of individualised clinical decision making based on data imputation. Furthermore, data missingness in the two retrospective external validations was substantially lower, suggesting that most of the missing data are subsequently entered into EHR by clinicians. Dynamic refinements of risk calculators may allow incorporating new predictors as soon as they are recorded in EHRs, as a similar approach has recently demonstrated.\textsuperscript{77} Furthermore, we have been unable to qualitatively collect reasons for non-response from clinicians.

**CONCLUSIONS**

This is the first implementation study to demonstrate that it is feasible to combine precision psychiatry and digital health to improve the detection of individuals with emerging psychosis. Future implementation research in this field is urgently needed.
Author contributions

DO and GS had full access to all data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: PF-P

Acquisition, analysis or interpretation of the data: DO, GS, PF-P, RP, RS, RD

Drafting of the manuscript: DO, PF-P

Critical revision of the manuscript for important intellectual content: All authors

Statistical analysis: DO, DS

Obtained funding: PF-P

Administrative, technical, or material support: RP, RS, RD

Supervision: PF-P, PM, RS, DS

Conflict of interest disclosures

PF-P has received advisory consultancy fees from Lundbeck outside of this work. RS has received research support from Roche, Janssen, GSK and Takeda outside of this work. The authors have declared that there are no conflicts of interest in relation to the subject of this study.

Funding/support

This study was 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. MB, HB, RS and RD are part-funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at the South London and Maudsley NHS Foundation Trust and King’s College London. RP has received support from an MRC Health Data Research UK Fellowship (MR/S003118/1) and a Starter Grant for Clinical Lecturers (SGL015/1020) supported by the Academy of Medical Sciences, The Wellcome Trust, MRC, British Heart Foundation, Arthritis Research UK, the Royal College of Physicians and Diabetes UK.
3.6. REFERENCES


3.7. SUPPLEMENTARY DATA

Oliver D, Spada G, Colling C et al. Real-World Implementation of Precision Psychiatry: Transdiagnostic Risk Calculator for the Automatic Detection of Individuals At-Risk of Psychosis

Presented in this thesis in part for brevity

Link to full supplementary

Supplementary Table 3-3 Model performance of the Transdiagnostic Individualised Clinically-based Risk Calculator for the Automatic Detection of Individuals at-Risk and the Prediction of Psychosis (revised version) in the original SLaM derivation and validation datasets*

Supplementary Table 3-4 STROBE Statement – Checklist of items that should be included in reports of cohort studies

Supplementary Methods 3-1 Explanation of the definition of acute and transient psychotic disorders as non-psychotic disorders*

Supplementary Methods 3-2 List of additional inclusion criteria following referral*

Supplementary Methods 3-3 List of ICD10 diagnoses of psychotic disorders considered as secondary outcome endpoints*

Supplementary Results 3-1 Missing data*

Supplementary Results 3-2 Sociodemographic differences between screened and detected patients*

Supplementary Results 3-3 Reasons for exclusion following referral*

Supplementary Results 3-4 Reasons for non-referral*

Supplementary Figure 3-4 Cumulative incidence (Kaplan–Meier failure functions) for risk of development of psychotic disorders in individuals detected and screened*

Supplementary Figure 3-5 Log-rank test comparing the cumulative incidence (Kaplan–Meier failure functions) for risk of development of psychotic disorders in individuals referred and not referred*

Supplementary References

*Included here
Supplementary Table 3-3 Performance of the Transdiagnostic Individualised Clinically-based Risk Calculator for the Automatic Detection of Individuals at-Risk and the Prediction of Psychosis (revised version) in the original SLaM derivation and validation datasets, refined without a three month lag after the index diagnosis.

<table>
<thead>
<tr>
<th>Performance measure</th>
<th>Derivation (n = 34,209)(a)</th>
<th>Validation (n = 54,716)(b)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean 95%CI</td>
<td>Mean 95%CI</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brier (c)</td>
<td>0.028</td>
<td>0.021</td>
</tr>
<tr>
<td>R²</td>
<td>0.746 0.704 0.785</td>
<td>0.719 0.673 0.761</td>
</tr>
<tr>
<td>Discrimination</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harrell's C</td>
<td>0.809 0.795 0.822</td>
<td>0.790 0.775 0.806</td>
</tr>
<tr>
<td>Calibration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calibration slope</td>
<td>1</td>
<td>0.968 0.929 1.015</td>
</tr>
</tbody>
</table>

a) The sample size is larger than the original 33,820 because individuals with psychosis onset within the first 3 months since baseline were not excluded.
b) The sample size of the external validation matches that used in the original study to facilitate comparability.
c) at 10-years.
Supplementary Methods 3-1 Explanation of definition of acute and transient psychotic disorders as non-psychotic disorders

The Acute and Transient Psychotic Disorder (ATPD) group was not considered a psychotic disorder because it is diagnostically and prognostically similar to the Brief Limited Intermittent Psychotic Symptoms (BLIPS) subgroup of the CHR-P construct (for details on these competing operationalisations, see eTable 1 in 2).

Supplementary Methods 3-2 List of additional inclusion criteria following referral

(i) sufficient understanding of English language, (ii) patient remained within the geographical catchment served by SLaM, and (iii) absence of ICD-10 organic psychiatric disorders, substance-induced psychotic disorders (below) or psychotic disorder confirmed by collateral clinical information:

F0x – Organic, including symptomatic, mental disorders
F1x.4 – Mental and behavioural disorders due to psychoactive substance use (Withdrawal state with delirium)
F1x.5 – Mental and behavioural disorders due to psychoactive substance use (Psychotic disorder)
F1x.7 – Mental and behavioural disorders due to psychoactive substance use (Residual and late-onset psychotic disorder)
Supplementary Methods 3-3 List of ICD10 diagnoses of psychotic disorders considered as secondary outcome endpoints

F20.0 - Paranoid schizophrenia
F20.1 - Hebephrenic schizophrenia
F20.2 - Catatonic schizophrenia
F20.3 - Undifferentiated schizophrenia
F20.4 - Post-schizophrenic depression
F20.5 - Residual schizophrenia
F20.6 - Simple schizophrenia
F20.8 - Other schizophrenia
F20.9 - Schizophrenia, unspecified
F25.0 - Schizoaffective disorder, manic type
F25.1 - Schizoaffective disorder, depressive type
F25.2 - Schizoaffective disorder, mixed type
F25.8 - Other schizoaffective disorders
F25.9 - Schizoaffective disorder, unspecified
F22.0 - Delusional disorder
F22.8 - Other persistent delusional disorders
F22.9 - Persistent delusional disorder, unspecified
F24 - Induced delusional disorder
F28 - Other non-organic psychotic disorders
F29 - Unspecified non-organic psychosis
Any [F10-F19].4 or any [F10-F19].5 or any [F10-F19].7 - Mental and behavioural disorders due to psychoactive substance use with psychotic symptoms or delirium
F30.2 - Mania with psychotic symptoms
F31.2 - Bipolar affective disorder, current episode manic with psychotic symptoms
F31.5 - Bipolar affective disorder, current episode severe depression with psychotic symptoms
F32.3 - Severe depressive episode with psychotic symptoms
F33.3 - Recurrent depressive disorder, current episode severe with psychotic symptoms
F53.1 - Severe mental and behavioural disorders associated with the puerperium, not elsewhere classified (post-partum psychosis)
Supplementary Results 3-1 Missing data

In addition to the 3,722 patients screened, there were a total of 2,678 patients who were given a non-organic, non-psychotic primary index diagnosis during the study period but had one of more missing predictors at the point of initial screening. For 647 of these patients, one or more predictors were missing at the point of initial screening but were subsequently entered during the in-vivo study period and therefore were screened by the risk calculator. The remaining 2,031 patients did not have their missing predictors updated during the in-vivo study period meaning they were unable to be screened by the risk calculator (2,027 ethnicity missing, two gender missing, two ethnicity and gender both missing).

Supplementary Results 3-2 Sociodemographic differences between patients screened and detected.

There were significant differences between those screened and detected in terms of ethnicity and index diagnosis (p < 0.001), with no significant differences in age (p = 0.437) or gender (p = 0.912).

Supplementary Results 3-3 Reasons for exclusion following referral

A total of 18 patients were excluded following referral:

- Three (3.4%) for lack of good English
- Four (4.5%) due to an organic psychiatric condition
- Nine (10.1%) as patients had moved out of SLaM
- One (1.1%) due to psychotic diagnosis emerged from collateral clinical information
- One (1.1%) as the patient declined participation
Supplementary Results 3-4 Qualitative reasons given by responsible clinicians either by phone or email for non-referring individuals detected by the transdiagnostic risk calculator.

13 patients were experiencing acute phases of mental disorders with high symptomatic severity:
Nine patients sectioned in inpatient units
Four patients admitted to outpatient units
Two patients were experiencing high severity of comorbid physical health conditions
Five patients were not engaging with the SLaM services
Two patients: the responsible clinician decided on the basis of their clinical judgement that the patients were not at-risk for psychosis
Three patients: the patient did not wish to participate
One patient: the patient had moved out of SLaM
Six patients: no given reason
**Supplementary Figure 3-4** Cumulative incidence (Kaplan–Meier failure function) for risk of development of psychotic disorders in individuals screened and detected by the transdiagnostic risk calculator and followed-up over a period of six-months

**A:** Cumulative incidence (Kaplan–Meier failure function) for risk of development of psychotic disorders in 3,640 individuals screened by the transdiagnostic risk calculator and followed-up over a period of six-months. The average follow-up time was 154.78 days (SD = 182.54). There were a total of 38 events (transition to psychosis), with the last event observed at 176 days, when 1,302 individuals were still at risk (average time to transition to psychosis 60.90 days, SD = 55.98). The cumulative incidence of psychosis was 0.016 (95%CI: 0.010-0.022) at six-month follow-up.

**B:** Cumulative incidence (Kaplan–Meier failure function) for risk of development of psychotic disorders in 101 individuals detected by the transdiagnostic risk calculator and followed-up over a period of six months. The average follow-up time was 187.66 days (SD = 188.76). There were a total of nine events (transition to psychosis), with the last event observed at 128 days, when 56 individuals were still at risk (average time to transition to psychosis was 55.89 days, SD = 45.81). The cumulative incidence of psychosis was 0.12 (95%CI: 0.04-0.19) at six-month follow-up.
Supplementary Figure 3-5 Log-rank test comparing the cumulative incidence (Kaplan-Meier failure function) for risk of development of psychotic disorders in 52 individuals detected and referred vs 49 individuals detected but not referred.

The cumulative incidence of psychosis in those detected but not referred was 0.147 (95%CI: 0.030-0.249, when 32 individuals were still at risk) at six-months and comparable to that observed in those detected and referred (0.094, 95%CI: 0-0.191, p=0.40).
3.8. SUPPLEMENTARY REFERENCES


PART B: IMPROVING PROGNOSTICATION OF OUTCOMES IN CHR-P INDIVIDUALS
4. INTRODUCTION

4.1. IMPROVING PROGNOSTICATION OF OUTCOMES

The second key step in providing effective psychosis prevention is accurate prognostication of outcomes, particularly transition to psychosis. This introduction will explore how psychosis risk is currently estimated, the extent to which we can stratify individuals according to their underlying psychosis risk and potential information that could improve this in the future, with a particular focus on non-genetic risk and protective factors. Finally, this introduction will explore the concept and development of the Psychosis Polyrisk Score (PPS), which I developed, digitally implemented and piloted over the course of this PhD.

Currently there is limited ability to chart the clinical path of individuals beyond the group level designation of being “at-risk” or “not at-risk” for psychosis. CHR-P assessments such as the SIPS or CAARMS have high sensitivity (95% and 86%, respectively) but sub-optimal specificity (45% and 55%, respectively).\(^1\)\(^2\) This means that while CHR-P services are able to effectively rule out risk of psychosis, they are less adept at ruling it in, with only 20% of “at risk” individuals transitioning to psychosis within two years.\(^3\)

Beyond the group-level, there is additionally subgroup-level prognostication, with three CHR-P subgroups, which each have their corresponding level of psychosis risk. Firstly, individuals who have a first-degree relative with a psychotic disorder in addition to a 30% drop in functioning in the past 12 months (measured by the Social and Occupational Functioning Assessment Scale) meet criteria for the Genetic Risk and Deterioration (GRD) subgroup. These individuals have the lowest risk of transition of the three subgroups (8% over two years),\(^4\) potentially owing to the lack of requisite present attenuated psychotic symptomatology. Secondly, individuals who present with relatively frequent attenuated positive psychotic symptoms (three to six times per week if more than an hour per occasion, or daily if less than an hour per occasion) are classed as Attenuated Psychotic Syndrome (APS) and have a transition risk of 24% over two years.\(^4\) Thirdly, individuals who present with positive psychotic symptoms of psychotic-level severity, which persist for less than one week, are classed as Brief Limited Intermittent Psychotic Symptoms (BLIPS) and have the highest transition risk of 38% over two years.\(^4\) While stratification into these three
subgroups has some clinical utility, this is somewhat limited as the vast majority (85%)\(^4\) of CHR-P individuals meet APS criteria.

Recent developments in healthcare are increasingly allowing for the opportunity of precision medicine, tailoring clinical recommendations and/or treatments to the individual.\(^5,6\) While this has largely been developed in fields such as oncology, there is potential clinical benefit in applying these methods to psychiatry.\(^7\) We are currently unable to refine our estimates of psychosis risk further than subgroup-level prognostication distinguishing between APS, BLIPS and GRD. Therefore, to advance precision psychiatry, we need to move beyond symptomatology alone. Supplementing the CHR-P assessment with additional information from other approaches (e.g. proteomics,\(^8\) neuroimaging,\(^9-11\) electrophysiology\(^12-19\) and clinical/neurocognitive\(^20-26\) data) could better inform prognostication of clinical and functional outcomes.

Our current methods of risk stratification and prognostication of outcomes is reliant on the severity and frequency of symptoms. Symptoms are not the underlying cause of psychosis but are instead epiphenomena of underlying gene-by-environment interactions.\(^27-29\) As genetic and environmental factors are more closely linked to aetipathology, these are more robust indicators of psychosis risk. This contrasts with the approaches mentioned above. For example, neuroimaging approaches can detect structural and functional abnormalities to inform prognosis, but these are also secondary to the interplay of genetic and environmental risk and protective factors. Additionally, many of these approaches necessitate additional logistical challenges, as well as financial and/or labour costs. This, in part, could explain the dearth of prediction models that have been replicated and implemented in clinical routine.\(^30\) Assessment of genetic and environmental risk and protective factors can mitigate these issues by addressing underlying aetipathology and facilitating implementation through reduced logistical and resource demands.

### 4.2. GENETICS IN PSYCHOSIS

Genetic and environmental risk and protective factors for psychosis closely relate to the causal framework that underpins psychosis onset. Due to the relatively high heritability of schizophrenia,\(^31\) using genetic information to predict the onset of psychotic disorders has been extensively researched.\(^32\) The genetic risk of psychosis cannot be explained by a single gene, instead genome-wide association studies (GWAS) can elucidate the small increases
in psychosis risk associated with alleles across the entire genome. Polygenic risk scores have been developed to use this information to predict the disorder through genetic variants “en masse”, summarising the additive risk across many associated loci. In this case, the polygenic risk score (PRS) for schizophrenia was based on a meta-analysis of GWAS studies in schizophrenia conducted by the Schizophrenia Working Group of the Psychiatric Genomics Consortium. This meta-analysis identified that despite the small effect sizes of single loci, the cumulative effect of thousands of schizophrenia-associated loci expressed a polygenic risk score that explained up to 18% of variance between cases of schizophrenia and controls in GWAS studies and 7% of the variance on the underlying liability to schizophrenia in the general population. However, as heritability of schizophrenia is 64% (95% CI: 62% - 68%), a large proportion of the variance remains unaccounted. Similarly, polygenic risk scores have explained 9% of the variance in predicting case-control status at the time of a first episode psychosis and 0.5% of the variance in predicting general mental health. The variance explained is currently too small for individual risk prediction. Therefore, the clinical utility of polygenic risk scores in psychosis is insufficient on a stand-alone basis.

Currently, applicability is another key issue with genetic prediction models. The PRS is only viable in individuals of white European origin, with low discriminative ability in other ethnic groups. This is particularly limiting considering the case-mix of FEP patients in SLaM, wherein a substantial majority (65%) are non-white. Efforts are required to bridge this gap.

**4.3. NON-GENETIC RISK AND PROTECTIVE FACTORS IN PSYCHOSIS**

In this section, I will outline how non-genetic factors that modulate psychosis risk can strengthen estimates provided by genetics, our current understanding of these non-genetic factors, where key knowledge is lacking and how meta-analysis can be used to both rectify this and inform precision psychiatry.

Given the small proportion of variance explained by genetics and the restricted population the PRS is available for, risk prediction needs to be boosted by supplementing the polygenic risk scores with additional information. The model that has received some empirical support indicates that the aetiology of psychotic disorders like schizophrenia involves direct genetic and environmental effects, along with their interaction. In reality, some of the most
predictive factors, including family history of mental illness and socioeconomic status, include both a genetic and environmental component and hence a distinction between genetic and environmental factors may be spurious.

Risk factors contributing to the psychosis risk enrichment observed in CHR-P samples are not entirely known. A recent meta-analysis has summarised the available evidence across 54 putative risk factors investigated in CHR-P samples, in comparison to controls, showing that CHR-P subjects are more likely to show obstetric complications, tobacco use, physical inactivity, childhood trauma, high perceived stress, childhood and adolescent low functioning, affective comorbidities, male gender, single status, unemployment and low educational level as compared to controls. Overall, this study suggests that risk enrichment of CHR-P samples can be attributed to demographic and environmental risk factors like childhood trauma, adverse life events and affective dysfunction.

The differential combination of risk/protective factors in each CHR-P individual is likely to account for the distinct clinical outcomes observed in these samples: psychosis onset, recovery or ongoing disability. However, the factors that modulate transition risk in CHR-P individuals may differ from those that distinguish CHR-P individuals from healthy controls. This is particularly prescient as the same factors that differentiate CHR-P individuals from healthy controls will be enriched in CHR-P samples. Efforts to elucidate our understanding of risk and protective factors for transition have been extensive but have been restricted to small samples and are often inconsistent in their results. These factors are heterogeneous and their association with clinical outcome has not been established. The lack of pathophysiological knowledge is a rate-limiting step in developing accurate prognostic models that replicate the underlying aetiopathology of the disorder. Meta-analysis would aid greater understanding of the specific factors that modify transition risk through the quantitative synthesis of available evidence. By utilising the literature as a whole, it is possible to ascertain the strength of the available evidence and their respective biases, with a view to future refinement of precision psychiatry. In my PhD, I aim to use meta-analysis to synthesise the existing literature to identify robust factors that modulate transition risk in CHR-P individuals. Through this, I will determine whether there is a significant effect of each factor on transition to psychosis and, for those that are significant, I will ascertain the direction and magnitude of the effect and the strength of the evidence for the association.
4.4. PSYCHOSIS POLYRISK SCORE (PPS)

In this section, I will explain the rationale for a Psychosis Polyrisk Score (PPS), the evidence for potential factors and the importance of digital implementation. The PPS aims to assess exposure to risk and protective factors for psychosis in a multivariate fashion to provide an estimate of individualised psychosis risk that is easily implementable in clinical routine.

4.4.1. RATIONALE AND EVIDENCE FOR A PSYCHOSIS POLYRISK SCORE (PPS)

The evidence base of non-genetic risk and protective factors for psychosis has been hampered by the focus on univariate associations. This is in spite of psychosis onset being reliant of the interaction of these factors on a multivariate level.40 Similar to a PRS, non-genetic risk and protective factors can be assessed “en masse” to provide a measure of their additive influence on psychosis risk in the form of a polyrisk score, such as the PPS. Polyrisk scores could help inform risk stratification by reflecting the exposure of an individual to non-genetic psychosis risk, aiding prognostication of outcomes and furthering the potential of precision psychiatry.

The inclusion of non-genetic factors in the development of polyrisk scores is not a conceptually novel approach, but it has been limited to date by the lack of established and robust a priori knowledge on the association of non-genetic factors and psychotic disorders. Such a limitation has been recently overcome by an umbrella review, which is a meta-analysis of meta-analyses or reviews, investigating several non-genetic risk/protective factors of psychosis that operate at an individual level. The umbrella review further classified these factors into convincing (class I), highly suggestive (class II), suggestive (class III), weak (class IV) and non-significant (ns) evidence, according to a standardised classification already widely adopted in other branches of clinical medicine41 to control for potential biases. For instance, sensitivity analyses restricted to prospective studies assessed whether there was pre-existing evidence for risk factor prior to disorder onset, therefore controlling for reverse causation.41 By providing the required gold-standard a priori knowledge, the core results of this meta-analysis place the groundwork for the development of a PPS.

The substantial role of sociodemographic risk/protective factors

Most aetiopathogenic models for psychotic disorders have focused on genetic and
environmental risk factors, while demographic factors have been investigated to a lesser degree, presumably in light of the fact that these factors are not strictly modifiable. Nevertheless, the recent umbrella review found a main effect for male gender, a main effect for 15-35 years of age and an association between psychotic disorders and being a male aged between 15-40 years. Age older than 35 was found to be a protective factor. The additional risk factor that was consistently associated with psychosis was ethnicity, variously defined as being an ethnic minority or as having an immigrant status or through specific categories of ethnicity. For instance, being of Black Caribbean (OR = 4.87, 95% CI = 3.96-6.00, class I), Black African (OR = 4.72, 95% CI = 3.30-6.77), Asian (OR = 2.83, 95% CI = 1.59-5.05) or mixed (OR = 2.19, 95% CI = 1.08-4.44) ethnicity in England or North African in Europe (OR = 2.22, 95% CI = 1.58-3.12) was associated with an increased liability to psychosis. These findings are of significant value for the development of polyrisk scores as they suggest that these factors should always be assessed and considered for the prediction of psychosis onset. In other branches of medicine, age and gender are consistently used in individualised risk scores for predicting cardiovascular diseases (QRISK), diabetes (AUSDRISK) or stroke (CHA2DS2-VASc score). Clinical utility of demographic variables for predicting psychosis onset was supported by the transdiagnostic psychosis risk calculator discussed in Part A that included age, gender, age by gender and ethnicity in an individualised risk estimation tool for predicting psychosis in secondary mental health care.

**Parental and perinatal risk/protective factors**

Psychotic syndromes are disorders of adapting to the environment, which include parental, perinatal and later risk factors, along with antecedents. The umbrella review identified that parental factors such as paternal age (>35 OR = 1.22, 95% CI = 1.06-1.41, >45 OR = 2.36, 95% CI = 1.35-4.11), low paternal socioeconomic status (OR = 1.30, 95% CI = 1.02-1.65) and parental history of severe mental disorder (OR = 5.94, 95% CI = 2.99-11.79) were all associated with psychosis. Polygenic studies controlling for the effect of parental risk factors found that parental socioeconomic status accounted for 45.8% (95% CI = 36.1-55.5) of cases with schizophrenia. Assuming social causation, this indicates that the impact of the environment is actually higher than the genetic factors. Similarly, a recent study indicated that the predictive value of polygenic risk scores can be improved, explaining 17.4% variance if used in cases with a family history of schizophrenia/psychoses (i.e. prediction by PRS including more genetic variants). These findings concur with the need
for integrating genetic and parental risk factors for psychosis in a polyrisk score. Some studies have already supplemented the polygenic score profile with information on family history for psychotic disorders. Other risk factors could be considered for the development of a polyrisk assessment, including urbanicity at birth (OR = 2.19, 95% CI = 1.55-3.09). As this factor was robust and survived sensitivity analyses (class I), it should always be measured and considered in polyrisk assessment approaches. Finally, a series of perinatal risk factors were shown to be useful for the polyrisk score. The most robust of them was winter/spring season of birth in northern hemisphere (OR = 1.04, 95% CI = 1.02-1.06, class III), followed by diabetes in pregnancy (OR = 10.12, 95% CI = 1.84-55.72), emergency caesarean section (OR = 3.36, 95% CI = 1.48-7.63), low birth weight (<2000 OR = 2.46, 95% CI = 1.11-5.46, <2500 OR = 1.57, 95% CI = 1.20-2.07), congenital malformations (OR = 2.31, 95% CI = 1.29-4.13), use of incubator or resuscitation (OR = 2.12, 95% CI = 1.29-3.47), threatened premature delivery (OR = 2.05, 95% CI = 1.02-4.10), maternal obesity (OR = 1.99, 95% CI = 1.26-3.14), uterine atony (OR = 1.93, 95% CI = 1.35-2.76), antepartum haemorrhage (OR = 1.62, 95% CI = 1.12-2.38) and small head circumference (OR = 1.41, 95% CI = 1.00-1.97). To the best of our knowledge, no studies have attempted to combine polygenic risk assessment with these risk factors, and this may prove to be a promising avenue of research.

Later risk/protective factors

Later risk factors that have been associated with psychosis include a variety of environmental risk factors, such as urbanicity at upbringing (OR = 2.19, 95% CI = 1.55-3.09), childhood trauma (OR = 2.87, 95% CI = 2.07-3.98), problems in parental communication (OR = 11.55, 95% CI = 5.81-23.06), pollution (OR = 5.55, 95% CI = 1.63-18.87), environmental benzenes (OR = 3.20, 95% CI = 1.01-10.12), adult life events (OR = 5.34, 95% CI = 3.84-7.43), substance abuse such as heavy cannabis use (OR = 5.17, 95% CI = 3.64-7.36), or tobacco use (OR = 2.19, 95% CI = 1.36-3.53) and traumatic brain injury (OR = 1.49, 95% CI = 1.09-2.05). Later risk factors also include a series of infective agents, such as IgG Toxoplasma gondii (OR = 1.82, 95% CI = 1.51-2.18), Toxocara (OR = 41.61, 95% CI = 9.71-178.32), Chlamydia Psittaci (OR = 29.05, 95% CI = 8.91-94.69), retroviruses type W (OR = 19.78, 95% CI = 6.50-60.22), Chlamydia pneumoniae (OR = 6.02, 95% CI = 2.86-12.66), Borna disease virus (OR = 1.94, 95% CI = 1.30-2.91) and herpes virus 2 (OR = 1.44, 95% CI = 1.14-1.81). Exposures to childhood trauma and Toxoplasma gondii were most robustly associated with increased risk of psychosis (class
III), while the other later factors showed weak association.\textsuperscript{41}

\textbf{Antecedents}
While antecedent factors (which include the CHR-P criteria per se as well as other early manifestations of non-psychotic psychopathology) are mentioned here, they are not discussed in full because antecedent features may actually represent an early manifestation of the disorder and may, therefore, play a separate role in the prediction of psychosis that is distinct from polyrisk scoring.

As a result of this umbrella review, we have robust evidence of the magnitude of effect and strength of association of non-genetic risk and protective factors with psychosis. This provides us with \textit{a priori} evidence to prespecify predictors for a multivariate model of psychosis risk (a PPS), thereby avoiding bias in predictor selection.\textsuperscript{49} Therefore, this evidence alongside the PPS allows for systematic assessment of exposure to psychosis risk and protective factors “en masse”, enabling prognostication beyond symptomatology, closer modelling of aetiopathology and facilitating precision psychiatry.

\textbf{4.4.2. DIGITAL IMPLEMENTATION IN CLINICAL ROUTINE}

Digital implementation of the PPS is a key aim of this thesis to enhance feasibility of assessment and facilitate its real-world use in clinical routine. Prognostic tools, such as the PPS, are dependent on real-world usability for clinical utility, regardless of their accuracy.\textsuperscript{50} The concurrent assessment of several demographic and environmental risk factors for psychosis may not appear logistically viable in clinical practice. A digital self-assessment that could be completed relatively quickly, either at home or in a clinical setting, would reduce the burden on both patients and clinicians. This would be composed of a combination of self-administered questions and automatic extraction of location-based variables from postcode data (e.g. urbanicity, pollution). The latter allows for scoring of multiple variables with the same answer, reducing the assessment time. Acceptability of digital assessments and interventions is generally high in psychiatric fields,\textsuperscript{51-53} with high response rates seen in online screening for CHR-P specifically.\textsuperscript{54,55} Additionally, digital implementation reduces staffing and financial demands, with the assessment being self-administered. Overall, the digital implementation of the PPS aids the facilitation of implementing the tool in clinical routine.
4.5. SUMMARY, AIMS AND OBJECTIVES

Precision psychiatry can allow us to prognosticate outcomes on an individual level and inform clinical decision making. A focus on non-genetic risk and protective factors allow us to better understand the aetiopathology of psychosis, while advancing the field of precision psychiatry. However, we are currently limited by a lack of understanding of the association of non-genetic risk and protective factors and clinical outcome in CHR-P, which my PhD seeks to rectify. Further to this, I aim to use the gold standard of existing evidence to develop, digitally implement and pilot a self-assessment of exposure to non-genetic risk and protective factors for psychosis, a Psychosis Polyrisk Score (PPS), to refine prognostic estimates. More specifically, I aimed to achieve the following:

**Chapter 5**: To synthesise all available information on risk and protective factors for transition to psychosis within the clinical high risk for psychosis population and stratify these according to the strength of available evidence.

**Chapter 6**: To present the digital implementation and feasibility of administration of the PPS. We also present pilot data to underline the theoretical potential of the PPS in a stepped risk assessment framework.
REFERENCES


5. WHAT CAUSES THE ONSET OF PSYCHOSIS IN INDIVIDUALS AT CLINICAL HIGH RISK? A META-ANALYSIS OF RISK AND PROTECTIVE FACTORS

Published in Schizophrenia Bulletin in 2019
doi: https://doi.org/10.1093/schbul/sbz039

Dominic Oliver1*, Thomas Reilly2*, Ottone Baccaredda Boy2, Natalia Petros2, Cathy Davies1, Stefan Borgwardt1, Philip McGuire2, Paolo Fusar-Poli1,4,5

*DO and TR contributed equally to this manuscript
1 Early Psychosis: Interventions & Clinical detection (EPIC) lab, Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, London, SE5 8AF, UK,
2 Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, London, SE5 8AF, UK,
3 Department of Psychiatry, University of Basel, Switzerland,
4 OASIS Service, South London and the Maudsley NHS National Health Service Foundation Trust, London, UK,
5 Department of Brain and Behavioural Sciences, University of Pavia, Pavia, Italy.

Permissions
Copyright Oliver et al., 2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0). The license can be viewed at: https://creativecommons.org/licenses/by/4.0/legalcode. No changes were made in the reproduction of this article.
5.1. ABSTRACT
Twenty percent of individuals at clinical high risk for psychosis (CHR-P) develop the disorder within two years. Extensive research has explored the factors that differentiate those who develop psychosis and those who do not, but the results are conflicting. The current systematic review and meta-analysis comprehensively addresses the consistency and magnitude of evidence for non-purely genetic risk and protective factors associated with the risk of developing psychosis in CHR-P individuals. Random effects meta-analyses, Standardised Mean Difference (SMD) and Odds Ratio (OR) were used, in combination with an established stratification of evidence that assesses the association of each factor and the onset of psychotic disorders (from class I, convincing evidence to class IV weak evidence), whilst controlling for several types of biases. 128 original controlled studies relating to 26 factors were retrieved. No factors showed class I-convincing evidence. Two further factors were associated with class II-highly suggestive evidence: attenuated positive psychotic symptoms (SMD = 0.348, 95%CI: 0.280, 0.415) and global functioning (SMD = -0.291, 95%CI: -0.370, -0.211). There was class III-suggestive evidence for negative psychotic symptoms (SMD = 0.393, 95%CI: 0.317, 0.469). There was either class IV-weak or no evidence for all other factors. Our findings suggest that despite the large number of putative risk factors investigated in the literature, only attenuated positive psychotic symptoms, global functioning and negative psychotic symptoms show suggestive evidence or greater for association with transition to psychosis. The current findings may inform the refinement of clinical prediction models and precision medicine in this field.
5.2. INTRODUCTION

The introduction of the first Clinical High Risk for Psychosis (CHR-P)\(^1\) service, the PACE clinic,\(^2\) has stimulated extensive research into psychosis prevention, to the point that the CHR-P construct has become a key component of clinical services for early intervention\(^3\) (e.g. NICE guidelines;\(^4\) NHS England Access and Waiting Time standard).\(^5\) Simultaneously, some challenges have emerged, such as the need to refine the prediction of outcomes.\(^6\) A key limitation is that the level of risk observed in CHR-P individuals is mostly accounted for by their sampling.\(^7\) For example, when CHR-P criteria are applied to the general population, the level of risk of individuals meeting them is very low.\(^8\)–\(^10\) An additional problem is that there is poor knowledge in factors that modulate the level of risk in these individuals, because their identification and outcomes are entirely predicated on the basis of symptoms. However, symptoms represent an epiphenomenon of an underlying etiopathology. In fact, the overarching model underlying the development of psychosis involves the culmination of genetic and environmental factors that can increase (risk factors) or decrease (protective factors) the likelihood of developing psychosis, as well as the interaction between them.\(^11,12\) It is therefore essential to better understand the role of specific risk and protective factors in this area. Accordingly, we have recently published an umbrella review (a review of reviews) to quantitatively synthesise the existing literature on risk/protective factors for psychosis in the general population.\(^12\) In a companion meta-analysis we confirmed that CHR-P individuals accumulate several environmental risk factors for psychosis, like childhood trauma, adverse life events and affective dysfunction, compared to controls, while the role of genetic and epigenetic risk factors in this group awaits clarification.\(^13\) The effect of different risk/protective factors on the risk of developing and later transition to psychosis within individuals who have met CHR-P criteria has yet to be clarified at a meta-analytical level.

Despite much research into risk/protective factors potentially associated with transition to psychosis in CHR-P samples, the studies are often small, underpowered or inconsistent in their results. Meta-analytical methods can address these issues. Reviewing risk/protective factors for the development of psychosis within CHR-P samples is relevant twofold. Firstly, while we know that 20% of CHR-P individuals transition to psychosis within two years,\(^14\) we are currently unable to predict who will transition and who will not. Greater understanding of the specific risk/protective factors that increase risk of transition at the
individual subject level will allow for improved prognostication. Secondly, some factors may be potentially modifiable, therefore allowing for novel, individualised therapeutic strategies, thereby improving primary indicated prevention of psychosis.

To our knowledge, this study is the first meta-analysis to quantitatively synthesise the evidence for risk/protective factors for developing psychosis in CHR-P individuals. The primary aim is to systematically review the evidence for risk/protective factors within the CHR-P population and to provide a meta-analytical summary of their magnitude, direction of effect and consistency, controlling for several biases (e.g. small study effect and excess significance bias). The latter point will be achieved by complementing the standard pairwise meta-analysis with the use of validated criteria that have been developed for umbrella reviews\textsuperscript{12,15–17} to stratify the evidence of association between risk/protective factors and outcomes.
5.3. METHODS

5.3.1. SEARCH STRATEGIES

The Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA)\textsuperscript{18} and Meta-analysis of Observational Studies in Epidemiology (MOOSE) guidelines\textsuperscript{19} were adhered to throughout to achieve high quality of reporting (Supplementary Tables 5-5 & 5-6). Details of the protocol for systematic review were registered on PROSPERO (CRD42017077470).

A two-step systematic search of the literature was performed by two independent researchers (T.R. & O.B.B.) to identify relevant studies investigating the effect of risk and protective factors for transition to psychosis in CHR-P individuals.

The Ovid database by Wolters Kluwer (including MEDLINE, EMBASE and PsycINFO) was searched. Full search strategy including keywords can be seen in Supplementary Methods 5-1. The search was extended from inception until 13th May 2018.

5.3.2. INCLUSION CRITERIA

Articles meeting the inclusion criteria for the current systematic review and meta-analysis were: a) original articles, written in English b) cohort studies examining the association between risk/protective factors and psychotic disorders in the CHR-P population c) included CHR-P individuals defined by standard psychometric instruments: Comprehensive Assessment of At Risk Mental States (CAARMS);\textsuperscript{20} Brief Psychiatric Rating Scale (BPRS);\textsuperscript{21} Structured Interview for Psychosis-Risk Syndromes (SIPS);\textsuperscript{22} Basel Screening Instrument for Psychosis (BSIP)\textsuperscript{23} d) reported transitions to a psychotic disorder as a key outcome measure, defined according to standard international Diagnostic and Statistical Manual of Mental Disorders (DSM) or International Statistical Classification of Diseases and Related Health Problems (ICD) criteria -any version- e) reported follow-up of at least one year, based on meta-analytical evidence suggesting that shorter follow-up times may be associated with infrequent events\textsuperscript{24} (transitions to psychosis) resulting in underpowered studies.
5.3.3. **EXCLUSION CRITERIA**

In line with our previous work\textsuperscript{12,13} we excluded biomarkers, purely genetic factors and cognitive factors, because these would require a different and specific meta-analytical approach. Furthermore, despite advances in genetic understanding in this field (e.g. polygenic risk scores), our understanding is still relatively limited, while the role of biomarkers\textsuperscript{25} and cognition\textsuperscript{26} has already been meta-analysed by our group. As such, we excluded: (a) conference abstracts, reviews, case-reports, cross-sectional studies and case-control studies, (c) purely genetic factors, (d) biomarkers or cognitive factors, (d) studies using CHR-P definitions other than those listed above.

5.3.4. **DATA EXTRACTION**

Data extraction was done independently by two investigators (T.R. & O.B.B.). Any discrepancies were resolved in consensus meetings with another author (D.O.) under the supervision of a senior researcher (P.F.P.). Data selection and extraction was based on a systematic approach that is further detailed in Supplementary Methods 5-2. For continuous factors we also considered the mean baseline value in the transition and mean baseline value in the non-transition group. The factors were grouped in the following domains that had no influence on the statistical analyses, in line with previous studies in this area: sociodemographic/parental factors, later factors, antecedents and symptom scores/clinical factors.\textsuperscript{12,27,28} Details of risk of bias assessment can be found in Supplementary Methods 5-3.

5.3.5. **STATISTICAL ANALYSIS**

*Standard pairwise meta-analysis*

The meta-analytical effect-size measure was Odds Ratio (OR) for dichotomous factors and Standardised Mean Difference (SMD) for continuous factors. An OR greater than 1 or an SMD greater than 0 indicated that the factor was associated with an increased likelihood of psychotic disorders. OR lower than 1 or SMD lower than 0 indicated that the factor was associated with a reduced likelihood of psychotic disorders, i.e. it was protective.

The meta-analysis investigated each specific risk/protective factor without providing pooled estimates (within-subgroup summary effects) as they were felt to be clinically uninterpretable. In the case of studies reporting different definitions of the same outcome
measure (e.g. reporting both CAARMS and BPRS for symptom scores), a mean effect size and an estimate of the variance based on the calculated weight of the included definitions was computed. Random-effects models were used to control for heterogeneity.

Hierarchical classification of the evidence

In line with previous studies employing umbrella review criteria for classifying the evidence of association between risk/protective factors and health disorders, analyses included: a) an Egger test to assess small-study effects that lead to potential reporting or publication bias; and b) a test of excess significance bias. The test of excess significance bias consisted of a binomial test to compare the observed vs. the expected number of studies yielding statistically significant results. This expected number was calculated as the sum of the statistical power of the studies. Small-study effects and excess significance bias were claimed at one-sided p values <0.05, as in previous studies.

The levels of evidence of the associations between putative risk/protective factors and transition to psychotic disorder were then classified according to the guidelines for umbrella reviews: convincing (class I) when number of cases >1000, p < 10^{-6}, I^2 < 50%, 95% prediction interval excluding the null, no small-study effects, and no excess significance bias; highly suggestive (class II) when number of cases >1000, p < 10^{-6}, largest study with a statistically significant effect, and class I criteria not met; suggestive (class III) when number of cases >1000, p<10^{-3}, and class I-II criteria not met; weak (class IV) when p < 0.05 and class I-III criteria not met; non-significant when p > 0.05.

Finally, a sensitivity analysis was conducted for the factors classified as class I-III by using only prospective studies. Prospective studies allow one to address the temporality of the association, thus dealing with the problem of reverse causation. Analyses were carried out using Comprehensive Meta-Analysis Version 3 and STATA 14.
5.4. RESULTS

5.4.1. DATABASE
Overall, 77,045 records were searched, 259 were screened and 128 were eligible (see Figure 5-1). The eligible articles were published between 1998 (shortly after the first CHR-P service was established) and 13th May 2018.

Overall, the 128 eligible studies comprising 17,967 patients reported on 26 putative risk/protective factors of transition to psychotic disorders for CHR-P individuals (Supplementary Table 5-7). These 26 putative risk/protective factors were separated for descriptive purposes into four categories: sociodemographic/parental factors, later factors, antecedents and symptom scores/clinical factors.

The number of cases was greater than 1,000 for five factors (19.2%). 11 out of the 26 analysed factors had significant associations with psychosis (34.6%), with eight (30.8%) reaching p<0.001 (Tables 5-1 – 5-4, Figure 5-2). Nine factors (34.6%) presented a large heterogeneity (I^2>50%). Additionally, the evidence for small-study effects was noted for ten out of the 23 factors (43.5%) with enough studies for this to be conducted.
Records identified through database and manual searching (n = 77,045)

Records screened (n = 1,076)

Records excluded (n = 75,969)

Full-text articles assessed for eligibility (n = 259)

Full-text articles excluded for not fitting inclusion criteria (n = 126)
Overlapping (n = 5)

Studies included in the meta-analysis (n = 128)

**Figure 5-1** PRISMA flowchart outlining study selection process
Figure 5-2 Graphical summary of risk/protective factors for psychosis onset in the CHR-P state. No factors met criteria for a convincing level of evidence (class I), two factors for a highly suggestive level of evidence (class II), one factor for a suggestive level of evidence (class III) and eight for a weak level of evidence (class IV).
5.4.2. **CONVINCING EVIDENCE FOR ASSOCIATION WITH TRANSITION TO PSYCHOSIS**

There were no risk or protective factors with a convincing level of evidence (class I: number of cases >1000, p < 10^{-6}, I^2 < 50%, 95% prediction interval excluding the null, no small-study effects, and no excess significance bias) for an association with risk of transition to psychosis (Tables 5-1 – 5-4, Figure 5-2).

5.4.3. **HIGHLY SUGGESTIVE EVIDENCE FOR ASSOCIATION WITH TRANSITION TO PSYCHOSIS**

There was highly suggestive evidence (class II: >1000 cases, p<0.001, largest study with statistically significant effect, and class I criteria not met) that two further factions are associated with increased (attenuated positive psychotic symptoms; SMD = 0.348, 95%CI: 0.280, 0.415) and decreased (global functioning; SMD = -0.291, 95%CI: -0.370, -0.211) transition risk, respectively (Table 5-4, Figure 5-2).

5.4.4. **SUGGESTIVE EVIDENCE FOR ASSOCIATION WITH TRANSITION TO PSYCHOSIS**

There was suggestive evidence (class III: >1000 cases, p<0.01, class I/II criteria not met) for negative psychotic symptoms (SMD = 0.393, 95%CI: 0.317, 0.469) increasing risk of transition to psychosis (Table 5-4, Figure 5-2). This changed little when analyses were run without total negative SIPS/SOPS scores or questionable negative SIPS/SOPS items (SMD = 0.369, 95%CI: 0.280,0.458)
5.4.5. **WEAK EVIDENCE OF ASSOCIATION WITH TRANSITION TO PSYCHOSIS**

There was weak evidence (class IV: p < 0.05 and class I-III criteria not met) of an association with increased risk of transition to psychosis for one sociodemographic/parental factor (male gender, OR = 1.178, 95%CI: 1.034, 1.341), three later factors (stress/trauma (OR = 1.146, 95%CI: 1.038, 1.265), living status (OR = 1.557, 95%CI: 1.085, 2.232), employment (OR = 0.553, 95%CI: 0.400, 0.765)), one antecedent (right handedness (OR = 1.602, 95%CI: 1.041, 2.465) and three symptom scores/clinical factors (disorganised/cognitive symptoms (SMD = 0.317, 95%CI: 0.172, 0.461), general symptoms (SMD = 0.227, 95%CI: 0.122, 0.332), total symptom scores (SMD = 0.307, 95%CI: 0.148, 0.467) (Tables 5-1 – 5-4, Figure 5-2).

There was no evidence of association with transition to psychotic disorders for all other 16 factors (see Tables 5-1 – 5-4)
<table>
<thead>
<tr>
<th>Risk factor</th>
<th>K</th>
<th>Random effects measures, ES (95% CI)</th>
<th>N</th>
<th>P random effects</th>
<th>I² (p)</th>
<th>PI (95% CI)</th>
<th>LS</th>
<th>SSE/ESB</th>
<th>eOR</th>
<th>CE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male gender</td>
<td>66</td>
<td>OR, 1.178 (1.034, 1.341)</td>
<td>1732</td>
<td>0.014</td>
<td>13.983</td>
<td>0.7810 – 1.7760</td>
<td>Yes</td>
<td>No/No</td>
<td>1.178</td>
<td>IV</td>
</tr>
<tr>
<td>Urbanicity</td>
<td>4</td>
<td>OR, 1.548 (0.584, 4.104)</td>
<td>59</td>
<td>0.380</td>
<td>69.435</td>
<td>-5.0916 – 8.1876</td>
<td>Yes</td>
<td>Yes/No</td>
<td>1.548</td>
<td>ns</td>
</tr>
<tr>
<td>Age</td>
<td>61</td>
<td>SMD, -0.035 (-0.102, 0.033)</td>
<td>1776</td>
<td>0.313</td>
<td>32.012</td>
<td>-0.3260 – 0.2560</td>
<td>No</td>
<td>No/No</td>
<td>0.939</td>
<td>ns</td>
</tr>
<tr>
<td>Parental socioeconomic status</td>
<td>14</td>
<td>OR, 0.955 (0.739, 1.234)</td>
<td>444</td>
<td>0.725</td>
<td>37.519</td>
<td>-0.2389 – 2.1489</td>
<td>No</td>
<td>No/No</td>
<td>0.955</td>
<td>ns</td>
</tr>
<tr>
<td>Migrant status</td>
<td>2</td>
<td>OR, 0.932 (0.544, 1.596)</td>
<td>113</td>
<td>0.797</td>
<td>33.457</td>
<td>N/A</td>
<td>No</td>
<td>N/A/No</td>
<td>0.932</td>
<td>ns</td>
</tr>
<tr>
<td>Non-white ethnicity</td>
<td>19</td>
<td>OR, 0.949 (0.604, 1.203)</td>
<td>714</td>
<td>0.665</td>
<td>25.641</td>
<td>-0.4990 – 1.8070</td>
<td>No</td>
<td>No/No</td>
<td>0.949</td>
<td>ns</td>
</tr>
<tr>
<td>Education</td>
<td>25</td>
<td>OR, 0.872 (0.718, 1.059)</td>
<td>795</td>
<td>0.167</td>
<td>40.038</td>
<td>-0.4570 – 1.6630</td>
<td>No</td>
<td>No/No</td>
<td>0.872</td>
<td>ns</td>
</tr>
</tbody>
</table>

k – number of samples for each factor, ES – effect size, N – number of cases, PI – prediction interval, CI – confidence interval, SSE – small study effect, ESB – excess significance bias, LS – largest study with significant effect, eOR – equivalent odds ratio, CE – class of evidence, OR – odds ratio, SMD – standardised mean difference, NA – not assessable, ns – not significant

Higher classes of evidence for associations are emphasised with darker blue. Bold text is indicative of why factors are not a higher CE.
<table>
<thead>
<tr>
<th>Risk factor</th>
<th>K</th>
<th>Random effects measures, ES (95% CI)</th>
<th>N</th>
<th>P</th>
<th>I² (p)</th>
<th>PI</th>
<th>LS</th>
<th>SSE/ESB</th>
<th>eOR</th>
<th>CE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stress/trauma</td>
<td>11</td>
<td>OR, 1.146 (1.038, 1.265)</td>
<td>454</td>
<td>&lt;10⁻⁶</td>
<td>35.681 (0.113)</td>
<td>0.9015 – 1.3905</td>
<td>No</td>
<td>No/No</td>
<td>1.146</td>
<td>IV</td>
</tr>
<tr>
<td>Living status</td>
<td>10</td>
<td>OR, 1.557 (1.085, 2.232)</td>
<td>289</td>
<td>0.016</td>
<td>0.000 (0.537)</td>
<td>0.6547 – 2.4593</td>
<td>No</td>
<td>No/No</td>
<td>1.557</td>
<td>IV</td>
</tr>
<tr>
<td>Employment</td>
<td>7</td>
<td>OR, 0.553 (0.400, 0.765)</td>
<td>268</td>
<td>&lt;10⁻⁴</td>
<td>0.000 (0.870)</td>
<td>0.4000 – 0.7650</td>
<td>No</td>
<td>No/No</td>
<td>0.553</td>
<td>IV</td>
</tr>
<tr>
<td>Stigma</td>
<td>2</td>
<td>OR, 4.604 (0.825, 25.701)</td>
<td>21</td>
<td>0.082</td>
<td>70.619 (0.065)</td>
<td>N/A</td>
<td>Yes</td>
<td>N/A/Yes</td>
<td>4.604</td>
<td>ns</td>
</tr>
<tr>
<td>Substance misuse¹</td>
<td>12</td>
<td>OR, 1.322 (0.965, 1.813)</td>
<td>382</td>
<td>0.082</td>
<td>13.760 (0.310)</td>
<td>0.1734 - 2.4706</td>
<td>No</td>
<td>No/No</td>
<td>1.322</td>
<td>ns</td>
</tr>
<tr>
<td>Tobacco use</td>
<td>10</td>
<td>OR, 1.285 (0.904, 1.826)</td>
<td>233</td>
<td>0.162</td>
<td>14.907 (0.306)</td>
<td>0.0342 – 2.5358</td>
<td>No</td>
<td>No/No</td>
<td>1.285</td>
<td>ns</td>
</tr>
<tr>
<td>Cannabis use</td>
<td>23</td>
<td>OR, 1.189 (0.954, 1.480)</td>
<td>759</td>
<td>0.123</td>
<td>35.848 (0.046)</td>
<td>0.0217 – 2.563</td>
<td>No</td>
<td>Yes/No</td>
<td>1.189</td>
<td>ns</td>
</tr>
<tr>
<td>Brain injury</td>
<td>2</td>
<td>OR, 0.888 (0.561, 1.405)</td>
<td>104</td>
<td>0.611</td>
<td>0.000 (0.665)</td>
<td>N/A</td>
<td>No</td>
<td>N/A/No</td>
<td>0.888</td>
<td>ns</td>
</tr>
<tr>
<td>Alcohol</td>
<td>10</td>
<td>OR, 0.834 (0.626, 1.110)</td>
<td>472</td>
<td>0.212</td>
<td>29.747 (0.171)</td>
<td>-0.3278 – 1.9958</td>
<td>Yes</td>
<td>No/No</td>
<td>0.834</td>
<td>ns</td>
</tr>
</tbody>
</table>

k – number of samples for each factor, ES – effect size, N – number of cases, PI – prediction interval, CI – confidence interval, SSE – small study effect, ESB – excess significance bias, LS – largest study with significant effect, eOR – equivalent odds ratio, CE – class of evidence, OR – odds ratio, SMD – standardised mean difference, NA – not assessable, ns – not significant

¹ Substance misuse refers to substances not covered by other factors i.e. does not refer to alcohol, cannabis or tobacco use

Higher classes of evidence for associations are emphasised with darker blue. Bold text is indicative of why factors are not a higher CE.
Table 5-3 Level of evidence for the association of antecedents and psychotic disorders

<table>
<thead>
<tr>
<th>Risk factor</th>
<th>K</th>
<th>Random effects measures, ES (95% CI)</th>
<th>N</th>
<th>P random effects I² (p)</th>
<th>PI</th>
<th>LS</th>
<th>SSE/ESB</th>
<th>eOR</th>
<th>CE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right handedness</td>
<td>16</td>
<td>OR, 1.602 (1.041, 2.465)</td>
<td>354</td>
<td>0.032</td>
<td>0.000</td>
<td></td>
<td></td>
<td>1.602</td>
<td>IV</td>
</tr>
<tr>
<td>Perinatal complications</td>
<td>6</td>
<td>OR, 2.058 (0.893, 4.746)</td>
<td>129</td>
<td>0.090</td>
<td>87.785</td>
<td></td>
<td></td>
<td>2.058</td>
<td>ns</td>
</tr>
<tr>
<td>Height</td>
<td>5</td>
<td>SMD, 0.157 (-0.047, 0.361)</td>
<td>138</td>
<td>0.132</td>
<td>0.000</td>
<td></td>
<td></td>
<td>1.329</td>
<td>ns</td>
</tr>
<tr>
<td>BMI</td>
<td>3</td>
<td>SMD, -0.060 (-0.440, 0.320)</td>
<td>26</td>
<td>0.756</td>
<td>0.000</td>
<td></td>
<td></td>
<td>0.897</td>
<td>ns</td>
</tr>
</tbody>
</table>

k – number of samples for each factor, ES – effect size, N – number of cases, PI – prediction interval, CI – confidence interval, SSE – small study effect, ESB – excess significance bias, LS – largest study with significant effect, eOR – equivalent odds ratio, CE – class of evidence, OR – odds ratio, SMD – standardised mean difference, NA – not assessable, ns – not significant

Higher classes of evidence for associations are emphasised with darker blue. Bold text is indicative of why factors are not a higher CE.
Table 5-4 Level of evidence for the association of symptom scores/clinical factors and psychotic disorders

<table>
<thead>
<tr>
<th>Risk factor</th>
<th>K</th>
<th>Random effects measures, ES (95% CI)</th>
<th>N</th>
<th>P random effects</th>
<th>I² (p)</th>
<th>PI</th>
<th>LS</th>
<th>SSE/ESB</th>
<th>eOR</th>
<th>CE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attenuated positive psychotic symptoms</td>
<td>49</td>
<td>SMD, 0.348 (0.280, 0.415)</td>
<td>1163</td>
<td>&lt;10⁻⁶</td>
<td>69.344</td>
<td>-0.0010 – 0.6970</td>
<td>Yes</td>
<td>Yes/No</td>
<td>2.563</td>
<td>II</td>
</tr>
<tr>
<td>Global functioning</td>
<td>49</td>
<td>SMD, -0.291 (-0.370, -0.211)</td>
<td>1560</td>
<td>&lt;10⁻⁶</td>
<td>76.205</td>
<td>-0.7146 – 0.1330</td>
<td>Yes</td>
<td>Yes/No</td>
<td>0.590</td>
<td>II</td>
</tr>
<tr>
<td>Negative psychotic symptoms</td>
<td>49</td>
<td>SMD, 0.393 (0.317, 0.469)</td>
<td>1374</td>
<td>&lt;10⁻⁶</td>
<td>62.872</td>
<td>-0.0090 – 0.7770</td>
<td>No</td>
<td>Yes/No</td>
<td>2.681</td>
<td>III</td>
</tr>
<tr>
<td>Disorganised/cognitive symptoms</td>
<td>18</td>
<td>SMD, 0.317 (0.172, 0.461)</td>
<td>503</td>
<td>&lt;10⁻⁶</td>
<td>77.067</td>
<td>-0.1810 – 0.8150</td>
<td>No</td>
<td>Yes/No</td>
<td>2.485</td>
<td>IV</td>
</tr>
<tr>
<td>Total symptoms score</td>
<td>29</td>
<td>SMD, 0.307 (0.148, 0.467)</td>
<td>675</td>
<td>&lt;10⁻⁶</td>
<td>72.282</td>
<td>-0.4403 – 1.0543</td>
<td>No</td>
<td>Yes/No</td>
<td>1.743</td>
<td>IV</td>
</tr>
<tr>
<td>General symptoms</td>
<td>21</td>
<td>SMD, 0.227 (0.122, 0.332)</td>
<td>541</td>
<td>&lt;10⁻⁴</td>
<td>62.307</td>
<td>-0.1190 – 0.5730</td>
<td>No</td>
<td>Yes/No</td>
<td>2.271</td>
<td>IV</td>
</tr>
<tr>
<td>Co-morbidity</td>
<td>19</td>
<td>OR, 1.134 (0.926, 1.389)</td>
<td>587</td>
<td>0.223</td>
<td>54.470</td>
<td>0.4282 – 1.8392</td>
<td>No</td>
<td>No/No</td>
<td>1.134</td>
<td>ns</td>
</tr>
<tr>
<td>Basic symptoms</td>
<td>2</td>
<td>SMD, 0.267 (-0.027, 0.562)</td>
<td>115</td>
<td>0.075</td>
<td>43.119</td>
<td>N/A</td>
<td>Yes</td>
<td>N/A/No</td>
<td>1.621</td>
<td>ns</td>
</tr>
</tbody>
</table>

k – number of samples for each factor, ES – effect size, N – number of cases, PI – prediction interval, CI – confidence interval, SSE – small study effect, ESB – excess significance bias, LS – largest study with significant effect, eOR – equivalent odds ratio, CE – class of evidence, OR – odds ratio, SMD – standardised mean difference, NA – not assessable, ns – not significant

Higher classes of evidence for associations are emphasised with darker blue. Bold text is indicative of why factors are not a higher CE.
5.4.6. **NO CHANGE IN CLASSIFICATION OF EVIDENCE OF ASSOCIATIONS AFTER SENSITIVITY ANALYSIS**

No factors with suggestive evidence or greater (attenuated positive psychotic symptoms, global functioning and negative psychotic symptoms) were downgraded following removal of studies with retrospective designs (Supplementary Table 5-10) or studies not employing ICD or DSM criteria to determine transition status in addition to CHR-P instruments. Only one study was considered to have a retrospective design with all other studies having prospective designs.

5.5. **DISCUSSION**

To the best of our knowledge, this is the first meta-analysis of risk and protective factors for transition to psychotic disorders in CHR-P individuals that includes a robust hierarchical classification of the published evidence. After two decades of CHR-P research, it was imperative to advance knowledge by screening the available evidence against robust criteria. Overall, 128 individual studies comprising 17,967 patients and 26 factors potentially associated with transition to psychosis were included. There were no factors with convincing evidence (class I) for an association with risk of transition. Attenuated positive psychotic symptoms and global functioning were characterised by highly suggestive evidence (class II) with negative psychotic symptoms supported by suggestive evidence (class III).

The main finding of this meta-analysis is that, although a large number of risk/protective factors for transition to psychotic disorders have been evaluated in numerous CHR-P studies, none show convincing evidence with few having suggestive or stronger support. This likely reflects a research field which is fragmented, heterogeneous and that still represents a small niche to display a scalable impact. For example, the availability of
different CHR-P assessment instruments is associated with disagreement in the designation of cases or definition of their outcomes.\textsuperscript{32} The recent introduction of the DSM-5 category of Attenuated Psychosis Syndrome has further complicated the psychometric comparability of CHR-P cases.\textsuperscript{33} On a similar note, the heterogeneity of the CHR-P group has already been demonstrated at both diagnostic\textsuperscript{32,34} and prognostic\textsuperscript{14,35} level, to the point that stratification of this group has been suggested in a previous issue of this journal.\textsuperscript{1,36} The limited scalability and impact of the CHR-P field has also received empirical demonstration on several lines of evidence. Since the CHR-P literature is characterised by relatively small studies with infrequent events (transition to psychosis), the meta-analytical findings did not survive the strict criteria for the classification of evidence, with it being particularly rare for factors to have over 1000 cases. While this criterion is intended to identify robust epidemiological risk/protective factors, the CHR-P field is per se epidemiologically weak,\textsuperscript{37} because it is characterised by substantial sampling biases.\textsuperscript{7} A striking example of these points is the recent evidence showing that only about 5\% of individuals who will develop psychosis can be detected at their CHR-P stage in secondary mental health care.\textsuperscript{38,39} Overall, this finding clearly indicates that future CHR-P research should be collaborative, scalable and better harmonised in terms of assessment of intake criteria and outcomes. Ongoing international projects such as PSYSCAN,\textsuperscript{40} PRONIA\textsuperscript{41} and North American Prodrome Longitudinal Study (NAPLS)\textsuperscript{42} which have all been integrated in the HARMONY project may reach the critical mass that is needed to better identify risk/protective factors that modulate transition to psychosis with convincing evidence.

Despite these caveats, we found highly suggestive evidence that attenuated positive psychotic symptoms and global functioning are directly and inversely associated with the risk of transitioning to psychosis, respectively. These findings are unsurprising. First,
severity of positive symptoms is the main factor in deciding whether an individual meets CHR-P criteria and develops a first episode of psychosis. CHR-P individuals with higher attenuated positive psychotic symptom scores at baseline are closer to the threshold of transition and therefore do not require the same degree of symptom progression as others with less severe symptoms. Although there is consensus that current CHR-P tools are biased towards detecting attenuated positive psychotic symptoms, the P1-P4 subscales on the CAARMS and the P1-P5 subscales on the SIPS actually contain a variety of attenuated symptoms beyond positive ones. For example, obsessive thoughts, derealisation and depersonalisation experiences as well as time perception alterations. Fine-grained data is not available: most studies did not report the single severity and frequency scores of the specific CAARMS/SIPS subscales (see Supplementary Table 5-9). Moreover, this is true of randomised controlled trials. When these data were available, sensitivity analyses confirmed that all individual attenuated positive psychotic symptoms remained significant with the exception of grandiose ideas (Supplementary Table 5-10). Moreover, even when splitting attenuated positive psychotic symptoms into individual items on assessments, this may not be fine-grained enough for optimal prediction. Previous important studies have shown that auditory hallucinations may be highly predictive of transition to psychosis, while visual hallucinations may be associated with a reduced risk. Unfortunately, this level of detail in data is rarely reported in primary literature, so further analysis was not possible. Second, previous research has already shown that higher functioning at baseline is associated with reduced risk of transition. Although impaired global functioning is variably ascertained by CHR-P assessment instruments, it is one of the most robust predictors in this field. Machine-learning prediction models determined social outcomes at 1 year in up to 83% of patients in clinical high-risk states and 70% of patients with recent-onset depression. We also found suggestive evidence for a direct association of negative
psychotic symptom severity and risk of transition to psychosis. This factor would have met the criteria for highly suggestive evidence, however, the largest study\textsuperscript{49} did not show a statistically significant effect. Negative psychotic symptoms of at least moderate severity are incredibly common among CHR-P individuals with 82\% endorsing at least one negative psychotic symptom\textsuperscript{50} and with high prevalence (41\%) of comorbid affective disorders.\textsuperscript{51} Negative psychotic symptoms along with impaired baseline global functioning, are typically the driving force for individuals seeking help at CHR-P services\textsuperscript{52} and their persistence leads to poor outcomes.\textsuperscript{53}

A number of other factors were found to have weak evidence of an association with transition to psychosis in CHR-P individuals, with the key restriction for a greater class of evidence being fewer than 1000 cases. Stress/trauma increased risk of psychosis within CHR-P individuals. Our previous meta-analyses found that trauma is a key risk factor for psychosis in the general population\textsuperscript{12} and a risk factor for CHR-P status.\textsuperscript{13} Male gender was also seen to increase psychosis risk within CHR-P samples. Our previous meta-analyses found it to be a risk factor for psychosis in the general population\textsuperscript{12} and for CHR-P status,\textsuperscript{13} however with greater effect sizes than in this analysis. One potential explanation for this lies in the fact that the current analysis focuses on an enriched sample for these factors, thus diluting the variance. This is likely to be true of other factors traditionally associated with psychosis, such as cannabis use, that were found to have non-significant associations in this analysis. Moreover, cannabis use has typically been assessed in a binary fashion, measuring if individuals have ever used before, despite degree of exposure seemingly being key to the association with psychosis in both the general population\textsuperscript{54} and in CHR-P.\textsuperscript{55} We also found that employment is protective, reducing the risk of transitioning to psychosis in CHR-P individuals. Employment is an indirect measure, contingent on other factors such as
symptoms and global functioning. Right handedness also had a weak association with increased psychosis risk within CHR-P individuals. This effect was in the opposite direction to in the general population. However, as many included studies were fairly small, interpretation should be taken cautiously. Other clinical factors, particularly disorganised and cognitive symptoms, were found to have a weak association with psychosis. Their impact can be particularly relevant within the clinical subgroup of brief limited intermittent psychotic symptoms (BLIPS), where disorganising or dangerous features have been associated with an extremely high risk of transition.

The above findings can advance clinical knowledge in this area. First, they can be used to improve the prognosis of outcomes. At present, CHR-P assessment tools have outstanding sensitivity but lack specificity i.e. they are adept at ruling out psychosis risk but are inefficient at ruling it in. Accordingly, recent studies have suggested using refined clinical prediction models to improve prognostic accuracy. The risk and protective factors identified in class II and III of evidence in the current meta-analysis could represent core benchmarks for developing future clinical prediction models. Some of the factors identified by our analysis have already been incorporated into risk calculators for CHR-P individuals. For example, the NAPLS calculator includes higher levels of attenuated positive psychotic symptoms (unusual thought content and suspiciousness) and greater decline in social functioning, while another calculator similarly includes attenuated positive psychotic symptoms (unusual thought content, visual perceptual abnormalities and disorganisation), negative psychotic symptoms (social anhedonia and ideational richness) and global functioning. Prognostic accuracy can be further improved when clinical prediction models are combined with biomarker or cognitive data in a sequential assessment framework. Stepped assessments offer the advantage to optimise the resources reserving more complex
assessment to those already filtered through simpler procedures. Our analysis also reveals key risk/protective factors that at the moment present with weak evidence for association and that awaits further validation through larger cohort studies. Improved understanding of which CHR-P individuals are more likely to transition to psychosis would also lead to some potential clinical benefits such as easiest detection of those more at risk and faster referrals to early detection services, thereby reducing the duration of untreated psychosis and improving outcomes. Finally, advancing knowledge on factors that modulate the onset of psychosis within CHR-P samples can inform preventive interventions, as some of these may be potentially modifiable. Available preventive interventions do not seem to be more effective, compared to each other, nor benefit the severity of attenuated positive psychotic symptoms, negative psychotic symptoms or global functioning that have been identified as class II-III. While the meta-analytical picture is currently bleak, due to the infancy of the field there have been very few randomised controlled trials in CHR-P. Further studies and increased focus on the effects of these treatments on the severity of attenuated positive psychotic symptoms, negative symptoms and global functioning are key to the progression of the field. Since there is no evidence that current preventive treatments can reliably modify the risk of developing psychosis in CHR-P samples, experimental therapeutics in this area are urgently needed and should be the focus of the next generation of research.

The main limitation of the current analysis is that the CHR-P literature is still relatively small compared to other areas of psychiatry. For example, our umbrella review assessing risk and protective factors for psychosis was able to draw on 50 years of evidence, and yet was only able to find two factors with a convincing level of evidence. Although the CHR-P field only has the past 20 years to draw evidence from, there are still two factors with highly
suggestive levels of evidence. This may be due to the fact that the CHR-P paradigm is intrinsically embedded in prospective cohort studies. Future studies in this area have the potential to move class III factors into higher classes and therefore to progress and improve the evidence base. Similarly, as already noted in the umbrella review,\textsuperscript{12} the vast majority of factors assessed in the current literature are risk factors, rather than protective factors. Protective factors like self-esteem, social support and resilience may be better assessed in future primary research studies to identify what may aid psychosis prevention. Another limitation is the clinical heterogeneity of the CHR-P population. Within this, there are people with attenuated psychotic symptoms (APS), BLIPS and genetic risk and deterioration syndrome (GRD).\textsuperscript{14} Furthermore, there are differences between APSS (attenuated positive symptom syndrome) as defined by the SIPS, APS as defined by the CAARMS, DSM-5 APS, as well as others.\textsuperscript{33} However, these differences are limited at the meta-analytical level\textsuperscript{8} with the majority of risk (around 60%) for developing psychosis in CHR-P individuals being accumulated before the assessment is performed.\textsuperscript{7,71} As such, there is a high degree of variance within the CHR-P cohorts in these studies, which can dilute the effect of certain risk factors as they can affect these subgroups differently. Future studies would be wise to subdivide and sufficiently power their samples to ascertain the differential effects of risk/protective factors on these subgroups.\textsuperscript{1} Finally, there were only a few studies available to contribute data for some factors, and as for any other meta-analysis that has adopted our classification criteria of evidence, we cannot exclude that other risk/protective factors may be identified and published in the near future.

**CONCLUSIONS**

Severity of attenuated positive psychotic symptoms and low global functioning show convincing evidence, while severity of negative psychotic symptoms shows suggestive
evidence for increasing transition risk in CHR-P individuals. These factors should be considered as benchmarks by future clinical prediction models and key targets of new experimental therapeutics.

Acknowledgments

Dominic Oliver 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.
5.6. REFERENCES


47. Bearden CE, Wu KN, Caplan R, Cannon TD. Thought disorder and communication deviance as


69. Devoe DJ, Farris MS, Townes P, Addington J. Interventions and social functioning in youth at risk of


5.7. SUPPLEMENTARY DATA

Oliver D, Reilly T, Baccaredda Boy O et al. What causes the onset of psychosis in individuals at clinical high risk? A meta-analysis of risk and protective factors.

Omitted from thesis for brevity

[Link to supplementary]

**Supplementary Methods 5-1** Search keywords
**Supplementary Methods 5-2** Extracted variables
**Supplementary Methods 5-3** Risk of bias (quality) assessment
**Supplementary Table 5-5** MOOSE guidelines for meta-analysis and systematic reviews
**Supplementary Table 5-6** PRISMA guidelines for meta-analysis and systematic reviews
**Supplementary Table 5-7** Factors examined by study
**Supplementary Table 5-8** Sensitivity analyses of attenuated positive psychotic symptoms at the symptom level
**Supplementary Table 5-9** Scales used to assess attenuated positive psychotic symptoms, negative psychotic symptoms and global functioning
**Supplementary Table 5-10** Sensitivity analysis for the associations of symptom scores/clinical factors and psychotic disorders within individual prospective studies of class I-III factors
**Supplementary Table 5-11** Risk of bias (quality assessment) using modified Newcastle Ottawa Scale
**Supplementary Figure 3-50** Funnel plot of standard error/precision against standardised mean difference/log odds ratio
**Supplementary References**
6. REAL-WORLD DIGITAL IMPLEMENTATION OF THE PSYCHOSIS POLYRISK SCORE (PPS): A PILOT FEASIBILITY STUDY

Published in Schizophrenia Research in 2020
doi: https://doi.org/10.1016/j.schres.2020.04.015

Dominic Oliver1,2, Giulia Spada1,2, Amir Englund3, Edward Chesney3, Joaquim Radua1,4,5, Abraham Reichenberg3,6,7,8, Rudolf Uher9, Philip McGuire3, Paolo Fusar-Poli1,2,10

1 Early Psychosis: Interventions and Clinical Detection (EPIC) Lab, Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom
2 OASIS Service, South London and the Maudsley NHS Foundation Trust, London, United Kingdom
3 Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom
4 Imaging of Mood- and Anxiety-Related Disorders (IMARD), Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), CIBERSAM, Barcelona, Spain
5 Department of Clinical Neuroscience, Centre for Psychiatry Research, Karolinska Institutet, Stockholm, Sweden
6 Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, United States
7 Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY, United States
8 Frieman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, United States
9 Department of Psychiatry, Dalhousie University, Halifax, NS, Canada
10 Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy

Permissions
Copyright Oliver et al., 2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0). The license can be viewed at: https://creativecommons.org/licenses/by/4.0/legalcode. No changes were made in the reproduction of this article.
6.1. ABSTRACT

**Background:** The Psychosis Polyrisk Score (PPS) is a potential biomarker integrating non-purely genetic risk/protective factors for psychosis that may improve identification of individuals at risk and prediction of their outcomes at the individual subject level. Biomarkers that are easy to administer are direly needed in early psychosis to facilitate clinical implementation. This study digitally implements the PPS and pilots its feasibility of use in the real world.

**Methods:** The PPS was implemented digitally and prospectively piloted across individuals referred for a CHR-P assessment (n=16) and healthy controls (n=66). Distributions of PPS scores was further simulated in the general population.

**Results:** 98.8% of individuals referred for a CHR-P assessment and healthy controls completed the PPS assessment with only one drop-out. 96.3% of participants completed the assessment in under 15 minutes. Individuals referred for a CHR-P assessment had higher PPS scores (mean=6.2, SD=7.23) than healthy controls (mean=-1.79, SD=6.78, p < 0.001). In simulated general population data, scores were normally distributed ranging from -15 (lowest risk, RR=0.03) to 39.5 (highest risk, RR=8912.51).

**Discussion:** The PPS is a promising biomarker which has been implemented digitally. The PPS can be easily administered to both healthy controls and individuals at potential risk for psychosis on a range of devices. It is feasible to use the PPS in real world settings to assess individuals with emerging mental disorders. The next phase of research should be to include the PPS in large-scale international cohort studies to evaluate its ability to refine the prognostication of outcomes.
6.2. INTRODUCTION

Primary indicated prevention in individuals at Clinical High Risk for Psychosis (CHR-P)\(^1\) has the potential to alter the course of psychotic disorders.\(^2\) Current assessment tools, like the Comprehensive Assessment for At-Risk Mental States (CAARMS) and the Structured Interview for Psychosis-risk Syndromes (SIPS) have very good prognostic accuracy (area under the curve at 3 years: 0.9) which is comparable to that of other tests used in clinical medicine. However, while these instruments are effective in ruling out psychosis risk, they are sub-optimal for predicting psychosis risk,\(^3,4\) leading to a 22% transition risk within three years.\(^5\) Furthermore, these psychometric tools alone are limited to the ascertainment of symptoms and are unable to accurately predict the disease course at an individual level.\(^6,7\)

There is therefore a clear need to improve prognostic accuracy of the CHR-P assessment by supplementing the clinical assessment with additional information from electrophysiology, neurocognition, blood-based, neuroimaging, genetics or environmental risk/protective biomarkers.

A number of studies have demonstrated that structural neuroimaging can predict both transition to psychosis (accuracy 82%;\(^8\) accuracy 84.2%)\(^9\) and global functioning (accuracy 76.9%)\(^10\) at the individual subject level. However, the implementation of these biomarkers in clinical routine has been limited by the cost of attaining sufficiently large datasets\(^7\) and of associated logistical challenges.

Genetic biomarkers for psychosis risk have also been investigated: psychotic disorders such as schizophrenia are highly genetic conditions, with a first-degree heritability of 64% (95%CI: 62-68%).\(^11\) Single genetic alterations are unlikely to be useful to predict clinical outcomes in psychosis: polygenic risk scores encompassing several genetic alterations have been developed using data from large genome-wide association studies (GWAS). However, polygenic risk scores have so far managed to only explain 18% of the variance between cases of established schizophrenia and healthy controls\(^12\) and only 9% of the variance between cases of first-episode psychosis and healthy controls.\(^13\) In CHR-P individuals, polygenic risk scores have been tested for stratified predictions,\(^14\) but again the proportion of variance explained still remain modest (12.3%). Therefore, the prognostic ability of polygenic risk scores needs to be supplemented by additional information before it can be considered for real-world individualised risk prediction.
Environmental data may be the most appropriate modality to supplement polygenic risk scores, because the aetiology of psychotic disorders involves direct genetic and environmental effects as well as their interaction. Numerous factors can modulate an individual’s risk of developing psychosis in the general population, or in those meeting CHR-P criteria compared to healthy controls or in CHR-P individuals transitioning to psychosis compared to those who do not transition. Emulating the polygenic risk score approach, environmental risk/protective factors for psychosis have been combined into a single measure: the Psychosis Polyrisk Score (PPS).

The conceptual and methodological underpinnings for developing the PPS have been detailed previously in a conceptual review. The PPS incorporates environmental risk factors for psychosis (sociodemographic, social, parental, perinatal, later risk/protective factors and antecedents) which have been systematically appraised in an umbrella review and stratified according to their level of evidence for association with psychotic disorders. Demographic, parental, social, and perinatal risk factors are generally thought to exert their role during the early developmental phases, while later risk/protective factors and antecedents can modulate psychosis risk from late childhood up to shortly prior psychosis onset. Later risk factors indicate a passive exposure to socio-environmental factors, whereas antecedents involve active risk-modifying processes involved in psychosis onset and premorbid deviations in functioning. However, these categories are only descriptive and may overlap.

The current study extends this line of research by digitally implementing the PPS and by conducting the first pilot feasibility study for its use in the real world. We also compared the distribution of PPS scores across three different groups: individuals referred for a CHR-P assessment, healthy controls and a simulated dataset representing the general population. We hypothesised that individuals referred for a CHR-P assessment would have the highest PPS scores and healthy controls the lowest. The findings of this study will inform subsequent research in this area, in particular the incorporation of the PPS in large scale prospective CHR-P cohort studies that are about to start.
6.3. MATERIALS AND METHODS

6.3.1. PARTICIPANTS

The study included two participant groups: individuals referred for a CHR-P assessment and healthy controls.

6.3.2. INDIVIDUALS REFERRED FOR A CHR-P ASSESSMENT

These individuals were sampled from those referred to Outreach and Support in South London (OASIS)\textsuperscript{22,23} a CHR-P service, part of the South London and Maudsley NHS Foundation Trust (SLaM) from secondary mental health services on suspicion of psychosis risk. They were all under SLaM care for a non-psychotic mental disorder. These individuals were included if they i) had been referred for a CHR-P assessment by SLaM mental health services; ii) were older than 14 years of age, with no upper age limit (to capture pathways to care within SLaM early intervention services in a lifespan-inclusive approach); iii) were willing and able to provide written informed consent; iv) had sufficient understanding of the English language; v) had no diagnoses of organic psychiatric disorders, substance-induced psychotic disorders or psychotic disorders (with the exception of Acute and Transient Psychotic Disorders; ATPD; Supplementary Methods 6-1), as ascertained by SLaM clinicians. In addition to the PPS, standard psychometric tools including the CAARMS and Global Assessment of Functioning (GAF) were administered.

6.3.3. HEALTHY CONTROLS

Healthy controls were recruited from a separate study completed at King’s College London via university circular emails. Participants met inclusion criteria if they i) were aged over 14, with no upper age limit; ii) were willing and able to provide written informed consent; iii) had sufficient understanding of the English language, iv) did not have a past or present mental disorder (any type) or psychotic-like symptoms or physical illness, as detected through the mini-Structured Clinical Interview for DSM-IV (mini-SCID), Adolescent Psychotic-like Symptom Screener (APSS)\textsuperscript{24} or physical examination; iv) had never been treated with antipsychotic or antidepressant medication at any dosage; v) did not have a first degree relative affected by a psychotic illness (assessed by the mini-SCID).
6.3.4. OPERATIONALISATION OF THE PPS

Selection of risk factors

To develop the PPS, we initially considered all 17 risk factors that were meeting the highest hierarchy of evidence (i.e. class I–III) for association with psychotic disorders, as detailed in the umbrella review.\textsuperscript{15} We then excluded those factors (n=6) that could not easily be measured at scale due to cost (Toxoplasma Gondii IgG), extended assessment time (premorbid IQ) or limited reliability of self-report (olfactory identification ability, minor physical anomalies). We also excluded the CHR-P state to investigate the PPS independent from this construct. This left 12 remaining class I-III risk factors (Supplementary Methods 6-2). In addition to these risk factors, we additionally included some class IV risk factors that could be recorded at low cost, high reliability and limited assessment time (n=11, Supplementary Methods 6-2). A final number of 22 factors were eventually operationalised in the PPS (Supplementary Methods 6-2).

To ensure accurate scoring that was reflective of the risk factor as reported in the umbrella review,\textsuperscript{15} where possible the same tools were used to assess the presence/absence of the risk factor as the studies from which the risk was reported. The tools used to ascertain the presence/absence of each risk factor as well as details of operationalisations and the cut-offs for defining each respective Odds Ratio (OR) of included factors can be seen in Supplementary Table 6-3.

Integration of different risk factors in the PPS

Similar to the polygenic risk score, the PPS is a weighted sum of exposure to risk and protective factors, using the ORs associated with each factor\textsuperscript{15} (see Supplementary Methods 6-3).

Simple adaptations were added to the model to account for interdependencies in exposures, such as the multiple risk/protective factors associated with immigration status (see Table 6-1). For example, individuals cannot be exposed to certain immigration-based risk/protective factors in conjunction with each other. Immigrants cannot be both first-generation and second-generation, and North African immigrants have to be either first- or second-generation immigrants. We combined these factors following this logic and assuming that the proportion and extra risk of North African immigrants is similar in first- and second-generation immigrants.\textsuperscript{25} Factors related to ethnicity have similar logical dependencies
between them, e.g. black Caribbean is a non-white ethnicity, and individuals cannot be from a low ethnic density area, from a medium density area and from a high ethnic density area at the same time. We combined these factors again following this logic and assumed that the proportion and extra risk of black Caribbean individuals between non-white ethnicity individuals is similar in low, medium and high ethnic density areas. Following these combinations, the final scoring of these factors can be seen in Table 6-1). Limitations of the PPS were previously presented.  

6.3.5. REAL-WORLD DIGITAL IMPLEMENTATION OF THE PPS ASSESSMENT

Feasibility is one of the core deliverables of the PPS, as risk estimation systems are of little value without real-world usability. To enhance this aspect, the PPS was designed as a self-report assessment, minimising researcher/clinician burden. Moreover, the PPS was implemented online to facilitate its administration on tablets (as it was in this study), computers or phones. This provides options for users and, alongside a progress saving feature, increases likelihood of completed assessments as individuals can complete them on any device across multiple sessions.

To characterise the administration properties of the PPS, at the end of the assessment, participants were asked how long the assessment took to complete (less than fifteen minutes, 15-30 minutes, 30-45 minutes or more than 45 minutes) and how distressing they found the process of completing the questionnaire (not distressing at all, mildly distressing, very distressing or extremely distressing).

6.3.6. STATISTICAL ANALYSIS

Baseline clinical and sociodemographic characteristics of the participants (individuals referred for a CHR-P assessment and healthy controls) were described by means and standard deviations for continuous variables, and absolute and relative frequencies for categorical variables. Differences between continuous sociodemographic/clinical variables in the two participant groups were assessed using independent sample t-tests;
Table 6-1: Scoring system for the Psychosis Polyrisk Score (PPS)

<table>
<thead>
<tr>
<th>Risk/Protective Factor</th>
<th>Exposure/Conditional</th>
<th>PPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Childhood Trauma</td>
<td>Yes</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>-0.5</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>White</td>
<td>-2</td>
</tr>
<tr>
<td></td>
<td>Black Caribbean</td>
<td></td>
</tr>
<tr>
<td></td>
<td>In low ethnic density area</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>In medium ethnic density area</td>
<td>5.5</td>
</tr>
<tr>
<td></td>
<td>In high ethnic density area</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td></td>
</tr>
<tr>
<td></td>
<td>In low ethnic density area</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>In medium ethnic density area</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>In high ethnic density area</td>
<td>1</td>
</tr>
<tr>
<td>Immigration</td>
<td>Not immigrant</td>
<td>-0.5</td>
</tr>
<tr>
<td></td>
<td>1st generation immigrant</td>
<td></td>
</tr>
<tr>
<td></td>
<td>From North Africa</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>From other regions</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2nd generation immigrant</td>
<td></td>
</tr>
<tr>
<td></td>
<td>From North Africa</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>From other regions</td>
<td>1.5</td>
</tr>
<tr>
<td>Non-right-handedness</td>
<td>Yes</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Pollution</td>
<td>Yes</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>-5.5</td>
</tr>
<tr>
<td>Urbanicity</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>-2.5</td>
</tr>
<tr>
<td>Winter or spring birth in</td>
<td>Yes</td>
<td>0</td>
</tr>
<tr>
<td>northern hemisphere</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Paternal age</td>
<td>&lt;35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;45</td>
<td></td>
</tr>
<tr>
<td>Low paternal socioeconomic status</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Parental severe mental illness</td>
<td>Yes</td>
<td>5.5</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>-2</td>
</tr>
<tr>
<td>Adult life events</td>
<td>Yes</td>
<td>5.5</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>-2</td>
</tr>
<tr>
<td>Daily smoker</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>-0.5</td>
</tr>
<tr>
<td>Heavy cannabis use</td>
<td>Yes</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Hearing problems in past 12 months</td>
<td>Yes</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Trait anhedonia</td>
<td>Yes</td>
<td>6.5</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Male &amp; 25-35yo</td>
<td>Yes</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
</tr>
</tbody>
</table>
differences between categorical sociodemographic/clinical variables were assessed using Fisher’s exact test.

To further estimate the usability of the PPS in the wider scenario, we built a simulated dataset to investigate the range and distributions of PPS scores in the general population. This additionally allows for greater comparability with the pilot PPS presented in our previous paper.\textsuperscript{18} In a first step, we ran 10,000,000 permutations for each PPS risk factor using general population prevalence data that best represented the sample from which the ORs were generated from (Supplementary Table 6-4). In a second step, a PPS score was generated for each permutation, enabling investigation of the range and distribution of PPS scores in the general population. Normality of distribution was investigated using an adjusted Jarque-Bera test.

We then tested the differences in PPS scores between the three groups (individuals referred for a CHR-P assessment, healthy controls and simulated general population) using one-way analysis of variance (ANOVA) and post-hoc Tukey Honest Significant Differences test. We also performed correlation tests in individuals referred for a CHR-P assessment using Pearson’s R to investigate potential correlations between PPS and CAARMS total scores. All analyses were conducted in R version 3.3.2\textsuperscript{27} and significance was set to $P < .05$.

### 6.4. RESULTS

#### 6.4.1. CHARACTERISTICS OF STUDY PARTICIPANTS

Sixteen individuals referred for a CHR-P assessment were recruited onto the study. One of them was unable to complete the assessment due to fatigue, leaving a final sample of fifteen individuals. Four individuals (26.7\%) were presenting with a diagnosis of ATPD and the remaining 11 (73.3\%) with bipolar mood disorders. Two individuals (13.3\%) met CHR-P criteria with a further two (13.3\%) meeting CAARMS criteria for psychosis. Mean age was 36.7 (SD=12.7) and 26.7\% were male (Table 6-2).

Sixty-six healthy controls were recruited onto the study. Mean age was 25.9 (SD = 4.9) and 45.5\% were male (Table 6-2).
Table 6-2 Sociodemographics of study participants: individuals referred for a CHR-P assessment and healthy controls.

<table>
<thead>
<tr>
<th></th>
<th>Individuals referred for a CHR-P assessment (n = 15)</th>
<th>Healthy controls (n = 66)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>36.7 (12.7)</td>
<td>25.9 (4.9)</td>
<td>0.006</td>
</tr>
<tr>
<td>Ethnicity No. (%White)</td>
<td>6 (40%)</td>
<td>52 (78.8%)</td>
<td>0.008</td>
</tr>
<tr>
<td>Gender No. (%Male)</td>
<td>4 (26.7%)</td>
<td>30 (45.5%)</td>
<td>0.25</td>
</tr>
<tr>
<td><strong>Index diagnosis</strong> No. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute and transient psychotic disorders</td>
<td>3 (25.0%)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Bipolar mood disorders</td>
<td>12 (75.0%)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>CAARMS P1-P4</strong></td>
<td>11.29 (5.59)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>No. meeting CHR-P criteria (%)</td>
<td>2 (13.3%)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>No. meeting psychosis criteria (%)</td>
<td>2 (13.3%)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>GAF</strong></td>
<td>71 (11.71)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Abbreviations: CAARMS Comprehensive Assessment for At-Risk Mental States P1-P4 domains; GAF Global Assessment of Functioning
6.4.2. **DIGITAL IMPLEMENTATION AND FEASIBILITY OF ADMINISTRATION**

The PPS has been integrated online ([https://www.youngspace.org](https://www.youngspace.org)) as a digital assessment tool that can be presented on multiple devices. 15/16 individuals referred for a CHR-P assessment (93.7%) were able to complete the digital self-report PPS with a single participant unable to complete due to fatigue. All 66 healthy controls completed the PPS assessment. Overall 81/82 participants (98.8%) were able to complete the assessment. 78 (96.3%) participants completed the assessment in under 15 minutes. No participants found the content of the assessment distressing.

6.4.3. **PPS SCORES IN INDIVIDUALS REFERRED FOR A CHR-P ASSESSMENT**

PPS scores in this group ranged from -3 (RR = 0.5) to 18 (RR = 63.10). Mean PPS score was 6.2 (SD = 7.23) The median PPS score in this group was 9 (RR = 7.94). The distributions in the quartile ranges were Q1 (-3 to -1.5): 26.7%; Q2 (-1 to 9): 33.3%; Q3 (9.5 to 11.5): 20%; Q4 (12 to 18): 20%. 60.0% of participants in this group had a PPS > 5 (RR > 3) with 6.7% having a PPS > 15 (RR > 30). Correlations between PPS and CAARMS total scores were non-significant (p=0.26; Supplementary Figure 6-3).

6.4.4. **PPS SCORES IN HEALTHY CONTROLS**

PPS scores in this group ranged from -14 (RR = 0.04) to 14.5 (RR = 28.18). Mean PPS score was -1.79 (SD = 6.78). The median PPS score in this group was -1.75 (RR = 0.67). The distributions in the quartile ranges were Q1 (-14 to -7): 27.3%; Q2 (-6.5 to -1.75): 22.7%; Q3 (-1.5 to 1.875): 24.2%; Q4 (2 to 14.5): 25.8%.

19.7% of participants in this group has a PPS > 5 (RR > 3) with 0% having a PPS > 15 (RR > 30).

6.4.5. **PPS SCORES IN THE SIMULATED GENERAL POPULATION**

PPS scores in this simulated group ranged from -15 (lowest risk; RR = 0.03) to 39.5 (highest risk; RR = 8912.51) (Figure 6-1). These scores were normally distributed (AJB = 104030, p < 0.001). Mean PPS score in this simulated group was 0.817 (SD = 6.87). The median PPS score was 0 (RR = 1). The distributions in the quartile ranges were Q1 (-15 to -4): 29.0%; Q2 (-3.5 to 0): 21.0%; Q3 (0.5 to 5.5): 26.5%; Q4 (6 to 39.5): 23.5%.
Figure 6-1 Distribution of PPS scores in the simulated general population. Blue bars indicate the proportion of individuals receiving each PPS score based on the prevalence of risk factors and 10,000,000 permutations.

Red line indicates the density curve to highlight normality.
Figure 6-2 Violin plot comparing PPS scores distribution in individuals referred for a CHR-P assessment, controls and the (simulated) general population.

PAT, patients referred for a CHR-P assessment; POP, simulated general population; CON, healthy controls.
26.7% of individuals in this simulated group had a PPS > 5 (RR > 3) with 2.7% of individuals having a PPS > 15 (RR > 30).

6.4.6. COMPARISON OF PPS SCORE DISTRIBUTION BETWEEN INDIVIDUALS REFERRED FOR A CHR-P ASSESSMENT, HEALTHY CONTROLS AND THE GENERAL POPULATION

There was a main effect of group on PPS scores (F2,10000000 = 9.357, p = 0.001). Individuals referred for a CHR-P assessment had higher PPS scores (mean = 6.2, SD = 7.23) than healthy controls (mean = -1.79, SD = 6.78) (p < 0.001) and the simulated general population dataset (mean = 0.817, SD = 6.87; p = 0.007). PPS scores in the simulated general population were higher than healthy controls (p = 0.006) (Figure 6-2).

6.5. DISCUSSION

This pilot study digitally implemented the PPS and illustrates its real-world feasibility of use in different scenarios. Additionally, it supports its potential as a biomarker to complement current assessment tools and improve both identification of individuals at risk of psychosis and prediction of their clinical outcomes.

The results of this pilot feasibility study reinforce the theoretical strengths of the PPS biomarker on a pragmatic implementation level. The digital implementation of the PPS highlights its future real-world usability, particularly through its ease of administration. However, further large-scale studies are needed to confirm whether the PPS can reliably identify individuals at risk for psychosis or improve the prediction of their outcomes. Digital assessments reduce costs and increase data completeness,28 take less time than paper assessments29 and are usually preferred by participants.30 Additionally, simulated data in the general population displayed a broad range of scores, suggesting a high level of potential variance in future studies, which is imperative for individualised prediction. There is currently high heterogeneity within the CHR-P construct with some CHR-P individuals at extremely high risk of transition to psychosis (e.g. Brief Limited Intermittent Psychotic Symptoms [BLIPS])31,32 and others who barely differ from the general population (e.g. Genetic Risk and Deterioration [GRD])31. While these subgroups can be used to stratify people to some degree, this approach is sub-optimal as high heterogeneity remains within the largest subgroup (85% of CHR-P individuals; (Attenuated Psychotic Symptoms [APS])).31 The high variance in PPS scores in this simulated dataset and independence of PPS and CAARMS total scores support the potential of the PPS to refine estimates of
psychosis risk, however this needs to be confirmed by future studies. This allows for transdiagnostic detection of individuals at risk for psychosis beyond the CHR-P approach.\textsuperscript{33,34} Further to this, the PPS was not age-restricted, instead adopting a lifespan-inclusive approach, allowing for an extended CHR-P phenotype unrestricted by age. The PPS concept was built on the assumption that as individuals accumulate environmental risk factors for psychosis through the various stages of help-seeking and disorder progression,\textsuperscript{15,16} which is supported by the higher scores being seen in individuals referred for a CHR-P assessment than in healthy controls. The PPS is also robust, as risk factors were selected systematically \textit{a priori}. As the weighting of the PPS risk factors was determined \textit{a priori} by umbrella review, the gold standard of evidence,\textsuperscript{7} the PPS is therefore based on data from a combination of different samples in different settings, reducing risk of overfitting issues and improving generalisability. The PPS is also optimisable: the knowledge base is dynamic and changing, which allows for weights of existing factors to be adjusted as new evidence becomes available or for new predictors to be added to the model, particularly more dynamic factors, e.g. symptom severity at multiple time points,\textsuperscript{35} digital phenotyping measures\textsuperscript{36} or automated speech analysis,\textsuperscript{37} which allow for repeated use and dynamic mapping. Similarly, if those factors that were previously considered to be impractical to measure exposure to (e.g. Toxoplasma Gondii) become more feasible, they could also be integrated.

The main application of the PPS biomarker may be to improve the detection of individuals at risk for psychosis.\textsuperscript{38} Despite considerable effort to improve the detection of CHR-P individuals, only 5-12\% of first episode of psychosis cases are detected prior to psychosis onset.\textsuperscript{39,40} Furthermore, up to one third of first episode cases do not experience a prodrome.\textsuperscript{41–43} The PPS could potentially then provide an adjunct to current symptom-based risk ascertainment strategies to identify individuals at risk who do not experience CHR-P features such as attenuated or intermittent psychotic symptoms before their first episode. The PPS could also supplement digital detection strategies such as the Youth-Mental Risk and Resilience study (YouR-Study).\textsuperscript{44} The YouR-Study integrated an online screening tool encompassing the Prodromal Questionnaire (PQ-16)\textsuperscript{45} and a nine-item scale of perceptual and cognitive aberrations. Out of 2,279 participants completing the assessment, 78\% met a risk threshold were invited to attend a clinical assessment, 356 interviews were completed, 28\% of these individuals met CHR-P criteria and 2\% met criteria for first episode psychosis at this assessment.\textsuperscript{44} This study provided the first evidence of feasibility of digital detection tools improving identification of psychosis in the general population but performance in
terms of sensitivity/specificity (81%/57%) of the screener could be improved upon. The YouR-Study approach could be combined with the PPS and the additional information could produce a more accurate screening tool and improve detection of CHR-P individuals. Integration of the PPS within YouR-Study is being planned as part of an ongoing project.

The above targets can be fully accomplished within a stepped assessment strategy which stratifies individuals’ risk of developing psychosis. Leveraging different assessment types in sequence could be a pragmatic method that greatly benefits risk stratification. For example, while Electronic Health Records are not data rich enough to allow for automated screening of PPS variables on a large scale—a goal achieved using risk prediction models that incorporate predictors collected routinely in clinical care\(^{40,46,47}\)—the use of natural language processing methods may fill this gap. The PPS biomarker, if integrated into Electronic Health Records or website approaches (such as YouR-Study) could then become the first entry point in a stepped risk stratification framework. More labour- and time-intensive biomarkers, such as those that rely on cognitive assessments\(^ {48}\), clinical assessments\(^ {48,49,50}\) or neuroimaging\(^ {8,51,52,10}\) could be reserved to those individuals initially detected through PPS screening.\(^ {46}\) However, more work needs to be done to assess the prognostic ability of the PPS. A previous study has demonstrated the potential predictive gain in sequential testing in this manner, with clinical and electrophysiological biomarkers following an initial CHR-P assessment, followed by subsequent structural MRI and blood biomarkers, with an individual only progressing to the next risk testing stage if the previous test was positive.\(^ {53}\)

The main limitation of this study is the small sample and lack of external validation of the PPS. Since we did not report data on transition to psychotic disorders, we were unable to produce measures of prognostic accuracy; the small sample size would have prevented meaningful analyses. The study is similarly underpowered to compare referrals who met CHR-P criteria and those who did not or to test correlations with psychopathology or other clinically relevant variables. Future efforts are clearly needed to further dissect the heterogeneity of CHR-P outcomes. This would clearly require establishing large scale datasets with enough power and follow-up time to assess the validity of the PPS. A potential research framework that may validate the PPS biomarker is the proposed 26-site ProNET cohort study. Other large-scale international collaborations that have recently been completed include the HARMONY project, incorporating NAPLS (https://campuspress.yale.edu/napls/),\(^ {54}\) PRONIA (https://www.pronia.eu/) and PSYSCAN
Another crucial limitation is that it is also unlikely that the risk factors used in the PPS are independent as, assumed. Further work emerging from the above international consortia will clarify how the PPS risk factors interact.

CONCLUSIONS
The digital implementation of the PPS as a self-assessed biomarker facilitates its real-world use in diverse scenarios. The PPS biomarker holds theoretical potential for improving the detection of individuals at risk and prediction of their outcomes, in particular if used within a stepped risk assessment framework. Future large-scale international consortia are needed to validate the PPS prospectively.

Contributions
D.O., G.S., A.E. and E.C. collected the data. D.O. performed all data analysis and drafted the first version of the manuscript. J.R. advised on data analysis. D.O., A.R., R.U., P.M. and P.F.-P. advised on study design and interpretation of results. P.F.-P. designed and supervised the study. All authors read and edited the manuscript.

Acknowledgements
This study was supported by the King’s College London Confidence in Concept award from the Medical Research Council (MRC) (MC_PC_16048) to P.F.-P.

D.O. 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.

Declarations of competing interests
P.F.-P. has received advisory consultancy fees from Lundbeck outside of this work. The authors have declared that there are no conflicts of interest in relation to the subject of this study.
6.6. REFERENCES


6.7. SUPPLEMENTARY DATA

Oliver D, Spada G, Englund A et al. Psychosis Polyrisk Score (PPS): a pilot study to improve detection of individuals at-risk and prediction of clinical outcomes

Presented in this thesis in part for brevity

Link to full supplementary

Supplementary Methods 6-1 Rationale for inclusion of acute and transient psychotic disorders
Supplementary Methods 6-2 List of included risk/protective factors from umbrella review*
Supplementary Methods 6-3 Details of PPS construction*
Supplementary Table 6-3 List of included factors, along with their definitions and the tentative cut-offs for defining each respective Risk Ratio*
Supplementary Table 6-4 Prevalence of factors used in PPS in the general population
Supplementary Figure 6-3 Correlation of PPS scores and CAARMS Total
Supplementary References

*Included here
Supplementary Methods 6-2 List of included risk/protective factors from umbrella review

Class I
1. Black-Caribbean ethnicity in England

Class II
2. Ethnic minority in low ethnic density area
3. Second generation immigrants
4. Trait anhedonia

Class III
5. Childhood trauma
6. Ethnic minority in high ethnic density area
7. First generation immigrants
8. Non-right handedness
10. Urbanicity
11. Winter/spring season of birth in Northern hemisphere

Class IV
12. Adult life events
13. Other ethnicity (Asian/Black African/Mixed/Other White ethnicity in England)
14. Heavy cannabis use
15. Low paternal socio-economic status
16. Parental severe mental illness
17. Tobacco use
18. Traffic
19. Paternal age >35
20. Paternal age >45
21. Hearing problems in past 12 months
22. Male & aged 25-35 years old
Supplementary Methods 6-3 Details of PPS construction

To construct the PPS, we first estimated a raw score for each factor as the 10-base logarithm of its OR. For example, the OR of psychosis in individuals living in urban settings is 2.2, and thus the raw score of the urbanicity factor was \( \log_{10}(2.2) = 0.34 \). We then subtracted the population average of this raw score, so that individuals at risk would have positive scores and the remaining individuals would have negative scores, with an average of zero. For example, given that \(~73.6\%\) individuals live in urban settings worldwide (and thus \(~26.4\%\) in rural settings with a raw score of 0), the population average of the urbanicity factor should be \((73.6\% \times 0.34) + (26.4\% \times 0) = 0.25\). We subtracted this average from the raw scores, i.e., the subtracted score was \(0.34 - 0.25 = 0.09\) for individuals in urban settings and \(0 - 0.25 = -0.25\) for individuals in rural settings. Further information about prevalence data used can be seen in Supplementary Table 6-4. Finally, for the ease of use we multiplied the subtracted scores by 10 and rounded them to the nearest half integer. In the example, the final scores were \(0.09 \times 10 \approx 1\) for individuals in urban settings and \(-0.25 \times 10 = -2.5\) for individuals in rural settings.
<table>
<thead>
<tr>
<th>Risk factor assessed (Classification of evidence; Meta-analytical reference)</th>
<th>Assessment tool; Reference</th>
<th>Scoring (Classification of evidence; meta-analytical reference)</th>
<th>Cutoff</th>
<th>PPS scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age/Gender&lt;sup&gt;1,2&lt;/sup&gt; (IV)</td>
<td></td>
<td>Aged 25-35 &amp; Male</td>
<td>Yes: 2 No: 0</td>
<td></td>
</tr>
<tr>
<td>Non-right-handedness&lt;sup&gt;3&lt;/sup&gt; (III)</td>
<td></td>
<td>90&lt;sup&gt;th&lt;/sup&gt; percentile of per capita emissions at local authority level</td>
<td>Yes: 2 No: -5.5</td>
<td></td>
</tr>
<tr>
<td>Pollution&lt;sup&gt;4&lt;/sup&gt; (IV)</td>
<td>UK census data</td>
<td>Over 50% population in urban areas</td>
<td>Yes: 1 No: -2.5</td>
<td></td>
</tr>
<tr>
<td>Urbanicity&lt;sup&gt;5&lt;/sup&gt; (III)</td>
<td>UK census data</td>
<td>Low: Lowest 33&lt;sup&gt;rd&lt;/sup&gt; percentile Medium: Middle 33&lt;sup&gt;rd&lt;/sup&gt; percentile High: Highest 33&lt;sup&gt;rd&lt;/sup&gt; percentile</td>
<td>N/A – combine with ethnicity below</td>
<td></td>
</tr>
<tr>
<td>Ethnic density</td>
<td>UK census data</td>
<td>In low ethnic density area&lt;sup&gt;7&lt;/sup&gt; (II)</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Black Caribbean Ethnicity&lt;sup&gt;6&lt;/sup&gt; (I)</td>
<td></td>
<td>In medium ethnic density area</td>
<td>5.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>In high ethnic density area&lt;sup&gt;7&lt;/sup&gt; (III)</td>
<td>3.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>In low ethnic density area&lt;sup&gt;7&lt;/sup&gt; (II)</td>
<td>3.5</td>
<td></td>
</tr>
<tr>
<td>Other Ethnicity&lt;sup&gt;2&lt;/sup&gt; (IV)</td>
<td></td>
<td>In medium ethnic density area</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>In high ethnic density area&lt;sup&gt;7&lt;/sup&gt; (III)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td></td>
<td></td>
<td>-2</td>
<td></td>
</tr>
<tr>
<td>Not immigrant</td>
<td></td>
<td></td>
<td>-0.5</td>
<td></td>
</tr>
<tr>
<td>1st generation immigrant&lt;sup&gt;3&lt;/sup&gt; (III)</td>
<td></td>
<td>From North Africa&lt;sup&gt;7&lt;/sup&gt; (III)</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>From other regions</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2nd generation immigrant&lt;sup&gt;3&lt;/sup&gt; (II)</td>
<td></td>
<td>From North Africa&lt;sup&gt;7&lt;/sup&gt; (III)</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>From other regions</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>Paternal age &lt;35</td>
<td></td>
<td></td>
<td>-0.5</td>
<td></td>
</tr>
<tr>
<td>Variable</td>
<td>Definition</td>
<td>Value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>------------</td>
<td>-------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paternal age &gt; 35(\star) (IV)</td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paternal age &gt; 45(\star) (IV)</td>
<td>3.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paternal SES(\star) (IV)</td>
<td>Higher executive, proprietor of large businesses, major professional = 0&lt;br&gt;Administrators, lesser professionals or proprietor of medium-sized business = 0.125&lt;br&gt;Smaller business owner, farm owner, manager or minor professional = 0.25&lt;br&gt;Technician, semi-professional or small business owner (business valued at £50,000-70,000) = 0.325&lt;br&gt;Clerical and sales worker, small farm or business owner (business valued at £25,000-50,000) = 0.5&lt;br&gt;Smaller business owner (&lt; £25,000), skilled manual labourer, craftsman or tenant farmer = 0.625&lt;br&gt;Machine operator or semi-skilled worker = 0.75&lt;br&gt;Unskilled worker = 0.875&lt;br&gt;Farm labourer, menial service worker, student, dependent on welfare or no regular occupation = 1&lt;br&gt;Not applicable or unknown = 1</td>
<td>≥0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paternal Severe Mental Illness(\star) (IV)</td>
<td>Family Interview for Genetic Studies [FIGS](\star)</td>
<td>Yes: 5.5&lt;br&gt;No: -2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adult Life Events(\star) (IV)</td>
<td>Life Threatening Events Questionnaire [LTE-Q](\star)</td>
<td>Yes: 5.5&lt;br&gt;No: -2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tobacco use(\star) (IV)</td>
<td>Daily smoker</td>
<td>Yes: 3&lt;br&gt;No: -0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heavy cannabis use(\star) (IV)</td>
<td>Childhood Trauma Questionnaire Short Form [CTQ-SF](\star)</td>
<td>Never true = 1, rarely true = 2, sometimes true = 3, often true = 4, Very often true = 5. None = 5-40; Low = 41-55; Moderate = 56-72; Severe = 73+</td>
<td>Moderate and above</td>
<td>Yes: 4&lt;br&gt;No: -0.5</td>
</tr>
<tr>
<td>Childhood Trauma(\star) (III)</td>
<td>Temporal Experience of Pleasure Scale [TEPS](\star)</td>
<td>Very true for me = 1, Often true for me = 2, Sometimes true for me = 3, Sometimes false for me = 4, Often false for me = 5, Very false for me = 6, items 51 is reverse scored</td>
<td>&gt;35</td>
<td>Yes: 6.5&lt;br&gt;No: 0</td>
</tr>
</tbody>
</table>
6.8. SUPPLEMENTARY REFERENCES


167
PART C: GENERAL DISCUSSION
7. GENERAL DISCUSSION

In this section, I will present a summary of the findings of my PhD, how they advance knowledge in the field, the limitations of these findings and provide recommendations for future research.

7.1. SUMMARY OF FINDINGS

The work presented in this thesis is intended to improve the detection of individuals at risk of developing psychosis (Part A) and the prognostication of their outcomes (Part B). In Part A, I presented the first international replication of an existing transdiagnostic risk calculator to detect psychosis risk, which aimed to demonstrate its generalisability and clinical transportability, as well as an implementation study to examine feasibility of its potential clinical utility of the same transdiagnostic risk calculator. Together, my work has contributed to improving the real-world detection of individuals at risk for psychosis through strengthening the evidence for the use of the transdiagnostic risk calculator in clinical routine. In Part B, I used meta-analysis to synthesise the available evidence for risk and protective factors for psychosis onset in CHR-P individuals, to identify robust prognostic factors for transition to psychosis, alongside the magnitude of effect and strength of evidence for the association of each factor. I also introduced the concept of a Psychosis Polyrisk Score (PPS), implemented the PPS assessment digitally, displayed the range and distribution of scores in a simulated general population dataset as well as piloting the feasibility of its use in individuals referred for a CHR-P assessment and healthy controls. Together, my work has facilitated prognostication of CHR-P individuals by identifying robust prognostic factors associated with transition and developing, digitally implementing and piloting a novel multivariate prognostic model assessing non-genetic risk for psychosis.

The main findings from Part A advance knowledge by demonstrating that the transdiagnostic risk calculator developed by our group\(^1\) can improve detection of individuals at risk of developing psychosis, in that it is transportable to mental healthcare settings outside the UK (Chapter 2) and can be implemented in real-world clinical practice (Chapter 3). In Chapter 2, I performed an external validation of a transdiagnostic risk calculator to detect psychosis risk developed in a secondary mental healthcare setting (SLaM). This was the first replication of the transdiagnostic risk calculator outside of the UK and the largest replication study of a risk prediction model in psychiatry. The external validation dataset was from a commercial insurance database with several key differences
from the derivation dataset, in particular country (US vs UK), care pathway (combination of primary and secondary care vs secondary care alone) and case-mix. In addition to this, individual-level ethnicity data, one of the key predictors in the model, were not available. Instead, a composite ethnicity score was imputed on the level of the Metropolitan State Area (MSA) of each patient. Despite these differences, the transdiagnostic risk calculator performed significantly better than chance (Harrell’s C = 0.68), highlighting the transportability of the model. In Chapter 3, I performed the first feasibility implementation study of a risk prediction model in psychiatry, to demonstrate the potential for systematic detection of psychosis risk at scale. In the in-vitro phase, I identified implementation barriers and overcame them following clinician and service user engagement. Following this, the transdiagnostic risk calculator was integrated into the local EHR. In the in-vivo phase, 3,722 individuals accessing secondary mental healthcare with non-organic, non-psychotic disorders were automatically screened by the transdiagnostic risk calculator. 115 individuals were detected as being at risk, defined as ≥5% risk of developing a psychotic disorder within two years of index diagnosis by the risk calculator, and their responsible clinicians were contacted to recommend referral for a refined psychosis assessment. 77% of clinicians responded to alerts sent by the risk calculator and 85% with outreach. 55% of these responses resulted in a referral for a refined psychosis assessment. Further to this, the incidence of psychosis in those detected by the risk calculator was 12% within six months of individuals being detected, comparable to the level of risk seen in CHR-P individuals (10% at six months).\(^2\) It is also important to note that the incidence of psychosis at six months was 14.7% in those not referred by clinicians, comparable to those referred. These findings add further support for the use of the transdiagnostic psychosis risk calculator in clinical routine as an automatic approach to systematically screen large scale datasets and improve the detection of individuals at risk of developing psychosis.

The findings from Part B have contributed to knowledge by identifying robust prognostic factors for transition to psychosis in CHR-P individuals (Chapter 5) and presenting the development, digital implementation and piloting of the PPS, which has the potential to refine prognostication of outcomes in CHR-P individuals (Chapter 6). Through my work, I have advanced our knowledge into prognostication of outcomes in CHR-P individuals. In Chapter 5, I retrieved 128 original controlled studies investigating CHR-P individuals relating to 26 risk and protective factors for psychosis onset. Random effects meta-analyses were used with Standardised Mean Difference (SMD) and Odds Ratio (OR) selected as outcome summary measures for continuous and dichotomous outcomes respectively. I
combined this approach with an established stratification of evidence to assess the strength of evidence for the association of each factor with the onset of psychotic disorders, categorising each factor into (class I, convincing; class II, highly suggestive; class III, suggestive; class IV, weak; or non-significant). No factors met criteria for convincing (class I) evidence, with two clinical factors (attenuated positive psychotic symptoms and global functioning) meeting criteria for highly suggestive (class II) evidence. Negative psychotic symptoms met criteria for suggestive (class III) evidence. Eight factors (male gender, stress/trauma, living status, employment, right handedness, disorganised/cognitive symptoms, general symptoms, total symptom scores) met criteria for weak (class IV) evidence, with no evidence of association for the remaining 16 factors. This has identified robust factors associated with transition risk in CHR-P individuals, which could potentially be used as predictors in future prognostic models. In Chapter 6, I presented the development of the PPS, the first multivariate prognostic model based on *a priori* epidemiological evidence to improve prognostication of outcomes in CHR-P individuals. Further to this, I implemented the PPS digitally, allowing any individual to complete the self-assessment online. 98.8% of individuals (both individuals referred for a CHR-P assessment and healthy controls) completed the assessment with 96.3% completing it in under 15 minutes and no participants finding the content distressing. The distribution of PPS scores in the general population was also investigated through a simulated dataset, showing a normal distribution with scores ranging from -15 (lowest risk, RR = 0.03) to 39.5 (highest risk, RR = 8912.51). Pilot data showed that PPS scores in individuals referred for a CHR-P assessment ranged from -3 to 18 (mean = 6.2, SD = 7.23), with healthy controls’ scores ranging from -14 to 14.5 (mean = -1.79, SD = 6.78). Individuals referred for a CHR-P assessment had higher PPS scores than healthy controls. This work has advanced knowledge by developing, testing and implementing the first multivariate prediction model that is based on robust epidemiological knowledge and therefore goes beyond symptomatic prognostication in CHR-P individuals.

**7.2. IMPACT OF FINDINGS**

**7.2.1. THE IMPORTANCE OF REPLICATION AND IMPLEMENTATION IN PSYCHIATRY**

Precision psychiatry has enormous potential to improve patient care and outcomes. Although a relatively large number of individualised diagnostic and prognostic models for CHR-P individuals have been developed, very few of the initially positive findings from these have been replicated. I identified only seven prognostic models that have been
replicated in early psychosis, one of which was the transdiagnostic psychosis risk calculator featured in this thesis (Table 2-2). While discovery of new prediction models with strong discrimination remains important, replication has arguably become even more critical than discovery, particularly due to the scarcity of replications.\(^5\) I have shown that the transdiagnostic risk calculator performs well outside the setting in which it was developed. This replication took place in a different country (USA vs UK) with a different service configuration (primary and secondary care combined vs secondary healthcare alone) and different population in terms of case-mix. In combination with other replication efforts,\(^6\) this demonstrates the clinical transportability of the transdiagnostic risk calculator, highlighting that there is potential clinical utility in a number of settings outside of SLaM alone.

To date, while there has been enormous progress in psychiatric research, relatively little of the findings have led to improvements in clinical care. The implementation of the transdiagnostic risk calculator (\textbf{Chapter 3}) was the first attempt to implement an individualised risk prediction model in psychiatric clinical routine\(^7\) and was therefore completed with limited guidance for best practice for conducting an implementation study. Consequently, my PhD advances empirical methodological knowledge in the field of implementation science, with my study being informative for researchers in the future. In particular, the use of \textit{in-vitro} and \textit{in-vivo} phases allows for better addressing conceptual and pragmatic barriers to implementation separately, which could aid future implementation attempts. Moreover, I have advanced knowledge by demonstrating that this tool can be used by clinicians to inform clinical decision making in a real-world setting. As stated before, clinician adherence of the recommendations made by the transdiagnostic calculator was high, suggesting feasibility of implementation and use in clinical routine. Additionally, this work raised novel considerations for the long-term use of the transdiagnostic risk calculator in clinical care. Firstly, I identified several barriers to implementation that will be improved with the conception and utilisation of information governance frameworks and guidance for implementing prediction models in EHRs. Secondly, of patients detected by the calculator there was a similar incidence of psychosis in non-referrals as those referred, suggesting that more work is needed to improve clinician adherence. Outreach resulted in a non-significant increase in adherence, although the sample was small and underpowered. Clinician adherence would likely benefit further from additional efforts in training for future clinicians. Thirdly, I have recognised the limitations of non-automated communication with clinicians, in particular, the consequent increased labour demands. As a result, we have streamlined this process to send automated alerts to clinicians when one of their patients is
detected as over the threshold for psychosis risk. Further work is needed, however, to optimise these alerts for clinicians and better integrate these alerts into local EHRs. The studies presented in Part A of this thesis have demonstrably improved the detection of individuals at risk for psychosis through replication and implementation, providing future avenues for impacting clinical care.

7.2.2. IDENTIFYING ROBUST PROGNOSTIC FACTORS FOR PSYCHOSIS RISK

The work presented in Chapter 5 addressed a clear limitation in our ability to prognosticate outcomes in CHR-P individuals by using meta-analysis to identify a robust set of prognostic factors that modulate transition risk. There have been numerous longitudinal studies investigating the impact of these factors on transition to psychosis in CHR-P individuals. Due to limitations in detection and, more critically, the low transition rate, many of these studies are underpowered when considered alone. As such, by synthesising the available evidence, I was able to identify non-genetic risk and protective risk factors associated with transition to psychosis in CHR-P individuals, accounting for small study effects and similar biases. For those factors that were significant, we were also able to quantify the magnitude and direction of effect. Finally, by implementing validated criteria that have been developed for umbrella reviews, I was able to stratify the evidence of association between risk/protective factors and outcomes.

Through this, I was able to highlight i) the lack of research into non-genetic risk/protective factors that modulate transition, compared to those that modulate psychosis onset in the general population; ii) that the strongest evidence for associated factors does not extend far beyond symptomatology (particularly attenuated positive psychotic symptoms and global functioning) and iii) that these symptomatological factors should be the benchmark of future prognostic models.

7.2.3. DEVELOPMENT, DIGITAL IMPLEMENTATION AND PILOTING OF A SYSTEMATIC, MULTIVARIATE ASSESSMENT OF NON-GENETIC PSYCHOSIS RISK

The work presented in Chapter 6 displayed the development, digital implementation and piloting of the first systematic, multivariate assessment of non-genetic risk and protective factors for psychosis. Previous research has been restricted by a narrow, univariate focus on individual non-genetic risk/protective factors when the effect of these factors on psychosis risk is never in isolation. As such, a multivariate approach is essential to capture the relative
positive and negative impacts that factors have on psychosis risk as well as each other. While other multivariate assessments exist, predictor selection was not unbiased or systematic. In contrast, the Psychosis Polyrisk Score (PPS) systematically used a priori gold standard evidence to inform predictor selection and weighting, avoiding selection bias. I have developed a single assessment to provide a representative estimate of exposure to multiple, evidence-based non-genetic risk and protective factors for psychosis that could inform prognostication of outcomes in CHR-P individuals. I have used pilot data to further support these conceptual strengths, showing high variance in PPS scores, independence from symptom scores and higher scores in individuals referred for a CHR-P assessment than healthy controls.

In addition, I have digitally implemented the PPS and demonstrated feasibility of its administration. As stated above, implementation of prediction models in psychiatry is incredibly important but rarely achieved. Digital assessments have a number of benefits, including increased data completeness, higher patient preference, lower costs and reduced time compared to paper assessments. Digital implementation of the PPS thereby greatly increases its potential for implementation in clinical routine. Further enhancing the implementation potential of the PPS, I showed that it is also feasible to administer, with 99% of participants completing the assessment, with 96% completing in under 15 minutes. Making the PPS assessment available online allows it to be completed by patients in their own homes, and at their own pace with a progress saving feature allowing for completion in multiple sessions. Alongside this, the feasibility of patients completing the assessment at home was further supported with no participants finding the content of the assessment distressing, despite the coverage of several potentially distressing topics (e.g. childhood trauma).

7.3. LIMITATIONS
7.3.1. ETHNICITY IN THE REPLICATION STUDY
The key limitation of the international replication of the transdiagnostic psychosis risk calculator in Chapter 2 was the lack of individual-level ethnicity data. We attempted to mitigate this by imputing individuals’ weighted ethnicity coefficients based on census data. However, individual-level ethnicity is a key element of the predictions provided by the calculator, with the imputed ethnicity scores representing a different concept. Ethnicity as a predictor is unlikely to stem from a genetic predisposition to psychosis, but probably reflects the interaction of social factors that are independently associated with both
psychosis and non-white ethnicity,\textsuperscript{18} such as discrimination,\textsuperscript{19} traumatic life events\textsuperscript{20} and deprivation.\textsuperscript{21} While some of these factors may also be captured in some degree by MSA-level ethnicity, the model is calibrated for individual-level ethnicity, so the associated beta coefficients will likely vary. When aggregate ethnicity coefficients were used in the SLaM development dataset, there was a decrement in performance (0.80 to 0.76), emphasising the benefit of this predictor being defined at the individual level. As such, to determine the true international clinical transportability of the risk calculator, an international external validation study should be performed with a dataset with individual-level ethnicity present.

7.3.2. \textit{FOLLOW-UP TIME}

The replication presented in \textbf{Chapter 2} was limited by a short time window for follow-up (mean follow-up time was 461 days compared to 1581 in SLaM), as ICD-10 was only integrated into United States healthcare on 1\textsuperscript{st} October 2015. ICD-9 diagnoses were considered to enable an extended follow-up time, however, converting the diagnostic clusters used by the transdiagnostic risk calculator proved inexact and inappropriate. As such, a future international replication study should prioritise a longer follow-up time. The fact that the performance of the risk calculator was still greater than chance in spite of these issues does support its clinical transportability and its robustness but the true extent of this remains unknown.

Similarly, in \textbf{Chapter 3}, due to a limited follow-up time (six months following screening), incidence of psychosis was low in our sample. Extending this follow up to a minimum of two years would be desirable for comparing actual incidence of psychotic disorders to that predicted by the model. This would be the basis of any prospective, longitudinal study of clinical utility.

7.3.3. \textit{META-ANALYTICAL LIMITATIONS}

Meta-analyses inherit the issues of the literature they synthesise. The meta-analysis presented in \textbf{Chapter 5} was largely limited by the low number of transitions across the studies that contributed data. Given the rate of transition to psychosis in CHR-P individuals is on average 20\% within two years, for the requisite 1000 cases required for a factor to be considered suggestive evidence (class III) or above,\textsuperscript{22} a total sample of 5000 CHR-P individuals would therefore be required. Only five of the 26 factors (19\%) in this meta-analysis had the required 1000 cases. Similarly, there was limited assessment in the
literature of protective factors for psychosis onset, such as resilience or insight, which could be informative for optimal psychosis prevention methods. Future prospective longitudinal studies in CHR-P individuals will help provide an evidence base that would allow for more factors, including protective factors, being suitably powered to be classified higher in the hierarchy of evidence. In addition to this, there was heterogeneity both in measurement of these factors, with different assessment tools used across studies, and in populations with different detection strategies employed. While this heterogeneity can affect the magnitude and variance of effect size estimates, it similarly should improve durability and generalisability when considering these factors for prognostic use.

7.3.4. SPECIFICITY, UNIVERSALITY AND DURABILITY OF PPS FACTORS
A crucial step towards the development of a PPS is to deconstruct and standardise the specificity of non-genetic risk factors. While polygenic risk scores build on variation in specific single nucleotides in exact positions in the genome, and thus are unambiguously defined at all ages for all individuals and thus across all studies, specificity of most non-genetic risk factors is not completely determined. For example, some of them may be ascertained through a multitude of instruments of questionable comparability. Others may require contextual specifiers (e.g. Black Caribbean ethnicity in England), since their predictive validity may depend on their universality in different cultural scenarios. Furthermore, other factors may be influenced by changes in the contextual environment (e.g. socioeconomic status) and therefore their durability over time periods may be questionable. An additional problem is that many factors are affected by both genetic and non-genetic influences; therefore, the specific components of these risk factors should also be better elucidated. For instance, the effect of parental history of schizophrenia/psychoses is only partly mediated through the individual’s genetic liability.\(^{23}\) The impact of shared environmental influences in the context of the parental history of severe mental illness on liability to schizophrenia amounts to nearly 11%.\(^{24}\) The umbrella review had adopted a pragmatic approach to partially mitigate the above concerns. First, it included several meta-analyses that were conducted worldwide and that were representative of different contextual environments (universality). These studies were also published over two decades, minimising the confounding role of time (durability). Finally, the umbrella review indicated that despite heterogeneous measurements (specificity) and spurious risk factors (encompassing genetic and non-genetic components), the factors analysed were robustly associated with psychosis onset.
7.4. FUTURE RESEARCH

7.4.1. ASSESSING THE PROSPECTIVE REAL-WORLD CLINICAL UTILITY OF THE TRANSDIAGNOSTIC PSYCHOSIS RISK CALCULATOR

My research during this PhD has set the building blocks for a prospective longitudinal study of the clinical utility of the transdiagnostic risk calculator in real-world clinical routine. Previous research has shown that the transdiagnostic risk calculator has shown good-to-moderate prognostic performance in its external validations\textsuperscript{1,6} and theoretical clinical benefit by decision curve analysis.\textsuperscript{1} While my PhD has demonstrated its further clinical transportability (Chapter 2), its implementation in clinical care (Chapter 3) and feasibility of use (Chapter 3), assessing the real-world clinical utility of the transdiagnostic risk calculator in clinical routine is imperative. The work done in the feasibility implementation study, such as workshops with clinicians and service users during the in-vitro phase, will inform the design of the protocol. In addition, the system of detection and alerting has been improved since the study through the migration to the CogStack system.\textsuperscript{8} This permits real-time detection that updates every ten minutes,\textsuperscript{8} compared to a once-per-week basis with the previous system. Additionally, whereas previously a member of the research team had to manually send alerts via email, CogStack can send alerts to clinicians automatically, reducing the logistical demands. This will be further enhanced with planned improvements to the existing EHR interface in SLaM, which should enable direct alerts to the relevant clinician on their personal dashboard. These advances can help further increase clinician adherence to the recommendations made by the transdiagnostic risk calculator and increase referrals of individuals detected to be at risk for psychosis. Prior work has laid the foundation for a prospective longitudinal study assessing the impact of the transdiagnostic risk calculator on total number of FEP cases, early identification of FEP cases and the DUP in those detected to evaluate its real-world clinical utility.

7.4.2. COMPLEMENTARY PRIMARY CARE APPROACH

The international replication of the transdiagnostic risk calculator for psychosis presented in Chapter 2 was the first opportunity to assess the prognostic accuracy of the calculator in a setting that incorporated both primary and secondary mental healthcare. The potential for primary care as a setting for psychosis prevention is high with 60% of adolescents and young adults seen by GPs at least once a year, rising to 90% when individuals have at least one medical condition.\textsuperscript{25} Recognising this, the Royal College of General Practitioners guidelines stress the importance of identifying the early signs and symptoms of psychosis. Currently,
the second most common referral source (21% of referrals) for our local CHR-P service in SLaM (OASIS) is from primary care and a greater number of primary care visits prior to a diagnosis of a psychotic disorder may result in a shorter DUP in FEP patients. Despite this, a qualitative study found that GPs perceive that they may not have the relevant skills to identify individuals who may meet CHR-P criteria, with some not familiar with the CHR-P construct. This emphasises the importance of outreach to primary care services to promote understanding of attenuated psychotic symptomatology. In NHS trusts like SLaM, intensive outreach schemes to inform GPs and promote referrals are feasible, however this may not be the case in all settings, either in the UK or worldwide. Prognostic models using clinical EHR variables could be a useful aid for GPs, with the early stages of developing a prognostic model for psychosis in primary care having already been completed using data from the Clinical Practice Research Datalink (CPRD). These data are high quality with limited missing data and are therefore an ideal source for developing and validating a primary care model. Integrating a primary care-focused approach to complement the transdiagnostic risk calculator in secondary mental healthcare could be extremely useful to aid GPs in identifying appropriate referrals to CHR-P services and extending the benefits of primary indicated prevention. However, efforts must be made to mitigate against potential dilution of pre-test risk enrichment.

7.4.3. INTEGRATING THE PPS AND POLYGENIC RISK SCORES

The PPS mostly includes non-genetic risk factors and therefore can be integrated with the polygenic risk score acquired from the same individuals. Integration of genetic and non-genetic information may benefit from considering gene-by-environment interactions, although there is no consensus on the most effective interaction model. The original GWAS meta-analysis found no epistatic or non-additive effects between the candidate loci and other studies largely did not find interactions between polygenic risk score and environmental risk factors. However, an interaction between polygenic risk score and demographic factors has been demonstrated in individuals of African ancestry (poor predictive accuracy) or with a family history of psychosis (high predictive accuracy).

Since the vast majority of potential interactions across genetic and non-genetic risk have not been tested yet, at present, an additive model that sums all known genetic and non-genetic risks is a pragmatic approximation. An additive approach combined with weighted summation to account for interactions has recently shown promise. A recent review of
gene-by-environment interactions confirmed that polymorphisms of catechol-O-methyltransferase (COMT), brain-derived neurotrophic factor (BDNF) and FK506-binding protein 5 (FKBP5) genes might interact with early life stress and cannabis abuse or dependence, influencing various outcomes of schizophrenia spectrum disorders. In the future, robust gene-by-environment interactions can be incorporated in the same way as other combinations of risk factors were already incorporated in the umbrella review. This would be facilitated by the proposed comprehensive approach that assesses several candidate risk factors and analyses them in a multivariate fashion. While this would be the ideal target for advancing the development of these integrated scores, with the evidence currently available to us, the most pragmatic approach would be an additive model.

7.4.4. **SEQUENTIAL TESTING, CLINICAL STAGING AND DYNAMIC MAPPING OF DEVELOPMENTAL RISK TRAJECTORIES**

Through the work in my PhD, I have advanced research into the feasibility of real-world use of two tools (transdiagnostic psychosis risk calculator and PPS), which can both be integrated into sequential testing and clinical staging frameworks. The transdiagnostic psychosis risk calculator can produce estimates of individualised psychosis risk to aid detection of individuals at risk of developing psychosis. Following this, if an individual is above a psychosis risk threshold, this estimate can be refined using a CHR-P assessment and the PPS in a sequential testing framework. The PPS approach combined with a polygenic risk score would allow researchers to control and replicate CHR-P risk enrichment in a controlled manner, while at the same time facilitating identification of at-risk cases on the basis of a determinate accumulation of risk factors. This would improve the detection of at-risk cases and refine the prediction of psychosis. In addition to genetic information, other information could be used to supplement the estimates made by the PPS and inform risk stratification. As discussed earlier, neuroimaging and neurocognitive models, while informative, can be respectively cost- and labour-intensive, thereby limiting their use. One method of mitigating the high logistical and financial costs is by restricting the use of biological assessments to the subgroup of subjects identified as being at high risk through clinical assessment, in a sequential approach, in line with similar stepped risk enrichment assessments that are used in physical healthcare. The PPS assessment accommodates a clinical staging framework (Figure 7-1) for the development of psychosis, which has recently been reviewed elsewhere. For this aim, it will be important to draw a distinction between individually stable factors (genes, prenatal and early childhood) that can
be carried forward and developmental/state factors that will require reassessment over the
life course (e.g. substance use, adult life events). For instance, the PPS assessment could be
administered during the preclinical phase in non-clinical samples (time 1), such as screening
programmes for schools or non-help-seeking youths in the community for identifying at-
risk groups and facilitate selective preventative focused interventions.\footnote{42} Screening these
populations could be key with a recent population-based cohort study showing that 60% of
individuals meeting criteria for a psychotic disorder at age 24 had a self-reported psychotic
experience at age 12.\footnote{43} Meanwhile 30% of those who met psychosis criteria had not sought
professional help for these experiences,\footnote{43} highlighting a significant clinical opportunity for
enhancing detection. This assessment could be followed by further testing in those
individuals who present with subtle symptoms of psychosis-like CHR-P features when
accessing secondary mental health services (time 2).\footnote{41} Child and adolescent mental health
services and early intervention services may be particularly suited for such an assessment.\footnote{38}

The systematic incorporation of a temporal dimension\footnote{44} in the PPS assessment is consistent
with a developmental framework for mental disorders that has been recommended for
advancing aetiological knowledge,\footnote{45} and better captures the changes in psychosis risk over
time. For example, Dynamic ElecTronic hEalth record deTection (DETECT),\footnote{46} combined
machine learning and EHR data (demographics, diagnoses, prescriptions, procedures,
encounters and admissions, observations and laboratory test results) to forecast FEP
diagnosis one year prior. These predictions are dynamic, updating its predictions at every
new event. Combining a similar approach with the PPS or refining the PPS to dynamically
update its risk estimates upon exposure to new, time-sensitive risk/protective factors would
greatly improve its potential clinical utility.

7.4.5. MODEL UPDATING AND REFINING EXISTING PREDICTORS
Research in the field is continually developing, both in terms of statistical methods and
evidence base, which provides us with the opportunity for improving prognostic models. As
well as developing new models with revised information, it is also important to update
existing models either by adding new predictors or refining existing predictors before re-
validating the model.\footnote{47}

For example, Natural Language Processing (NLP) can be used to extract data from free text
within EHRs instead of pre-set categorical variables.\footnote{48,49} This allows us to retrieve more
Figure 7-1 Putative PPS assessment for the detection of at-risk individuals and the prediction of psychosis.
Risk or protective factors that are diluted during the pre-clinical stages may accumulate as the individual progresses across different stages until they trigger signs or symptoms and functional impairment that are associated with help-seeking behaviour and access to mental health care. In the later stages, specific aggregations of risk and protective factors may be associated with specific clinical outcomes.
detailed information from EHRs (e.g. symptomatology, prescribed medication, suicide attempts) and potentially use these as new predictors for psychosis onset. In terms of the PPS, new research could present evidence for new risk/protective factors and/or increase statistical power to allow for improved status in the hierarchy of evidence (e.g. non-significant to weak [class IV] evidence). Similarly, as the weightings of PPS factors are based on the magnitude of their effect, as new studies are published these may change and this should be reflected in revised weightings, optimising the PPS. Additionally, as the PPS is currently based on evidence for factors that modulate psychosis risk in the general population, it can be refined with factors that modulate risk in other, indicated populations, for example in CHR-P individuals. Evidence syntheses such as that presented in Chapter 5 can help inform a more targeted version of the PPS.

Further to this, existing factors can be refined. In the transdiagnostic psychosis risk calculator, the non-linear modelling of age may more accurately represent the time course of psychosis risk over the lifespan, improving the discrimination (Harrell’s C) from 0.79 to 0.81. Additionally, the PPS factors were selected based on a priori evidence and assumed to be independent. More work is needed to understand the interdependencies between these factors and adjustments should be made to the model accordingly. Moreover, a factor analysis could be performed to assess the possibility of some factors not providing any increment in prognostic accuracy and therefore could be pruned from the model.

7.4.6. TRANSDIAGNOSTIC POTENTIAL FOR THE PREDICTION OF NON-PSYCHOTIC MENTAL DISORDERS

There is emerging evidence that the same risk factors may be associated with multiple types of disorders, beyond psychosis (pleiotropy). For instance, a recent umbrella review has indicated that childhood adversity, exposure to Toxoplasma gondii and a history of head injury are also linked to bipolar disorders. These findings do not eliminate the possibility that even if these risk factors are shared between bipolar disorder and psychosis, the loading and combination of factors that results in either of the two disorders may still be constituted of unique dimensions. While the risk factors themselves may be shared with other psychiatric disorders, the weighting of these factors will be different i.e. the same factor could have a differential impact on risk for different disorders. In addition, in combining transdiagnostic and lifespan-inclusive approaches, this enables a more holistic approach to
mental healthcare that cuts across symptom domains as well as adolescents and adults. This could be enhanced further through dynamic prediction methods to better model the trajectory of mental health throughout the lifespan. What is evident is that there is great potential for transdiagnostic research that focuses on broad and heterogeneous samples of mental disorders.

7.5. CONCLUSIONS

In this thesis, I have presented work that has aided our ability to detect individuals at risk for psychosis and prognosticate their clinical outcomes. Firstly, I have performed the largest replication study in psychiatry, replicating an existing transdiagnostic risk calculator for the systematic and automatic detection of psychosis risk in a US dataset, highlighting its clinical transportability outside the UK for the first time. The impact of this work is the facilitation of the automatic screening for psychosis risk and overcoming the current limitations associated with unstructured recruitment. Secondly, I have implemented this risk calculator in the local EHR and shown that its real-world use is feasible in day-to-day clinical routine. The impact of this work is producing the first implementation attempt of an individualised risk prediction model for CHR-P individuals in clinical practice. Thirdly, I have synthesised evidence for risk and protective factors for transition to psychosis in CHR-P individuals. The impact of this work is an updated evidence-based summary of the most robust prognostic factors for CHR-P individuals, thus guiding their incorporation into multivariate prediction models that go beyond symptomatic assessment alone. Finally, I have developed and piloted an assessment of non-genetic psychosis risk, the Psychosis Polyrisk Score (PPS). The impact of this work is not only the development but, again, implementation of the first multivariate assessment instrument that captures several core non-genetic, largely environmental, prognostic factors in CHR-P samples. Together, my research has progressed the CHR-P field and can be further built on to improve detection, prognostication and, consequently, clinical outcomes and the lives of many young people at risk for psychosis.


22. Fusar-Poli P, Radua J. Ten simple rules for conducting umbrella reviews. *Evid Based Ment Health*


8. APPENDIX: WORK PUBLISHED DURING PHD

1. Using Natural Language Processing on Electronic Health Records to Enhance Detection and Prediction of Psychosis Risk
   Irving, J., Patel, R., Oliver, D., Colling, C., Pritchard, M., Broadbent, M., Baldwin, H., Stahl, D., Stewart, R., Fusar-Poli, P.

2. The Case for Improved Detection of First-Episode Psychosis: Electronic Health Record Cohort Study
   Fusar-Poli, P., Oliver, D., Spada, G., Estrade, A., McGuire, P.

3. Implementing Precision Psychiatry: A Systematic Review and Meta-Regression Analysis of Individualised Prediction Models for Clinical Practice

4. Precuneus and insular hypoactivation during cognitive processing in first-episode psychosis: systematic review and meta-analysis of fMRI studies
   Revista de Psiquiatría y Salud Mental, 2020 doi: 10.1016/j.rpsm.2020.08.001

5. Psychiatric and neuropsychiatric presentations associated with severe coronavirus infections: a systematic review and meta-analysis with comparison to the COVID-19 pandemic
   Lancet Psychiatry, 2020 doi: 10.1016/ S2215-0366(20)30203-0

6. Universal and Selective Interventions to Prevent Poor Mental Health Outcomes in Young People: Systematic Review and Meta-Analysis
   European Neuropsychopharmacology, 2020 doi: 10.1016/j.euroneuro.2020.10.007
7. Real-world implementation of precision psychiatry: Transdiagnostic risk calculator for the automatic detection of individuals at-risk of psychosis

Schizophrenia Research 2020 doi: 10.1016/j.schres.2020.05.007

8. Transdiagnostic Individualised Clinically-based Risk Calculator for the Automatic Detection of Individuals at-Risk and the Prediction of Psychosis: External Replication in 2,430,333 US Patients

Translational Psychiatry, 2020 doi: 10.1038/s41398-020-01032-9

9. Real-world digital implementation of the Psychosis Polyrisk Score (PPS): a pilot feasibility study

Oliver, D., Spada, G., Englund, A., Chesney, E., Radua, J., Reichenberg, A., Uher, R., McGuire, P., Fusar-Poli, P.


Martins, D., Davies, C., De Micheli, A., Oliver, D., Krawczun-Rygmaczewska, A., Fusar-Poli, P., Paloyelis, Y.
Translational Psychiatry, 2020 doi: 10.1038/s41398-020-00890-7

11. Acute oxytocin effects in inferring others’ beliefs and social emotions in people at clinical high risk for psychosis

Translational Psychiatry, 2020 doi: 10.1038/s41398-020-00885-4

12. Adverse effects of cannabidiol: a systematic review and meta-analysis

Chesney, E., Oliver, D., Green, A., Sovi, S., Wilson, J., Englund, A., Freeman, T., McGuire, P.
Neuropsychopharmacology, 2020 doi: 10.1038/s41386-020-0667-2
13. How experimental cannabinoid studies will inform the standardised THC unit
(Commentary on Freeman & Lorenzetti: 'Standard THC Units': a proposal to standardise
dose across all cannabis products and methods of administration)
Chester, L., Chesney, E., Oliver, D., Wilson, J., Englund, A.

14. Prenatal and perinatal risk factors for psychosis: a meta-analysis
Davies, C., Segre, G., Estradé, A., Radua, J., De Michele, A., Provenzani, U., Oliver, D.,
Ramella-Cravaro, V., Besozzi, M., Dazzan, P., Miele, M., Caputo, G., Spallarossa, C.,

15. A real-time psychosis risk detection and alerting system based on electronic health records using CogStack
Wang, T., Oliver, D., Msosa, Y. J., Colling, C., Spada, G., Roguski, L., Folarin, A.,
Stewart, R. J., Roberts, A., Dobson, R. J. B., Fusar-Poli, P.

16. Worldwide implementation of clinical services for the prevention of psychosis: the IEPA early intervention in mental health survey
Kotlicka-Antczak, M., Podgorski, M., Oliver, D., Marik, N., Valmaggia, L., Fusar-Poli, P.

17. Lower speech connectedness linked to incidence of psychosis in people at clinical high risk
Spencer, T., Thompson, B., Oliver, D., Dierderen, K., Demjaha, A., Weinstein, S., Morgan,
S. E., Day, F., Valmaggia, L., Rutigliano, G., De Micheli, A., Mota, N. B., Fusar-Poli, P.,
McGuire, P.

18. Neurochemical Effects of Oxytocin in People at Clinical High Risk for Psychosis
Davies, C., Rutigliano, G., De Micheli, A. I., Stone, J. M., Ramella-Cravaro, V.,
Provenzani, U., Cappucciati, M., Scutt, E., Paloyelis, Y., Oliver, D., Murguia, S., Zelaya,
D. J., McGuire, P., Fusar-Poli, P.
European Neuropsychopharmacology, 2019 doi: 10.1016/j.euroneuro.2019.03.008

19. Neurochemical models of psychosis risk and onset
Oliver, D., Modinos, G., McGuire, P.
Chapter In: Risk Factors for Psychosis: Paradigms, Mechanisms, and Prevention, 2020
20. Psychosis Polyrisk Score (PPS) for the detection of individuals at-risk and the prediction of their outcomes

   Oliver, D., Radua, J., Reichenberg, A., Uher, R., Fusar-Poli, P.

21. What causes the onset of psychosis in individuals at clinical high risk? A meta-analysis of risk and protective factors

   Oliver, D., Reilly, T., Baccaredda Boy, O., Petros, N., Davies, C., Borgwardt, S., McGuire, P., Fusar-Poli, P.

22. Oxytocin Modulates Hippocampal Perfusion in People at Clinical High Risk for Psychosis

   Neuropsychopharmacology, 2019 doi: 10.1038/s41386-018-0311-6

23. Real world implementation of a transdiagnostic risk calculator for the automatic detection of individuals at risk of psychosis in clinical routine: Study protocol

   Fusar-Poli, P., Oliver, D., Spada, G., Patel, R., Stewart, R., Dobson, R., McGuire, P.

24. Unmet needs in patients with brief psychotic disorders: Too ill for clinical high risk services and not ill enough for first episode services

   Minichino, A., Rutigliano, G., Merlino, S., Davies, C., Oliver, D., De Micheli, A. I., Patel, R., McGuire, P., Fusar-Poli, P.

25. Are you on my wavelength? Interpersonal coordination in naturalistic conversations

   Journal of Nonverbal Behavior, 2019 doi: 10.31234/osf.io/5r4mj

26. Long term outcomes of acute and transient psychotic disorders: The missed opportunity of preventive interventions

   Rutigliano, G., Merlino, S., Minichino, A., Patel, R., Davies, C., Oliver, D., De Micheli, A., McGuire, P., Fusar-Poli, P.
   European Psychiatry, 2018 doi: 10.1016/j.eurpsy.2018.05.004
27. 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
Fusar-Poli, P., Werbeloff, N., Rutigliano, G., Oliver, D., Davies, C., Stahl, D., McGuire, P., Osborn, D

28. Can we reduce the duration of untreated psychosis? A systematic review and meta-analysis of controlled interventional studies
Oliver, D., Davies, C., Crossland, G., Lim, S., Gifford, G. W. G., McGuire, P., Fusar-Poli, P.

29. Meta-analytical prognostic accuracy of the Comprehensive Assessment for At Risk Mental States (CAARMS): the need for refined prediction
Oliver, D., Kotlicka-Antczak, M., Minichino, A., Spada, G., McGuire, P., Fusar-Poli, P.

World Psychiatry, 2018 doi: 10.1002/wps.20490

Bhavsar, V., McGuire, P., Fusar-Poli, P., Oliver, D., MacCabe, J.

32. Why transition risk to psychosis is not declining at the OASIS ultra high risk service: the hidden role of stable pretest risk enrichment.
Fusar-Poli, P., Palombini, E., Davies, C., Oliver, D., Bonoldi, I., Ramella-Cravarro, V., McGuire, P.

33. Diagnostic and prognostic significance of DSM-5 attenuated psychosis syndrome in services for individuals at ultra high risk for psychosis.
Fusar-Poli, P., De Micheli, A., Cappucciati, M., Rutigliano, G., Davies, C., Ramella-Cravaro, V., Oliver, D., Bonoldi, I., Rocchetti, M., Gavaghan, L., Patel, R., McGuire, P.
34. The role of parietal cortex in overimitation: a study with fNIRS.

**Oliver, D.**, Tachtsidis, I., Hamilton, A. F. de C.


35. Deconstructing vulnerability for psychosis: systematic review of genetic and environmental risk factors for psychosis in subjects at ultra high risk
