European Psychiatry 49 (2018) 62–68

Contents lists available at ScienceDirect

European Psychiatry

journal homepage: http://www.europsy-journal.com

Review

Meta-analytical prognostic accuracy of the Comprehensive Assessment of at Risk Mental States (CAARMS): The need for refined prediction

D. Oliver [a,*], M. Kotlicka-Antczak [b], A. Minichino [c], G. Spada [a], P. McGuire [d,e,f], P. Fusar-Poli [a,e,f]

[a] Early Psychosis: Interventions & Clinical-detection (EPIC) lab, Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London SE5 8AF, United Kingdom
[b] Medical University of Lodz, Department of Affective and Psychotic Disorders, Lodz, Poland
[c] Department of Neurology and Psychiatry, Sapienza University of Rome, Rome, Italy
[d] Department of Psychosis Studies, IoPPN, King's College London, London SE5 8AF, United Kingdom
[e] OASIS Service, South London and the Maudsley NHS National Health Service Foundation Trust, London, United Kingdom
[f] National Institute for Health Research (NIHR) Biomedical Research Centre for Mental Health, IoPPN, King's College London, SE5 8AF, United Kingdom

ARTICLE INFO

Article history:
Received 1st September 2017
Received in revised form 4 October 2017
Accepted 4 October 2017
Available online 3 February 2018

Keywords:
Psychosis
CAARMS
Prevention
Prognostic accuracy
Clinical utility

ABSTRACT

Primary indicated prevention is reliant on accurate tools to predict the onset of psychosis. The gold standard assessment for detecting individuals at clinical high risk (CHR-P) for psychosis in the UK and many other countries is the Comprehensive Assessment for At Risk Mental States (CAARMS). While the prognostic accuracy of CHR-P instruments has been assessed in general, this is the first study to specifically analyse that of the CAARMS. As such, the CAARMS was used as the index test, with the reference index being psychosis onset within 2 years. Six independent studies were analysed using MIDAS (STATA 14), with a total of 1876 help-seeking subjects referred to high risk services (CHR-P+: $n = 892$; CHR-P–: $n = 984$). Area under the curve (AUC), summary receiver operating characteristic curves (SROC), quality assessment, likelihood ratios, and probability modified plots were computed, along with sensitivity analyses and meta-regressions. The current meta-analysis confirmed that the 2-year prognostic accuracy of the CAARMS is only acceptable (AUC = 0.79 95% CI: 0.75–0.83) and not outstanding as previously reported. In particular, specificity was poor. Sensitivity of the CAARMS is inferior compared to the SIPS, while specificity is comparably low. However, due to the difficulties in performing these types of studies, power in this meta-analysis was low. These results indicate that refining and improving the prognostic accuracy of the CAARMS should be the mainstream area of research for the next era. Avenues of prediction improvement are critically discussed and presented to better benefit patients and improve outcomes of first episode psychosis.

1. Introduction

Psychosis is a severe psychiatric condition and there is limited evidence that treatments are successful in improving patients' functioning once the disorder is established [1]. Intervening in the earlier phases is therefore the only viable possibility to substantially alter the course of the disorder [2,3]. Within early intervention, a key focus for improving the outcome has been primary indicated prevention [2,4,5]. Primary indicated prevention allows for early intervention for those at clinical high risk of developing psychosis (CHR-P), with greater scope for improving outcomes. To do this effectively, the first necessary step is to reach an accurate, robust prognostic identification of individuals meeting CHR-P criteria who will subsequently develop psychosis or not. Ideally, all subjects who will actually develop psychosis should be classified as "at risk" (CHR-P+) while those not developing an established psychosis should be classified as "not at risk" (CHR-P–). These key concepts involved in prognostic reasoning in the CHR-P have been detailed and presented in a recent paper by our group [6].

Prognostic prediction is used in many branches of medicine to identify individuals who may develop a particular disease [7]. For example, fasting glucose, oral glucose tolerance test and glycated

* Corresponding author.
E-mail address: dominic.a.oliver@kcl.ac.uk (D. Oliver).

https://doi.org/10.1016/j.eurpsy.2017.10.001

haemoglobin are used to detect individuals at high risk for developing diabetes (pre-diabetes or intermediate hyperglycaemia) [8] and systolic blood pressure and ratio of total serum cholesterol to high density lipoprotein cholesterol levels are used to detect individuals at high risk for developing cardiovascular disease [9]. However, unlike these other fields, there are no biological tests to assess the risk of developing mental disorders [10], which is instead reliant on semi-structured CHR-P psychometric interviews, such as the CAARMS (Comprehensive Assessment for At Risk Mental States) [11]. Recently, the CAARMS has become the mainstream tool to detect CHR-P individuals in the UK, recommended by international bodies, such as NICE [12]. Therefore, understanding its exact psychometric properties is of paramount clinical relevance. The CAARMS shows excellent inter-rater reliability when performed by trained raters (0.85) [13]. However, its prognostic accuracy is uncertain. A recent meta-analysis by our lab [14] investigated the prognostic accuracy of CHR-P instruments, showing generally excellent prognostic performance of these instruments. However, CHR-P tools were grouped together including the CAARMS [11], the SIPS (Structured Interview for Prodromal Syndromes) [15] and the SPI-A (Schizophrenia Proneness Instrument-Adult Version) [16]. This was due to the fact that there were not enough studies contributing data to assess the meta-analytical prognostic accuracy of the CAARMS specifically. Given the marked differences between the CAARMS and other CHR-P instruments [17], in particular with respect to the functional deterioration criterion [18], it is possible that the previously reported meta-analytical prognostic accuracy is not completely accurate. In addition, the previous meta-analysis combined multiple follow-up time points, and even though meta-regressions of this variable found no significant effect, validity of the prognostic accuracy results would be improved by using a more defined and consistent follow-up time [14].

The current study tackles these caveats and advances knowledge in the psychometric properties of the CAARMS. We capitalize on recently published CAARMS studies reporting useful and innovative meta-analytical data to conduct a meta-analytical prognostic accuracy analysis of the CAARMS at two-year follow-up. This is the period of time during which most transitions to psychosis occur [19]. The results will hopefully support the refinement of psychosis prediction and therefore facilitate indicated primary prevention in CHR-P individuals.

2. Methods

2.1. Search strategy

Two investigators (DO, PFP) conducted a two-step literature search. At a first step, the Web of Knowledge database was searched, incorporating both the Web of Science and Medline. The search was extended until August 2017, only including abstracts in English. The electronic research adopted several combinations of the following keywords: "at risk mental state", "psychosis risk", "prodrome", "prodromal psychosis", "ultra-high risk", "high risk", "help-seeking", "diagnostic accuracy", "sensitivity", "specificity", "psychosis prediction", "psychosis onset". The second step involved the use of Scopus to investigate citations of previous systematic reviews on transition outcomes in CHR-P subjects and a manual search of the reference lists of the retrieved articles.

Articles identified through these two steps were then screened for the selection criteria on the basis of abstract reading. The articles surviving this selection were assessed for eligibility on the basis of full text reading. To achieve a high standard of reporting, we adopted the Meta-analysis Of Observational Studies in Epidemiology (MOOSE) checklist [20].

2.2. Selection criteria

Studies were eligible for inclusion if:

- they were reported in original articles, written in English;
- they had used the CAARMS (index test) in the same pool of referrals;
- they had followed up both CHR-P+ and CHR-P– subjects for psychosis onset (reference index) using established international diagnostic manuals (ICD or DSM);
- they had reported sufficient prognostic accuracy data at 2-year follow-up.

With respect to this last point, when data were not directly presented, they were indirectly extracted from associated data. Additionally, we contacted all corresponding authors to request additional data when needed.

We excluded:

- abstracts, reviews, articles in a language other than English;
- studies in which interviews were not conducted in the same pool of referrals or that used an external CHR-P group of healthy controls;
- studies with overlapping datasets.

In case of multiple publications deriving from the same study population, we selected the article reporting the largest and most recent data set. The literature search was summarized according to PRISMA guidelines [21].

2.3. Recorded variables

Data extraction was independently performed by two investigators (DO, PFP). Data included author, year of publication, characteristics of subject samples (baseline sample sizes, mean age and age range, proportion of females), diagnostic criteria used at follow-ups to assess the psychotic outcome, prognostic accuracy data (number of true and false positives, true and false negatives or associated data) and quality assessment conducted with the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) checklist [22].

2.4. Statistical analysis

The statistical analysis followed the Cochrane Guidelines for Systematic Reviews of Diagnostic Test Accuracy, Version 1.0 [23] and the Methods Guide for Authors of Systematic Reviews of Medical Tests by the Agency for Healthcare Research and Quality (chapter 8) [24]. Evaluating test accuracy requires knowledge of two quantities: the test's sensitivity (Se) and specificity (Sp). Meta-analysis methods for diagnostic test accuracy thus have to deal with two summary statistics simultaneously rather than one [23]. Methods for undertaking analyses, which account for both Se and Sp, the relationship between them, and the heterogeneity in test accuracy, require fitting advanced hierarchical random effects models [23].

For each study, we constructed a two-by-two table, which included true positive, false positive, true negative, and false negative values. The baseline sample size was conservatively used as the base reference.

Data were then analysed with MIDAS (Meta-analytical Integration of Diagnostic Accuracy Studies) [25], a comprehensive program of statistical and graphical routines for undertaking meta-analysis of diagnostic/prognostic test performance in STATA 14 software [26]. The index tests of CHR-P status (CHR-P+ or CHR-P–) and reference tests of transition to psychosis according to international