National Institute for Health Research

Improving prediction of psychosis in ARMS using a clinically useful prognostic tool (IPPACT)

Protocol

1. Background and Rationale

Psychotic disorders such as schizophrenia have a lifetime prevalence of over 1% and are among the world's leading causes of disability [1]. They have a major detrimental effect on individuals' personal, social and educational functioning [2], and a high economic impact, costing the UK economy £11.8bn per year in direct healthcare costs, lost productivity and costs to carers [3].

It is now possible to identify individuals at high and imminent risk of developing a first episode of psychosis through specific criteria known as "At Risk Mental State" (ARMS) or "Ultra High Risk" (UHR) criteria [4, 5]. People meeting the ARMS criteria have low grade "psychotic –like symptoms" that cause distress. About 15 - 22% of ARMS individuals develop a full psychotic disorder such as schizophrenia within 12 months [5, 6]. Identification of ARMS individuals therefore presents the opportunity for early intervention to prevent of onset of full psychotic disorder.

It is estimated that over 12,000 people in England per year present to clinical services with distressing psychotic-like ARMS symptoms [7]. In recognition of the scale of the problem and the potential benefits of early detection and intervention in the ARMS group, assessment and management of ARMS individuals is included in NICE guidelines for children, young people and adults [8, 9], and since April 2016 NHS England has required clinical services to treat people with the ARMS [10, 11]. A specialised instrument, the Comprehensive Assessment of At Risk Mental States (CAARMS) [12], developed by Principal Investigator Yung, is freely available to all clinical services in the UK to detect and monitor the ARMS.

However, currently it is not possible to predict which ARMS individuals will develop a psychotic disorder and which will not. This means that some ARMS individuals are having unnecessary treatment, and may be told that they are at high risk of psychosis when they are not, creating possible fear and stigma. It also means that the NHS may be using a costly treatment, (such as Cognitive Behaviour Therapy; CBT), in people who may not need it. For this reason, NIHR has commissioned research to improve the ability to predict people most at risk of psychosis, over and above the existing ARMS criteria, through the development of a refined prognostic tool. Such a tool will enable prediction of outcome for individual service users, enhance informed decision making, and to stratify the ARMS group into different risk profiles for the purposes of resource allocation and clinical trials. The prognostic tool has the potential to directly benefit service users through more focused services and treatments and could lead to the development of more cost-effective management plans in the NHS.

This project will involve a three-phase design:

Phase 1 will consist of an individual participant data evidence synthesis to identify factors associated with transition to psychosis. At the end of the evidence synthesis, decisions will be made about the best factors that will be included in the refined prognostic tool. Successful implementation of a prognostic model requires support by leading professionals in the field [13]. Measures also need to

be acceptable to service users. Thus, at the end of the evidence synthesis we will run focus groups to gauge opinion about the use of the individual measures in the prognostic tool. This will help ensure that the chosen measures are acceptable and feasible for routine use in the NHS. A major strength of the development of our tool is that it will only include prognostic factors which are measurable in an easy, quick and reliable way during routine care (i.e. without the need for expensive brain scanning or laboratory tests), thus optimising it as an efficient and cost-effective "bedside" tool for use in the NHS.

Based on the above findings, in **Phase 2** we will develop and internally validate a prognostic tool to better predict the later development of psychosis.

Finally, **Phase 3** will be a cohort study of ARMS individuals with 12 month follow up to assess the external validity of the new prognostic tool. In addition, a valid prognostic tool could lead to more cost-effective pathways and management plans. To this end the project also includes health economic modelling.

2. Why this Research is Needed Now

Since April 2016 NHS England has required all Early Intervention in Psychosis Services to assess and manage individuals with an ARMS. This has resulted in a huge expansion in number of assessments and individuals accessing care at these services. While NHS England allocated additional funding for services, the extra demand has far exceeded that which was expected, resulting in larger caseloads and a strain on resources. Research to better stratify ARMS patients according to levels of risk of psychosis could reduce this strain, through more efficient use of NHS resources. Thus, those at highest risk could be offered CBT, while lower risk patients could be offered less costly and less intensive regular mental state monitoring. In this way, a prognostic tool that can be used in routine practice in Early Intervention Services will lead to the development of more cost-effective pathways and management plans.

While much ARMS research has attempted to identify the predictors of development of psychosis, including the development of prognostic models, no prognostic tool has yet been implemented in the NHS. Two recent systematic reviews summarised these ARMS studies [14, 15]. Both reviews found that only 7 studies of prognostic models in the ARMS group had any validation procedure. All 7 of these used internal validation only; that is, used the same sample that the model was derived from to test its predictive ability. None had any external validation. Furthermore, all 7 studies used pre-defined prognostic factors and so may have missed assessing other factors likely to be predictive of psychosis. Other problems included small sample sizes, poor reporting of missing data and low number of individuals developing psychosis relative to the number of predictor variables (i.e. low event to predictor variable (EPV) ratio, which might increase the risk of overfitting and overestimating the performance of the model [16, 17]). The development and external validation of one prognostic model has been published [18, 19] (not included in the above reviews).

However, this new model also suffers from methodological weaknesses, (e.g. external validation without assessment of calibration, i.e. how closely the probability of the event predicted by the model agrees with the observed probability).

The current project will address the above weaknesses by exploiting the strengths of state-of the-art evidence syntheses methods, in particular individual patient data (IPD) meta-analysis [20], to synthesise all the current literature on risk factors for psychosis in the ARMS group. Following this, we will develop a prognostic tool [13] to improve prediction of psychosis in the ARMS group and externally validate this is a new cohort of ARMS individuals.

3. Aims and Objectives

Aims

- To synthesise the evidence about predictive and protective factors for development of psychosis in the At Risk Mental State (ARMS) group.
- To develop a prognostic tool that predicts ARMS individuals at highest and lowest risk of psychosis that is feasible and acceptable to mental health staff and service users.
- To assess the external validity of this tool.
- To assess the health economic impact of implementing this tool in clinical practice.

Objectives

1. Evidence synthesis: To achieve this aim we will conduct a systematic review and IPD meta-analysis of prospective studies of individuals meeting ARMS criteria, including review of existing prognostic models. We will focus on risk factors for development of psychosis within the 12 months after an individual has been identified as meeting ARMS criteria. This first 12 months is the highest risk period for psychosis onset [21, 22], and the time when individuals are most distressed [23] and most likely to engage with services.

2.Development of a prognostic tool suitable for use in the NHS:

2a. Factors and models that show consistent evidence of prognostic value in the evidence synthesis phase will be discussed at focus groups with early intervention staff and service users. This is to ensure that the measures in the tool are acceptable to service users and feasible for NHS staff to conduct or arrange.

2b. Informed by the prognostic models and prognostic factors identified above, and feedback from service users and staff, state-of-the-art methodology will be used, including assessing apparent (calibration and discrimination of the model) and internal validation i.e. bootstrapped resampled estimates of calibration (how closely the probability of the event predicted by the model agrees with the observed probability [24]) and discrimination (the discriminative ability of the model as assessed by the c index [25]).

3. External validation of the model: predictive performance and acceptability

3a. Acceptability: At baseline assessment with the prognostic tool, acceptability will be examined by assessing whether any measures included in the tool were not completed and we will record reasons for this.

3b. The predictive performance of the tool: After 12 month follow up of the ARMS cohort we will evaluate the discrimination and the calibration of the tool and report our results in line with the best practice Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) guidelines [17].

3c. Through seeking consent at baseline to access health records for 2,3 and 5 year follow up, the long term predictive validity of the prognostic tool can be assessed in the whole sample in future studies.

4. To achieve Aim 4 we will conduct an exploratory cost-effectiveness analysis. We will explore the potential cost-effectiveness of the new prognostic tool compared to usual care (no prognostic tool) from the perspectives of the NHS and Social Care (costs) and service users (health benefits) over 1, 2 and 5 year time-horizons. We will use a cost-effectiveness acceptability approach to estimate the probability the new prognostic tool is cost-effective. A decision analytic model will capture transition between at risk, treated health states and the longer-term recovery/relapse cycle of psychosis. Probabilistic sensitivity analysis will account for data uncertainty. Structural uncertainty will be assessed in sensitivity, threshold and scenario analyses.

4. Research Plan

This project will (i) synthesise evidence, (ii) develop a prognostic tool for prediction of psychosis in ARMS individuals and (iii) test the tool's external validity and acceptability, and (iv) model the health economic impact of implementing this tool in clinical practice

4.1 Phase 1: Evidence synthesis

Our evidence synthesis will comprise three components: (1) an IPD meta-analysis of prospective studies considering individuals meeting ARMS criteria to identify factors that show consistent evidence of prognostic value; (2) a narrative systematic review and quality appraisal of existing prognostic models to inform subsequent model updating and aggregation activities at Phase 2; and (3) a focused review for economic evaluations of health care for people at risk of psychosis and for first episode of psychosis, to inform the health economic model evaluated at Phase 3.

Our synthesis will follow state-of-the-art guidelines for IPD meta-analytic syntheses in prognostic research [26, 27], and our outputs will comply as a minimum with the PRISMA statement for the reporting of IPD meta- analysis [28] and the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) reporting guidelines [29, 30]. Our review protocol will be published on the HTA website and on a specific register of systematic review protocols (PROSPERO) to allow for independent scrutiny of all outputs against our planned review and analytic strategy.

4.1.1 IPD meta-analysis of prognostic factors

Study selection criteria: The following selection criteria will be used:

Population: Individuals meeting ARMS criteria, defined as 1) attenuated psychotic symptoms, 2) full-blown intermittent psychotic symptoms and 3) genetic/familial risk for schizophrenia in conjunction with a significant decrease in functioning and operationalised using suitable measures such as the Comprehensive Assessment of At-Risk Mental States (CAARMS) [31] or the Structured Interview for Prodromal Syndromes (SIPS) [32].

Unlike previous aggregate data meta-analyses of this research area [e.g. 33], we will not consider prospective studies that only assessed for the presence of "basic symptoms" of psychosis i.e. subtle subjective cognitive and perceptual disturbances shown to predict onset of first episode psychosis [34, 35]. Basic symptoms are in fact distinct from the ARMS criteria and are not used in the NHS.

Study design: Any prospective study (i.e. cohort studies as well as randomised controlled trials of preventive interventions) with participants meeting the above-mentioned ARMS criteria will be eligible. Studies must include at least 12-month follow-up longitudinal assessments and collected data on psychosis transition in ARMS individuals.

Outcomes: Our primary outcome will be psychosis transition at 12-month follow-up assessments defined using standard diagnostic classification systems (DSM-III, DSM-IV, DSM-5, ICD-10, ICD-11) or commonly used ARMS assessment schedules (e.g. SIPS or CAARMS). Secondary outcomes will include psychosis transition at subsequent available follow-up assessments (18-month, 24-month etc) and time to transition (time to event).

Identifying studies - information sources: Database searches will be conducted on PsychInfo, PubMed, EMBASE and CINAHL. No restriction will be placed on language of publication but the searches will be restricted to 1994 onwards, the initial year of the first prospective study using ARMS criteria [36, 37]. Database searches will be supplemented by 1) inspection of studies included in previous systematic reviews and meta-analyses of psychosis transition studies; 2) inspection of reference lists of psychosis transition studies identified through the database searches; 3) inspection of citations of psychosis transition studies identified through the database searches.

Identifying studies - electronic searches: Titles, abstracts and keywords will be searched in the publication databases using search terms adapted from previous systematic reviews and meta-analyses of this research area [33, 38-41]:

['psychosis'] AND ["clinically at high risk' OR 'clinically at risk' OR 'clinical high risk' OR 'ultra-high risk' OR prodrom* OR 'at risk mental state' OR 'risk of psychosis' OR '' OR 'ARMS' OR 'prodromal psychosis']

Study selection process: Titles and abstracts will be screened for relevance, and full-text reports will be assessed for eligibility against the abovementioned study selection criteria. When required, additional information to ascertain eligibility will be requested from study authors. Discrepancies in selection decisions will be discussed, and arbitration by another member of the research team sought to resolve such discrepancies.

IPD collection process: IPD will be collected from principal investigators and data custodians of past and on-going ARMS prospective studies identified using the search strategy described above. We have already set up a collaborative network of active researchers in this area to support the retrieval of relevant IPD and have approached the principal investigators/data custodians of several of the most relevant ARMS prospective studies internationally. We will continue to expand the collaborative network by sending invitation emails to all relevant researchers and invite them to share available IPD for this IPD meta-analysis.

Prognostic factors: This evidence synthesis will focus on prognostic factors that can be feasibly assessed in routine clinical practice in the NHS. Our data requests will consider a range of variables that have been examined in previous primary research aimed at identifying predictors of psychosis transition (for a review see [41]), including:

1) Socio-demographic characteristics (e.g. age, gender, ethnicity, SES, education);
2) Clinical characteristics, including baseline measures of specific ARMS risk criteria, general/social/occupational functioning, general psychopathology, positive symptoms, negative symptoms, disorganisation, anxiety, depression, duration of illness, medications, history of trauma, family history of psychosis/schizophrenia and substance use/abuse;
3) Baseline neurocognitive variables (e.g. working memory, verbal memory, processing speed, sustained attention)

All data received will be systematically recoded to ensure consistent predictor definitions and maximize the use of common scales of measurement across studies. We will liaise with principal investigators and statisticians of the primary studies to resolve any data issues and prepare the dataset for IPD analysis.

Risk of bias assessment in individual studies: As per recent guidelines on the conduct and execution of IPD meta-analysis of prognostic studies [26], we will assess risk of bias using the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (CHARMS) checklist [42] and, if available prior to the onset of the project, the Prediction Study Risk of Bias Assessment Tool (PROBAST) [43].

IPD data synthesis of prognostic factors associated with psychosis transition: A series of random effects IPD meta-analyses will be used to examine the contribution of prognostic factors for psychosis transition at 12-month follow-up assessments. Initially, all IPD sets will be re-analysed separately and the original authors asked to confirm accuracy of the individual study results, with any discrepancies resolved. Then, for each prognostic factor, we will perform either a one-step or a two-step IPD meta-analysis to obtain the pooled effect. The one-step approach analyses the IPD from all studies simultaneously, whilst accounting for the clustering of patients within studies. In contrast, the two-step approach first estimates prognostic effects from the IPD in each study separately, and then pools them using a conventional meta-analysis. We will employ a random effects approach, and examine statistical heterogeneity using the 'tau-squared' statistics (which provides an estimate of between-study variance) and I^2 (which provides the proportion of total variance that is due to between-study variance). A similar analytic approach will be followed for our

analyses of secondary outcomes (psychosis transition at subsequent follow-up assessment points, e.g. 18-month, 24-month etc., and time to transition). For any time-to-event outcome, we will aim to fit a Cox regression model and synthesise the estimated hazard ratios obtained. For binomial outcomes, we will fit a logistic regression model to synthetize Odds Ratios. If sufficient data will be available, subgroup analyses will be conducted to compare studies employing different research designs (trials of preventive interventions vs prospective studies) and studies conducted in different countries.

4.1.2 Systematic review of prognostic models for transition to psychosis:

Based on the literature searches conducted for the IPD meta-analysis, we will also identify and summarise studies examining prognostic models (and clinical decision rules based on such models) that utilise multiple prognostic factors in combination to predict transition to psychosis in ARMS individuals. The primary outcome for the review will be the predictive accuracy of prognostic models in relation to transition to psychosis.

Data extraction specifically related to prognostic models will include the final model (its specification, included factors, values of regression coefficients and standard errors), how it was developed, and any internal or external validation performance statistics for discrimination (such as the c-statistics or area under the curve) or for calibration (such as the expected/observed events ratio), together with their confidence intervals. This will be informed by the CHARMS checklist [42].

The quality (risk of bias) of any studies developing or evaluating a prognostic model will be assessed using the criteria described by Altman et al [44] and by specific tools including the Quality in Prognostic Studies tool (QUIPS) and PROBAST (Prediction Study Risk of Bias Assessment Tool) if it is published within time [43]. Elements to be considered include:

- Study design (e.g. whether it was a prospective design, and whether prognostic factor and outcome measures were reliable);
- Sample size (e.g. whether there was a pre-specified sample size consideration accounting for numbers of events and multiple comparison in selection of factors, and how much data was available for external validation).
- Missing data (such as adequate reporting on completeness of data, and whether imputation was used)
- Statistical analysis (e.g. handling of continuous variables, selection of possible factors, and use of bootstrapping or shrinkage).
- Internal and external model validation (e.g. whether model validations are reported and how these were carried out).

Any studies reporting the development of a prognostic model will be summarised narratively, in particular what variables (prognostic factors) were included in the final model; how the included variables were coded; what the specification of the model was and how it produces an individual outcome probability or risk score; the reported predictive accuracy of the model; and whether the model was validated internally and/or externally, and if so how.

If multiple studies are found that externally validate the same prognostic model, then calibration statistics (such as expected/observed events) and discriminatory statistics (such as the c-statistic or area under the curve) will be synthesised using the random-effects meta-analysis of DerSimonian and Laird [45, 46] to summarise the model's average performance across different settings and its predicted performance in a future setting [47, 48]. If we identify multiple prognostic models that have been adequately externally validated, we will compare their performance narratively, taking into account the different case mix, how this relates to our own setting, and also the quality of studies according to QUIPS [49] and PROBAST [43]. Candidate models will be taken forward to Phase 2.

4.1.3 Review of economic evaluations for health economic modelling

In addition to the reviews above, we will run a systematic, focussed search for economic evaluations of health care for people at risk of psychosis and for first episode of psychosis, to inform the economic model considered in subsequent components of the project.

Search strategy: The search strategy used for our IPD meta-analysis will be modified to specifically identify relevant health economics literature. More specifically, titles, abstracts and keywords will be searched in the publication databases using the search filter developed to identify economic evaluations for NHS EED with the clinical terms identified for the reviews above (an example search strategy for EMBASE is shown in Appendix 1) [50]. The databases searched will include EMBASE, CINAHL, EconLit, All EBM Review, Medline and PsychInfo.

Study selection criteria

Population: We will use the same population criteria as the IPD meta-analysis.

Study design: Full and partial economic evaluations (including cost of illness/economic burden of disease studies) published in the previous 20 years at the time of the search will be included. These can be prospective, retrospective and economic model study designs as well as systematic reviews of economic evidence. Studies must include formal cost and/or health benefit measures and analyses. Editorials and other descriptive studies (e.g. historical discussion and case reports) will be excluded.

Intervention and comparators: No specific interventions or comparators will be specified since the review will be used to identify data about the costs and health benefits of care following transition to psychosis.

Outcomes: Studies must include one or more of the following measures: health and social care service use, direct or indirect costs, preference based utility or health status measures or quality adjusted life years (QALYs).

Only studies published in English will be included.

Review strategy and strategy for reviewing literature: Two researchers will screen titles and abstracts for relevance and subsequently assess eligibility by examining the full-text reports against the above study selection criteria. When required, additional information to ascertain eligibility will be requested from study authors. Discrepancies in selection decisions will be discussed, and arbitration by another member of the research team (Varese or Yung) sought to resolve such discrepancies. Data will be extracted by two reviewers using an adapted version of the NHS EED critical appraisal checklist [50, 51] and results cross-checked. Any disagreement will be resolved by discussion with the assistance of a third reviewer (Varese or Yung). Costs will be converted to a common currency (UK pound sterling) and price year using purchasing power parity indices. A high level of heterogeneity in objectives, design and measures is anticipated. Accordingly, narrative methods will be used to summarize and tabulate the data by objectives, methods and results for each measure of service use, cost and health benefit (Appendix 2 lists the key information to be extracted). A permutation index will be used as a framework to present the outcomes of the evaluations from the review, by type of economic analysis. If there are groups of studies using similar designs and measures we will explore the feasibility of quantitative synthesis (e.g. weighted means; meta regression) to further summarise the data and facilitate estimation of parameters for the economic model.

4.1.4 Acceptability assessment in preparation of Phase 2

Focus groups with service users and NHS clinicians: At the end of the evidence synthesis, decisions will be made about the best models and factors to take forward to Phase 2. To inform Phase 2, we will run focus groups with early intervention staff (managers and clinicians) and separate service user focus groups. At these groups, we will present the Phase 1 results and gauge opinion about the use of the individual measures to be included in the refined prognostic tool. As a key factor for

successful implementation of a prognostic model is whether a model is supported by leading professionals in the field [13] this is an important step. If any measures are thought to be too difficult to assess/arrange in routine practice or unacceptable to service users then we not include these factors in the prognostic tool. This will help ensure that the chosen measures are acceptable and feasible for routine use in the NHS.

Two focus groups with NHS staff will be run, with 5-8 participants per focus group. To ensure representations of the views of staff working in both urban and rural areas and therefore the usability of the prognostic tool in services located in geographically and demographically diverse areas of the country, these focus groups will take place in the Midlands (Birmingham) and the North East (Gateshead).

Three focus groups will consider the views of service users, with 5 participants per focus group. These focus groups will take place in the North West (Manchester), the Midlands (Birmingham) and the North East (Gateshead).

All focus groups will be audio-recorded and transcribed verbatim.. We will assume a realist perspective and report the experiences of participants. Themes will be coded inductively at a manifest level to inform the design of the prognostic tool. Transcripts will be reviewed and coded by Co-applicant Byrne, our service user researcher, with input from our service user reference group, the Lead Applicant Yung and Co-Lead Varese. Coding will be conducted systematically and iteratively.

4.2 - Phase 2: Development of a refined prognostic and risk stratification tool

Development of a refined prognostic model will include two components: (1) an updated prognostic model informed by the prognostic models and prognostic factors identified in Phase 1, and (2) development and internal validation of a novel prognostic model for transition to psychosis according to state-of-the art guidelines to avoid known methodological flaws.

4.2.1 Updating an existing model, to potentially include new prognostic factors

To avoid excess prognostic models with the same aim within the same clinical field we will begin Phase 2 by updating any models identified as statistically valid according to QUIPS [53] and PROBAST [43], and clinically relevant by the focus groups in Phase 1, potentially with additional prognostic factors identified in Phase 1 also.

There are two main approaches for updating clinical prognostic models – regression coefficient updating, and meta-model updating. Regression coefficient updating focuses on updating some or all coefficients from an existing prognostic model, while meta-model updating synchronises multiple existing prognostic models into one new meta-prognostic model [54]. If only small changes exist between the IPD used to develop the model (“original data”) and the additional IPD collected as part of Phase 1 (“updating data”), the simple regression coefficient updating method will be sufficient.[54] However, if larger changes are observed more involved methods such as a meta-model approach may be required.[27]

4.2.1.1 Regression Coefficient Updating

There is a hierarchical approach to updating the regression coefficients of an existing model, based on the available information. The most straightforward strategy is to adjust a model’s intercept such that the mean predicted probability of the event according to the prognostic model becomes equal to the observed event rate in the updating data (intercept update)[54]. Additional updating methods vary from overall adjustment of the model intercept and the overall calibration slope (logistic calibration), adjustment of a particular regression coefficient, to the re-estimation of included, or the addition of completely new, predictors to the existing model (model revision). The most appropriate

strategy will be determined once we know the results of the Phase 1 evidence synthesis and the initial feedback from the focus groups.

4.2.1.2 Meta-Model Updating

If we identify multiple historical prognostic models for the same or similar endpoints and populations in Phase 1, we might apply relevant meta-analysis techniques to synchronise them into one meta-model, in the presence of the updating data identified in Phase 1 – the IPD not used to develop the historic prognostic model(s). Stacked regressions will be used for this purpose.[27] Stacked regressions simultaneously update, discover and estimate the best combination of literature models in the updating data. They treat the predictions from each model as a predictor variable of the meta-model and subsequently create a linear combination of model prognostics. In particular, stacked regressions directly combine all original literature models into a meta-model by estimating unknown parameters from the updating data.

This approach offers a unique opportunity to update and combine any existing models using large datasets across a wide range of populations and settings [55-57]. This will enhance the potential generalisability of the resulting prognostic model across subgroups, settings and countries [58].

Aggregation techniques will be most appropriate if the validation samples are relatively small. If large amounts of data are available and patient populations from literature models are too heterogeneous with the validation population, developing a novel model may be the best strategy, using the available IPD from Phase 1[27]. Extensive updating strategies use more information from the updating dataset than the original dataset and may therefore lead to overfitting. Consequently, if extensive strategies are required, shrinkage will be applied towards the original prognostic model according to the relative sizes of the original and updated datasets to adjust for potential overfitting [59].

4.2.1.3 Model Assessment

Explained variation will be examined by Cox and Snell's R^2. The updated model(s) will be internally assessed according to discrimination and calibration. Calibration refers to how closely the probability of the event predicted by the model agrees with the observed probability. It will be assessed graphically [60] – if predicted and observed probabilities agree over the whole range of probabilities, the plots will show a 45° line. Discrimination refers to the ability of the prognostic model to differentiate between those who experience the event during the study and those who do not. The discriminative ability of the model will be measured with the *c*-index which is equivalent to the area under the ROC curve [61]. The *c*-index is measured on a scale ranging from 0.5 (no better than chance) to 1 (perfect prognostic) [62].

4.2.2 Developing a novel prognostic model

As described above, a recent systematic review identified several prognostic models and many prognostic factors described as associated with transition to psychosis in patients with ARMS. The authors of that review highlighted the poor methodological quality of the included studies and specifically stated that future studies should more strictly adhere to current checklists and guidelines on clinical prognostic models, such as the recently published TRIPOD statement [63]. Other recommendations included using an appropriate sample size, undertaking internal and external validation, and reporting model performance.

Given the poor quality of existing prognostic models and factors within this field [41], it is likely that model updating is still going to lead to a prognostic model which demonstrates poor discrimination in particular on internal validation. Even though the IPD will help to avoid some issues such as poor reporting of missing data and categorisation of continuous prognostic factors, model updating may not be sufficient to produce a well-performing prognostic model. In that case, we will develop a

novel prognostic model using the IPD from Phase 1, which will not suffer from the specified methodological flaws so frequently seen within prognosis research as outlined above.

4.2.2.1 Data

Given that the purpose of our model is to stratify patients at risk of transitioning to psychosis within the NHS, we will only use IPD from Phase 1 from prognostic models developed within the United Kingdom, or from countries with an equivalent healthcare system. Use of the IPD will enable standard analysis methods to be used across the IPD datasets and consistent predictor and outcome definitions [26]. The transportability of the novel model will be assessed by considering the discrimination and calibration of our model within the other IPD, obtained from outside of the UK.

We will make ensure that our dataset reports at least 10 transition events per variable (EPV)[64]. This may restrict the number of prognostic factors we can consider for inclusion in the model, but we will assess this once we have produced our list of candidate predictors and are aware of the relevant data from Phase 1.

4.2.2.2 Methodology – overall

In order to ensure transparency, we will develop and validate our model following the recommendations of the PROGRESS Group [65]. We will conform to the TRIPOD statement, a consensus-based guidance for improving the quality of reporting of multivariable prognostic model studies that includes 37 items covering 22 topics that should be included in any article describing the development or validation of a prognostic model[63].

There will be two stages to our prognostic modelling – model development, and internal validation (external validation of the novel model will take place in Phase 3 with a specifically-recruited UK-based cohort study).

4.2.2.3 Methodology – model development

The pool of potential prognostic factors for inclusion in our model will be based on the published systematic review [41], our own review undertaken in Phase I, and the output of the focus group work from Phase 1. All continuous covariates will be modelled as continuous covariates to avoid reducing the power to detect relationships and a loss of predictor information which can arise with categorisation. Linear and non-linear relationships will be considered including simple log transformations and more complex restricted cubic spline [66] and fractional polynomial options [67].

Given the difficulty of collecting certain information it is likely that some values may be missing for individuals which will cause a loss of power and precision. There is no established cut-off regarding an acceptable percentage of missing data. However, there is some evidence to suggest that that statistical analysis is likely to be biased when more than 10% of data are missing [68]. Therefore, in the case of more than 10% missing data for any covariate, multiple imputation will be used to impute missing values invoking Rubin's rule to combine estimates as necessary [69].

Logistic regression prognostic models should be built with a sample size of at least 10 events per variable (EPV) [64]. If our pool of prognostic factors together with relevant IPD from Phase 1 exceeds this sample size requirement, data reduction will be used [66]. In the first step variables with very narrow distributions will be removed. We will also consider whether it is sensible to remove variables with large amounts of missing data as it may be difficult to record these in practice. Backwards elimination according to Akaike's Information Criteria will then be used as an automatic selection procedure to determine variables in the final multivariable model [67].

4.2.2.4 Methodology – internal validation

Overall model fit will once again be examined by Cox and Snell's R^2. Internal validation is required to obtain an unbiased estimate of the developed model's predictive performance and, if necessary, adjust the developed model for optimism. Bootstrap resampling with 1000 replications will be used as it leads to accurate estimates of model performance [70]. Bootstrapping will also enable us to examine model consistency across bootstrap samples in terms of the selected predictors and the functional form of the continuous predictors. If necessary, we will revise our model if the original predictors or functional forms are not commonly chosen across bootstrap samples [71]. The final model will also be adjusted for optimism using uniform shrinkage [72].

4.2.3 Selection of model to be tested for external validation in Phase 3

Depending on the findings of phase 2, the model taken forward for external validation in phase 3 will either be (1) an existing model which will be updated using data from the IPD meta-analysis (as per section 4.2.1) or (2) a new model developed based on available IPD data (as per 4.2.2). However, if the updated model (4.2.1) or the new model (4.2.2) are not satisfactory due to limitations with available data, and an existing model with promising properties from previous internal and external validation studies is identified through stage 1 literature searches, this will be also evaluated in phase 3 in addition to any new model developed as part of the phase 2 work.

4.3 Phase 3: External validation of the prognostic tool

4.3.1 Design:

Prospective cohort study of individuals meeting the ARMS criteria, with 12 month follow up.

4.3.2 Aim:

To assess the external validity of the new prognostic tool

4.3.3 Hypothesis:

That the new prognostic tool will have better predictive validity for onset of psychosis in the ARMS group than the existing method of using the CAARMS.

4.3.4 Setting

Secondary NHS services that manage ARMS individuals in the North West, North East and West Midlands, covering the NHS Trusts of Greater Manchester, Lancashire, Birmingham, Solihull, South Staffordshire, Black Country, Worcester, Warwick, Coventry, Northumberland Tyne and Wear, Tees Esk and Wear Valley and Cumbria. Additionally, we include "Forward Thinking Birmingham", a youth mental health service that provides NHS services to people aged 0-25, including ARMS patients. All services already use the CAARMS to assess for an ARMS. Clinicians at all services have received training in use of the CAARMS funded by NHS England and developed by Lead Applicant Yung.

4.3.5 Target population

Inclusion criteria:

- Individuals attending secondary care services in the NHS aged 16 to 35 who meet ARMS criteria, as defined by the CAARMS. These validated operationalised criteria have been published previously and consist of cut-offs regarding intensity, frequency duration of

psychotic-like experiences and functional impairment [73]. The age range is included in the ARMS criteria and reflects that most psychotic disorders have their onset in late adolescence early adulthood [74]. The ARMS criteria have not been validated in older populations [75].

- In contact with NHS Early Intervention Services or Child and Adolescent Mental Health Services
- Individuals must be able to give informed consent for the follow up subsample.

Exclusion criteria:

- History of a treated or untreated psychotic episode of one week's duration or longer;
- Previous or current treatment with antipsychotics at dose of over 5 mf of haloperidol or equivalent for over 3 weeks

4.3.6 Sampling

Consecutive sampling will be used.

Sample size: We aim to recruit 798 ARMS individuals over 18 months, with an anticipated attrition rate of 30% leaving a target follow-up of 558 participants at 12 months. Using an estimated 20% rate of development of psychosis ("transition rate") within 12 months we expect to see 112 cases of psychosis ("transitions"). A more conservative estimate would be 100 transitions (18% transition rate). This estimated transition rate is based on recent data of an over 25% rate in Early Intervention for Psychosis Services (EIPS) in both Manchester and Warwick. This rate, which is higher than those recently reported [76], is to be expected given the change in the way that many ARMS patients are now coming to clinical attention following implementation of the Access and Waiting Times Standard in Early Psychosis [11]. This Standard requires all individuals referred to EIPS in England to be assessed within 2 weeks. If they are found not to have a psychosis then they must be assessed for presence of an ARMS. This means that many ARMS individuals have high levels of symptoms (they were thought to actually have psychosis by referrers). This pathway is similar to the first ever ARMS service established in Melbourne by Lead Applicant Yung. This service (the PACE Clinic [36]) managed individuals who were referred to the EIPS but were found to be below threshold for psychosis. Reflecting this, the rate of development of psychosis in early studies was 35-41% within 12 months [5, 77].

Justification of Sample Size: We expect to see 100 - 112 transition events during the study. Latest research regarding sample size calculations for external validation studies suggests at least 100 events and 100 non-events to access calibration-in-the-large and calibration slope, and at least 200 events and 200 non-events to derive flexible calibration curves for logistic regression [78, 79]. This was demonstrated via assessment of multiple measures of calibration including the c-statistic, calibration-in-the-large, and calibration slope using five established simulated examples [80].

Whilst flexible calibration curves are desirable for individual risk communication, they are unrealistic in empirical medical research which must provide value for money research outputs. Prediction models that are calibrated via calibration-in-the-large and calibration slopes guarantee that clinically non-harmful decisions are made based on the model [78]. Therefore, calibration-in-the-large and calibration slope will be sufficient to determine this level of external validation of our prognostic tool. Therefore, a dataset with at least 100 transition events is required.

Feasibility of recruitment: In the 12 months from July 2016 to June 2017, our included Trusts managed 887 ARMS individuals (420 Midlands, 294 North West, 163 North East and 10 Cumbria). We will recruit for 18 months. Even after allowing for year to year fluctuations and potential for less ARMS cases we will still be able to recruit 798 ARMS individuals. Allowing for 30% attrition over 12 months results in 558 individuals retained.

4.3.7 Data Collection

All data will be collected by clinicians working in NHS clinical services or by a research assistant employed by the clinical teams to support data collection and extraction. Data will be anonymously transferred by the research assistant in the clinical team to the research database which contains a unique participant identifier.

Participants in the main prospective observational cohort will have the option to 'opt out' of their routinely collected data being used for the research study using a combined information sheet and opt out form.

5a) Main cohort

Baseline: CAARMS assessment will have been routinely completed by NHS clinicians, or a research assistant employed within the clinical team, to determine ARMS status at point of entry to their service. The Research Assistants within the clinical teams will conduct the brief set of non-invasive measures to inform the refined prognostic model. If the CAARMS assessment was done over 4 weeks before the additional measures, it will be repeated by the Research Assistant in the clinical team to ensure that the individual still meets ARMS criteria.

12 and 24 month follow up of routine data: All participants in the cohort will be followed up via the clinical notes at 12 and 24 months. The research assistant within the clinical team will use the clinical records to determine if transition to a first episode of psychosis has occurred. This will be informed by clinician's notes which may include the following:

- Standard diagnostic classification systems (DSM-III, DSM-IV, DSM-5, ICD-10, ICD-11) or commonly used ARMS assessment schedules (e.g. SIPS or CAARMS)
- Standard treatment pathway decisions by clinicians- i.e. moved to a first episode psychosis treatment pathway
- A treated or untreated psychotic episode of one week's duration or longer;
- Treatment with antipsychotics at dose of over 5 mg of haloperidol or equivalent for over 3 weeks

5b) Cohort subsample

Participants in the main cohort will be invited to opt in to an additional phase of the study, where they will meet with the research assistant employed by the clinical team at 12 months after they entered the clinical service, to repeat the CAARMS assessment and some additional measures as described below. If participants decide they would like to opt in, they will be given a leaflet explaining the study and complete a consent form before their assessment.

12 month follow up for cohort subsample: CAARMS assessment will be completed by Research Assistants at 12 month follow up for a subsample of participants (minimum of 100 participants) who consent to this part of the study at their baseline assessment with the service. The 12-month time frame has been chosen as our previous cohort studies have found that most transitions to psychosis occur during this period [5, 31]. We will also include the EQ-5D-5L [32] and a core service use form [33] adapted from previous economic evaluations in mental health. These data will be used to generate estimates to populate the decision analytic model described below.

Minimising attrition: Attrition will be minimised by offering participants reimbursement of £25 per assessment. We will ask participants for the contact details of someone (relative or friend) who might know their whereabouts if they move during the follow up period. This strategy has received ethical approval in previous studies and is effective at reducing attrition.

Additionally, we will contact participants at 6 months using a method of their choice (phone, text, email, letter etc.) to 'check in'. This will be an opportunity to check contact details are up to date and

check if the participant is still with the clinical service. We will send participants a £5 voucher to thank them for keeping in touch.

We will also apply usual good practice in trial management by continuously monitoring completion of data at all visits and identify any sites that need help in achieving and maintaining high rates of return. Weekly management supervision of Research Assistants will monitor compliance to follow-up rates and solve problems relating to attrition as they arise.

Long term follow up of routine data in future studies: Additionally, through seeking consent with subsample participants to access health records up to 5 year follow up, the long term predictive validity of the prognostic model can be assessed in the sample in future research and potentially linked to other downstream physical and mental health outcomes available via electronic record records and other routine databases. Participation in any future studies and consent to contact about future studies will be completely optional.

4.3.8 Instruments

In addition to the baseline CAARMS assessment [12], assessments will include those that measure variables identified as potential predictors and included in the prognostic tool. The exact variables will depend on Phase 1 and Phase 2 of the study. We will only include factors that are feasible for routine use in the NHS. We will also include the EQ-5D-5L [84] and a core service use form [85] adapted from previous economic evaluations in mental health. These data will be used in the economic model.

4.3.9 Assessment of primary outcome

The primary outcome is development of psychosis by 12 month follow up. Psychosis onset will be assessed using the standard definition of psychosis included in the CAARMS. (This is essentially 7 days or more of full threshold psychotic symptoms occurring at least several times per week) or by clinician's notes on transition events which may include the following:

- Standard diagnostic classification systems (DSM-III, DSM-IV, DSM-5, ICD-10, ICD-11) or commonly used ARMS assessment schedules (e.g. SIPS or CAARMS)
- Standard treatment pathway decisions by clinicians- i.e. moved to a first episode psychosis treatment pathway
- A treated or untreated psychotic episode of one week's duration or longer;
- Treatment with antipsychotics at dose of over 5 mg of haloperidol or equivalent for over 3 weeks.

4.3.10 Assessment of acceptability

Descriptive statistics will be used to summarise assessments of feasibility and acceptability in terms of recruitment (willingness to take part and eligibility) and drop-out. Where possible, reasons for drop-out at 12 months will be recorded. Additionally, as a check on acceptability of the prognostic tool, Research Assistants will check participants' responses on measures included in the tool for completeness at initial assessment and the subsample follow up. If any are non-complete, participants' reasons for refusing to complete them will be recorded.

4.3.11 Assessment of treatments received

We will utilise electronic case notes to document any treatments received during the study. These will be recorded by the Research Assistant at follow up. In addition the core service use form [83] will be completed at follow up assessments with the consenting subsample participants. This instrument was developed by health economists with experience of UK-based economic evaluations

and consists of 10 items such as medications received, community-based health appointments and Accident and Emergency visits.

4.3.12 Analysis

Predictive validity: Effect of applying the prognostic tool versus usual practice will be assessed in the analysis. Treatments received (e.g. number of sessions of cognitive therapy) will be added as covariates in the analysis. External validation of the model will be undertaken by evaluating the discrimination and the calibration of the model developed in Phase 2 in the cohort of patients collected in Phase 3. Calibration refers to how closely the probability of the event predicted by the model agrees with the observed probability. It will be assessed graphically [60] – if predicted and observed probabilities agree over the whole range of probabilities, the plot shows a 45° line. The discriminative ability of the model will be measured with the *c* index. The *c* index is similar to the area under the receiver operating characteristic curve for logistic models and is measured on a scale ranging from 0.5 (no better than chance) to 1 (perfect prediction)[62].

Acceptability: Descriptive statistics will be used to summarise recruitment and attrition. Reasons for non-completion of any measures will be recorded and summarised.

Controlling for treatment received: Treatments received (e.g. number of sessions of cognitive therapy) will be added as covariates in the analysis.

Exploratory cost-effectiveness analysis: We will explore the potential cost-effectiveness of the new prognostic tool compared to usual care (no prognostic tool) from the perspectives of the NHS and Social Care (costs) and service users (health benefits) over 1, 2 and 5 year time-horizons. The primary and sensitivity analyses will use a cost-effectiveness acceptability approach to estimate incremental cost-effectiveness ratios, probability the new prognostic tool is cost-effective and net benefit statistic and generate cost-effectiveness-acceptability curves (CEAC). The range of hypothetical willingness to pay thresholds (WTPT) to gain a unit of health benefit values for the analyses will be £0 to £30,000, with a mid-estimate of £15,000 for the main analyses [86, 87]. For the primary analysis, the measure of health benefit for the primary analysis will be the quality adjusted Life year (QALY) and direct costs of health and social care services will be estimated (primary, secondary and community physical and mental health care; social care). Sensitivity analyses will explore the impact of using (i) alternative measures of health benefit (ii) including indirect costs of lost productive activity in the cost estimates. Threshold and scenario analyses will be used to explore the minimum level of effectiveness the prognostic tool needs to achieve to be cost-effective.

A decision analytic model will be used for all the analyses. The decision model structure, assumptions and data estimates will be developed iteratively from the reviews and analyses in Phases 1, 2 and 3 and structured discussions with key stakeholders (service users and health care professionals) and the research team at Investigator and Project Steering Committee meetings and through established service user networks in Manchester, the North East and Birmingham (face validity) [88]. Two researchers (Health Economics Research Fellow and Davies) will independently build and populate the model (internal validity). Sections of the model will be used to predict events/health states, costs and QALYs observed in the prospective cohort study and published literature (external validity).

It is anticipated that a Markov cohort model [89] will be an appropriate vehicle to capture transition between at risk, treated health states and the longer term recovery/relapse cycle of psychosis. The model will synthesise data from several sources to estimate the probability, costs and health benefits of events/health states. These include the data collected in Phases 1-3 to estimate probability of transition between risk states, with and without the prognostic tool, the costs and health benefits of the different risk states with and without appropriate treatment, the costs and health benefits of recovery and relapse states for psychosis. Data from systematic reviews, meta-analysis and controlled trials will be preferred to data from other sources. Probabilistic sensitivity

analysis will be used to account for data uncertainty in all analyses. Structural uncertainty will be assessed in the sensitivity analyses, threshold and scenario analyses outlined above.

5. Research timetable

The overall project will be undertaken over 51 months (01/10/18 – 31/12/22).

Phase 1 (evidence synthesis) will be completed within 16 months (01/10/18 – 01/02/2020). Prior to the start of the project, we will continue to consolidate our collaborative network and invite principal investigators of relevant ARMS cohort studies to join this collaboration. Between 01/10/18 to 01/01/19 we will finalise our review protocols and literature searches to identify all relevant IPD sets. The retrieval and processing of IPD will be completed by 31/11/19. In parallel, we will conduct the systematic review of multivariable prognostic models in the ARMS group (completed by 31/12/2019) and the focused health economic review. Data-analysis will be completed in parallel with the Phase 2 analyses (01/12/19 – 31/03/20). The Phase 1 findings will be written-up for publication by the 31/06/2020. In terms of health economic analyses, development and validation of decision analytic model structure and assumptions will be completed by 31/03/20. Estimation and validation of probability, cost and health benefit variables, initial analyses and model verification will take place between 01/04/20 and 31/10/22.

Phase 2 (development of prognostic tool) will start on 01/12/19. By 31/03/20 we will complete (1) the statistical analyses updating, and if necessary developing and internally validating a novel prognostic model and (2) the focus groups to assess feasibility and acceptability of prognostic factors identified in the evidence synthesis – to be conducted between 01/12/19 and 28/2/20.

Phase 3 (cohort study) will be undertaken over 33 months (01/11/20 – 01/05/22), and will be preceded by a 3-month set up period (01/08/20 – 31/10/2020) for gaining all necessary ethics/research governance approvals and ensure the timely recruitment/training of the RAs at all sites. Recruitment will commence on 01/11/20 and continue for 18 months. 12 month follow-up assessments will commence on 01/11/2021 and completed by 01/05/22. Data analysis will commence on 01/05/23 and, as findings become available, we will start the write-up of reports, journal submissions and our planned dissemination activities.

Final health economic analysis will occur from 01/05/23 to 30/04/23. The final HTA report will be submitted on the 31/07/23.

6. Outputs

The project will lead to the following outputs:

- A refined prognostic tool suitable for use in the NHS that improves the prediction of development of psychosis over and above the existing ARMS criteria.
- A Risk Calculator website: The prognostic tool will be made free of charge to NHS clinicians working with ARMS individual via a website, hosted by an NHS Trust. The website will include the measures that have been externally validated as predictive of development of psychosis and will provide an individualised estimated risk level for based on scores for each measure inputted by the clinician. In this way, it will be similar to risk calculators such as the online QRisk calculator that estimates an individual's 10 year risk of a myocardial infarction or stroke according to their level of known risk factors (see https://qrisk.org/three/index.php). Unlike the QRisk calculator, we plan that our website will be password protected to help ensure that it is only accessed by service providers who will be able to interpret the results appropriately.
- A published IPD meta-analysis of prognostic factors associated with transition to psychosis in the ARMS group;
- A published systematic review of ARMS prognostic models;

- A refined prognostic model and tool, complemented by an online risk stratification calculator that will be accessible through an NHS-hosted website and suitable for immediate implementation at the end of the research;
- A training package that supports the use of the tool and website
- A published paper illustrating the validity and performance of the new prognostic tool in a large UK sample of ARMS individuals;
- A health economic model that will guide the implementation tool across NHS services, and will be published in a peer-reviewed journal.

7. Research management arrangements

Effective project management will be established and maintained throughout the project. The applicants will establish an effective system of support to ensure high quality research governance across all stages of the research, and that this project is undertaken in a timely and effective way.

Investigator Meetings: 5 meetings will be held in Year 1, the first a videoconference meeting of all investigators and then 4 meetings of investigators closely involved in the evidence synthesis (2 face to face, 2 by videoconference). In Years 2 - 5 there will be an all-investigator face to face meeting each year and videoconference meetings every 2 months to monitor progress.

We will also set up a Project Steering Committee (PSC) prior to the start of the study. The PSC will comprise the study applicants, representatives of service users and sponsor, and have an independent chairman. It will meet annually. The PSC will monitor and supervise progress, consider reports and recommendations and responsible for approving relevant protocols and standard operating procedures.

An Independent Data Monitoring and Ethics Committee (IDMC) will also be established prior to the start of Phase 3 to monitor (1) participant recruitment, (2) ethical issues of consent, (3) quality of data (4) the incidence of adverse events, and (5) any other factors that might compromise the progress and satisfactory completion of the project. This will also have an independent chairman, and include an independent statistician. It will meet 6-monthly during the cohort study.

Prior to the start of Phase 3, we will ensure researchers at all sites receive training in Good Clinical Practice. Each site will have a weekly team meeting and the Manchester-based project manager, who will conduct weekly telephone supervision with all RAs that will focus on recruitment, liaison with referrers, compliance to follow-ups, and scoring queries for assessment measures. Lead Applicant Yung, who developed the CAARMS and associated training material, will provider regular monthly CAARMS supervision for Research Assistants and the Project manager and will be available for ad hoc scoring queries. She already provides monthly CAARMS supervision for Manchester-based clinicians and will extend this to other sites (via videoconferencing/ teleconference). These meetings will be supplemented by local supervision by site-PIs, focusing on problem solving and adherence to local policies and procedures.

8. Publication policy

In addition to the HTA report, the findings of this project will be published as series of papers in peer-reviewed journals. Group authorship on all outputs arising from the Phase 1 IPD meta-analysis will be offered to collaborators who will provide relevant IPD for the evidence synthesis conducted as part of this project.

9. Conflicts of interest

The applicants have no conflicts of interest to declare

10. Funding acknowledgement:

This project was funded by the NIHR Health Technology Assessment Programme (project number 17/31/05)

11. Department of Health disclaimer:

The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

References

1. Global Burden of Disease Study, C., *Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013.* Lancet, 2015. 386(9995): p. 743-800.
2. Häfner, H., et al., *The ABC schizophrenia study: a preliminary overview of the results.* Social Psychiatry and Psychiatric Epidemiology, 1998. 33(8): p. 380-386.
3. Schizophrenia Commission, *Schizophrenia - The Abandoned Illness*, R. Murray, Editor 2012.
4. Yung, A., et al., *Monitoring and care of young people at incipient risk of psychosis.* Schizophrenia Bulletin, 1996. 22(2): p. 283-303.
5. Yung, A.R., et al., *Risk factors for psychosis in an ultra high-risk group: Psychopathology and clinical features.* Schizophrenia Research, 2004. 67(2-3): p. 131-142.
6. Fusar-Poli, P., et al., *Predicting psychosis: a meta-analysis of evidence.* Archives of General Psychiatry, 2012. 69(3): p. 220-229.
7. McCrone, P., et al., *Paying the Price - The cost of mental health care in England to 2026*, 2008, King's Fund: London.
8. National Institute of Clinical Excellence, *NICE clinical guideline 155: Psychosis and schizophrenia in children and young people: Recognition and management*, 2013, National Institute for Health and Clinical Excellence,.
9. National Institute of Clinical Excellence, *NICE guidelines 178: Psychosis and schizophrenia in adults: treatment and management* http://www.nice.org.uk/guidance/CG178, 2014.
10. NHS England, *Guidance to support the introduction of access and waiting time standards for mental health services in 2015/16*, 2015, NHS England: London.
11. NHS England, National Collaborating Centre for Mental Health, and National Institute of Clinical Excellence, *Implementing the Early Intervention in Psychosis Access and Waiting Time Standard: Guidance.* NHS England Publications Gateway Reference 04294, 2016.
12. Yung, A., et al., *Mapping the onset of psychosis: the Comprehensive Assessment of At-Risk Mental States.* Australian and New Zealand Journal of Psychiatry, 2005. 39: p. 964-971.
13. Steyerberg, E.W., et al., *Prognosis Research Strategy (PROGRESS) 3: Prognostic Model Research.* PLOS Medicine, 2013. 10(2): p. e1001381.
14. Schmidt, A., et al., *Improving prognostic accuracy in subjects at clinical high risk for psychosis: systematic review of predictive models and meta-analytical sequential testing simulation.* Schizophrenia Bulletin, 2017. 43: p. 375–388.
15. Studerus, E., A. Ramyead, and A. Riecher-Rössler, *Prediction of transition to psychosis in patients with a clinical high risk for psychosis: a systematic review of methodology and reporting.* Psychological Medicine, 2017. 47: p. 1163-1178.
16. Moons, K.G., et al., *Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist.* PLoS Medicine, 2014. 11: p. e1001744.
17. Collins, G., D. Altman, and K. Moons, *Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement* Annals of Internal Medicine, 2015. 162: p. 55-63.
18. Cannon, T., et al., *An individualized risk calculator for research in prodromal psychosis.* American Journal of Psychiatry, 2016. 173: p. 980-988.
19. Carrión, R., et al., *Personalized prediction of psychosis: external validation of the NAPLS-2 psychosis risk calculator with the EDIPPP project.* American Journal of Psychiatry, 2016. 173: p. 989-996.

20. Lyman, G.H. and N.M. Kuderer, *The strengths and limitations of meta-analyses based on aggregate data.* BMC Medical Research Methodology, 2005. 5: p. 14.
21. Fusar-Poli, P., et al., *Predicting psychosis: Meta-analysis of transition outcomes in individuals at high clinical risk.* Archives of General Psychiatry, 2012. 69: p. 220-229.
22. Nelson, B., et al., *Long term follow up of a group at ultra high risk ("prodromal") for psychosis: the PACE 400 study. .* JAMA Psychiatry, 2013. 70: p. 793-802.
23. van der Gaag, M., et al., *Cognitive Behavioral Therapy for subjects at Ultra High Risk for developing psychosis: a randomized controlled clinical trial.* Schizophrenia Bulletin, 2012. 38: p. 1180–1188.
24. Steyerberg, E.W., et al., *Assessing the performance of prediction models: a framework for some traditional and novel measures.* Epidemiology, 2010. 21(1): p. 128-138.
25. Pencina, M.J. and R.B. D'Agostino, *Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation.* Statistics in Medicine, 2004. 23: p. 2109-2123.
26. Debray, T.P., et al., *Individual participant data (IPD) meta-analyses of diagnostic and prognostic modeling studies: guidance on their use.* PLoS medicine, 2015. 12(10): p. e1001886.
27. Debray, T.P.A., et al., *Meta-analysis and aggregation of multiple published prediction models.* Statistics in Medicine, 2014. 33(14): p. 2341-2362.
28. Stewart, L.A., et al., *Preferred reporting items for a systematic review and meta-analysis of individual participant data: The prisma-ipd statement.* JAMA, 2015. 313(16): p. 1657-1665.
29. Collins, G.S., et al., *Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (tripod): The tripod statement.* Annals of Internal Medicine, 2015. 162(1): p. 55-63.
30. Moons, K.M., et al., *Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (tripod): Explanation and elaboration.* Annals of Internal Medicine, 2015. 162(1): p. W1-W73.
31. Yung, A.R., et al., *Mapping the onset of psychosis: The Comprehensive Assessment of At Risk Mental States (CAARMS).* Australian & New Zealand Journal of Psychiatry, 2005. 39: p. 964-971.
32. McGlashan, T.H., et al., *Instrument for the assessment of prodromal symptoms and states.* Early intervention in psychotic disorders, 2001: p. 135-149.
33. Fusar-Poli, P., et al., *Predicting psychosis: Meta-analysis of transition outcomes in individuals at high clinical risk.* Archives of General Psychiatry, 2012. 69(3): p. 220-229.
34. Schultze-Lutter, F., *Subjective symptoms of schizophrenia in research and the clinic: The basic symptom concept.* Schizophrenia Bulletin, 2009. 35(1): p. 5-8.
35. Schultze-Lutter, F., et al., *Schizophrenia proneness instrument, adult version (SPI-A).* Rome: Giovanni Fioriti, 2007.
36. Yung, A.R., et al., *The PACE Clinic: Development of a Clinical Service for young people at high risk of psychosis.* Australasian Psychiatry, 1995. 3(5): p. 345-349.
37. Yung, A.R., et al., *Monitoring and care of young people at incipient risk of psychosis.* Schizophr Bull, 1996. 22(2): p. 283-303.
38. Fusar-Poli, P., et al., *At Risk for Schizophrenic or Affective Psychoses? A Meta-Analysis of DSM/ICD Diagnostic Outcomes in Individuals at High Clinical Risk.* Schizophrenia Bulletin, 2013. 39(4): p. 923-932.
39. Fusar-Poli, P., et al., *At risk or not at risk? A meta-analysis of the prognostic accuracy of psychometric interviews for psychosis prediction.* World Psychiatry, 2015. 14(3): p. 322-332.
40. Simon, A.E., et al., *Moving beyond transition outcomes: Meta-analysis of remission rates in individuals at high clinical risk for psychosis.* Psychiatry Research, 2013. 209(3): p. 266-272.

41. Studerus, E., A. Ramyead, and A. Riecher-Rössler, *Prediction of transition to psychosis in patients with a clinical high risk for psychosis: a systematic review of methodology and reporting.* Psychological medicine, 2017. 47(7): p. 1163-1178.
42. Moons, K.G.M., et al., *Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist.* PLOS Medicine, 2014. 11(10): p. e1001744.
43. Wolff, R. *PROBAST*. 2017; Available from: http://s371539711.initial-website.co.uk/probast/.
44. Altman, D.G., *Prognostic Models: A Methodological Framework and Review of Models for Breast Cancer.* Cancer Investigation, 2009. 27(3): p. 235-243.
45. DerSimonian, R. and N. Laird, *Meta-analysis in clinical trials.* Control Clin Trials, 1986. 7(3): p. 177-88.
46. Riley, R.D., J.P.T. Higgins, and J.J. Deeks, *Interpretation of random effects meta-analyses.* BMJ, 2011. 342.
47. Steyerberg, E.W., et al., *Prognosis Research Strategy (PROGRESS) 3: Prognostic Model Research.* PLoS Medicine, 2013. 10(2): p. e1001381.
48. Altman, D.G., et al., *Prognosis and prognostic research: validating a prognostic model.* BMJ, 2009. 338: p. b605.
49. Hayden, J.A., et al., *ASsessing bias in studies of prognostic factors.* Annals of Internal Medicine, 2013. 158(4): p. 280-286.
50. Centre for Reviews and Dissemination, *NHS Economic Evaluation Database Handbook*. 2007: University of York, UK.
51. Shields, G.E., et al., *Cost-effectiveness of cardiac rehabilitation: a systematic review.* Heart, 2018.
52. Braun, V. and V. Clarke, *Using thematic analysis in psychology.* Qualitative Research in Psychology, 2006. 3(2): p. 77-101.
53. Hayden, J.A., P. Cote, and C. Bombardier, *Evaluation of the quality of prognosis studies in systematic reviews.* Ann Intern Med, 2006. 144(6): p. 427-37.
54. Su, T.-L., et al., *A review of statistical updating methods for clinical prediction models.* Statistical Methods in Medical Research, 2016. 0(0): p. 0962280215626466.
55. Debray, T.P., et al., *A framework for developing, implementing, and evaluating clinical prediction models in an individual participant data meta-analysis.* Stat Med, 2013. 32(18): p. 3158-80.
56. Royston, P., M.K.B. Parmar, and R. Sylvester, *Construction and validation of a prognostic model across several studies, with an application in superficial bladder cancer.* Statistics in Medicine, 2004. 23(6): p. 907-926.
57. Cai, T., et al., *Robust Prediction of t-Year Survival with Data from Multiple Studies.* Biometrics, 2011. 67(2): p. 436-444.
58. Ahmed, I., et al., *Developing and validating risk prediction models in an individual participant data meta-analysis.* BMC Medical Research Methodology, 2014. 14(1): p. 3.
59. Steyerberg, E.W., et al., *Validation and updating of predictive logistic regression models: a study on sample size and shrinkage.* Statistics in Medicine, 2004. 23(16): p. 2567-2586.
60. Steyerberg, E.W., et al., *Assessing the performance of prediction models: a framework for some traditional and novel measures.* Epidemiology (Cambridge, Mass.), 2010. 21(1): p. 128-138.
61. Austin, P.C. and E.W. Steyerberg, *Interpreting the concordance statistic of a logistic regression model: relation to the variance and odds ratio of a continuous explanatory variable.* BMC medical research methodology, 2012. 12(1): p. 82.
62. Pencina, M.J. and R.B. D'Agostino, *Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation.* Stat Med, 2004. 23(13): p. 2109-23.

63. Collins, G.S., et al., *Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement.* BMC medicine, 2015. 13(1): p. 1.
64. Peduzzi, P., et al., *A simulation study of the number of events per variable in logistic regression analysis.* Journal of clinical epidemiology, 1996. 49(12): p. 1373-1379.
65. Peat, G., et al., *Improving the transparency of prognosis research: the role of reporting, data sharing, registration, and protocols.* PLoS medicine, 2014. 11(7): p. e1001671.
66. Harrell, F.E., in *Regression Modeling Strategies with Applications to Linear Models, Logistic Regression, and Survival Analysis*. 2001, Springer-Verlag: New York.
67. Royston, P. and W. Sauerbrei, *Multivariable Model-Building - A pragmatic approach to regression analysis based on fractional polynomials for modelling continuous variables*. 2008: Wiley.
68. Bennett, D.A., *How can I deal with missing data in my study?* Australian and New Zealand journal of public health, 2001. 25(5): p. 464-469.
69. Little, R.J.A. and D.B. Rubin, in *Statistical Analysis with Missing Data, Second edition*. 2002, John Wiley and Sons: New York.
70. Steyerberg, E.W., et al., *Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis.* Journal of Clinical Epidemiology, 2001. 54(8): p. 774-781.
71. Sauerbrei, W., P. Royston, and H. Binder, *Selection of important variables and determination of functional form for continuous predictors in multivariable model building.* Stat Med, 2007. 26(30): p. 5512-28.
72. Steyerberg, E.W., *Clinical Prediction Models*. Statistics for Biology and Health. 2009, New York: Springer Science+Business Media, LLC.
73. McGorry, P., et al., *Effect of ω-3 polyunsaturated fatty acids in young people at ultrahigh risk for psychotic disorders: The NEURAPRO randomized clinical trial* JAMA Psychiatry, 2017. 74: p. 19-27.
74. Kessler, R.C., et al., *Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National comorbidity Survey Replication.* Archives of General Psychiatry, 2005. 62: p. 593-602.
75. Fusar-Poli, P., et al., *At risk or not at risk? A meta-analysis of the prognostic accuracy of psychometric interviews for psychosis prediction.* World Psychaitry, 2015. 14: p. 322-331.
76. Hartmann, J.A., et al., *Declining transition rates to psychotic disorder in "ultra-high risk" clients: Investigation of a dilution effect.* Schizophr Res, 2016. 170: p. 130-136.
77. Yung, A.R., et al., *Psychosis prediction: 12-month follow up of a high-risk ("prodromal") group.* Schizophrenia Research, 2003. 60(1): p. 21-32.
78. Van Calster, B., et al., *A calibration hierarchy for risk models was defined: from utopia to empirical data.* Journal of clinical epidemiology, 2016. 74: p. 167-176.
79. Collins, G.S., E.O. Ogundimu, and D.G. Altman, *Sample size considerations for the external validation of a multivariable prognostic model: a resampling study.* Statistics in medicine, 2016. 35(2): p. 214-226.
80. Vach, W., *Calibration of clinical prediction rules does not just assess bias.* Journal of clinical epidemiology, 2013. 66(11): p. 1296-1301.
81. Ratheesh, A., et al., *Neurocognitive functioning in the prodrome of mania- an exploratory study.* Journal of Affective Disorders, 2013. 147: p. 441-445.
82. Janssen, M.F., et al., *Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study.* Quality of Life Research 2013. 22: p. 1717-1727.
83. Thorn, J.C., et al., *Core Items for a Standardized Resource Use Measure (ISRUM): Expert Delphi Consensus Survey.* Value in Health, Published online: September 1, 2017. DOI: http://dx.doi.org/10.1016/j.jval.2017.06.011.

84. Herdman, M., et al., *Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L).* Quality of Life Research, 2011. 20: p. 1727-1736
85. Thorn, J.C., et al., *Core Items for a Standardized Resource Use Measure (ISRUM): Expert Delphi Consensus Survey.* Value in Health. Published online: September 1, 2017. DOI: http://dx.doi.org/10.1016/j.jval.2017.06.011.
86. McCabe, C., Claxton, K., and A.J. Culyer, *The NICE cost-effectiveness threshold.* Pharmacoeconomics, 2008. 26: p. 733-744.
87. Claxton, K., et al., *Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold.* Health Technology Assessment, 2015. 19(14): p. 1-503.
88. Roberts, M., et al., *Conceptualizing a model: A report of the ISPOR-SMDM modeling good research practices task force-2.* Value Health, 2012. 15: p. 804-811.
89. Siebert, U., et al., *State-transition modeling: A report of the ISPOR-SMDM modeling good research practices task force-3.* Value Health, 2012. 15: p. 812-820.

Appendix 1 - EXAMPLE SEARCH FOR THE ECONOMIC EVAULUATIONS REVIEW (EMBASE)

1 psychosis/ or psychosis.mp.
2 (clinically at high risk or clinically at risk or clinical high risk or ultra-high risk or prodrom* or at risk mental state or risk of psychosis or ARMS or prodromal psychosis).mp.
3 1 and 2
4. Health Economics/
5. exp Economic Evaluation/
6. exp Health Care Cost/
7. exp Pharmacoeconomics/
8. 6 or 7 or 8 or 9
9. (econom$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$).ti,ab.
10. (expenditure$ not energy).ti,ab.
11. (value adj2 money).ti,ab.
12. budget$.ti,ab.
13. 11 or 12 or 13 or 14
14. 10 or 15
15. letter.pt.
16. editorial.pt.
17. note.pt.
18. 17 or 18 or 19
19. 16 not 20
20. (metabolic adj cost).ti,ab.
21. ((energy or oxygen) adj cost).ti,ab.
22. ((energy or oxygen) adj expenditure).ti,ab.
23. 22 or 23 or 24
24. 21 not 25
25. exp Animal/
26. exp Animal/
27. exp Animal Experiment/
28. Nonhuman/
29. (rat or rats or mouse or mice or hamster or hamsters or animal or animals or dog or dogs or cat or cats or bovine or sheep).ti,ab.
30. 27 or 28 or 29 or 30 or 31
31. exp Human/
32. exp Human Experiment/
33. 33 or 34
34. 32 not (32 and 35)
35. 26 not 36
36. 3 and 37

Appendix 2 DATA ITEMS TO BE EXTRACTED FOR THE ECONOMIC REVIEW

Study details
Title of paper
Author(s)
Journal title, Year, Vol. (issue): pages
Subject of study
Intervention(s)
Comparator(s)
Country and setting
Time horizons
Condition
Study question and perspective(s)
Key elements
Type of economic analysis
Study population
Modelling study? (y/n); If yes, basic model details
Were costing and effectiveness inputs derived from the same patient sample?
Clinical evidence
List the clinical inputs included
Data sources
Methods to obtain data
Measures of health benefit
Summary measure of health benefit
If no measure of health benefit: was an adequate cost-minimisation argument made? (y/n/NA)
Utility values: whose values?
Utility values: how were they elicited?
Discounting of health outcomes
Direct costs
List the types of direct cost included
Describe who bears these costs
Source of resource use data (primary, secondary, assumptions?)
Are resource use and cost inputs reported separately?
Sources of unit prices
Currency and price year
Price adjustments
Do the authors report excluding any direct cost items?
Discounting of direct costs
Indirect costs
Are any indirect costs included? (If no, NA for rest of section.)
List the types of indirect cost included
Describe who bears these costs
Source of resource use data (primary, secondary, assumptions?)
Are resource use and cost inputs reported separately? (y/n)
Sources of unit prices
Currency and price year
Price adjustments

Do the authors report excluding any indirect cost items?
Discounting of indirect costs
Statistical summary of costs
Descriptive statistics reported
Was a test used to determine whether there is a statistically significant difference in costs?
Was the study powered to detect cost differences?
Analysis of uncertainty
If model: was parameter uncertainty explored?
If model: were all parameters included?
If model: was structural uncertainty explored?
If not model: was variation in patient-level data explored?
All studies: were alternative subgroups / settings explored?
All studies: was value of information analysis performed?
Results: Estimated benefits (if applicable)
Average and incremental measures of benefit (bullet point per comparison)
Did the duration of benefit match the observed data?
Were adverse effects captured in the measure?
Results: Estimated costs
Total cost: intervention arm(s)
Total cost: comparator arm(s)
Incremental cost (column per comparison)
Were adverse effect costs included?
Result of statistical test for difference in costs
Did the duration of costs match the time horizon?
Synthesis of benefits & costs, and conclusions
How were benefits and costs combined?
Summary results (eg ICER; CEAC)
Important differences in results for subgroups or sensitivity analyses.
Summary of authors' conclusions
Critical review
Is the choice of comparator suitably justified?
If model: was the model structure suitable?
If model: was a model schematic presented?
If model: was the model adequately reported?
Validity of primary effectiveness data
Validity of secondary effectiveness data
Validity of estimated health benefit
Validity of estimated costs
Do the authors discuss the generalisability of their findings?
Do the authors compare their findings to previous studies?
Strengths mentioned by the authors
Limitations mentioned by the authors
Are the authors' conclusions justified?
Implications
Do the authors describe policy implications of their findings? Are they appropriate?
Do the authors describe research recommendations? Are they appropriate?