the sample size, multiple imputation outperformed complete-case analyses and should be considered when validating a prognostic model

## P9

### CHecklist for critical Appraisal and data extraction in systematic Reviews of clinical prediction Modelling Studies (CHARMS)

K. G. M. (Carl) Moons[1], Joris A. H. de Groot[1], Sue Mallett[2], Doug G. Altman[3], Johannes B. Reitsma[1,4], Gary S. Collins[3]
[1]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, 3584 CG Utrecht, Netherlands; [2]Institute of Applied Health Research, University of Birmingham, Birmingham, UK; [3]Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK; [4]Cochrane Netherlands, University Medical Center Utrecht, Utrecht, The Netherlands
**Correspondence:** K. G. M. (Carl) Moons (k.g.m.moons@umcutrecht.nl)
*Diagnostic and Prognostic Research* 2016, **1(Suppl 1):**P9

**Background**: Publications on multivariable clinical prediction models have become abundant for both prognostic and diagnostic purposes. Systematic reviews of these studies are increasingly required to identify and critically appraise the existing evidence. There is currently no checklist or tool providing guidance for systematic reviews of studies developing or validating prediction models that can assist reviewers to define the review objectives and appraise study methodology.
**Objective**: To develop a checklist to help reviewers framing a well-defined review question, and to determine which details to extract and critically appraise from primary studies on the development or validation of multivariable diagnostic or prognostic prediction models, with a view to assessing the risk of bias and sources of heterogeneity.
**Methods**: We critically examined existing reporting guidelines and quality assessment tools, key methodological publications on clinical prediction modelling, and tools used in published systematic reviews of multivariable prediction models, to identify the relevant characteristics and domains. The checklist was tested in various systematic reviews.
**Results**: We identified 7 items important for framing the review question (diagnostic versus prognostic model, intended scope of the review, type of prediction modelling studies, target population, outcome to be predicted, time span of prediction, intended moment of using the model), and 11 domains to critically appraise the primary included studies (source of data, participants, outcome, predictors, sample size, missing data, model development, model performance, model evaluation, results, interpretation). Both were combined into the CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies (CHARMS).
**Conclusions**: CHARMS is designed to assist reviewers to help systematic reviewers framing their review objectives, and to determine which data to extract and critically appraise from primary studies on the development and/or validation of (diagnostic and prognostic) prediction models.

## P10

### New Guideline for the Reporting of Studies Developing, Validating, or Updating a Prediction Model: the TRIPOD Statement

Karel G. M. Moons[1], Douglas G. Altman[2], Johannes B. Reitsma[1,3], Gary S. Collins[2]
[1]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, 3584 CG Utrecht, Netherlands; [2]Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK; [3]Cochrane Netherlands, University Medical Center Utrecht, Utrecht, The Netherlands
**Correspondence:** Karel G. M. Moons (k.g.m.moons@umcutrecht.nl)
*Diagnostic and Prognostic Research* 2016, **1(Suppl 1):**P10

**Background and objective.** Patients and care providers are confronted with making numerous decisions based on a probability; a probability that a specific disease or condition is present (diagnostic setting) or a specific event or outcome will occur in the future (prognostic setting). To guide practitioners and patients in these probability estimations, so-called multivariable prediction models are developed. Prediction models convert 2 or more pieces of information, i.e. predictors, from the participant - e.g., an individual's age, gender, symptoms, signs, laboratory and imaging test results - into a diagnostic or prognostic probability. Prediction models are becoming increasingly abundant. In virtually all medical domains, prediction models are being developed, evaluated (validated), extended and implemented. For some specific diseases, there are even an overwhelming number of competing prediction models for the same outcome or target population. It is therefore important that these clinical prediction models and the research done to develop, evaluate or extend these models be transparently reported. However, the overwhelming evidence shows that the quality of reporting of prediction model studies is poor. Only with full and clear reporting of information on all aspects of a prediction model can risk of bias and potential usefulness of prediction models be adequately assessed.
**Methods and results.** The Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) initiative, which has included numerous clinicians, statisticians, epidemiologists and journal editors, has produced a guideline for the reporting of studies developing, validating or updating a prediction model, whether for diagnostic or prognostic purposes. The TRIPOD Statement is a checklist of 22 items, deemed essential for transparent reporting of any prediction model study, and addresses model development, model validation and model extension studies, regardless of the study methods used. The TRIPOD Statement is accompanied by an Explanation and Elaboration article that describes the rationale for the checklist, clarifies the meaning of each item, and discusses why transparent reporting is important, with a view to assessing risk of bias and clinical usefulness of a prediction model. Each item is explained in detail and accompanied by published examples of good reporting. The document also provides a valuable reference of issues to consider when designing, conducting, and analyzing prediction model studies.
**Conclusions.** The endorsement and use this checklist by researchers and medical journal editors will help ensure that medical research findings are complete and accurately reported, understood by readers, and ultimately used by medical practitioners.

## P11

### Incorporating the time-dependency in ROC methodology for censored clinical event

Adina Najwa Kamarudin, Ruwanthi Kolamunnage-Dona, Trevor Cox
Department of Biostatistics, University of Liverpool, Liverpool, UK
**Correspondence:** Adina Najwa Kamarudin (adinajwa@liv.ac.uk)
*Diagnostic and Prognostic Research* 2016, **1(Suppl 1):**P11

The classical approach of ROC (receiver operating characteristic) analysis considers event (disease) status and biomarker of an individual as fixed over time; however in practice both the disease status and biomarker change over time. Individuals who are disease-free earlier may develop the disease later due to longer study follow-up, and also have their biomarker changed from baseline over follow-up. Thus, an ROC as function of time is more appropriate. The time-dependent sensitivity and specificity can be defined into three definitions which are cumulative/dynamic (C/D), incident/dynamic (I/D) and incident/static (I/S). We focus on I/D and I/S definitions in this presentation. Incident sensitivity and dynamic specificity use a pre-defined time point for discriminating between individuals who failed and individuals who remained disease-free while static specificity uses a time interval. Further, I/D definition is used for a single marker while I/S definition is used for longitudinal marker. We review the current estimation methods and compare their behaviour in practice using a real dataset in primary biliary cirrhosis.
**Keywords:** ROC, time-dependent, accuracy, biomarker, event-time, longitudinal data

## P12

### Bayesian network approach to assessment of medical technologies

Melody Ni, Jeremy Huddy, Simone Borsci, George Hanna
NIHR-Diagnostic Evidence Cooperative, Imperial College London, London, UK
**Correspondence:** Melody Ni (z.ni@imperial.ac.uk)
*Diagnostic and Prognostic Research* 2016, **1(Suppl 1):**P12