UMC Utrecht

TRIPOD-CLUSTER

Reporting of prediction model studies in IPD-MA, EHR and other clustered datasets

Thomas P. A. Debray, Kym Snell, Ben van Calster, Gary S. Collins, Richard D. Riley, Johannes B. Reitsma, Doug G. Altman, Karel G. Moons

Background

Poor conduct & reporting of prediction model studies

Poor conduct

- Use of small data sets
- Inappropriate statistical methods
- Lack of (external) validation

Poor reporting

- Examination of predictors
- Presence and handling of missing data
- Model-building strategy

UMC Utrecht

The TRIPOD Statement

Annals of Internal Medicine RESEARCH AND REPORTING METHODS

Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement

Gary S. Collins, PhD; Johannes B. Reitsma, MD, PhD; Douglas G. Altman, DSc; and Karel G.M. Moons, PhD

Prediction models are developed to aid health care providers in estimating the probability or risk that a specific disease or condition is present (diagnostic models) or that a specific event will occur in the future (prognostic models), to inform their decision making. 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. The Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) Initiative developed a set of recommendations for the reporting of studies developing, validating, or updating a prediction model, whether for diagnostic or prognostic purposes. This article describes how the TRIPOD Statement was developed. An extensive list of items based on a review of the literature was created, which was reduced after a Web-based survey and revised during a 3-day meeting in June 2011 with methodologists, health care professionals, and journal editors. The list was refined during several meetings of the steering group and in e-mail discussions with the wider group of TRIPOD contributors. The resulting TRIPOD Statement is a checklist of 22 items, deemed essential for transparent reporting of a prediction model study. The TRIPOD Statement aims to improve the transparency of the reporting of a prediction model study regardless of the study methods used. The TRIPOD Statement is best used in conjunction with the TRIPOD explanation and elaboration document. To aid the editorial process and readers of prediction model studies, it is recommended that authors include a completed checklist in their submission (also available at www.tripod-statement.org).

Ann Intern Med. 2015;162:55-63. doi:10.7326/M14-0697 **www.annals.org**

For author affiliations, see end of text.

For contributors to the TRIPOD Statement, see the **Appendix** (available at www.annals.org).

The TRIPOD Statement

Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis

- Reporting guidelines for studies developing, validating and/or updating diagnostic or prognostic models
- Checklist of 22 items
  (covering title and abstract, background and objectives, methods, results, discussion and other information)
- Focus on situations where studies are based on a single dataset.

http://www.tripod-statement.org/

TRIPOD in an era of “big data”

Growing access to large amounts of individual participant data (IPD), which allows for more complex prediction model development and validation strategies

- Meta-analyses based on IPD
- Registries with electronic health-care records (EHR)
- Prospective multicenter studies

TRIPOD in an era of “big data”

The use of large datasets exposes important new challenges. Individuals are often clustered within studies, hospitals, regions, or even countries

Heterogeneity across data sources

- Research protocols
- Patient characteristics (“case mix”)
- Measurement methods & variable definitions
- Data quality
- Predictive associations (e.g. regression coefficients)
- Predictive performance

TRIPOD in an era of “big data”

Key questions:

- Is pooling of data appropriate?
- How to combine the available data?
- How to configure (e.g. regression or machine learning) models?
- How to report estimated model parameters?
- How to evaluate model performance?
- How to generate predictions for new individuals?

Extending the TRIPOD Statement

Development process

- 2016: Forming of an executive group
- 2016: Reviewing of existing reporting guidelines
- 2016 – 2019: Conventions in Utrecht, Oxford & Keele
- 2019: Online Delphi surveys
- 2020: Manuscript preparation







Extending the TRIPOD Statement

Existing guidelines

- The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. PLOS Medicine. 2015.
- Preferred Reporting Items for Systematic Review and Meta-Analyses of individual participant data: the PRISMA-IPD Statement. JAMA. 2015.
- Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. PLoS Med. 2007.

Extending the TRIPOD Statement

Main changes in Methods items

- Describe how the data were identified, requested and collected
- Describe how the data were prepared for analysis, including any cleaning, harmonization and linkage
- Describe how potential sources of bias were assessed
- Describe how any heterogeneity (e.g., across data sources or settings) in model parameters was handled during model development
- Describe how any heterogeneity (e.g., across data sources or settings) in model performance was handled and quantified

Extending the TRIPOD Statement

Main changes in Results items

- Report the results from any bias assessments (e.g. PROBAST), for each data source or setting.
- Report results of any heterogeneity (e.g., across data sources or settings) in model parameters, and subsequent actions (e.g., inclusion or exclusion of particular predictors or data sources).
- Report results from any subgroup or sensitivity analysis.

Extending the TRIPOD Statement

Delphi Survey 1

Electronic questionnaire distributed to 77 participants

- Review of 27 modifications by statisticians, epidemiologists, physicians, and journal editors
- Rated as “agree”, “no opinion” or “disagree (please comment)”
- Survey opened in January 2019 (duration: 50 days)
- Feedback from 17 participants (22%)
  - 3 modifications with > 30% “disagree”
  - 3 modifications with 15 – 30% “disagree”
  - Agreement on remaining item changes

Extending the TRIPOD Statement

Delphi Survey 2

Electronic questionnaire distributed to 77 participants

- Review of 19 modifications
- Rated as “agree”, “no opinion” or “disagree (please comment)”
- Survey opened in March 2019 (duration: 46 days)
- Feedback from 30 participants (39%)
  - No modifications with > 30% “disagree”
  - 2 modifications with 15 – 30% scored “disagree”
  - Agreement on remaining item changes

Closing remarks

Full and transparent reporting of prediction model studies is critical

- The analysis of large datasets with clustering (often) requires extra efforts
- The implementation of prediction models derived from large datasets with clustering is more complex
- Reporting standards should reflect this, yet, guidelines are currently lacking

TRIPOD-Cluster has been submitted to Annals of Internal Medicine (revision underway)

More information: T.Debray@umcutrecht.nl (Twitter @TPA_Debray)