Journal of Clinical Epidemiology 102 (2018) 38–49

Journal of Clinical Epidemiology

REVIEW

# Measurement error and timing of predictor values for multivariable risk prediction models are poorly reported

Rebecca Whittle[a,*], George Peat[a], John Belcher[a], Gary S. Collins[b], Richard D. Riley[a]

[a]*Centre for Prognosis Research, Arthritis Research UK Primary Care Centre, Research Institute for Primary Care & Health Sciences, Keele University, Keele, Staffordshire, UK*

[b]*Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Botnar Research Centre, University of Oxford, Oxford, UK*

Accepted 14 May 2018; Published online 18 May 2018

## Abstract

**Objective:** Measurement error in predictor variables may threaten the validity of clinical prediction models. We sought to evaluate the possible extent of the problem. A secondary objective was to examine whether predictors are measured at the intended moment of model use.

**Methods:** A systematic search of Medline was used to identify a sample of articles reporting the development of a clinical prediction model published in 2015. After screening according to a predefined inclusion criteria, information on predictors, strategies to control for measurement error, and intended moment of model use were extracted. Susceptibility to measurement error for each predictor was classified into low and high risks.

**Results:** Thirty-three studies were reviewed, including 151 different predictors in the final prediction models. Fifty-one (33.7%) predictors were categorized as high risk of error; however, this was not accounted for in the model development. Only 8 (24.2%) studies explicitly stated the intended moment of model use and when the predictors were measured.

**Conclusion:** Reporting of measurement error and intended moment of model use is poor in prediction model studies. There is a need to identify circumstances where ignoring measurement error in prediction models is consequential and whether accounting for the error will improve the predictions. 

*Keywords:* Prediction models; Prediction; Prognosis; Diagnosis; Measurement error; Error

## 1. Background

Predicting a patient's future outcome risk is an important part of medical research as it guides treatment, informs clinical decision-making, and helps patients understand their risk. Prognosis research can be used to help predict future outcomes in patients with a particular disease or health condition by developing a prediction model [1]. The number of articles reporting clinical prediction models has been increasing steadily over time, with approximately 500 articles published in 2011 [2], and these models use values of multiple predictors to enable individualized risk prediction [3]. Such models are intended "to assist clinicians with their prediction of a patient's future outcome and to enhance informed decision-making with the patient" [4]. Therefore, the predictions from these models should have optimal performance when being practically implemented at the "intended moment of using the model" [5].

However, when developing such models, measurement error may affect the observed predictor values, which could potentially lead to biased or incorrect estimates of predictor-outcome associations [6–9]. Measurement error is a difference between the measured values of a predictor and the true values of the predictor, or if the predictor is categorical, it is the classification to an incorrect category (misclassification). The term measurement error will be used throughout this article to refer generally to measurement error in continuous predictors and misclassification of categorical predictors. Measurement error is common within clinical studies, particularly observational studies [10], and has been found to be commonly neglected within the medical literature [11]. Measurement error can occur

Conflict of interest: None.

Funding: R.W. was funded through an NIHR Research Professorship in General Practice for Christian Mallen (NIHR-RP-2014-04-026). G.S.C. was supported by the NIHR Biomedical Research Centre, Oxford. The views expressed in this article are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health.

* Corresponding author. Tel.: +44 1782 734848; fax: +44 1782 734719.

*E-mail address:* r.l.whittle@keele.ac.uk (R. Whittle).

https://doi.org/10.1016/j.jclinepi.2018.05.008

**What is New?**

**Key findings**

- Many published prediction models include predictors that are susceptible to measurement error, and this measurement error is not being acknowledged or accounted for in the development of the models.
- Most prediction model articles do not explicitly state the intended moment of model use or exactly when the predictors used in the model development were measured.

**What this adds to what was known?**

- Reporting of measurement error and intended moment of model use is poor in prediction model studies.

**What is the implication and what should change now?**

- There is a need to identify circumstances where ignoring measurement error in prediction models is consequential and whether accounting for the error will improve the predictions.
- Future prediction model research studies must clearly report the intended moment of use of the prediction model and be explicit about when the predictors were measured.

for many different reasons such as biological variability, inaccuracy of measurement instruments, imperfect recall, cost or resource limitations, the subjective nature of measures, laboratory or measurer error, and timing error. For example, measurement error in blood pressure commonly occurs due to biological variability [12]. Body mass index (BMI) is also commonly measured with error either due to the inaccuracy of measurement instruments (i.e., the scales not being calibrated correctly) or due to imperfect recall by the patient, and this measurement error could then cause misclassification into an incorrect category.

Prognosis research is becoming increasingly more important [1], but there has been little research into the impact that measurement error in the predictors used to develop a prediction model may have, both in terms of the predictions made and model performance. It is also unclear how accounting for measurement error within the statistical modeling may improve this. A recent study demonstrated that measurement error in the predictors can dramatically reduce the c-statistic and increase the Brier score [13], and another study found that both random and systematic errors in self-reported health data influence the calibration, discrimination, and predicted risks [14], but in general, the extent and impact of measurement error in prediction model research is often overlooked. However, the STRengthening Analytical Thinking for Observational Studies initiative (www.stratos-initiative.org) has identified measurement error as a common issue in observational studies, which is often ignored and for which guidance is needed. There is a vast amount of literature on the statistical effect of measurement error in general, but whether investigators consider measurement error when developing a prediction model has not previously been evaluated. Models developed with predictors containing measurement error could therefore provide inaccurate estimates of patient risk and the model may not perform as well as expected in practice. A summary of the most commonly used methods to correct for measurement error is given by Brakenhoff et al. [11] with more detailed reviews of these (and other) methods given by Caroll et al. [8] and Gustafson [9]. Several other methods that can be used to account for measurement error in the particular context of prediction research have been developed, including methods in a Bayesian framework, using an item response theory model to handle the measurement error [15] and bootstrap regression calibration [16], based on resampling techniques.

A particular aspect of measurement error in the predictors is timing error. This refers to whether the predictors used in the model development were measured at the moment the model is intended to be used in practice. When time-dependent predictors are not able to be measured at "baseline," this creates time-dependent bias, which has been shown to often have an impact on the estimates of key predictors and study conclusions [17]. In addition, the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis statement recommends to clearly define when the predictors used in the development of the model were measured [18] and states that "all predictors should be measured before or at the study time origin and known at the intended moment the model is intended to be used" [19]. Nevertheless, for a range of practical and ethical reasons, researchers may design prognosis studies that collect time-varying predictor information after the intended moment of use, which itself may lead to errors and misleading predictions [20].

The aim of this article is to present a systematic review of recent studies developing prediction models, to ascertain how susceptible to measurement error the predictors used in the final models are and how often the measurement error was acknowledged or accounted for within the development of the models. A secondary objective is to determine whether the predictors were measured at a different time point to the intended moment of using the prediction model.

## 2. Methods

### 2.1. Data source and search

A systematic search was carried out in Medline on 27th November 2015 to identify the 30 most recent articles reporting the development of a multivariable prediction model for either individualized diagnosis or prognosis. It