Measurement error in clinical research, yes it matters

Groenwold, R.H.H.; Dekkers, O.M.

Note: To cite this publication please use the final published version (if applicable).

Measurement error in clinical research, yes it matters

Rolf H H Groenwold[1,2] and **Olaf M Dekkers** [iD][1,3]

[1]Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands, [2]Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, the Netherlands, and [3]Department of Endocrinology, Leiden University Medical Center, Leiden, the Netherlands

Correspondence should be addressed to R H Groenwold
Email
R.H.H.Groenwold@lumc.nl

Abstract

The validity of any biomedical study is potentially affected by measurement error or misclassification. It can affect different variables included in a statistical analysis, such as the exposure, the outcome, and confounders, and can result in an overestimation as well as in an underestimation of the relation under investigation. We discuss various aspects of measurement error and argue that often an in-depth discussion is needed to appropriately assess the quality and validity of a study.

Introduction

Measurement error (or misclassification) potentially threatens the validity of medical studies. Many variables can be affected by measurement errors, which is especially the case for variables that are difficult to measure and classify (e.g. exposure to air pollution, environmental chemicals) or that rely on self-reporting (weight, smoking, alcohol intake). Nevertheless, in the discussion section of research articles, authors often make rather uninformative claims that measurement error probably did not result in any substantial bias; a claim that is often not supported by quantitative arguments (1). Here we will provide an introduction to measurement error, using various examples from endocrinological research, and argue that a more thorough discussion of it is often needed to be able to fully assess the quality of published results.

A few remarks on terminology

The term measurement error is often used to denote the situation where a continuous variable (e.g. HbA1c or BMI) is not perfectly measured. Measurement error of categorical variables is called misclassification, for

example, when a Cushing patient is incorrectly classified as a non-Cushing patient. Mind that measurement error and misclassification are not separate issues, especially when categorization (obesity yes/no) is based on error-prone continuous variables (self-reported weight). Random measurement error means that the error does not depend on other variables or on the true value of the variable itself and – on average – does not affect the value of the intended measurement; an example would be blood pressure measurement, where direction and magnitude of the error is probably not related to the true blood pressure (high or low) and also not related to other variables like BMI. In principle, measurement error and misclassification can affect all variables measured in a study, including the exposure (e.g. risk factor or treatment), the outcome, and confounding variables.

Bias due to measurement error

Consider an observational cohort study to assess the relation between BMI and the risk of type 2 diabetes mellitus, the latter being measured by means of HbA1c

Published by Bioscientifica Ltd.

European Journal of Endocrinology

levels. In this study, measurements of BMI may be based, for convenience, on self-reporting instead of a standardized measurement using scales. Note that the intended measurement according to a fixed protocol (BMI measurement using scales) need not correspond perfectly to the 'true' value of BMI. Nevertheless, the intended measurement is often referred to as the error-free measurement, whereas the measurement contaminated with error is referred to as the error-prone measurement.

The illustration in Fig. 1 shows the estimated relation between the error-free (standardized) measurements of BMI and HbA1c levels (grey points, grey line) and between the error-prone (self-reported) values of BMI and HbA1c levels (black points, black line). In this example, the relation between BMI and HbA1c is attenuated due to measurement error of the independent variable (BMI).
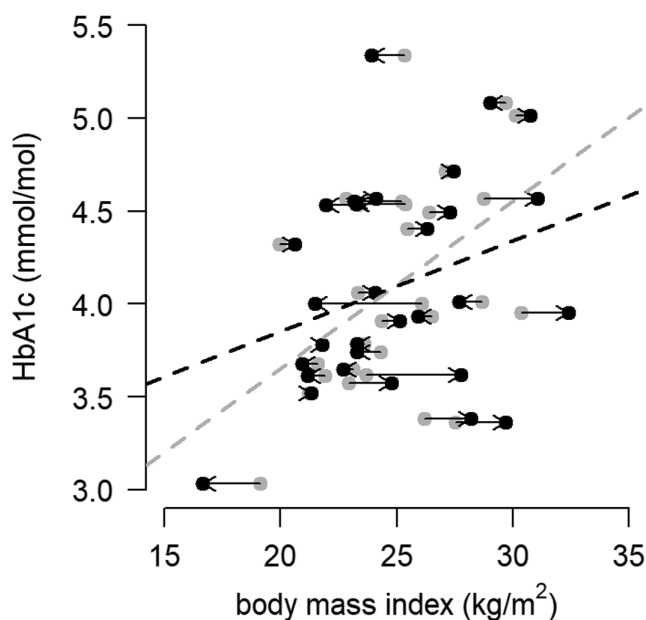

Figure 1

Relation between BMI and level of HbA1c in a hypothetical study of 30 subjects. Grey points represent observations based on standardized measurements of BMI. The slope of the grey line represents the estimated relation between BMI and HbA1c; it was estimated that BMI increases HbA1c by 0.09 mmol/mol (95%CI: 0.02–0.15) per unit increase in BMI. Black points represent observations based on self-reported values of BMI. The slope of the black line represents the estimated relation between BMI and HbA1c; it was estimated that BMI increases HbA1c by 0.05 mmol/mol (95%CI: −0.01 to 0.11) per unit increase in BMI. The arrows indicate the change in BMI value due to measurement error.

It shows that the true association is underestimated by almost 50% due to measurement error.

Although in this example measurement error results in an underestimation of the relation, such underestimation cannot be assumed in general (2). Measurement error can also lead to an overestimation of the relation, even if the misclassification is random. This is important, as it means that a remark in the discussion that 'any remaining bias due to measurement error will have caused an underestimation of the observed effect' is not necessarily true.

In a large cohort study, the relation between BMI and mortality was investigated. The cohort data contained information on both measured and self-reported BMI, which was categorized. Using self-reported data would have introduced misclassification in 36% (3); this misclassification occurred in both directions, resulting in both over and underestimations of the measured BMI. When comparing mortality risk between BMI categories, mortality was higher for overweight persons (BMI: 25–29.9) compared to the reference category (BMI: 22.6–24.9) when the analysis was based on self-reported data (hazard ratio: 1.24). However, when the analysis was based on measured BMI, the apparent increased risk disappeared (hazard ratio: 0.85), underlining that measurement error can overestimate an observed association.

The problem gets more complex, when confounders are also measured with error. In the example mentioned previously (BMI and HbA1c), one could consider lifestyle factors (diet, physical activity) as potential confounders; yet both variables are likely measured with error. Adding these two variables to the statistical model will result in incomplete adjustment for confounding. Therefore, also measurement error of confounders could lead to either an overestimation or to an underestimation of the relation of interest (4).

The direction and the magnitude of the bias due to measurement error depend on the discrepancy between the intended (error-free) measurement and the measurement that was made and used for the analysis (error prone). But, how measurement error plays out also depends on whether the error is made in an independent variable (exposure or covariate) or in the dependent (outcome) variable. In the example mentioned previously, the association between BMI and HbA1c was attenuated due to measurement error in the exposure (BMI). However, random measurement error in a continuous outcome variable in a randomised trial will generally not result in a bias of the estimated treatment effect (although it will lead to lower power). However, misclassification of treatment status will often

European Journal of Endocrinology

result in a bias of the treatment effect, even in otherwise well-conducted randomised trials. Another relevant aspect to consider is whether the error is dependent on other variables or other errors being made (5).

Measurement error may not only lead to bias and/or loss of power but also mask certain patterns in the data. For example, precise rhythmic patterns of cortisol secretion may go unnoticed if cortisol levels are measured imprecisely.

Various methods have been proposed to correct for measurement error, all of which require some information on the extent and direction of the errors. Such information may come from, for example, a subsample of a study, in which error-prone measurements are complemented by error-free measurements or from an external validation study. For an overview of the theory on measurement error and methods to correct for measurement error, we refer to the literature (5).

Concluding remarks

Measurement error is a potential source of bias that can occur in all types of clinical studies. However, because in randomised trials often many precautions are taken to prevent measurements are made with error, the impact in these studies probably is smaller than in observational studies based on, for example, routinely collected electronic health records data. Nevertheless, even randomised trials are not immune to bias from measurement error and misclassification.

We reiterate that measurement error can result in an underestimation as well as in an overestimation of the relation being investigated. Therefore, making unsubstantiated claims that measurement error did not materially affect the results of a particular study is jumping to conclusions. Instead, well-informed sensitivity analysis can be helpful to quantify the possible impact of this source of bias (6). The responsibility for this lies with the researchers; after all, they are best informed about the nature of potential measurement error in their study.

References

1 Brakenhoff TB, Mitroiu M, Keogh RH, Moons KGM, Groenwold RHH & van Smeden M. Measurement error is often neglected in medical literature: a systematic review. *Journal of Clinical Epidemiology* 2018 **98** 89–97. (https://doi.org/10.1016/j.jclinepi.2018.02.023)

2 van Smeden M, Lash TL & Groenwold RHH. Reflection on modern methods: five myths about measurement error in epidemiological research. *International Journal of Epidemiology* 2020 **49** 338–347. (https://doi.org/10.1093/ije/dyz251)

3 Flegal KM, Kit BK & Graubard BI. Bias in hazard ratios arising from misclassification according to self-reported weight and height in observational studies of body mass index and mortality. *American Journal of Epidemiology* 2018 **187** 125–134. (https://doi.org/10.1093/aje/kwx193)

4 Brakenhoff TB, van Smeden M, Visseren FLJ & Groenwold RHH. Random measurement error: why worry? An example of cardiovascular risk factors. *PLoS ONE* 2018 **13** e0192298. (https://doi.org/10.1371/journal.pone.0192298)

5 Keogh RH, Shaw PA, Gustafson P, Carroll RJ, Deffner V, Dodd KW, Küchenhoff H, Tooze JA, Wallace MP, Kipnis V *et al*. Stratos guidance document on measurement error and misclassification of variables in observational epidemiology: part 1-basic theory and simple methods of adjustment. *Statistics in Medicine* 2020 **39** 2197–2231. (https://doi.org/10.1002/sim.8532)

6 Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC & Greenland S. Good practices for quantitative bias analysis. *International Journal of Epidemiology* 2014 **43** 1969–1985. (https://doi.org/10.1093/ije/dyu149)