Correction methods for measurement error in epidemiologic research

Nab, L.

1

General introduction and outline of the thesis

1.1. Introduction

Measurement error affects the validity of many epidemiologic studies, that often rely on imperfect data [1]. Epidemiologic studies may for example rely on data obtained from electronic health records. These records are retrieved for other purposes than epidemiologic precision and may therefore be more subject to measurement error than data retrieved for answering specific research questions. Another example of error-prone data includes data collection based on self-reports by study participants [2]. Self-reports may come with (selective) reporting and recollection biases [3]. The inability to accurately measure variables of interest in epidemiologic research studies may result in failure to observe associations between a certain exposure and health outcome [4], or oppositely, the observation of spurious associations [5].

Epidemiologic studies often rely on the salient assumption of no measurement error. This assumption may be satisfied for some variables (e.g., age in years) but much harder to justify for others, such as variables subject to natural variation (e.g., blood pressure) [6] or laboratory error (e.g., Inhibin B) [7]. As an example, Figure 1.1 illustrates the discrepancy between two consecutive measurements of systolic blood pressure in the National Health and Nutrition Examination Survey (NHANES) [8].

Other epidemiologic studies may rely on self-reported measures, such as self-reported length or weight [9], physical activity [10] or diet [11]. A self-reported measure tends to be prone to error and generally does not perfectly correlate with the phenomenon it aims to measure. In Figure 1.2 it is illustrated that in the NHANES [8] self-reported weight was not perfectly correlated with weight measured by trained health technicians with a calibrated weight scale.

When measurement error is not accounted for in the design or the analysis of an epidemiologic study, measurement error can lead to considerable bias in exposure-outcome associations. The consequences of measurement error in exposure and outcome variables have been well established in the scientific literature [12–15]. The triple whammy of

measurement error describes the three consequences of measurement error: i) it may lead to bias in statistical parameter estimation, ii) it may lead to a loss of power, and iii) it may mask the functional form of a relationship between two variables [13]. For the first whammy, a common misconception is that the bias due to measurement error always attenuates exposure-outcome associations. This general statement can be true in case of random measurement error in the exposure, also known as 'classical' measurement error. For other forms of measurement error, e.g., systematic or differential measurement error, this simple heuristic may not apply [16].

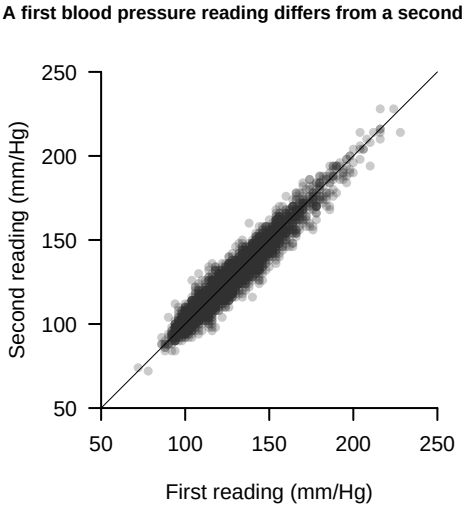

Figure 1.1: Discrepancy between two consecutive systolic blood pressure measurements in the National Health and Nutrition Examination Survey (NHANES) 2017-2018 cycle [8]

Various correction methods for measurement error exist. Examples include regression calibration [17, 18], simulation-extrapolation [19], moment reconstruction [20], non-parametric maximum likelihood estimation [21], imputation-based methods [22, 23], and Bayesian methods [15, 24]. Among these methods, regression calibration appears to be the one that is most commonly used in epidemiology [25, 26].

In spite of the abundance of literature on measurement error, and more specifically, on measurement error correction methods, correction for measurement error remains seldomly applied in epidemiologic research [25–27]. In most epidemiologic studies, the impact of measurement error is inadequately discussed [26] and often erroneously dismissed as leading to an underestimation of the exposure-outcome association [25, 26]. Importantly, this practice has not changed over the last decades [25–27]. This may, in part, be due to insufficient understanding of the impact of measurement error in settings that go beyond the classical example of attenuated exposure-outcome associations. An alternative explanation may be that researchers are unfamiliar with available measurement error correction methods and tools to quantitatively assess the impact of measurement error. In addition, researchers may not appreciate the added value of the collection of

(external) validation data for measurement error correction, hampering the inclusion of additional validation data within study designs when measurement error is suspected or anticipated.

The aim of this thesis is to improve the understanding of the impact of measurement error in epidemiologic studies, to facilitate the application of measurement error correction methods, to improve the design of epidemiologic studies when measurement error in a variable is suspected and to develop tools to quantitatively assess the impact of measurement error in epidemiologic studies.

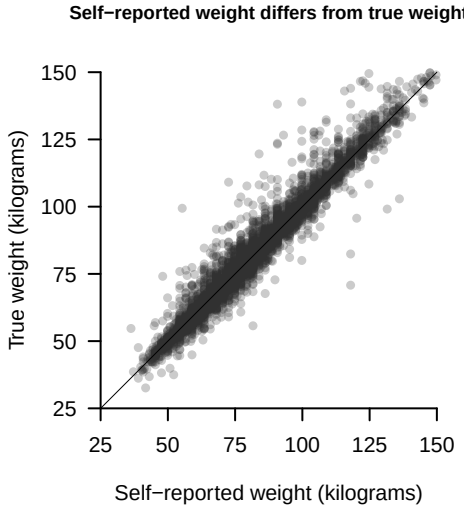

Figure 1.2: Discrepancy between weight in kilograms measured by health technicians using a calibrated weight scale (true weight) and self-reported weight in the National Health and Nutrition Examination Survey (NHANES) 2017-2018 cycle [8]

1.2. Outline

This thesis is organised as follows. To improve the understanding of the impact of measurement error, in Chapter 2, it is investigated how randomised controlled trials are affected by measurement error in a continuous endpoint. Three types of measurement error are distinguished, classical (or random) measurement error, systematic measurement error and differential measurement error.

To improve the application of measurement error correction methods, in Chapter 3 the `R` package `mecor` is described for measurement error correction in linear models with a continuous outcome. The `R` package `mecor` facilitates measurement error correction by means of regression calibration, method of moments and a maximum likelihood-based method. Information about the measurement error model and its parameters can be obtained from four types of validation studies: internal validation, replicates, calibration and external validation data. Each of these are discussed in detail. Chapter 4 provides

an exploration of the bias–variance trade off for the regression calibration estimator implemented in `mecor`, and an investigation of the performance of the estimator in settings where measurement error is relatively large.

To improve the design of epidemiologic studies when measurement error is suspected, guidance is provided for the collection of validation data needed for measurement error correction in Chapter 5. Here, sampling methods for validation data are studied and the assumptions required for the correct application of regression calibration for measurement error correction investigated. Deterministic and non-deterministic methods for validation data sampling are compared in terms of statistical efficiency. Next, in Chapter 6 reporting guidelines are proposed for studies on venous thromboembolism incidence in Corona disease patients. These studies on incidence report highly heterogeneous results. Different clinical and methodological sources of this heterogeneity are identified, including misclassification error in the diagnosis of venous thromboembolism and overall data quality. The proposed reporting guidelines guide future studies on venous thromboembolism incidence.

To quantitatively assess the impact of measurement error in the absence of validation data, sensitivity analysis or quantitative bias analysis could be used. In Chapter 7, two methods, regression calibration and simulation-extrapolation are compared for a sensitivity analysis for random exposure measurement error. In Chapter 8, a quantitative bias analysis for confounder misclassification is proposed. The quantitative bias analysis approach is described for traditional conditional regression and marginal structural models estimated using inverse probability weighting. This thesis ends with a general discussion including recommendations and directions for future research in Chapter 9.

References

[1] K. B. Michels, A renaissance for measurement error, International Journal of Epidemiology 30 (3) (2001) 421–422. doi:10.1093/ije/30.3.421.

[2] J. R. Marshall, Commentary: About that measurement problem, International Journal of Epidemiology 34 (6) (2005) 1376–1377. doi:10.1093/ije/dyi228.

[3] T. Johnson, M. Fendrich, Modeling sources of self-report bias in a survey of drug use epidemiology, Annals of Epidemiology 15 (5) (2005) 381–389. doi:10.1016/j.annepidem.2004.09.004.

[4] J. A. Hutcheon, A. Chiolero, J. A. Hanley, Random measurement error and regression dilution bias, BMJ 340 (c2289) (2010). doi:10.1136/bmj.c2289.

[5] J. R. Marshall, J. L. Hastrup, Mismeasurement and the resonance of strong confounders: Uncorrelated errors, American Journal of Epidemiology 143 (10) (1996) 1069–1078. doi:10.1093/oxfordjournals.aje.a008671.

[6] P. Burstyn, B. O'Donovan, I. Charlton, Blood pressure variability: The effects of repeated measurement, Postgraduate Medical Journal 57 (670) (1981) 488–491. doi:10.1136/pgmj.57.670.488.

[7] N. J. Perkins, J. Weck, S. L. Mumford, L. A. Sjaarda, E. M. Mitchell, A. Z. Pollack, E. F. Schisterman, Combining biomarker calibration data to reduce measurement error, Epidemiology 30 (2019) S3–S9. doi:10.1097/EDE.0000000000001094.

[8] Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS), National health and nutrition examination survey data (2017). URL https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2017

[9] D. Gunnell, L. Berney, P. Holland, M. Maynard, D. Blane, S. Frankel, G. D. Smith, How accurately are height, weight and leg length reported by the elderly, and how closely are they related to measurements recorded in childhood, International Journal of Epidemiology 29 (3) (2000) 456–464. doi:10.1093/intjepid/29.3.456.

[10] P. Ferrari, C. Friedenreich, C. E. Matthews, The role of measurement error in estimating levels of physical activity, American Journal of Epidemiology 166 (7) (2007) 832–840. doi:10.1093/aje/kwm148.

[11] W. Willett, Commentary: Dietary diaries versus food frequency questionnaires—A case of undigestible data, International Journal of Epidemiology 30 (2) (2001) 317–319. doi:10.1093/ije/30.2.317.

[12] J. Buonaccorsi, Measurement error: Models, methods, and applications, Chapman & Hall/CRC, Boca Raton, FL, 2010.

[13] R. Carroll, D. Ruppert, L. Stefanski, C. Crainiceanu, Measurement error in nonlinear models: A modern perspective, 2nd Edition, Chapman & Hall/CRC, Boca Raton, FL, 2006.

[14] W. Fuller, Measurement error models, John Wiley & Sons, New York, NY, 1987.

[15] P. Gustafson, Measurement error and misclassification in statistics and epidemiology: Impacts and Bayesian adjustments, Chapman & Hall/CRC, Boca Raton, FL, 2004.

[16] M. van Smeden, T. L. Lash, R. H. H. Groenwold, Reflection on modern methods: Five myths about measurement error in epidemiological research, International Journal of Epidemiology 49 (1) (2020) 338–347. doi:10.1093/ije/dyz251.

[17] R. J. Carroll, L. A. Stefanski, Approximate quasi-likelihood estimation in models with surrogate predictors, Journal of the American Statistical Association 85 (411) (1990) 652–663. doi:10.1080/01621459.1990.10474925.

[18] L. Gleser, Improvements of the naive approach to estimation in nonlinear errors-in-variables regression models, in: P. Brown, W. Fuller (Eds.), Statistical analysis of measurement error models, American Mathematics Society, Providence, 1990, pp. 99–114.

[19] J. R. Cook, L. A. Stefanski, Simulation-extrapolation estimation in parametric measurement error models, Journal of the American Statistical Association 89 (428) (1994) 1314–1328. doi:10.2307/2290994.

[20] L. S. Freedman, V. Fainberg, V. Kipnis, D. Midthune, R. J. Carroll, A new method for dealing with measurement error in explanatory variables of regression models, Biometrics 60 (1) (2004) 172–181. doi:10.1111/j.0006-341X.2004.00164.x.

[21] S. Rabe-Hesketh, A. Pickles, A. Skrondal, Correcting for covariate measurement error in logistic regression using nonparametric maximum likelihood estimation, Statistical Modelling 3 (3) (2003) 215–232. doi:10.1191/1471082X03st056oa.

[22] S. R. Cole, H. Chu, S. Greenland, Multiple-imputation for measurement-error correction, International Journal of Epidemiology 35 (4) (2006) 1074–1081. doi:10.1093/ije/dyl097.

[23] M. Blackwell, J. Honaker, G. King, A unified approach to measurement error and missing data: Overview and applications, Sociological Methods & Research 46 (3) (2017) 303–341. doi:10.1177/0049124115585360.

[24] J. W. Bartlett, R. H. Keogh, Bayesian correction for covariate measurement error: A frequentist evaluation and comparison with regression calibration, Statistical Methods in Medical Research 27 (6) (2018) 1695–1708. doi:10.1177/0962280216667764.

[25] T. B. Brakenhoff, M. Mitroiu, R. H. Keogh, K. G. M. Moons, R. H. H. Groenwold, M. van Smeden, Measurement error is often neglected in medical literature: A systematic review, Journal of Clinical Epidemiology 98 (2018) 89–97. doi:10.1016/j.jclinepi.2018.02.023.

[26] P. A. Shaw, V. Deffner, R. H. Keogh, J. A. Tooze, K. W. Dodd, H. Küchenhoff, V. Kipnis, L. S. Freedman, Epidemiologic analyses with error-prone exposures: Review of current practice and recommendations, Annals of Epidemiology 28 (11) (2018) 821–828. doi:10.1016/j.annepidem.2018.09.001.

[27] A. M. Jurek, G. Maldonado, S. Greenland, T. R. Church, Exposure-measurement error is frequently ignored when interpreting epidemiologic study results, European Journal of Epidemiology 21 (12) (2007) 871–876. doi:10.1007/s10654-006-9083-0.