Correction methods for measurement error in epidemiologic research

Nab, L.

9

Summary and general discussion

9.1. Summary

Measurement error is common in epidemiologic research and may affect the validity of research results. It is therefore important to scrutinise the effects of measurement error in epidemiologic research. Even simple forms of measurement error, for instance random measurement error in an exposure, can introduce bias in exposure-outcome associations. And even though there are situations in which measurement error does not introduce bias in the exposure-outcome association, for instance in case of random measurement error in a continuous outcome, it nearly always affects the precision and power of a study. In addition, other forms of measurement error, for example systematic measurement error or differential measurement error in an exposure, covariate or outcome, can affect exposure-outcome associations in complex ways that may not easily be anticipated. Adjusting for measurement error using measurement error correction methods may thus be necessary to obtain reliable estimates of exposure-outcome associations.

To facilitate measurement error correction, information about the underlying measurement error mechanism (i.e., model) and its parameters is needed. The measurement error model can sometimes be estimated from internal or external validation data, replicates data or calibration data. Collection and the use of such measurement error mechanism data will likely improve the quality of epidemiologic analyses in the presence of measurement error. This can be done through the application of measurement error correction methods, which adjust the analyses taking into account the information from the measurement error mechanism. Alternatively, in the absence of concrete data about the mechanisms or the parameters of measurement error, sensitivity analysis for measurement error can be used, in which the impact on the epidemiologic analyses of one or a range of hypothesized measurement error mechanisms or their parameters can be investigated.

The studies described in the thesis were set out to improve the *understanding* of the impact of measurement error, 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 research.

In Chapter 2, consequences were studied of measurement error in a continuous outcome in a randomized trial. Using an example of the efficacy of a low-dose iron supplement on haemoglobin levels in pregnant women, different forms of measurement error were discussed (i.e., random, systematic and differential measurement error). Using the example trial, it was shown that random measurement error in a trial outcome does not lead to bias in the effect estimator but can lead to a reduced precision and power. It was shown that systematic measurement error and differential measurement error in an outcome can lead to bias in the effect estimator and consequently, a null-hypothesis significance test for the treatment effect can deviate substantially from the nominal level. Subsequently, a regression calibration-like method was proposed to reduce bias in the treatment effect estimator and obtain confidence intervals with nominal coverage and tested in a Monte Carlo simulation study. The proposed method made use of external validation data to estimate the measurement error model and its parameters and four different methods for confidence interval construction were proposed. Different parameters for the measurement error model (i.e., systematic and differential measurement error) and explained variance of the measurement error model were tested. In our simulation study, it was shown that the regression calibration-like method was effective in improving trial inferences when an external validation dataset with at least 15 subjects was available.

In Chapter 3 the `R` package `mecor` for measurement error correction was introduced. The package facilitates measurement error correction in linear models with a continuous outcome if there is measurement error in the outcome or in a continuous covariate. The package accommodates measurement error correction methodology for a wide range of data structures: internal and external validation studies, replicates studies, and calibration studies. Various measurement error correction methods were implemented in the package: regression calibration, method of moments and correction based on maximum likelihood estimation. For standard error estimation and construction of confidence intervals, the delta method and bootstrap were implemented for all methods. The package also facilitates sensitivity analysis, when no data are available to estimate the parameters of the measurement error model. The package contains synthetic data based on examples from epidemiology following the structure of internal validation data, replicates data, calibration data and external validation data.

In Chapter 4 settings were studied in which application of regression calibration for exposure measurement error correction may not be appropriate. This was illustrated in a study of the association between active energy expenditure and lean body mass. A simulation study, based on the case study, showed that particularly in small samples the regression calibration estimator may be less efficient in terms of mean squared error than an estimator not correcting for the exposure measurement error. This phenomenon is an example of the commonly known bias–variance trade off. Particularly, when the measurement error is relatively large and sample sizes small, the simulation study showed that the performance of regression calibration was poor, indicated by biased estimates, large mean squared errors and large empirical standard errors in these settings.

In Chapter 5 three internal validation sampling strategies (i.e., random, stratified random and extremes sampling) were investigated in conjunction with regression calibration to correct for measurement error in a continuous covariate. This was illustrated in an example study of the investigation of the association between visceral adipose tissue and insulin resistance. The exposure measure visceral adipose tissue was only available in

40% of the population. Waist circumference was measured in all individuals and assumed an error-prone substitute measure of the reference measure visceral adipose tissue. In a setting where the reference measure is obtained in only 40% of the whole study, it was studied which individuals should be included in that subset and which not by means of Monte Carlo simulation. The simulation study showed a small efficiency gain in terms of mean squared error of stratified random and extremes sampling over a random sampling strategy for the internal validation restricted and regression calibration analyses, but only when measurement error was non-differential. For regression calibration, this gain in efficiency was at the cost of higher percentages bias and lower confidence interval coverage. It was therefore recommended that, in general, regression calibration using randomly sampled validation samples are preferred over stratified or extremes sampled samples.

The study described in Chapter 6 showed that studies on venous thromboembolism (VTE) incidence in Coronavirus disease 2019 (COVID-19) patients report highly heterogeneous results. Different sources of the observed heterogeneity were identified, notably, clinical and methodological sources, and illustrated using various examples. Clinical sources included the characteristics of study participants and testing for VTE. Methodological sources included inclusion types of the VTE endpoint, data quality and data analysis. Careful description was recommended of the elements that potentially affect VTE incidence and thus may cause heterogeneity in future VTE incidence studies and guidance was provided in the form of a list with reporting recommendations.

In Chapter 7 regression calibration and simulation-extrapolation were compared for sensitivity analysis for random measurement error in an exposure variable. These two random exposure measurement error correction methods were illustrated in two example studies. The first example study investigated the relation between the exposure blood pressure and , and the second example study investigated the relation between the exposure sodium intake and hypertension. These relations were modelled using linear and logistic regression, respectively. In both example studies the exposure variable was an error-prone version of an error-free exposure variable. Based on these two examples, a simulation study was conducted to study the relative performance of regression calibration and simulation-extrapolation in linear and logistic regression models. The simulation study showed that without extra data, but with correct assumptions about the variance of the measurement error, regression calibration was generally unbiased for linear and logistic regression, while simulation-extrapolation was biased. A small gain in efficiency in terms of mean squared error was seen for simulation-extrapolation in linear regression but not for logistic regression. The use of regression calibration for sensitivity analysis for random exposure measurement error was recommended and its use illustrated in the example study of the association between blood pressure and kidney function.

Inverse probability weighting and conditional models are both important and frequently used tools to adjust for confounding variables in observational studies. In Chapter 8, expressions were derived for the bias in the average treatment effect in a marginal structural model estimated using inverse probability weighting and a conditional model when a confounding variable is measured with error. Compared to bias in the average treatment effect estimator from a conditional model, the bias in a marginal structural model estimated using inverse probability weighting can be different in magnitude but is equal in sign. The derived bias expressions informed a quantitative bias analysis for bias due to a misclassified confounding variable. The use of a quantitative bias analysis was demonstrated in an

example study of the effect of using diuretics versus beta-blockers on blood pressure adjusted for the error-prone confounding variable self-reported body mass index category.

9.2. Discussion

This thesis provides an overview of correction methods for measurement error in epidemiologic research. The studies described in the thesis were set out to improve the *understanding* of the impact of measurement error and to facilitate the *application* of measurement error correction methods in epidemiologic studies. Guidance was provided to improve the *design* of epidemiologic studies when measurement error is suspected, and reporting guidelines proposed. All methods were demonstrated in case studies using empirical data (for an overview of case studies, see Table 9.1). Special attention was paid to sensitivity analysis for measurement error in settings where measurement error is suspected, but data about measurement error structure and its parameters, essential for measurement error correction methods, were not available. Here, we discuss the contribution of our work to this field and set out directions for future research.

9.2.1. Impact of measurement error in epidemiologic studies

The impact of measurement error often goes beyond the simple heuristic of 'attenuation to the null' [1]. This heuristic wrongfully suggests that estimates of effects in epidemiologic studies will only become smaller due to the measurement error. Unfortunately, this myth remains persistent despite a vast body of literature arguing against it [2–5]. Particularly, depending on the target of the analysis and the type of measurement error, the effects of measurement error can go in either direction and are therefore often unpredictable, as shown by Keogh et al. [6].

This thesis aimed at improving the *understanding* of the impact of measurement error in epidemiologic research. To evaluate the impact of measurement error in a specific study, four considerations are; i) what statistical model is used; ii) which of the variable(s) of the model is (are) error-prone and what is their role in the model; iii) what is the structure of the measurement error model; and iv) what are the parameters of the measurement error model (see Figure 9.1). All these components may affect *if* an epidemiologic study is affected by measurement error and if so, *how* an epidemiologic study is affected by measurement error. For example, random exposure measurement error introduces bias in the effect estimator of a linear regression model [4], and a logistic regression model [7] and leads to a so-called 'induced hazard function' for a Cox regression model [8]. In contrast, random measurement error in a continuous outcome does not introduce bias but reduces precision and power at a chosen sample size, and systematic and differential measurement error in such outcomes introduce bias in the effect estimator of a linear regression model that can go in either direction (Chapter 2). When exposure measurement error is suspected, restricting the analysis to the subset of individuals for whom the error-free exposure measurement is obtained, does not lead to biased inference. Yet, when that subset is sampled using information about an error-prone substitute exposure (e.g., when for all individuals exceeding a specific threshold of the substitute exposure, the error-free exposure is obtained), bias is introduced in the complete case analysis if the error in the substitute exposure is differential, but not if the error in the substitute exposure is non-differential (Chapter 5). When a confounding variable is misclassified,

marginal structural models estimated using inverse probability weighting were shown to be biased but affected differently than conditional models (Chapter 8). There are innumerable combinations of the considerations displayed in Figure 9.1 and, therefore, measurement error can affect estimated exposure-outcome associations in complex ways that may not easily be anticipated and need to be evaluated from one setting to another.

9.2.2. Measurement error correction methods in epidemiologic studies

There is an abundance of texts on measurement error correction methods [2–5]. Yet, correction methods remain seldomly applied in epidemiologic research [9–11]. Methods for measurement error corrections include, regression calibration [12, 13], simulation-extrapolation [14], moment reconstruction [15], non-parametric maximum likelihood estimation [16], imputation-based methods [17, 18] and Bayesian methods [5, 19]. Regression calibration is among the most commonly used methods in epidemiologic research [10, 11].

This thesis facilitated the *application* of measurement error correction in epidemiologic research with the development of the software package `mecor` for measurement error correction in linear models with a continuous outcome. In this software package for `R`, regression calibration [20], validation regression calibration, efficient regression calibration [21], method of moments [2] and maximum likelihood-based methods [22] were implemented for a wide range of validation data structures (Table 9.1). Notably, different methods for variance estimation of the corrected estimators were implemented in `mecor`. An informed choice for the variance estimation of the measurement error corrected estimators is important as was shown that the Zero Variance, Delta, Fieller and bootstrap methods had different performance in terms of coverage and average confidence interval width (Chapter 2 and 4). The methods implemented in `mecor` are consistent but not necessarily more statistically efficient than the uncorrected estimator nor unbiased. Particularly in small samples, the estimator not correcting for measurement error may be more efficient in terms of mean squared error compared to the regression calibration estimator (Chapter 4). A phenomenon referred to as the bias–variance trade off. Particularly when measurement error is relatively large, the performance of regression calibration can be poor in small samples, as was shown by high percentages bias and large mean squared errors in these settings. However, compared to regression calibration, the simulation-extrapolation estimator was even more prone to bias (Chapter 7). Regression calibration relies on the assumption of non-differential measurement error, and large biases can occur in the estimator if this assumption is not warranted, as was shown in Chapter 5. In conclusion, measurement error correction methods can correct for measurement error when extra data are available to estimate the measurement error model and its parameters provided sufficiently large sample size of the validation set and measurement error that is not extremely large. What constitutes 'sufficiently large' and 'not extremely large' will be study specific and can be informed by statistical simulation studies, as presented in Chapter 4.

9.2.3. Design of epidemiologic studies affected by measurement error

For measurement error correction, validation data are needed to estimate the measurement error model and its parameters. Collection of such data should preferably be included in the *design* of an epidemiologic study. Considerations include the data structure, size and the

Table 9.1: Empirical example case studies used for investigations of measurement error correction methods and sensitivity analysis in the thesis

Example (Chapter)	**Measurement Error**	**Data Structure**	**Availability in `mecor`**	**Solutions Investigated in the Thesis**
Low-dose iron suppletion and haemoglobin levels (Chapter 2)	The reference measure venous haemoglobin level was substituted by the error-prone measure of capillary haemoglobin level	External validation data	Available as dataset `haemoglobin` and `haemoglobin_ext`	• Method of moments (referred to as regression calibration in Chapter 2)
Dietary intervention and sodium intake (Chapter 3, based on an example from Keogh et al. [23])	Sodium intake measured in urine was assumed subject to random measurement error and repeatedly measured, and sodium intake measured by a questionnaire was assumed subject to systematic measurement error	Calibration data	Available as dataset `sodium`	• Method of moments • Efficient method of moments
Active energy expenditure and lean body mass (Chapter 4)	The reference measure active energy expenditure measured by Actiheart® was substituted by the error-prone measure based on the international physical activity questionnaire	Internal validation data	Not available	• Regression calibration
Visceral adipose tissue and insulin resistance (Chapter 5)	The reference measure visceral adipose tissue was substituted by the error-prone measure waist circumference in a subset of the study	Internal validation data	Available as dataset `vat`	• Regression calibration • Efficient regression calibration • Validation regression calibration • Internal validation restricted analysis
Blood pressure and kidney function (Chapter 7)	Blood pressure was assumed subject to random measurement error and measured repeatedly	Replicates data or sensitivity analysis	Available as dataset `bloodpressure`	• Regression calibration • Simulation-extrapolation Informed by the replicate measures or in a sensitivity
Sodium intake and Hypertension (Chapter 7)	Sodium intake was assumed subject to random measurement error and measured repeatedly	Replicates data or sensitivity analysis	Not available	• Regression calibration • Simulation-extrapolation Informed by the replicate measures or in a sensitivity analysis
Diuretics vs beta-blocker use and systolic blood pressure (Chapter 8)	The confounding variable self-reported body mass index was assumed subject to classification error	Sensitivity analysis	Not available	• Sensitivity analysis informed by the bias expressions derived in Chapter 8 visualised in a Shiny application

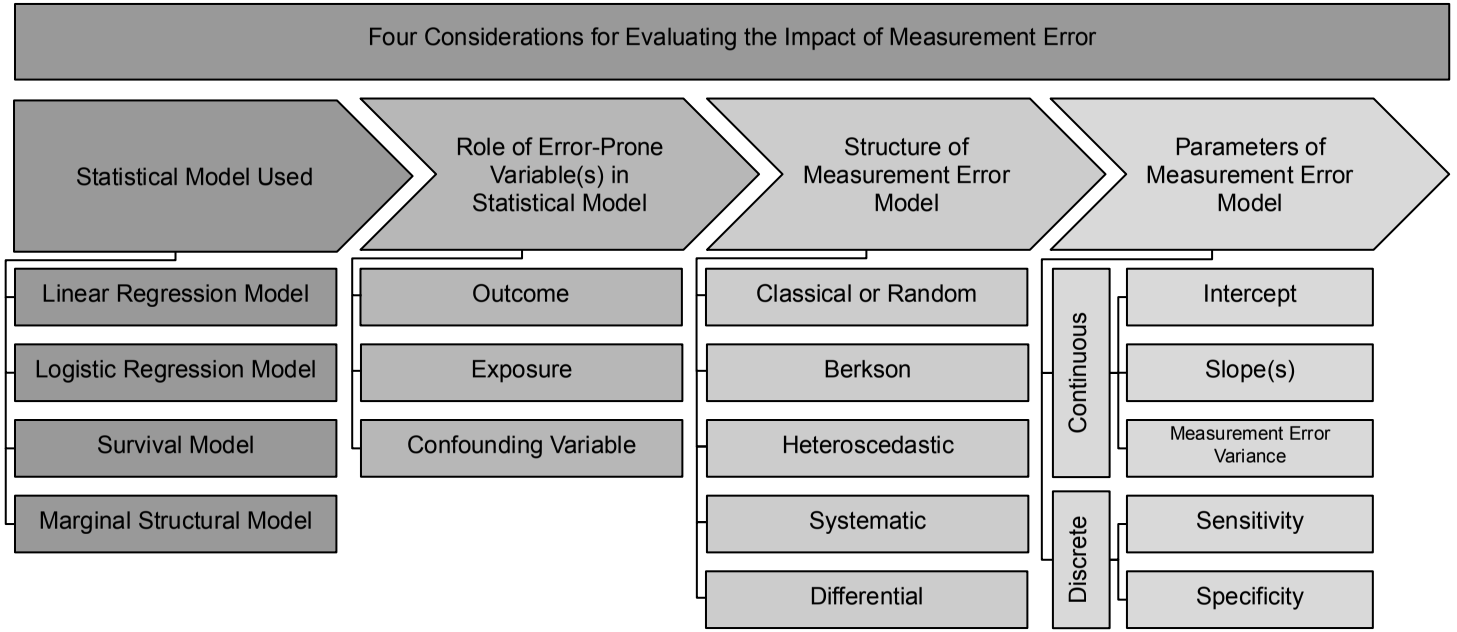

Figure 9.1: Four considerations for evaluating the impact of measurement error in an epidemiologic study. Focus lies on statistical models commonly used in epidemiologic research, measurement error model structures generally distinguished in the measurement error literature and, prevailing operationalisations of the parameters of the measurement error model.

sampling strategy of the validation data. For such considerations, the different components shown in Figure 9.1 needed to evaluate the impact of measurement error need to be taken into account. Particularly, certain structures of validation data (internal, external, replicates data or calibration data) are not suited for certain measurement error structures (e.g., replicates data can only be used if random measurement error is suspected). After making assumptions about the measurement error model structure and its parameters and deciding what type of validation data is suited, Monte Carlo simulation can be used to inform sample size and sampling strategy of the validation data. An example of a Monte Carlo simulation study to examine the optimal sampling strategy of an internal validation data set in the Netherlands Epidemiology of Obesity study [24] was described in Chapter 5. Here, sampling the extremes or stratified randomly showed a small gain in efficiency, but at the cost of bias and confidence interval coverage and should only be used when measurement error is strictly non-differential. A difficulty here is, however, that in studies like the Netherlands Epidemiology of Obesity study, the first two components that influence the impact of measurement error (described in first two columns in Figure 9.1) may differ across studies. Specifically, a variable can be an outcome in one study and an exposure in another study.

9.2.4. Sensitivity analysis for measurement error in epidemiologic studies

In epidemiologic research, it is commonly assumed (often implicitly) that all variables are measured without error; an assumption that is often not justified. Yet, when measurement error is suspected or anticipated, methods to correct for the measurement error rely on the availability of data on the measurement error mechanisms and parameters. Such data may not be available, maybe incomplete or be itself unreliable, in which case sensitivity analysis for measurement error can help to assess the sensitivity of research results to measurement error. In epidemiology, a sensitivity analysis may alternatively be referred to as quantitative bias analysis [25].

Sensitivity analysis for measurement error should be included in study protocols and valued independent of the outcome of the sensitivity analysis (i.e., results should not only be shown if the sensitivity analysis shows research results are *not* sensitive to the assumption of no measurement error). Sensitivity analysis can be informed by expert knowledge about the structure of the measurement error model and its parameters. Distributions of these parameters can be used to put more weight on the assumed most plausible values [25].

The sensitivity of research results to random exposure measurement error can be checked using regression calibration or simulation-extrapolation, of which regression calibration was shown most suited in Chapter 7. Graphical presentation of the results of a sensitivity analysis allows readers to judge the sensitivity of research results for the whole distribution of assumed parameters of the measurement error model, and may be preferred over a single summary number (see for example Figure 7.9 in Chapter 7). Alternatively, interactive *tools* may be designed to allow readers to test the sensitivity of research results to their own assumed parameters of the measurement error model, as was facilitated by the Shiny application demonstrated in Chapter 8.

9.2.5. Future research

The studies presented in the thesis aimed to improve the (application of) methods to limit the impact of measurement error in epidemiologic research. The application of

measurement error correction methods was facilitated through the development of the `R` package `mecor`. To aid the application of measurement error correction methods in epidemiology, numerous methods were illustrated in empirical data (Table 9.1). Extensive Monte Carlo studies were set up to study the performance of measurement error correction methods in epidemiologic studies based on the empirical data and have been made publicly available. The simulation code can easily be adapted by researchers to settings of intended use to improve the design and statistical analysis of epidemiologic studies when measurement error is suspected. However, we are not there yet. There are several topics that require future research to further develop the field of measurement error methodology.

First, the main focus of this thesis was on linear models with a continuous outcome and measurement error in one of the continuous variables of those models. In epidemiologic studies, measurement error may, however, be anticipated in more than one variable. In addition, other statistical models (e.g., logistic and survival analysis) are commonly used in epidemiologic research. For models with binary outcomes, the impact of covariate measurement error and classification error in the binary outcome has been studied by Carroll et al. in [7] and [26], respectively. Also, correction methods have been proposed for situations where one or multiple variables in a logistic regression model are measured with error [20]. For survival outcomes, the impact of covariate measurement error has been studied by Prentice et al. [8] and an investigation of measurement errors in the failure time outcome and correction methods for this setting were examined by Oh et al. [27]. Yet, the implications of a combination of complex forms of outcome measurement error and covariate measurement error need further study.

Second, this thesis only investigated the use of parametric measurement error models and it was generally assumed that the measurement error model was well specified. Future research may examine methods to test for the structure of the measurement error model in empirical data and study the impact of misspecification of the measurement error model structure on measurement error correction methods.

Third, the validation data structures discussed in the thesis that aid measurement error correction methods rely on certain assumptions. For an external data set, it is assumed that the measurement error model and its parameters are transportable from the main study to the external study. For a replicates study, it is assumed that measurement error in the subsequent replicate measurements is independent. Investigations are needed if information about the reliability of e.g., biomarkers can be transported to studies where these biomarkers are used and if the assumption of independent measurement error in such biomarkers is warranted.

Fourth, this thesis presents measurement error correction methods for measures of which a clear concept about the 'true' measure of a variable is needed and is in most instances assumed observable (except when random measurement is assumed in which case repeated measures of the error-prone measure are adequate). This assumption might be reasonable and applicable for measures such as an individual's weight in kilo grams or blood pressure, but may be difficult or even impossible to establish for constructs such as patient well-being or pain [28]. Future research may pay specific attention to the applicability of latent class analysis for the analysis of error-prone epidemiologic data, which does not rely on the assumption of observable 'true' measures. Instead, it is assumed that the true variable can be estimated by combining multiple imperfect measurements of the variable. These methods are widespread in psychology and the social sciences [29],

but received relatively little attention in epidemiologic research (exceptions include e.g., [30, 31]).

9.2.6. Conclusion

Measurement error in epidemiologic research is not uncommon and can hamper the validity of research results if ignored. The old saying "to prevent is better than to cure" also applies here, and therefore actions to improve the overall quality of measurement in epidemiologic analyses are likely to have a larger effect on the validity of epidemiologic studies than widespread application of measurement error correction methods. However, in settings where measurement error cannot be avoided, measurement error correction methods and sensitivity analysis for measurement error provide tools to correct for or quantitatively assess the impact of measurement error. In combination with reliable information about the measurement error model and its parameters, these methods can help to estimate relevant epidemiologic parameters that are more reliable than what would be obtained if estimated without taking account of possible measurement error.

References

[1] 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.

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

[3] 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.

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

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

[6] R. H. Keogh, P. A. Shaw, P. Gustafson, R. J. Carroll, V. Deffner, K. W. Dodd, H. Küchenhoff, J. A. Tooze, M. P. Wallace, V. Kipnis, L. S. Freedman, 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 39 (16) (2020) 2197–2231. doi:10.1002/sim.8532.

[7] R. J. Carroll, C. H. Spiegelman, K. K. G. Lan, K. T. Bailey, R. D. Abbott, On errors-in-variables for binary regression models, Biometrika 71 (1) (1984) 19. doi:10.2307/2336392.

[8] R. L. Prentice, Covariate measurement errors and parameter estimation in a failure time regression model, Biometrika 69 (2) (1982) 331–342. doi:10.2307/2335407.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] B. Rosner, D. Spiegelman, W. C. Willett, Correction of logistic regression relative risk estimates and confidence intervals for measurement error: The case of multiple covariates measured with error, American Journal of Epidemiology 132 (4) (1990) 734–745. doi:10.1093/oxfordjournals.aje.a115715.

[21] D. Spiegelman, R. J. Carroll, V. Kipnis, Efficient regression calibration for logistic regression in main study/internal validation study designs with an imperfect reference instrument, Statistics in Medicine 20 (1) (2001) 139–160. doi:10.1002/1097-0258(20010115)20:1<139::AID-SIM644>3.0.CO;2-K.

[22] J. W. Bartlett, B. L. De Stavola, C. Frost, Linear mixed models for replication data to efficiently allow for covariate measurement error, Statistics in Medicine 28 (25) (2009) 3158–3178. doi:10.1002/sim.3713.

[23] R. H. Keogh, R. J. Carroll, J. A. Tooze, S. I. Kirkpatrick, L. S. Freedman, Statistical issues related to dietary intake as the response variable in intervention trials, Statistics in Medicine 35 (25) (2016) 4493–4508. doi:10.1002/sim.7011.

[24] R. de Mutsert, M. den Heijer, T. J. Rabelink, J. W. A. Smit, J. A. Romijn, J. W. Jukema, A. de Roos, C. M. Cobbaert, M. Kloppenburg, S. le Cessie, S. Middeldorp,

F. R. Rosendaal, The Netherlands epidemiology of obesity (NEO) study: Study design and data collection, European Journal of Epidemiology 28 (6) (2013) 513–523. doi:10.1007/s10654-013-9801-3.

[25] T. Lash, M. Fox, A. Fink, Applying quantitative bias analysis to epidemiologic data, Springer, New York, NY, 2009.

[26] R. J. Carroll, D. Ruppert, L. A. Stefanski, C. M. Crainiceanu, Logistic regression with response error, in: Measurement error in nonlinear models, 2nd Edition, Chapman & Hall/CRC, Boca Raton, FL, 2006, Ch. 15, pp. 345–352.

[27] E. J. Oh, B. E. Shepherd, T. Lumley, P. A. Shaw, Considerations for analysis of time-to-event outcomes measured with error: Bias and correction with SIMEX, Statistics in Medicine 37 (8) (2018) 1276–1289. doi:10.1002/sim.7554.

[28] D. J. Hand, Measurement: A very short introduction, Oxford University Press, New York, NY, 2016.

[29] K. A. Bollen, Latent variables in psychology and the social sciences, Annual Review of Psychology 53 (1) (2002) 605–634. doi:10.1146/annurev.psych.53.100901.135239.

[30] J. Kaldor, D. Clayton, Latent class analysis in chronic disease epidemiology, Statistics in Medicine 4 (3) (1985) 327–335. doi:10.1002/sim.4780040312.

[31] M. van Smeden, C. A. Naaktgeboren, J. B. Reitsma, K. G. M. Moons, J. A. H. de Groot, Latent class models in diagnostic studies when there is no reference standard: A systematic review, American Journal of Epidemiology 179 (4) (2014) 423–431. doi:10.1093/aje/kwt286.