UMC Utrecht
Julius Center

Re CoD ID

Adjusting for exposure misclassification in an individual participant data meta-analysis of observational studies

Valentijn M.T. de Jong*,
Harlan Campbell,
Thomas Jaenisch,
Paul Gustafson,
Thomas P.A. Debray.

*Julius Center for Health Sciences and Primary Care,
University Medical Center Utrecht, The Netherlands

V.M.T.deJong-2@umcutrecht.nl

Measurement error

- Measurement error may bias estimates of parameters, even when the error is entirely random and independent. [Keys and Kihlberg, 1963, Carroll et al., 2006]
- Attenuation is only guaranteed to occur when
  - the misclassification is non-differential [Bross, 1954, Keys and Kihlberg, 1963]
  - the exposure has no more than two categories [Dosemeci et al., 1990, Birkett, 1992]
  - and all covariates are measured without error. [Carroll et al., 2006]
- Bias is difficult to quantify when covariate is measured with error. [Carroll et al., 2006]
- Extreme misclassification can reverse the sign of the observed association. [Weinberg et al., 1994]

Measurement error in IPD-MA

- In an individual participant data meta-analysis (IPD-MA), misclassification may be present in one or more studies.
  - For instance, a different (e.g. less accurate) measurement instrument for a certain exposure variable may have been used in some studies.
  - If one of these instruments is prone to misclassification, this will result in a biased estimate for the corresponding exposure's effect.
  - If different measurements methods are used across studies, this directly implies that a gold standard measurement are missing for entire studies.
- In IPD-MA, methods must also account for the effects of clustering in individual studies [Abo-Zaid et al., 2013] and should allow for heterogeneity of the effect of interest.

Measurement error in IPD-MA

- Traditional methods for misclassification in IPD-MA assume that the misclassification rate is transportable to other studies.
  - tenable when the measurement instruments, protocol, population and setting are the same in the included studies,
  - rare in the context of IPD-MA.
- We should apply a method that accounts for possible heterogeneity across studies in:
  - misclassification
  - exposure prevalence
  - outcome prevalence
  - exposure-outcome association

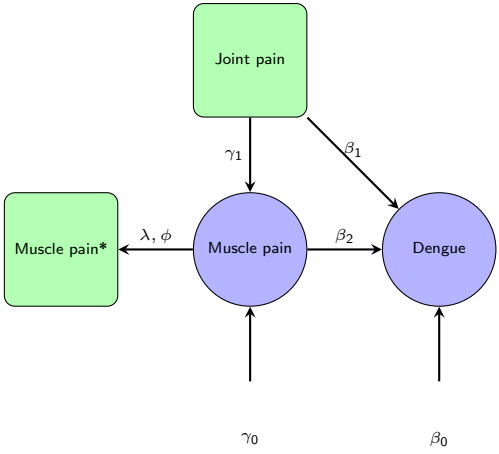
Fictional clinical IPD-MA example

- Outcome: presence of dengue virus (vs other febrile illness), always correctly classified.
- Exposure: muscle pain. Two measurement methods:
  - Always observed but may be misclassified,
  - Missing in some studies, but correctly classified when observed.
- Covariate: joint pain, always correctly classified.
- Source: simulated data based on real data gathered in three studies of the IDAMS consortium*

*Jaenisch T, Tam DTH, Kieu NTT, Van Ngoc T, Nam NT, Van Kinh N, et al. Clinical evaluation of dengue and identification of risk factors for severe disease: protocol for a multicentre study in 8 countries. BMC infectious diseases. 2016;16(1):120.

Fictional clinical IPD-MA example

- Green squares: fully observed data
- Blue circles: at least partially observed data
- Variances are omitted



Bayesian misclassification model

To account for all sources of heterogeneity and error, we fit three sub-models*:

- Outcome model
- Exposure model
- Measurement model

*Richardson S, Gilks WR. A Bayesian approach to measurement error problems in epidemiology using conditional independence models. American Journal of Epidemiology. 1993;138(6):430–442.

Submodel 1: outcome model

To allow the prevalence of the outcome to vary by study, we apply random intercepts [Abo-Zaid et al., 2013] and to account for heterogeneity in the exposure-outcome association, we apply random effects for the outcome [DerSimonian and Laird, 1986]:

$$\begin{aligned} y_{ij} &\sim \text{Bernoulli}(\pi_{ij}), \\ g(\pi_{ij}) &= \beta_{00} + \beta_{0j} + \beta_1 z_{ij} + \beta_{20} x_{ij} + \beta_{2j} x_{ij}, \end{aligned} \tag{1}$$

where g(.) is a link function, and
we applied Normal priors for the coefficients, and inverse Gamma priors for variances.

Submodel 2: Exposure model

We may model the varying prevalence of the gold standard measurement of the exposure x_{ij} by applying random intercepts and a covariate effect.

$$\begin{aligned} x_{ij} &\sim \text{Bernoulli}(p_{ij}), \\ g(p_{ij}) &= \gamma_{00} + \gamma_{0j} + \gamma_1 z_{ij}, \end{aligned} \tag{2}$$

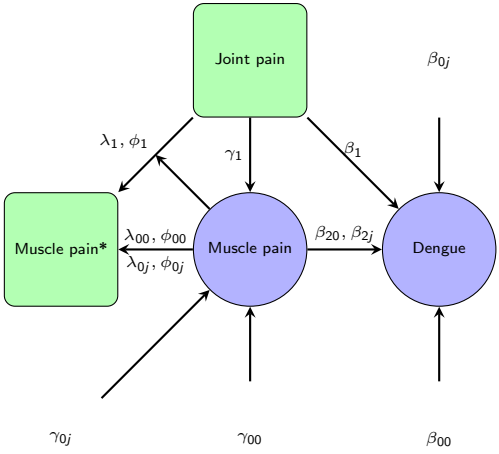
Submodel 3: measurement model

Participant-specific misclassification will lead to heterogeneity of the exposure-outcome association if the case-mix distributions vary across studies, as estimates of exposure-outcome associations will then be affected differently across studies.

$$
\begin{aligned}
x_{ij}^{*} &\sim \text{Bernoulli}(p_{ij}^{*}), \\
g(p_{ij}^{*}) &= (\lambda_{00} + \lambda_{0j} + \lambda_{1} z_{ij}) \text{ if } x_{ij} = 1, \\
&\quad (\phi_{00} + \phi_{0j} + \phi_{1} z_{ij}) \text{ if } x_{ij} = 0,
\end{aligned}
\tag{3}
$$

Diagram of the model

- Green squares: fully observed data
- Blue circles: at least partially observed data
- Variances are omitted



ReCoDID

Accounting for differential error

So far we have assumed the error in the measurement of the exposure is non-differential, that is that conditional on the gold standard measurement of the value of the exposure and on the perfectly measured covariates, the error in the measurement is unrelated to the outcome. In any other case the error is differential.

$$\begin{aligned} x_{ij}^* &\sim \text{Bernoulli}(p_{ij}^*), \\ g(p_{ij}^*) &= \lambda_{00} + \lambda_{0j} + \lambda_1 z_{ij} + \lambda_2 y_{ij} \text{ if } x_{ij} = 1, \\ g(p_{ij}^*) &= \phi_{00} + \phi_{0j} + \phi_1 z_{ij} + \phi_2 y_{ij} \text{ if } x_{ij} = 0, \end{aligned} \tag{4}$$

Alternative approach: stratify by y.

Application to clinical example: methods

- Data were sampled from the non-differential misclassification model, with parameters estimated from data from the IDAMS consortium. [Jaenisch et al., 2016]
- 700 participants per study, 10 studies.
- We applied 5 models:
  1. Analysis on the full data as if the gold standard exposure was observed for all (simulated) participants (for comparison).
  2. A complete cases analysis.
  3. A naive model in which the surrogate measurement of the muscle pain was used for participants for whom the gold standard measurement was not available.
  4. The misclassification models.

Application to clinical example: results

Model	β_{20}(95%CI)	$\tau_{\beta_{2j}}$(95%CI)
Full data	0.87 (0.60 : 1.14)	0.32 (0.18 : 0.61)
Complete cases	1.02 (0.67 : 1.38)	0.23 (0.05 : 0.73)
Naive	0.60 (0.31 : 0.89)	0.37 (0.21 : 0.69)
Non-differential misclassification model	0.79 (0.48 : 1.11)	0.35 (0.19 : 0.67)
Differential misclassification	0.80 (0.48 : 1.10)	0.35 (0.18 : 0.68)

Simulation methods

- Data were sampled from a non-differential misclassification model, with some parameters estimated from data from the IDAMS consortium. [Jaenisch et al., 2016]
- 700 participants per study, 10 studies. 1000 simulation repetitions.
- We applied 4 models:
  1. Analysis on the full data as if the gold standard exposure was observed for all participants (for comparison)
  2. A complete cases analysis,
  3. A naive model in which the surrogate measurement of the muscle pain was used for participants for whom the gold standard measurement was not available.
  4. An overspecified non-differential misclassification model.

Simulation results

Model	Mean	RMSE	% Bias	Power	Coverage
Full data	0.80	0.21	2.49	0.97	0.94
Complete cases	0.80	0.29	2.98	0.75	0.96
Naive	0.48	0.34	-37.98	0.80	0.62
Non-differential misclassification model	0.79	0.24	1.66	0.92	0.96

RMSE: Root mean square error.
Coverage: Coverage probability of the 95% Credibility Interval.

Key findings

- The Bayesian estimation framework is very flexible and allows adjusting for patient-level sources of measurement error.
- Modeling of patient-level data appears most fruitful when interested in multivariable models (e.g. prediction, etiology).
- Adjustment for (potential) measurement error is possible on data set level and/or on patient level.
- Over-specification for (potential) measurement error may be beneficial in IDP-MA where $n >> p$.

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under ReCoDID grant agreement No 825746.

We thank the IDAMS consortium for providing aggregate data on the diagnosis of dengue.

ReCoDID

Individual participant data meta-analyses should not ignore clustering.
Journal of Clinical Epidemiology, 66(8):865–873.e4.

Effect of Nondifferential Misclassification on Estimates of Odds Ratios with Multiple Levels of Exposure.
American Journal of Epidemiology, 136(3):356–362.

Misclassification in 2 x 2 tables.
Biometrics, 10(4):478–486.

Measurement error in nonlinear models.
A modern Perspective: Chapmn & Hall/CRC.

Meta-analysis in clinical trials.
Controlled Clinical Trials, 7(3):177–188.

Does nondifferential misclassification of exposure always bias a true effect toward the null value?
American journal of epidemiology, 132(4):746–748.

Clinical evaluation of dengue and identification of risk factors for severe disease: protocol for a multicentre study in 8 countries.
BMC infectious diseases, 16(1):120.

Effect of misclassification on estimated relative prevalence of a characteristic: Part I. Two populations infallibly distinguished. Part II. Errors in two variables.
American Journal of Public Health and the Nations Health, 53(10):1656–1665.

When will nondifferential misclassification of an exposure preserve the direction of a trend?
American Journal of Epidemiology, 140(6):565–571.