TG4 on measurement error and misclassification

STRATOS initiative

Veronika Deffner

Department of Statistics, LMU Munich

18.09.2017

STRATOS INITIATIVE

STABLAB Statistisches Beratungs Labor

Situations with one or several mismeasured variables

Two cases:

1. Analysis of the association between two variables:

$$
\begin{array}{ccc}
Y & \overset{\textbf{?}}{\longleftrightarrow} & X \\
 & & \\
Y^* & \overset{\text{Assessable association}}{\longleftrightarrow} & X^* \\
\text{(observed)} & & \text{(observed)}
\end{array}
$$

2. Analysis of the distribution of a variable:

$$X \sim \textbf{?}$$

$$X^* \sim \text{Distribution (assessable)}$$

- Impact on epidemiological analysis

Example: Association between air pollution and human health

Health outcome $\longleftrightarrow$ Individual particle number concentration (PNC)

Health outcome $\longleftrightarrow$ PNC*, e.g.

- Error-prone individual PNC measurement
- Ambient PNC

Results from Peters et al. (2015):

Table 4 Associations between ambient 1-hour average air pollution concentrations at the central monitoring site and 1-hour average ECG-measures

	HR		SDNN		RMSSD	
	%-change	**95% CI**	**%-change**	**95% CI**	**%-change**	**95% CI**
Personal PNC	0.13	−0.19; 0.45	−0.93†	−2.01; 0.16	0.53	−0.70;1.77
UFP	0.40	−0.16; 0.95	0.99	−0.66; 2.64	−0.12	−2.40; 2.21

Analyses considered concurrent exposures and adjusted for trend, meteorology and time of day. Effect estimates are shown for an increase in interquartile range as given in Table 2.

†p-value <0.1, *p-value <0.05, **p-value <0.01, CI: confidence interval, HR: heart rate, RMSSD: root mean square of successive differences, SDNN: standard deviation of NN-intervals, PNC: Particle number concentrations, PM_{10}: particulate matter with an aerodynamic diameter <10µm, $PM_{2.5}$: particulate matter with an aerodynamic diameter <2.5µm, UFP: ultrafine particles (10-100µm); ACP: accumulation mode particles (100-800 nm).

STRATOS TG4: Measurement error and misclassification

Chairs: Laurence Freedman, Victor Kipnis

Members:

Raymond Carroll	Ruth Keogh
Veronika Deffner	Helmut Küchenhoff
Kevin Dodd	Pamela Shaw
Paul Gustafson	Janet Tooze

Activities:

- Survey of current practice
- Development of guidance documents
  - For epidemiologists, on nutritional epidemiology
  - For statisticians with epidemiological background

Literature survey of current practice

Aims:

- Assess current practice for addressing measurement error in observational epidemiology
- Identification of knowledge gaps

Research areas:

- Dietary intake cohort studies (Pamela Shaw/Ruth Keogh)
- Dietary intake surveys (Kevin Dodd)
- Physical activity cohort studies (Janet Tooze)
- Air pollution cohort studies (Veronika Deffner/Helmut Kuechenhoff)

Survey procedure

- Separate literature searches for each research area
- Two stages:
  - A: general search terms related to the research area
  - B: (only cohort studies,) search terms related to measurement error in addition to the general search terms
- Data extraction via survey instruments

  / \

  General questions for all research areas | Research-area-specific questions

- Quality control

→ Number of articles reviewed (A/B):

Dietary intake cohort studies:	51	27
Dietary intake surveys:	67	
Physical activity cohort studies:	30	39
Air pollution cohort studies:	50	25

- Analysis of the survey data

Survey results

- Insufficient description of the measurement error, even if adequate data is available
- Inadequate discussion of the impact of measurement error on the study results
- Several incorrect claims about the possible direction of the bias

	Dietary intake cohort	Dietary intake survey	Physical activity cohort	Air pollution cohort
Mention ME as potential problem N (%)	48 (94%)	53 (79%)	17 (57%)	20 (40%)
Used a method to adjust for ME N (%)	5 (10%)	19 (28%)	0 (0%)	3 (6%)

- Rare use of methods which take measurement error into account in spite the availability of adequate methods
- Multiple error-prone exposures not acknowledged

Guidance document for statisticians

- General background on measurement error
- Effects of measurement error and misclassification on study results
- Guidance for taking measurement error and misclassification into account
  - Study design
  - Statistical analysis methods
  - Software
  - Special topics and practical advice

Common error structures: classical and Berkson error

Classical measurement error

$$X^* = X + U$$

- $E(U) = 0, X \perp U$
- Example: error, when measuring the concentration of air pollutants
- Extension: **linear** measurement error

$$X^* = \alpha_0 + \alpha_X X + U$$

Berkson error

$$X = X^* + U$$

- $E(U) = 0, X^* \perp U$
- Example: error, when assigning ambient air pollutant concentrations to individuals

Classification of error: differential and non-differential error

Regression of X on Y: $\mathbb{E}(Y|X) = f(X)$

Differential error of X^*

- The distribution of $Y|X$ does not equal the distribution of $Y|X^*, X$
- Example case-control studies: errors in the measurements (X^*) depend on the outcome (Y: case/control)

Differential error of Y^*

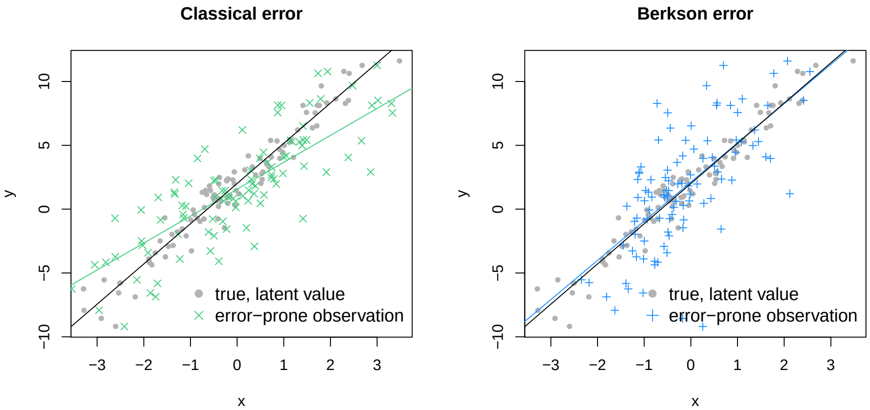
- The distribution of $Y^*|Y$ does not equal the distribution of $Y^*|Y, X$
- Example comparison of the dietary intake between two groups: error in reported dietary intake (Y^*) differs by the group (X)

Effects of measurement error on study results

Table 1: Effects of measurement error according to type of error and target of the analysis

Analysis	Target	Non-differential error			Differential error
		Classical	Linear	Berkson	Any
Regression with single error-prone covariate	Regression coefficient	Underestimated	Biased in either direction	Unbiased	Biased in either direction
	Test of null hypothesis	Valid	Valid	Valid	Invalid
Regression with multiple error-prone covariates	Regression coefficients	Biased in either direction	Biased in either direction	Unbiased	Biased in either direction
	Tests of null hypothesis	Invalid	Invalid	Valid	Invalid
Regression with error-prone outcome variable	Regression coefficients	Unbiased	Biased in either direction	Underestimated	Biased in either direction
	Tests of null hypothesis	Valid	Valid	Valid	Invalid
Distribution with an error-prone continuous variable	Mean	Unbiased	Biased in either direction	Unbiased	-
	Lower percentile	Underestimated	Biased in either direction	Overestimated	-
	Upper percentile	Overestimated	Biased in either direction	Underestimated	-

Effects of measurement error on study results



Guidance: study design

1. Obtain information about the measurement error model and its parameters by the use of validation studies:
   - true values of the variable (reference instrument) and
   - its error-prone values (test instrument)
2. Adaptation of the final design of the study to the presence of measurement error

   Classical covariate measurement error in a simple regression model:

$$n_{X^*} = \frac{1}{\mathrm{Corr}(X, X^*)^2} \cdot n_X$$

   Example: $\mathrm{Corr}(X, X^*) = 0.9 \Rightarrow$ 1.23 times higher sample size

Guidance: statistical analysis methods

Regression calibration

Regression using the predicted values of X based on X^* and Z

$$\mathbb{E}(Y|X^*, Z) = \beta_0 + \beta_X \mathbb{E}(X|X^*, Z) + \beta_Z Z$$

Moment reconstruction and moment-adjusted multiple imputation

Construction of a quantity with the same distribution as X based on the moments of the joint distribution of (X, Y)

$$X_M(X, Y) = f(\mathbb{E}^k(X, Y)), \quad k = 1, 2, \ldots$$

Multiple imputation

Consider the data (X, X^*, Z, Y), which include an internal validation subset, as a problem of missing data and impute $X|X^*, Z$

Original data

1	24.60
2	21.28
3	14.82
4	0.93
5	8.59
6	NA
7	NA
8	NA
9	NA
10	NA

$\Rightarrow$

Imputed Data

1	24.60
2	21.28
3	14.82
4	0.93
5	8.59
6	0.44
7	1.59
8	12.57
9	28.63
10	21.82

Guidance: statistical analysis methods

Likelihood methods

1. Model as if X were observable
2. Error model
3. Distribution for X (only in the case of classical measurement error)
4. Likelihood of (Y, X^*) through combining steps (1)-(3)

$$f(y, x^*|z, \boldsymbol{\theta}) = \int f(y|z, x, x^*, \boldsymbol{\theta_1}) \cdot f(x^*|z, x, \boldsymbol{\theta_2}) \cdot f(x|z, \boldsymbol{\theta_3}) d\mu(x)$$

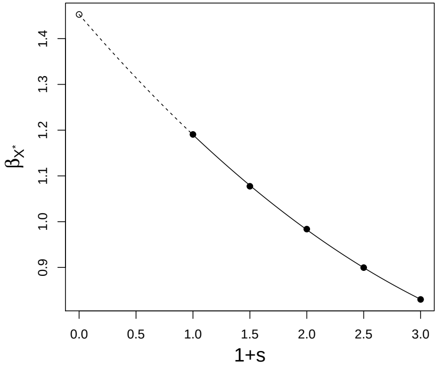
Bayesian methods

Specification of models like for likelihood methods and in addition, specification of prior distributions

$$f(\boldsymbol{\theta}|y, x^*, z) \propto \int f(y|z, x, x^*, \boldsymbol{\theta_1}) \cdot f(x^*|z, x, \boldsymbol{\theta_2}) \cdot f(x|z, \boldsymbol{\theta_3}) d\mu(x) \cdot p(\boldsymbol{\theta})$$

SIMEX (simulation and extrapolation)

Estimate the relationship between the size of the classical measurement error and the limits of the parameter estimates in naive regression and extrapolate to the error-free case



$1 + s$: scaling factor of the measurement error variance $(\mathrm{Var}(U) \cdot (1 + s))$

Guidance: software

Regression calibration

`rcal` (package `merror`)	STATA	Hardin et al. (2003a)
`eivregl`	STATA	Hardin et al. (2003a)
NCI macros	SAS	Kipnis et al. (2009)
`%blinplusl`	SAS	Rosner et al. (1990)
`%relibpls8l`	SAS	Rosner et al. (1992)
`%rrcl`	SAS	Liao et al. (2011)

SIMEX

`simex`, `simexplot` (package `merror`)	STATA	Hardin et al. (2003b)
package `simex`	R	Cook and Stefanski (1994), Küchenhoff et al. (2006), Lederer and Küchenhoff (2013)
package `simexaft`	R	Genz et al. (2011), He et al. (2007)
package `hSIMEXUnknown`	R	Delaigle and Hall (2008)

Bayesian methods

package BayesME	R	Sarkar et al. (2014a,b),
	BUGS	Lunn et al. (2000, 2009, 2012)
	Stan	Stan Development Team (2016a,b)

Conclusion

- Inadequate treatment of measurement error and misclassification in epidemiological analyses is commonplace
- Three steps of adequate treatment:
  1. Consideration of potential measurement error at the design stage
  2. Explicit statement of assumptions regarding measurement error and exploration of its potential impact on the study results
  3. Application of analysis methods which take measurement error into account
- STRATOS TG4 contributes to improving the consideration of measurement error and misclassification in the statistical analyses of observational studies:
  1. Overview of measurement error types and their impact
  2. Overview and introduction of methods which take measurement error into account

References

Cook JR, Stefanski LA. Simulation-Extrapolation Estimation in Parametric Measurement Error Models. Journal of the American Statistical Association 1994; 89:1314–1328.

Delaigle A, Hall P. Using SIMEX for smoothing parameter choice in errors-in-variables problems. JASA 2008; 103:280-287.

Genz A, Bretz F, Miwa T, Mi X, Leisch F, Scheipl F, Hothorn T. mvt-norm: Multivariate Normal and t Distributions. R package version 0.9-9991. 2011: URL http://CRAN. R-project.org/package=mvtnorm.

Hardin JW, Schmiediche H, Carroll RJ. The regression-calibration method for fitting generalized linear models with additive measurement error. The Stat Journal 2003a; 3:361-372.

Hardin JW, Schmiediche H, Carroll RJ. The simulation extrapolation method for fitting generalized linear models with additive measurement error. The Stat Journal 2003b; 3:373-385.

He W, Yi GY, Xiong J. Accelerated Failure Time Models with Covariates Subject to Measurement Error. Statistics in Medicine 2007; 26:4817-4832.

Kipnis V, Midthune D, Buckman DW, Dodd KW, Guenther PM, Krebs-Smith SM, Subar AF, Tooze JA, Carroll RJ, Freedman LS. Modeling data with excess zeros and measurement error: application to evaluating relationships between episodically consumed foods and health outcomes. Biometrics 2009; 65:1003-1010.

Küchenhoff H, Mwalili SM, Lesaffre E. A General Method for Dealing with Misclassification in Regression: The Misclassification SIMEX. Biometrics 2006; 62:85–96.

Lederer W, Küchenhoff H. Simex: SIMEX- and MCSIMEX-Algorithm for Measurement Error Models. 2013: http://CRAN.R-project.org/package=simex.

References

Liao X, Zucker D, Li Y, Spiegelman D. Survival analysis with error-prone time-varying covariates: a risk set calibration approach. Biometrics 2011; 67:50-58.

Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS-a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and computing. 2000 Oct 1; 10(4):325-37.

Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: Evolution, critique and future directions. Statistics in medicine. 2009 Nov 10; 28(25):3049-67.

Lunn D, Jackson C, Best N, Thomas A, Spiegelhalter D. The BUGS book: A practical introduction to Bayesian analysis. CRC press; 2012 Oct 2.

Peters A, Hampel R, Cyrys J, Breitner S, Geruschkat U, Kraus U, Zareba W, Schneider A. Elevated particle number concentrations induce immediate changes in heart rate variability: a panel study in individuals with impaired glucose metabolism or diabetes. Particle and Fibre Toxicology. 2015 12(1):7.

Rosner B, Spiegelman D, Willett W. 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 1990; 132:734-735.

Rosner B, Spiegelman D, Willett W. Correction of logistic regression relative risk estimates and confidence intervals for random within person measurement error. American Journal of Epidemiology 1992; 136:1400-1413.

Sarkar A, Mallick BK, Carroll RJ. Bayesian semiparametric regression in the presence of conditionally heteroscedastic measurement and regression errors. Biometrics 2014a; 70:823-834.

Sarkar A, Mallick BK, Staudenmayer J, Pati D, Carroll RJ. Bayesian semiparametric density deconvolution in the presence of conditionally heteroscedastic measurement errors. Journal of Computational and Graphical Statistics 2014b; 25:1101-1125.

References

Stan Development Team. (2016a). Stan modeling language users guide and reference manual, version 2.14.0 [Computer software manual]. Retrieved from http://mc-stan.org

Stan Development Team (2016b). RStan: the R interface to Stan. R package version 2.14.1. http://mc-stan.org/.