Guidance for key issues of design and analysis of observational studies

Raymond Carroll, Paul Gustafson, Victor Kipnis, Helmut Küchenhoff, Len Stefanski

TG4 Measurement error

Aug, 29, 2013

Sources of measurement error

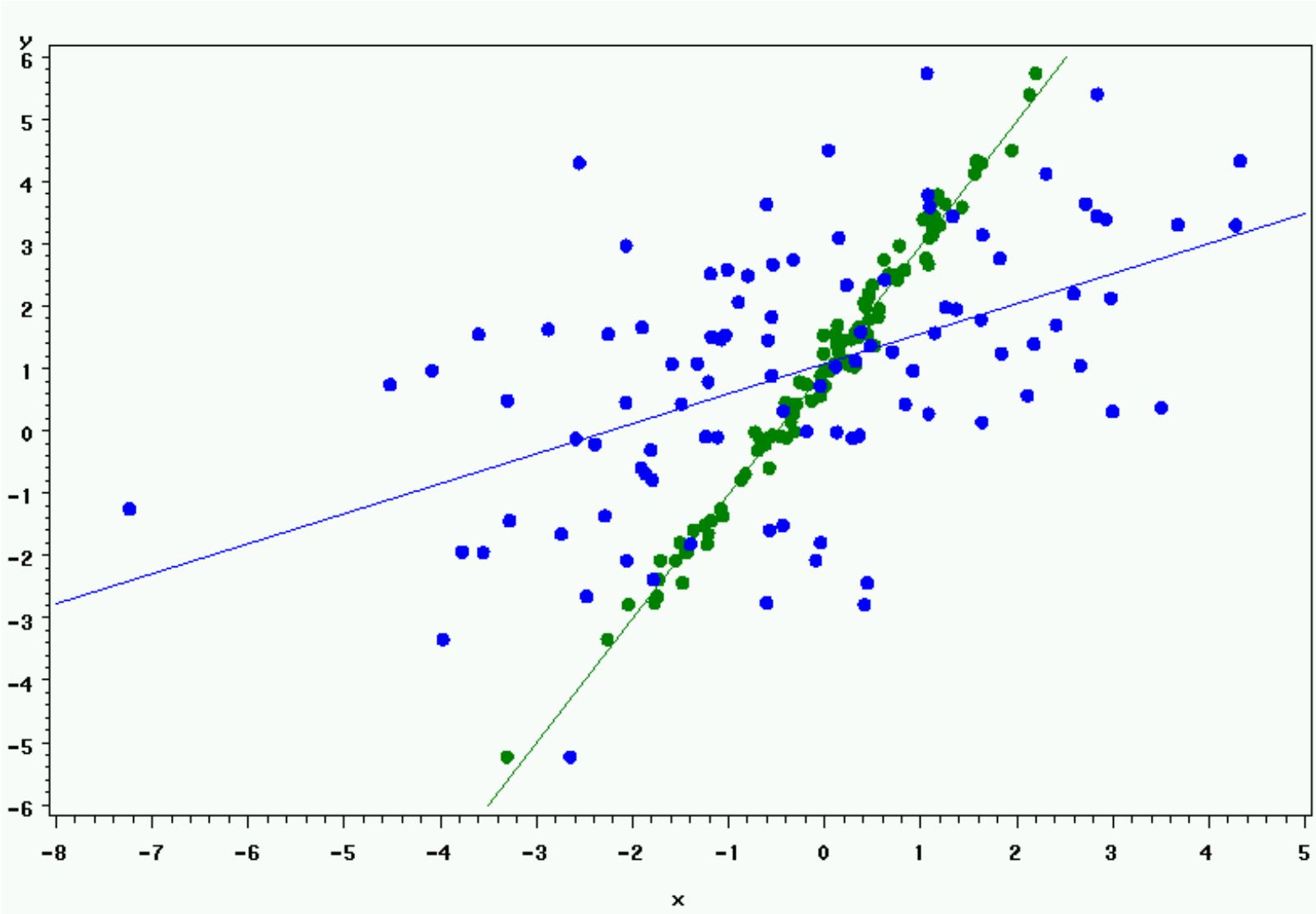
- Induced by an instrument (laboratory value, blood pressure)
- Induced by medical doctors or patients
- Wrong diagnosis (misclassification)
- Measurement error induced by definition, e.g. *long term mean of daily fat intake*
- Surrogate -Variables e.g. *mean of exposure in a region where the study participant lives* instead of individual exposure

The triple whammy effect of measurement error

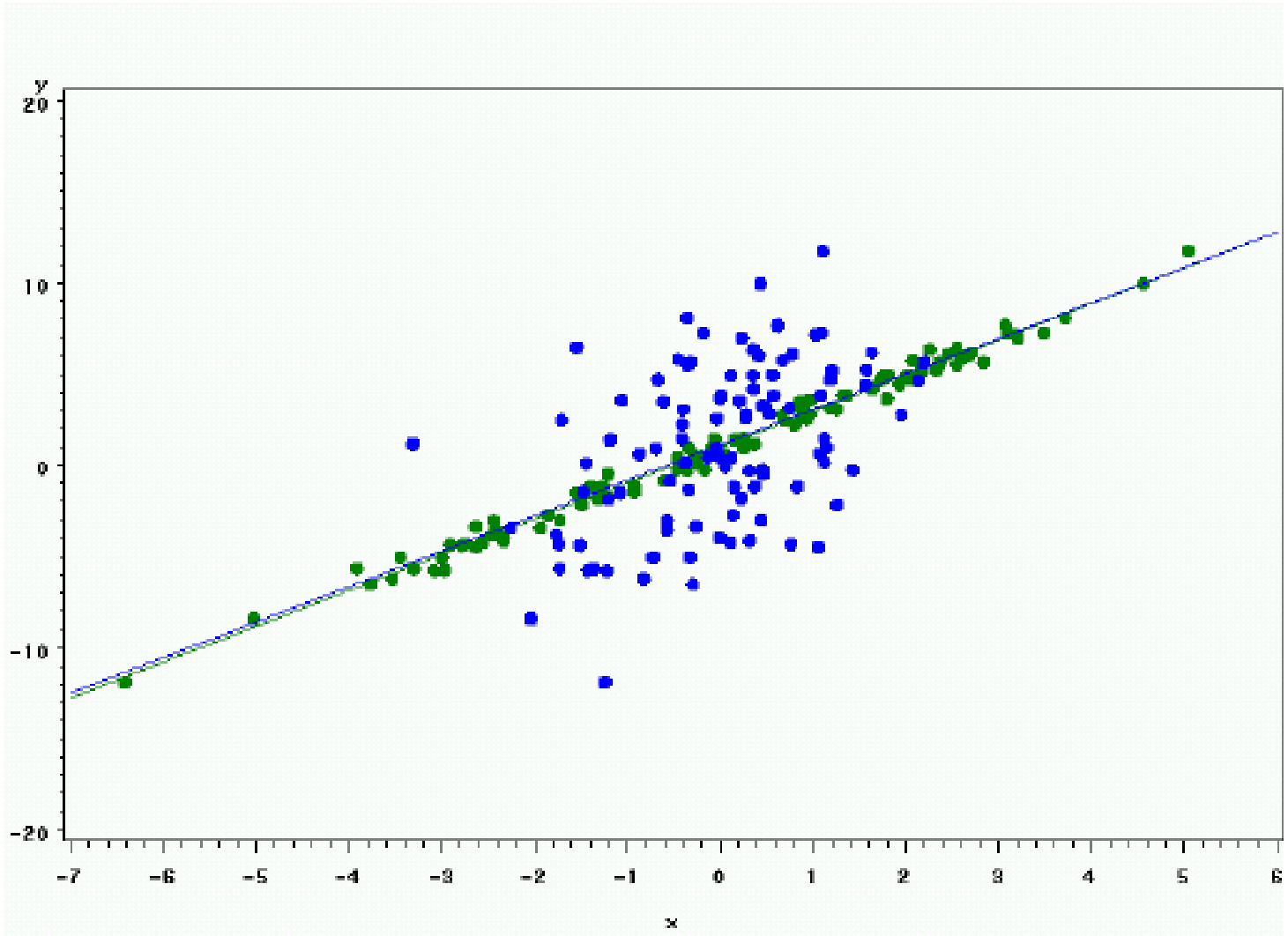
- Bias
- Masking of features
- Loss of power

Note: Also the parameter estimates of exactly measured covariates may be affected.

Effect of additive measurement error in linear regression



Effect of Berkson measurement error in linear regression



Correction for measurement error

General procedure

- Model measurement process (measurement error) using extra information (Validation study replicates)
- Include this model in the main model (estimation procedure)

Methods

- Correction and method of moments
- Regression calibration
- Simulation and extrapolation (SIMEX)
- Likelihood
- Bayes

Example 1: I^{131} exposure and thyroid disease

- The Nevada Test Site Study and the Hanford Thyroid Disease Study, both cohort studies with $\sim$ 2,000 subjects
- Various indicators of thyroid disease (cancer, thyroiditis)
- Unique Structure: Mixture of Berkson and Classical measurement error
- Focus: Risk estimation and the overestimation of the power to detect significant effects

Mallick, et al (2002, *Biometrics*), Li, et al (2007, *Biometrics*)
Carroll, et al (2006, Chapter 1.4)

Example 2: Perfluorinated acids and hypothyroxinemia in pregant women

- 96 cases matched to 175 controls (matched on age, physician)
- Measurement error in the continuous trivariate exposure (PFOA, PFOS, PFHxS)
- Modeling of a quality-control experiment yields information on the measurement error magnitude

Espino-Hernandez, Gustafson, Burstyn (2011)

Example 3: NIH AARP Diet and Health Study

- NIH-AARP is a prospective US cohort comprising 188736 postmenopausal women who completed a 124-item food frequency questionnaire (FFQ) in 1995-1996.
- Outcome: breast cancer development
- Exposure: percent of energy from usual (long-term average) fat intake
- Risk model: Cox proportional hazard regression
- Calibration substudy: about 1000 women with two 24-hour non-consecutive dietary recalls assumed to be unbiased for true usual intake
- Estimated hazard ratio for a twofold increase in percent energy from fat
- Using FFQ: $1.15(95\% CI = 1.05 to 1.26)$
- Using regression calibration with 24-hour dietary recall as the reference: $1.32(95\% CI = 1.11 to 1.58)$

Thiebaut ACM, Kipnis V et al.(2007)

Example 4: Radon radiation and lung cancer

- Matched case control (608 cases , 626 controls)
- Response: lung cancer
- Population controls (frequency matching by age gender region)
- Covariates : Smoking asbestos (binary)
- Analysis : Conditional logistic regression

Heid, Küchenhoff et al.(2002)

Example 5: Mortality and Blood Pressure

- Observational: Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients with Heart Failure (OPTIMIZE-HF) N = 48,612
- Responses: In-hospital and post-discharge mortality
- Analyses: Logistic regression (in-hospital mortality) Cox Proportional Hazard (post-discharge mortality)
- Error-prone predictors: Systolic and diastolic BP (via splines)
- Gheorghiade et al. (2006) Thomas et al. (2013)

Aims

- Reporting measurement error
- Criteria for necessity of using ME methods
- Adequate methods
- Sensitivity analysis