Recent Contributions of STRATOS Topic Group 4: Measurement Error and Misclassification

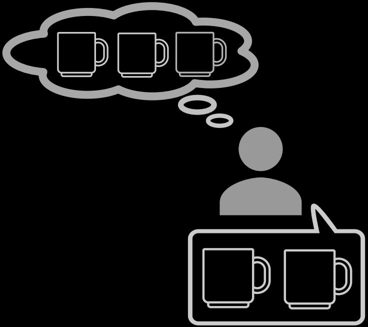
Michael Wallace, University of Waterloo

Slide deck available at: `mpwallace.github.io`

TG4 Chairs: Paul Gustafson, Pamela Shaw
TG4 Members: Jonathan Bartlett, Hendriek Boshuizen, Raymond Carroll, Veronika Deffner, Kevin Dodd, Laurence Freedman, Sabine Hoffmann, Ruth Keogh, Victor Kipnis, Helmut Küchenhoff, Douglas Midthune, Cécile Proust-Lima, Anne Thiebaut, Janet Tooze, Michael Wallace

Topic Group 4: Measurement Error and Misclassification

- Measurement error: When what we observe differs from what we want to observe.
- Impact unpredictable, and requires specialist methodology.



STRATOS Topic Group 4: Dedicated to exploration and education for all things measurement error.

Highlight: Two comprehensive ‘guidance papers’ published in Statistics in Medicine (2020).

TG4's recent work includes:

- Categorization of continuous error-prone observations.
- Post-prediction inference and Berkson error.
- Education through our website, R Shiny app, and videos.

Categorization of continuous variables occurs for various reasons:

- ‘Real-world’ interpretations: e.g., blood pressure (hypertensive vs. not); BMI (obese, overweight, etc.).
- Analytical decisions: e.g., to use more familiar methods, simplify assumptions.

Categorization: Not without caveats, but a common practice.

Categorization and Measurement Error

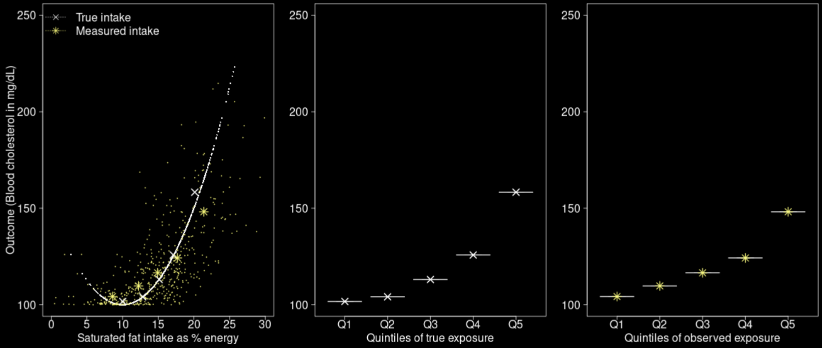
Categorization of error-prone variables can lead to *misclassification*:

- Truth: Average long-term blood pressure (X):
  - Not hypertensive: $X \leq 130; X_c = 0$
  - Hypertensive: $X > 130; X_c = 1$
- Observed: Single blood pressure measurement (X^*):
  - Not hypertensive: $X^* \leq 130; X_c^* = 0$
  - Hypertensive: $X^* > 130; X_c^* = 1$
- Possibility:
  - True $X = 125 \implies$ not hypertensive ($X_c = 0$)
  - Observed $X^* = 135 \implies$ hypertensive ($X_c^* = 1$)
- **Question**: What are the implications for analysis?

Led by Anne Thiébaut (Inserm, France).

- Many myths, misconceptions, and misunderstandings surround categorization.
- Measurement error adds to the list.
- Goal: Explore, explain (and dispel!) five common misconceptions.

"Categorizing a mismeasured exposure can help with finding the shape of the exposure-outcome relationship"



- Measurement error in X^*: *non-differential* w.r.t. outcome Y if

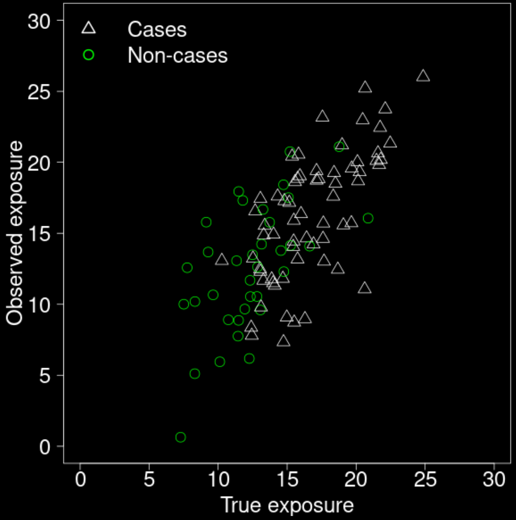
$$X^* \perp Y | X$$

- Differential error example: patients diagnosed with lung cancer report their smoking history with a different level of accuracy to those without lung cancer.
- **Question**: If X subject to non-differential error, will the misclassification in X_c^* also be non-differential?[1]

[1] Yes, but only in highly improbable scenarios: Flegal et al. (1991) Differential misclassification arising from non-differential errors in exposure measurement. doi:10.1093/oxfordjournals.aje.a116026

"Categorization of a continuous variable with non-differential error will produce non-differential misclassification"

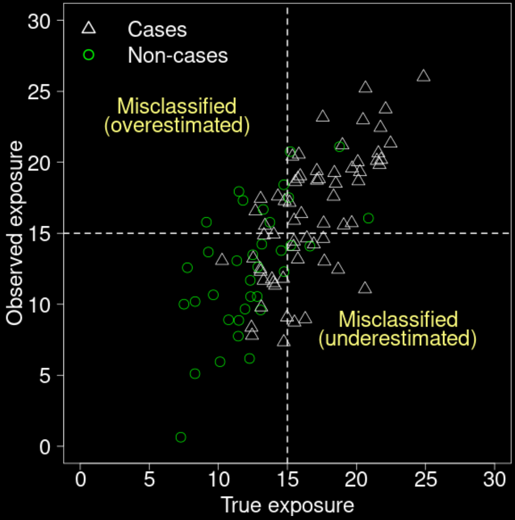
- Suppose: Binary outcome Y denoting presence/absence of disease, with an exposure positively associated with the outcome.



Misconception 2

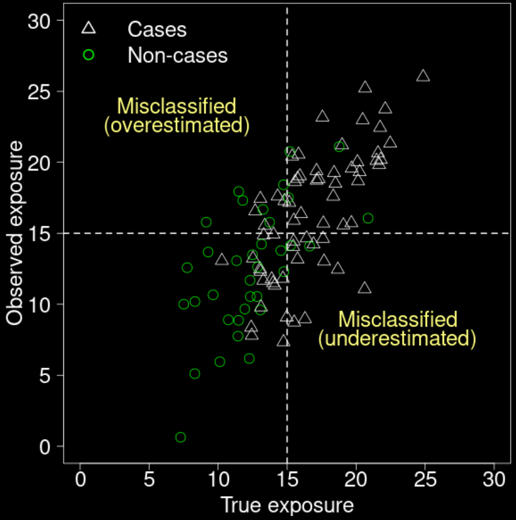
"Categorization of a continuous variable with non-differential error will produce non-differential misclassification"

- Probability of misclassification when categorizing X^* higher around boundary, so depends on true exposure X



"Categorization of a continuous variable with non-differential error will produce non-differential misclassification"

- In this example: sensitivity is higher amongst cases (75%) than amongst non-cases (67%).



Other Misconceptions

Other misconceptions covered:

- Categorizing an error-prone continuous exposure mitigates bias due to measurement error.
- The comparison of extreme quantiles involves less misclassification and therefore results in smaller bias.
- Misclassification of exposure always results in attenuated association.

Led by Hendriek Boshuizen (Wageningen University & Research, Netherlands).

A categorized error-prone continuous variable causes various problems:

- Biases effect estimates (both attenuation and not).
- Obfuscates shapes of relationships.
- Differential misclassification (even if the continuous variable error is non-differential).
- Goal: Develop a new approach to measurement error correction in this setting.

We propose a correction method using regression calibration (RC)[2]:

- Principle of RC: Estimate X using X^* and confounders Z:

$$\hat{X} = E[X|X^*, Z]$$

  and use $\hat{X}$ in place of X in standard analysis.
- Exact for linear models, good for many GLMs.
- Could be applied to categorical models by replacing X_c with $E[X_c|X^*, Z]$...
- ...but categorization of a non-differential errorred X results in differential misclassification, which violates RC.

---

[2]We extend MacMahon et al. Blood pressure, stroke, and coronary heart disease. Part 1, Prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. Lancet. 335(8692) 765-774

- Assume non-differential error: $X^* \perp Y | X, Z$
- Define C_k: set of values for category k (e.g. the k^{th} quintile).
- Define $\hat{X} = E[X|X^*, Z]$, then

$$E[\hat{X}|\hat{X} \in C_k, Z] = E[X|\hat{X} \in C_k, Z]$$

- Thus: $E[\hat{X}|\hat{X} \in C_k, Z]$ in category k can be interpreted as the mean exposure in the category defined by $\hat{X} \in C_k$.
- We can link $E[\hat{X}|\hat{X} \in C_k, Z]$ to the mean of Y in category k.

- If there is confounding, categorization means effect of confounder is not estimated correctly.
- Mitigation: use a residual exposure model:

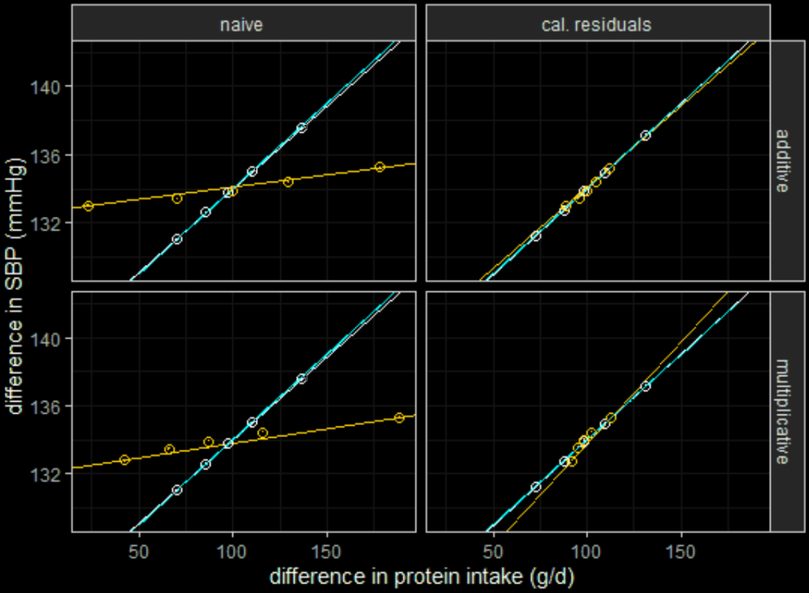
$$R = X - E[X|Z] \qquad R^* = X^* - E[X^*|Z]$$

- Like RC estimates of X, we have RC estimates of R:

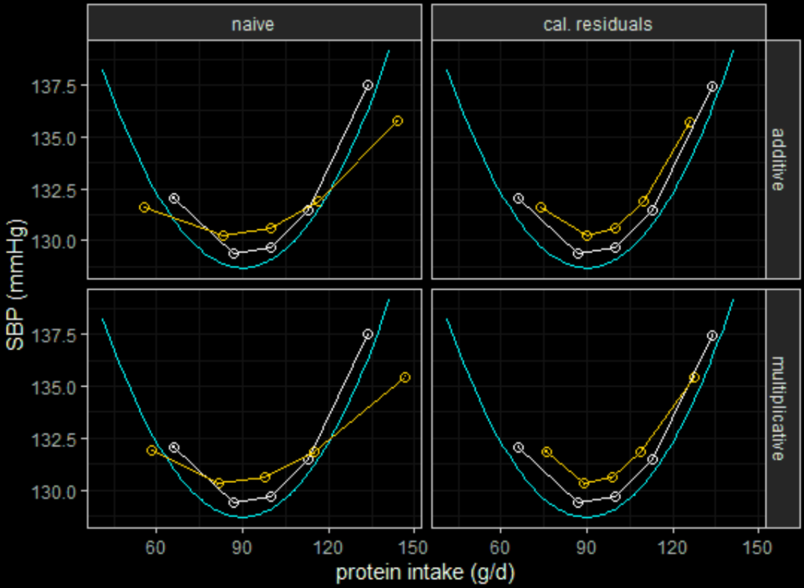
$$\hat{R} = \hat{E}[R|R^*, Z]$$

- We can then define $\hat{R}_c$ derived from sets C_k based on $\hat{R}$ and use these in our analysis.

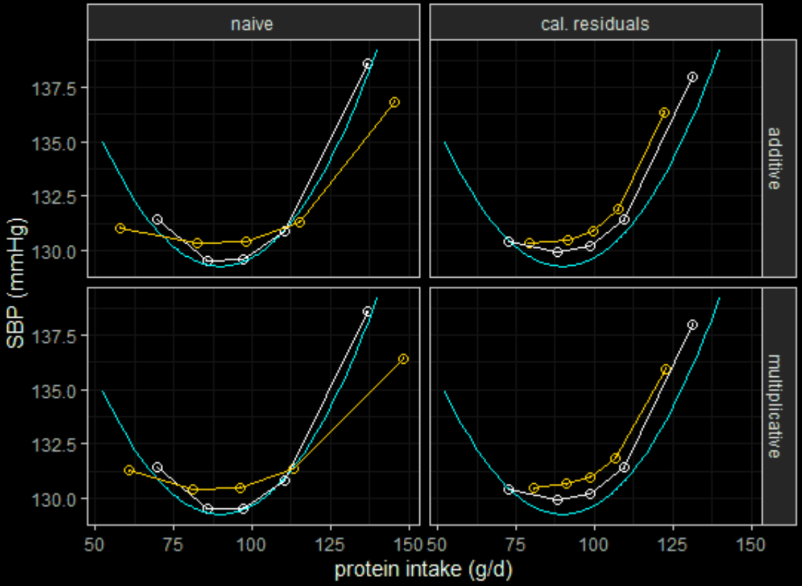
Simulations: Linear, Log-Normal X, confounding



Simulations: Quadratic, Log-Normal X, no confounding



Simulations: Quadratic, Log-Normal X, confounding



Project 2: Conclusions

Overall, our proposed method:

- Accurate for linear relationships, but in such cases using the continuous exposure variable would seem prudent.
- Less effective for strongly non-linear relationships the approach does not, but still an improvement over naive analysis.
- Could be used to help determine whether linear modelling is appropriate.

Preceding projects focused on *classical* measurement error.

e.g., Reported usual daily calorie intake X^* equals true intake X plus some random error U:

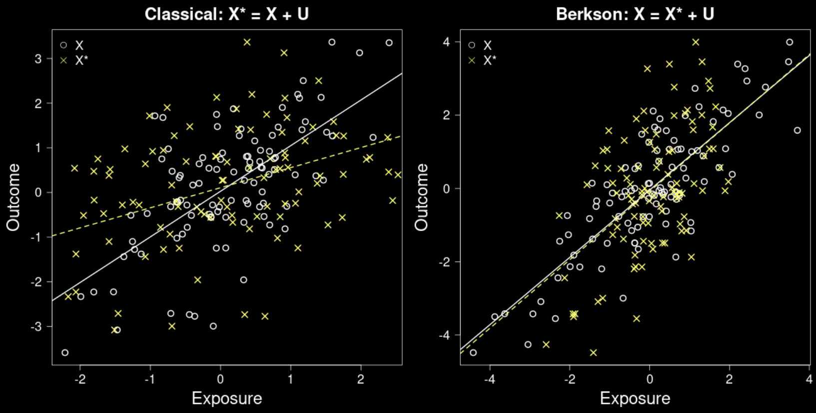
$$X^* = X + U; \qquad U \perp X$$

In contrast, there is *Berkson* error.

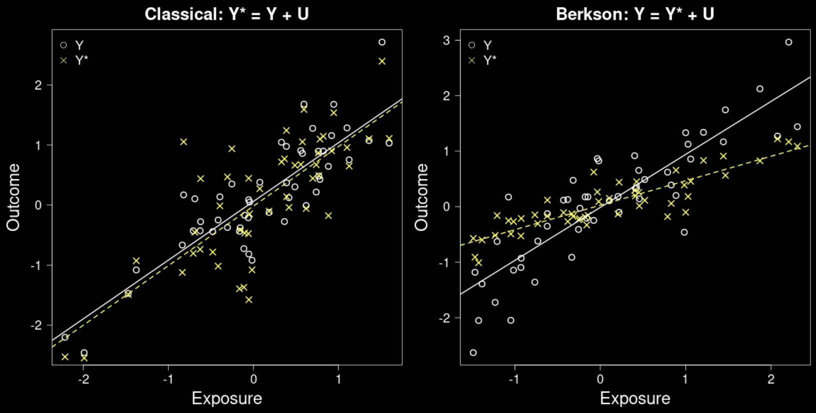
e.g., True nutrient absorption X equals nutrient intake X^* plus some random error U:

$$X = X^* + U; \qquad U \perp X^*$$

Classical vs. Berkson: Exposure



Classical vs. Berkson: Outcome



TG4 has three projects on Berkson error, addressing:

- An introduction to Berkson error in exposure and outcome variables variables.
- The impact of Berkson error on estimating distributional measures and a correction approach.
- The relationship between Berkson error and regression calibration.

Led by Lillian Boe (Memorial Sloan Kettering Cancer Center, NY).

- In some settings *predicted* values used to estimate true values.
- Example: Schofield's equation to predict basal metabolic rate as a function of body mass and activity level.
- Such predicted measures often subject to Berkson error.
- In particular: Regression calibration and Berkson error highly related.

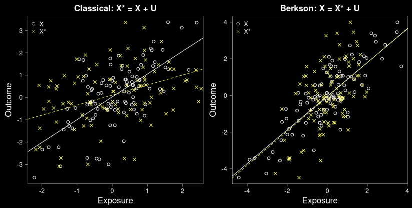
Suppose $X^* = X + U$ (classical error). RC tells us to use

$$\hat{X} = E[X|X^*, Z]$$

in place of X.

But: $\hat{X}$ is itself an error-prone measurement of X. How is this an improvement?

Answer: $\hat{X}$ has Berkson error, and may result in unbiased effect estimates in certain settings.



Viewing RC with a Berkson lens highlights limitations.

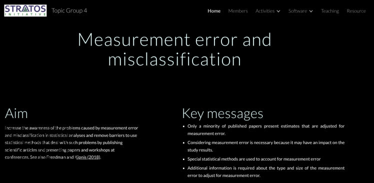
e.g., We must estimate $\hat{X} = E[X|X^*, Z]$ using the same variables Z in the calibration equation as in the outcome model.

This principle means that, for any given exposure, there is no single calibration equation that is appropriate for all analyses.

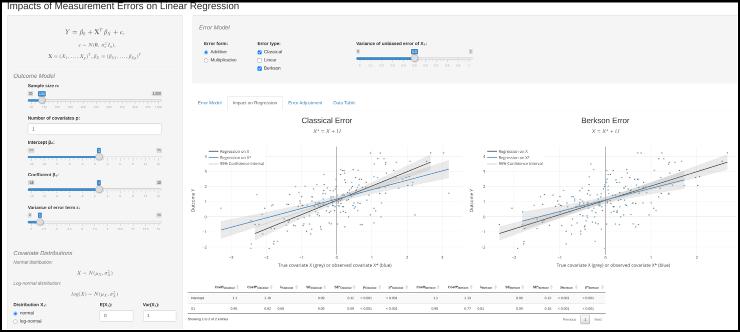
This project provides a checklist when implementing RC, including:

- Further modelling considerations for the outcome model.
- Where and how to source additional data to inform the calibration model.
- Advice on adjusting standard errors to account for calibration uncertainty.

- Website featuring previous presentations and other resources.
- General audience introductory video series.



- An R Shiny app for exploring measurement error.



- **Measurement Error Guidance**: P. A. Shaw et al. (2020). STRATOS guidance document on measurement error and misclassification of variables in observational epidemiology: Part 2 - sample size, more complex methods of adjustment and advanced topics. *Statistics in Medicine* **39(16)** 2197-2231.
- **Measurement Error Guidance**: R. H. Keogh et al. (2020). STRATOS guidance document on measurement error and misclassification of variables in observational epidemiology: Part 1 - basic theory, validation studies and simple methods of adjustment. *Statistics in Medicine* **39(16)** 2232-2263.
- **Epidemiologic Review Paper**: P. A. Shaw et al. (2018). Epidemiologic analyses with error-prone exposures: Review of current practice and recommendations. *Annals of Epidemiology* **28(11)** 821-828.
- **Misconceptions**: A. C. M. Thiébaut et al. (2025). Five misconceptions about categorizing exposure variables measured with error in epidemiological research. *In review.*
- **Regression Calibration and Berkson Error**: L. A. Boe et al. (2023). Issues in Implementing Regression Calibration Analyses. *Practice of Epidemiology* **192(8)** 1406-1414.
- **General Audience Article**: M. P. Wallace (2020). Analysis in an imperfect world. *Significance* **17(1)**.

- **TG4 website**: `http://www.stratostg4.statistik.uni-muenchen.de`
- **Shiny app**: `https://mem-explorer.shinyapps.io/MEMExplorer-v5`
- **Introductory videos**: `https://youtube.com/@TheSTRATOSinitiative`

michael.wallace@uwaterloo.ca
https://mpwallace.github.io