Measurement error as missing data: the case of epidemiologic assays

Roderick J. Little

Outline

- Discuss two related calibration topics where classical methods are deficient
- (A) “Limit of quantification” methods for analysis of calibration samples
  - Guo, Harel and Little (2010)
- (B) Regression analysis involving covariates with . measurement error
  - Guo and Little (2011a, 2011b)
- Proposed tools: simple Bayesian models, multiple imputation

LOQ methods

- Classical LOQ methods for analysis of calibration samples are deficient -- imply uncertainty of prediction is "too high to report" below LOQ, "zero" above LOQ.
- Model the Bayesian predictive distribution for the true concentration (x) given a measured value (y).
  - Classical approach models y given x, the wrong distribution
  - Our model allows for non-constant variance of measurement errors
  - Provides a better quantification of the uncertainty
- Apply to calibration data for fat soluble vitamins.
  - Suggests problems with the classical approach, for values both above and below the LOQ
- Discuss some implications for analyses involving estimates of X

Vitamin calibration data

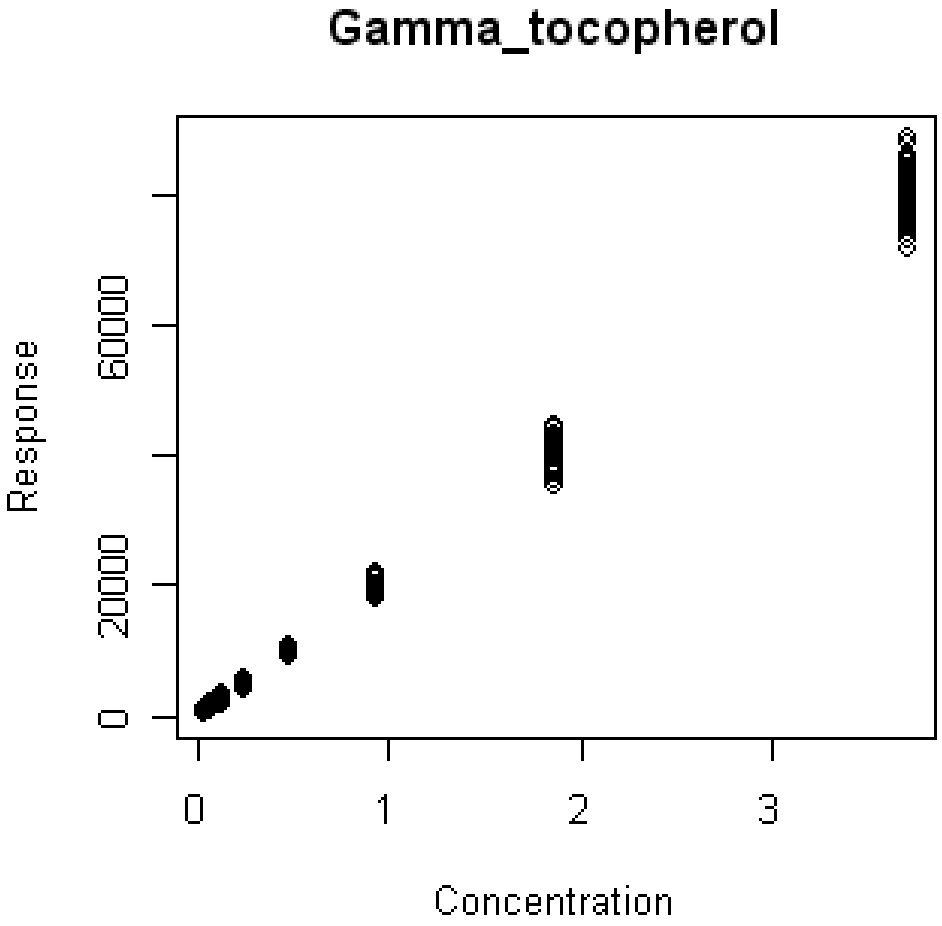
$y_{ij} =$ measured value for target concentration i, replicate j

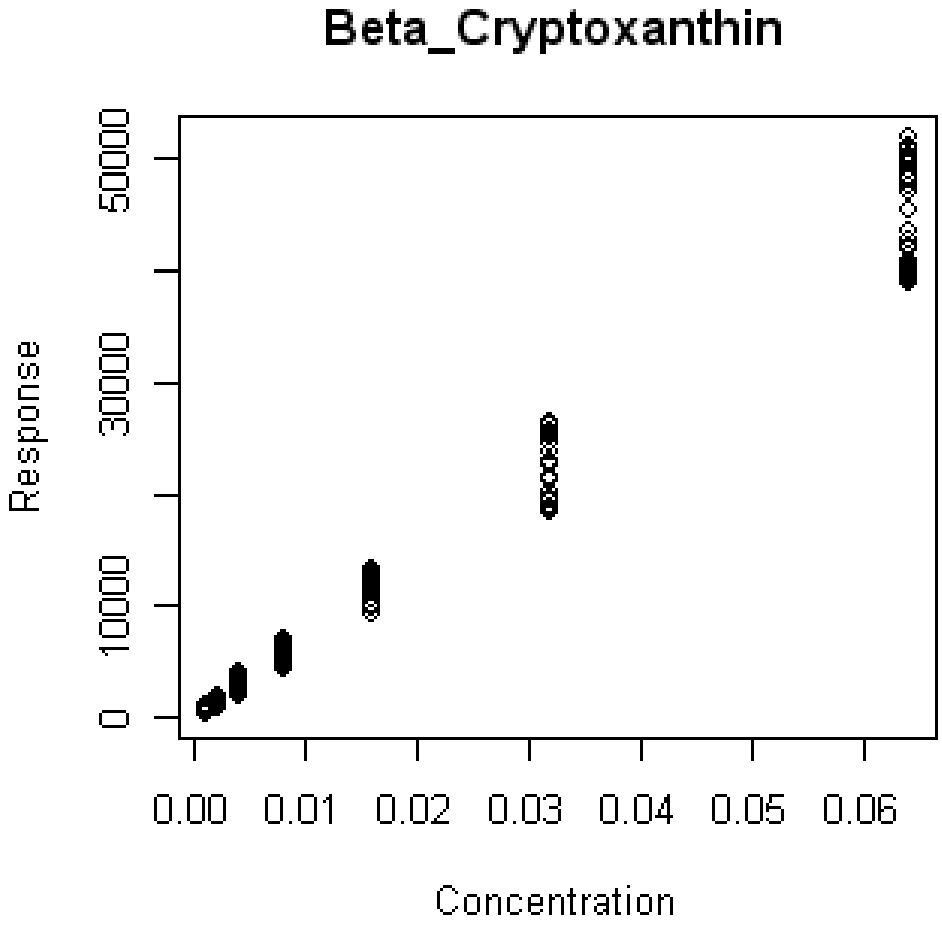
$x_i =$ known concentration of the substance

$i = 1, ..., I$ (here $I = 8$),

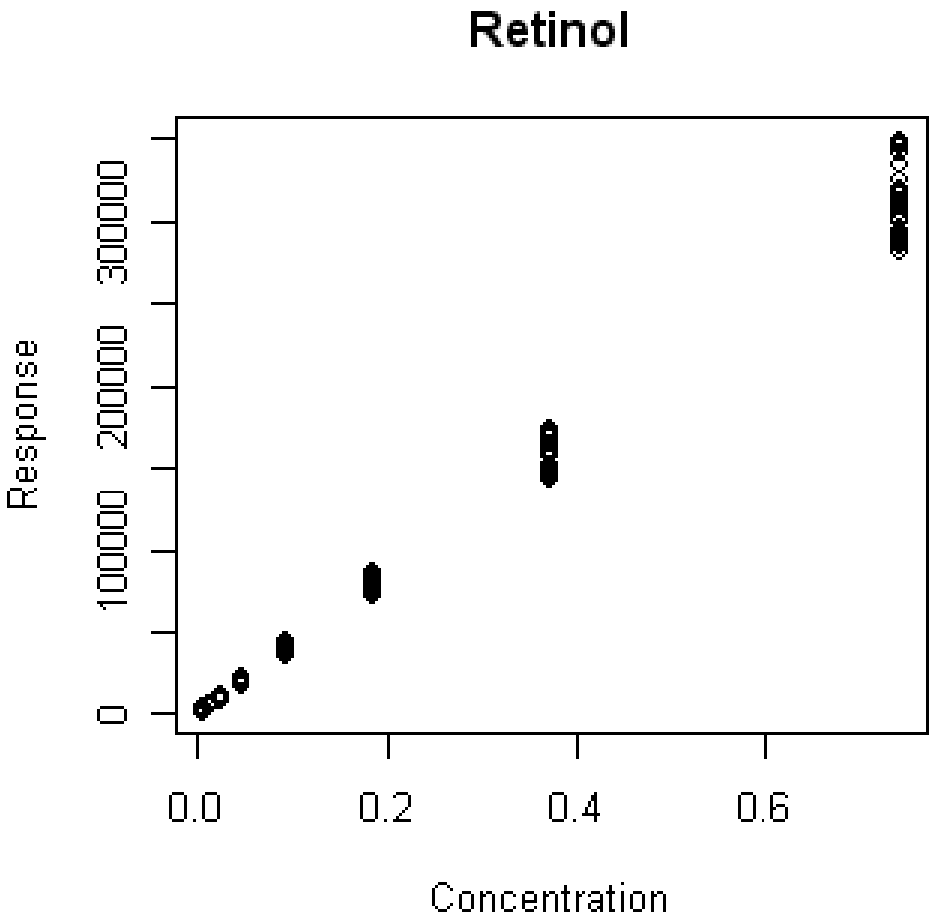
$j = 1, ..., J$ (here $J = 30$, pooled across 3 experiments)

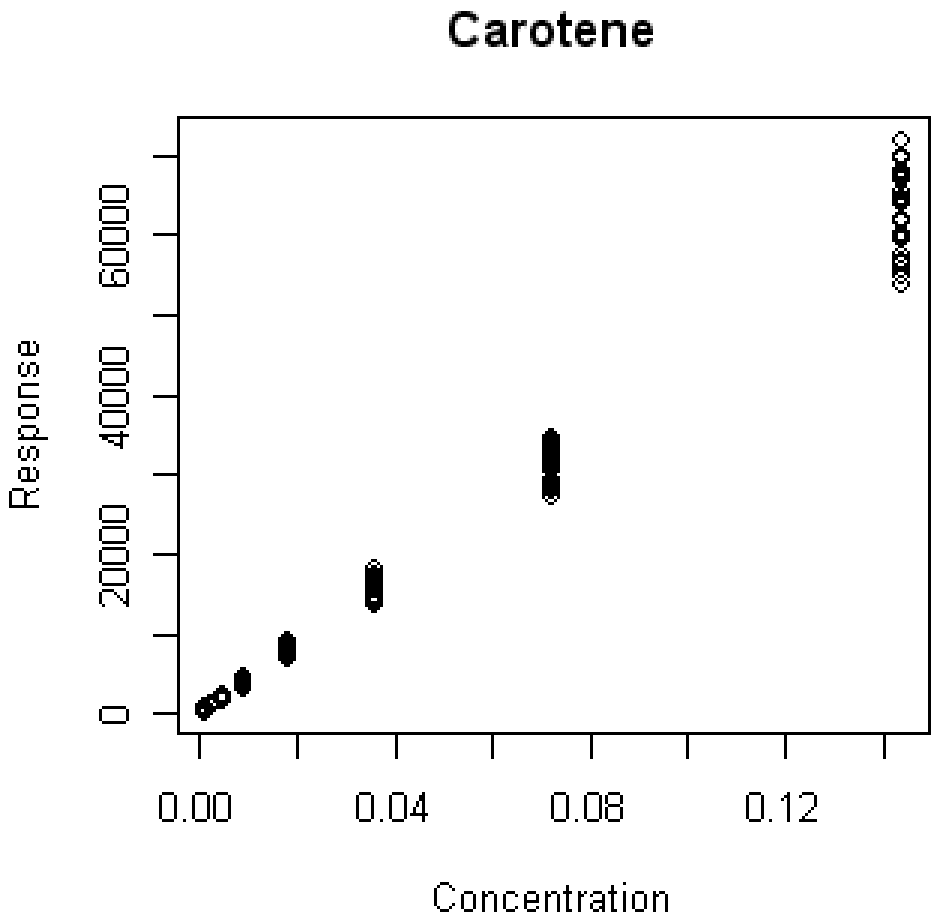
Data for four vitamins





Data for four vitamins





A conventional analysis

For each known target value of substance x_i,

$(\bar{y}_i, s_i, cv_i) =$ mean, sd, cv of measurements on Y

Least squares linear regression of Y on X yields

$E(y_{ij} \mid x_{ij}) = \hat{\beta}_0 + \hat{\beta}_1 x_{ij}$

Calibration prediction of X given Y : $\hat{x} = (y - \hat{\beta}_0) / \hat{\beta}_1$

LOQ: Estimate CV's of Y at each X

Estimated LOQ is the value of X where CV = 0.2

Then if y^* is a future measured value, the reported value of X is

$$x^* = \hat{x}(y^*) = \begin{cases} (y - \hat{\beta}_0) / \hat{\beta}_1, \text{ if } y^* > \text{LOQ} \\ \text{ND, if } y^* < \text{LOQ} \end{cases}$$

Proposed approach

Compute posterior predictive distribution of x^*

given measured value y^* and calibration data

$C=\{(x_i, y_{ij}), i=1,...,I; j=1,...,J\}$

using a calibration model for C

This predictive distribution can be used

(a) in a measurement error model, or

(b) to generate multiple imputations of X

Calibration model

$(y_{ij} \mid x_i) \sim N(\mu(x_i;\beta), \sigma^2 x_{ij}^{\alpha})$ $p(\beta, \alpha, \log\sigma^2) = const.$

$\mu(x_i;\beta) = \beta_0 + \beta_1 x_i$ (linear model) or

$\mu(x_i;\beta) = \beta_0 + \beta_1 x_{ij} + \beta_2 x_{ij}^2$ (quadratic model)

$x_i \sim \pi()$ = Prior distribution for X, uniform

(Similar results using a dispersed normal prior $N(0,1000)$)

Methods

For known parameters $\theta = (\beta_0, \beta_1, \alpha, \log \sigma^2)$,

predictive distribution of X given Y is

$$p(x^* \mid y^*, \theta) \propto p(x^*)\, p(y^* \mid x^*, \theta)$$

Given calibration data C, two approaches:

Empirical Bayes: $p(x^* \mid y^*, \hat{\theta}) \propto p(x^*)\, p(y^* \mid x^*, \hat{\theta})$, (1)

$\hat{\theta} =$ estimate of θ

Full Bayes: $p(x \mid y, C) \propto p(x) \times p(y \mid x, C)$, (2)

$$p(y \mid x, C) = \int p(y \mid x, \theta)\, p(\theta \mid C) d\theta$$

Predictions of x are draws from (1) or (2)

$(\hat{x}_L(y), \hat{x}_H(y)) = 95\%$ prediction interval for x given y

Methods continued

- Empirical Bayes (simple)

$\hat{\alpha} =$ slope of regression of $\log(y - \hat{y})^2$ on $\log X$, $\hat{y} =$ LS prediction

$(\hat{\beta}_0, \hat{\beta}_1, \hat{\sigma}^2)$ = weighted least squares estimates, $\text{weight}_i \propto x_i^{-\hat{\alpha}}$

  - Also computed ML estimates, similar results

- Full Bayes (more complex, propagates error in θ)

Prior distribution: $p(\theta) = const.$

Draws from posterior distribution using Gibbs' sampler with Metropolis step for α

Estimates of α for eight analytes, linear model.

Analytes	Residual Regression	ML	Bayes	
		Estimate	Post. Mean	95% HPD
Gamma tocopherol	0.56	0.56	0.56	(0.51,0.61)
Lutein	0.61	0.63	0.63	(0.58,0.68)
Alpha tocopherol	0.62	0.62	0.62	(0.57,0.67)
Delta tocopherol	0.63	0.65	0.65	(0.53,0.77)
Beta Cryptoxanthin (BC)	0.65	0.64	0.64	(0.58,0.71)
Lycopene	0.70	0.72	0.71	(0.67,0.76)
Retinol	0.71	0.68	0.67	(0.63,0.72)
Carotene	0.77	0.72	0.72	(0.67,0.77)

BC: ML and Bayes parameter estimates, linear model

Parameters	ML	Bayes	
	Mean	Mean	95% PI
Beta0	-0.0029	-0.0028	(-0.0115, 0.0056)
Beta1	0.7055	0.7054	(0.6893, 0.7216)
Sigma^2	0.0144	0.0149	(0.0123, 0.0181)
Alpha	0.6423	0.6425	(0.5778, 0.7066)

Intercept negative (evidence against linearity)

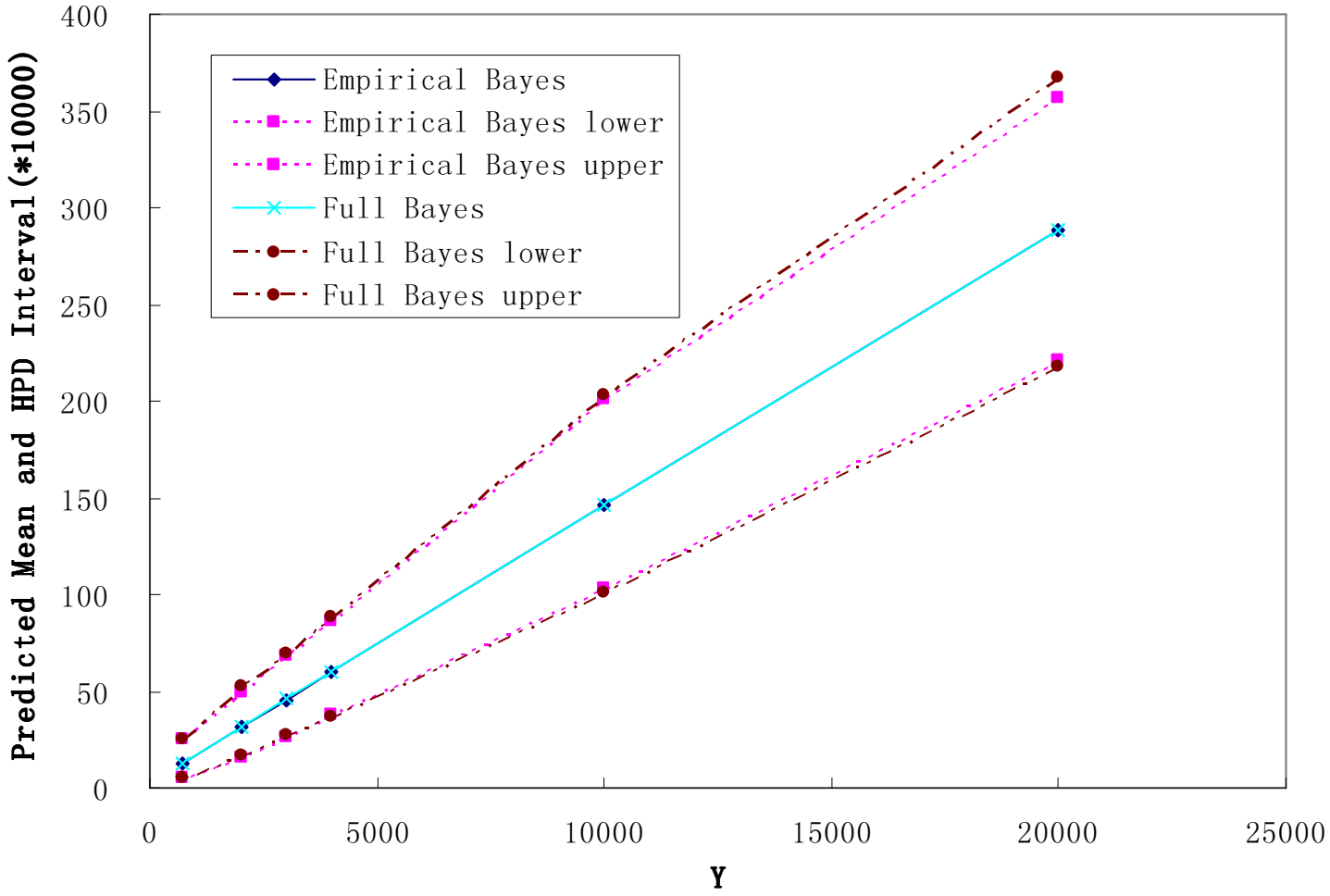
Alpha close to 0.64

Predictions of X* given Y*

Calibration	Full Bayes			
X*	Posterior mean	Posterior SD	Posterior CV	95% HPD interval
283.80	288.07	35.78	12.42	(218.00,367.00)
142.03	146.23	22.63	15.48	(101.00,203.00)
56.96	60.29	12.85	21.31	(36.55,88.75)
42.79	46.62	10.92	23.42	(27.20,69.50)
28.62	31.58	8.84	27.99	(16.75,52.25)
10.19	12.47	5.10	40.86	(4.75,25.25)

Beta Cryptoxanthin 986 C1 data (LOQ=61.89, LOD=13.1). Numbers are multiplied by 10,000

Figure 2. Predictions of the true values of *X* and 95% HPD interval, linear model, uniform assumption



Conclusions

- The mean of predictive distribution quite close but slightly higher than predictions using classical calibration.
- Error variance and width of the prediction intervals increases with X^*
- Substantial uncertainty in the predictions above the LOQ. For example, when $X^* = 142.0 * 10^{-4}$, a value more the twice the LOQ, the 95% HPD prediction interval for the linear model with normal prior is $(103 * 10^{-4}, 199 * 10^{-4})$
- There is considerable information concerning the values of X below the LOQ. For example, when $X^* = 28.6 * 10^{-4}$, a value less than half the LOQ, the 95% HPD prediction interval for the linear model with normal prior is $(17 * 10^{-4}, 49 * 10^{-4})$

  Conveys considerably more information than the statement that the value is less than LOQ.

Discussion

- Current approach is based on the CV of the measurement error of the regression of *Y* on *X* – when the CV is above the cutoff, uncertainty is essentially ignored, and when the CV is below the cutoff uncertainty results in a value not being reported.
- Two problems with this perspective.
  1. The error concerns the predictive distribution of *X* given *Y*, not the predictive distribution of *Y* given *X*, estimated in the conventional approach. Our Bayesian approach is focused on the right conditional distribution.
  2. The SD of the predictive distribution of *X* is more relevant to distortions in statistical analysis than the CV.

     -- SD directly determines distortion of marginal distribution of *X*

     -- Measurement error models involve the SD, not the CV.

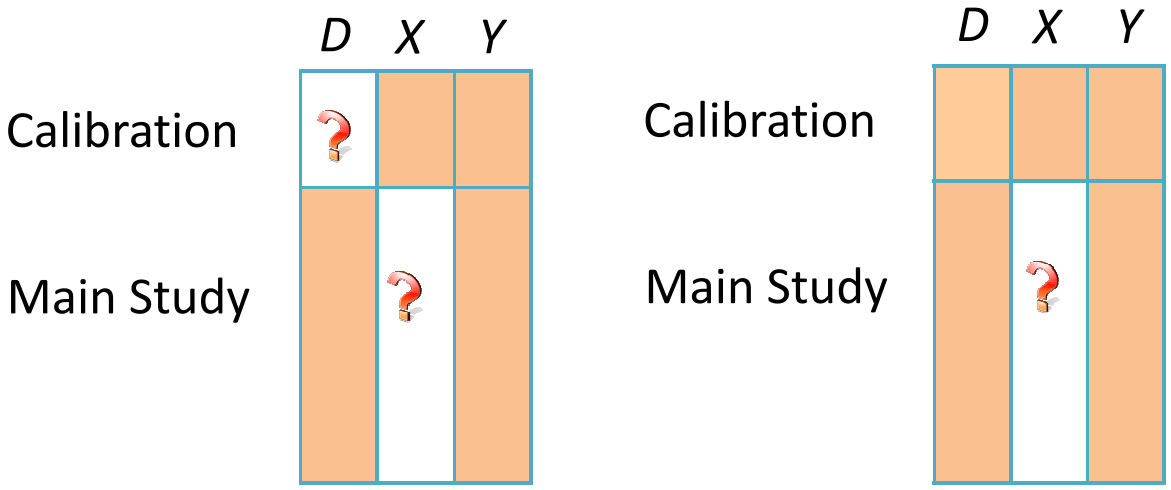
Discussion (contd)

- Logic behind reporting LOQ values is flawed
- Is a new paradigm in how these values should be reported needed?
- Our Bayesian model yields predictive distributions for the true values of the analyte that properly reflect uncertainty

Outline

- (A) "Limit of quantification" methods for analysis of calibration samples
  - Guo, Harel and Little (2009)
- (B) Regression analysis involving covariates with measurement error
  - Guo and Little (2011a, 2011b)
- Tools: simple Bayesian models, multiple imputation
- Compare with classical methods

(B) Regression on covariates with heteroscedastic measurement error



(a) External calibration (b) Internal calibration

X: covariate of interest but unobserved

Y: observed error-prone measurement related to X

D: response variable.

Measurement Error Model

This model links unobserved covariate *X* with error-prone measurement *Y*, considering potentially nonlinear mean functions and heteroscedastic measurement error

$$(y_i \mid x_i, \theta) \sim_{\text{ind}} N(\mu(x_i;\beta), \sigma^2 g(x_i;\beta,\alpha))$$

with $\theta = (\beta, \sigma^2)$, the function g to model heteroscedasticity. specifically, we assume that

$$\mu(x_i;\beta) = \beta_0 + \beta_1 x_i$$

$$g(x_i;\beta,\alpha) = x_i^{2\alpha}$$

Analysis Model

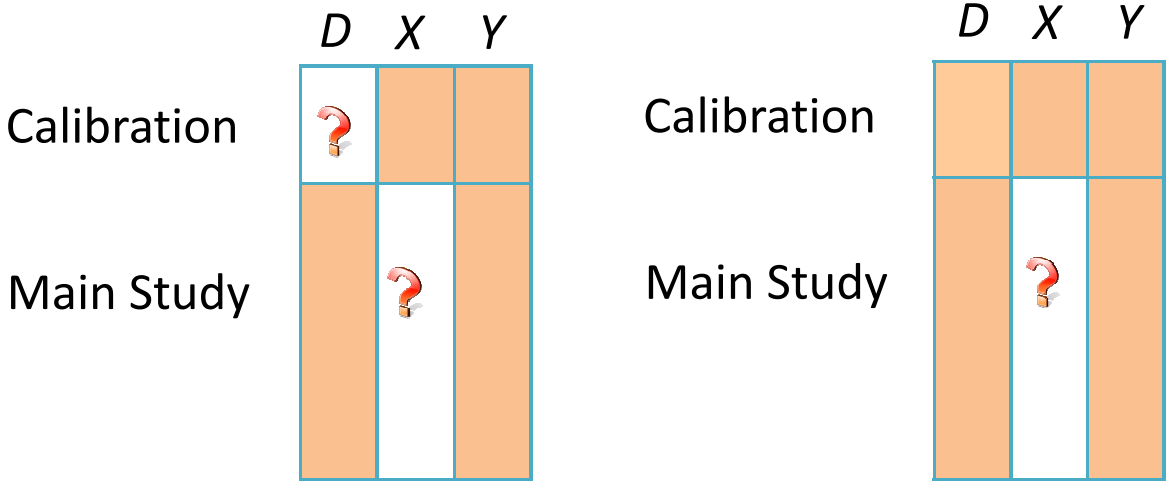
This model links unobserved covariate X with outcome D. For simplicity we assume the model

$$(d_i \mid x_i, \phi) \sim_{\text{ind}} N(\gamma_0 + \gamma_1 x_i, \tau^2)$$

where $\phi = (\gamma_0, \gamma_1, \tau^2)$, although more generally nonlinear relationships between Y and X can be modeled.

Our aim is to estimate the unknown regression parameters, taking into account the measurement error in X.

(B) Non-differential measurement error (NDME) assumption



(a) External study (b) Internal study

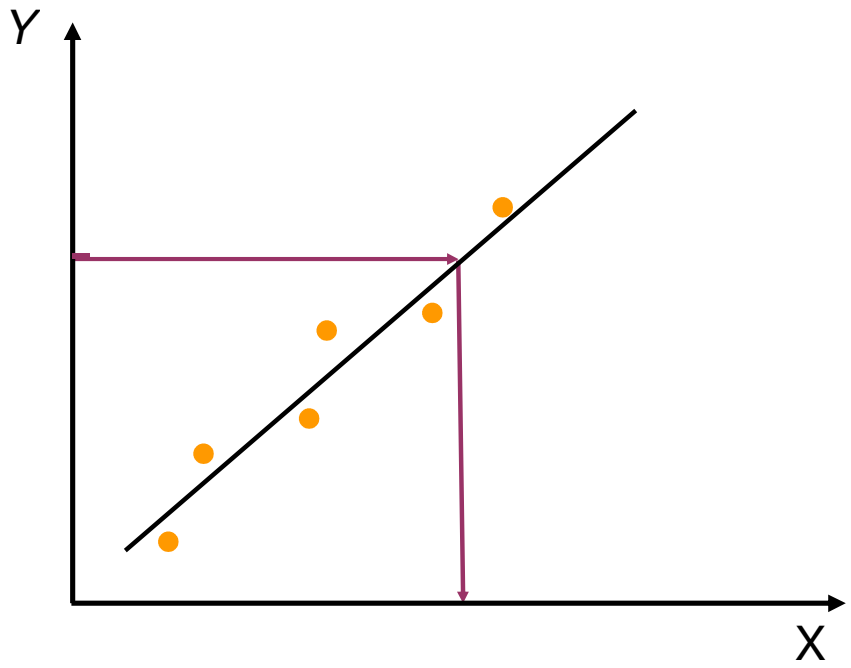
Non-differential measurement error: D is independent of Y given X

With external calibration, assumption is needed to identify parameters

With internal calibration, assumption is not needed but improves efficiency if assumed and true

Conventional Calibration (CA)

- fits an appropriate curve to the calibration data
- estimates the true value of *X* by inverting the fitted calibration curve (usually assumes a linear association)
- Regresses *D* on calibration estimates of *X*



Regression Calibration: RC

Estimates the regression of X on Y using the calibration data

Replaces the unknown values of X in main study with predictions

$$\hat{X} = E(X \mid Y)$$

Simple and easy to be applied (Carroll and Stefanski,1990)

Standard errors: Asymptotic formula or bootstrap both data sources

"Efficient" Regression Calibration (ERC)

When the internal calibration data is available, direct estimates of the regression D on X are available from the calibration sample

These can be combined with the RC estimates, weighting the two estimates by their precision

Spiegelman et al. (2001) call this efficient regression calibration

Multiple Imputation (MI)

Multiply impute all the values of X using draws from their predictive distribution given the observed data.

We develop MI methods based on a fully Bayes model and

$$p(X \mid D, Y, \beta, \sigma, \alpha, \gamma, \tau) \propto p(D, Y \mid X, \beta, \sigma, \alpha, \gamma, \tau)\, p(X, \beta, \sigma, \alpha, \gamma, \tau)$$

$$\propto p(D \mid X, Y, \beta, \sigma, \alpha, \gamma, \tau)\, p(Y \mid X, \gamma, \tau)\, p(X, \beta, \sigma, \alpha, \gamma, \tau)$$

$$\propto \underbrace{p(D \mid X, \beta, \sigma, \alpha, \gamma, \tau)}_{\text{Main study model}} \underbrace{p(Y \mid X, \gamma, \tau)}_{\text{measurement error model}} \underbrace{p(X, \beta, \sigma, \alpha, \gamma, \tau)}_{\text{prior distribution}}$$

Comparisons with constant measurement error variance

- Freedman et al. (2008) evaluate the performance of CA, RC, ERC and MI for the case of internal calibration data.
- CA biased, ERC better than RC, MI
- But: ERC assumes non-differential measurement error, and MI based on a model that does not make this assumption
  - This accounts for superiority of ERC (Guo and Little 2011b)
- A limitation of their work is that it assumes the variance of the measurement errors is constant. As discussed, in many real applications, the variance of the measurement error increases with the underlying true value.
- We compare methods when measurement variance is not constant

Weighted Regression Calibration (WRC)

An alternative to RC, taking into account heteroscedastic measurement error. We reformulate the measurement error model as

$$(x_i \mid y_i, \eta, \pi, \lambda) \sim_{ind} N(\eta_0 + \eta_1 y_i, \pi^2 y_i^{2\lambda})$$

- estimate λ as the slope of a simple regression of logarithm of the squared residuals of the regression on X on Y on the logarithm of the squared Y using the calibration data.
- estimating η_0 and η_1 by weighted least squares.
- substituting unknown values X in main study with estimates,

$$\hat{X}_{WRC} = \hat{\eta}_0 + \hat{\eta}_1 Y$$

Multiple Imputation (MI)

MI is applied with a measurement error model that incorporates non-constant variance of form $\sigma^2 x_i^{2\alpha}$.

(a) Full posterior distribution requires Metropolis step for draws of α

(b) Approximation based on weighted least squares avoids Metropolis step

Prior distributions

Noninformative prior distributions for the marginal distribution of X and the parameters. Specifically, we assumed

$$p(X,\beta,\sigma,\alpha,\gamma,\tau) = p(X)\,p(\beta,\sigma,\alpha,\gamma,\tau)$$

where the prior distribution of X is normal with mean 0 and variance 1000, and

$$p(\beta,\log(\sigma),\alpha,\gamma,\log(\tau)) = const., \quad -2 < \alpha < 2$$

Simulation Study

- Factors varied:
  - study design: external calibration and internal calibration
  - measurement error size
  - outcome-covariate relationship
- Twelve simulation scenarios were generated by combining the following choices of parameters:
  - analysis model: $\beta_0 = 0, \beta_1 = 0.3$ or 0.75
  - measurement error model:

$$\gamma_0 = 0, \gamma_1 = 0.6 \text{ or } 0.8$$

$$\text{Var}(Y \mid X) = 0.5, 0.8 \text{ or } 1;\ \alpha = 0.4$$

Main study sample size = 400, calibration data sample size = 80

Results: internal calibration

Table 1. Empirical means of the estimator of γ_x with internal calibration data based on 500 simulations when the variances of measurement error is heteroscedastic. Empirical standard deviation is given in parenthesis.

β_1	σ	γ_x	CA	RC	WRC	WERC	CalibRC	MI
0.8	0.5	0.3	0.136(0.039)	0.313(0.100)	0.310(0.095)	0.295(0.087)	0.305(0.116)	0.303(0.090)
0.8	0.8	0.3	0.074(0.030)	0.334(0.170)	0.326(0.143)	0.288(0.096)		0.295(0.095)
0.8	1	0.3	0.052(0.026)	0.372(0.528)	0.351(0.343)	0.281(0.105)		0.296(0.101)
0.8	0.5	0.75	0.341(0.055)	0.784(0.164)	0.782(0.152)	0.742(0.102)	0.754(0.118)	0.748(0.097)
0.8	0.8	0.75	0.186(0.047)	0.842(0.604)	0.823(0.463)	0.734(0.108)		0.747(0.103)
0.8	1	0.75	0.131(0.041)	0.945(1.843)	0.875(0.931)	0.728(0.113)		0.744(0.106)
0.6	0.5	0.3	0.098(0.033)	0.326(0.129)	0.317(0.121)	0.303(0.097)	0.312(0.117)	0.305(0.093)
0.6	0.8	0.3	0.049(0.025)	0.358(1.077)	0.329(0.520)	0.277(0.114)		0.293(0.100)
0.6	1	0.3	0.031(0.021)	0.424(2.111)	0.360(1.136)	0.275(0.118)		0.295(0.104)
0.6	0.5	0.75	0.248(0.053)	0.819(0.198)	0.811(0.185)	0.745(0.105)	0.748(0.118)	0.756(0.101)
0.6	0.8	0.75	0.117(0.041)	0.898(1.637)	0.870(0.988)	0.732(0.118)		0.743(0.105)
0.6	1	0.75	0.081(0.033)	1.110(3.265)	0.894(1.428)	0.726(0.123)		0.741(0.109)

Confidence coverage

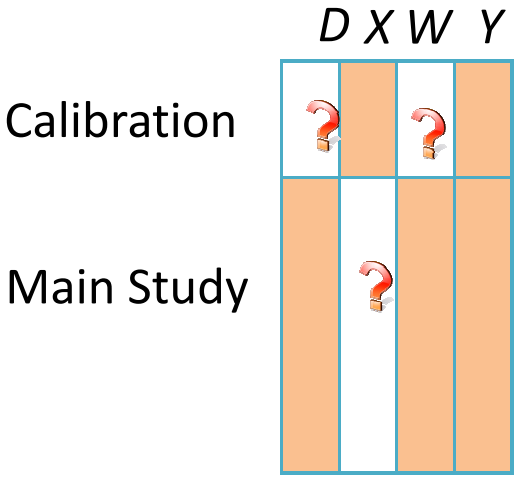
Table 2. Coverage of 95% confidence interval of the estimator of γ_x with the internal validation calibration data based on 500 simulations.

β_1	σ	γ_x	CA	RC	WRC	WERC	MI
0.8	0.5	0.3	3.0	71.6	72.4	76.4	94.5
0.8	0.8	0.3	0.2	77.2	80.4	80.2	94.5
0.8	1	0.3	0.0	80.2	82.4	82.6	94.1
0.8	0.5	0.75	0.0	89.2	90.6	90.0	94.0
0.8	0.8	0.75	0.0	90.0	90.0	90.2	93.8
0.8	1	0.75	0.0	89.2	90.8	89.6	93.5
0.6	0.5	0.3	0.0	74.2	78.2	78.6	94.3
0.6	0.8	0.3	0.0	80.0	81.4	81.0	94.5
0.6	1	0.3	0.0	83.4	83.0	82.6	94.0
0.6	0.5	0.75	0.0	90.0	91.0	91.0	94.2
0.6	0.8	0.75	0.0	89.6	90.6	90.0	94.0
0.6	1	0.75	0.0	90.8	91.4	91.0	93.5

Conclusions

- Convention approach is very biased and the estimate is attenuated when measurement error increase.
- Regression calibration works poorly in presence of heteroscedastic measurement error.
- Multiple imputation yields satisfactory results with small biases and good coverage. Shows flexibility of model-based methods
- Similar findings for external calibration, under NDME

External calibration with additional covariates, homoscedastic measurement error



(a) External calibration

Guo & Little (2011) provide simple MI algorithm based on summary statistics from calibration sample

X: covariate of interest but unobserved

Y: observed error-prone measurement related to *X*

D: response variable

W: other covariates.

Summary

- Bayes requires a prior distribution, but models the predictive distribution of interest.
- Multiple imputation – a useful approach for measurement error as well as nonresponse
  - Allows features such as nonconstant measurement error to be reflected in the inference.

References

- Carroll, RJ and Stefanski, L (1990). Approximate quasi-likelihood estimation in models with surrogate predictors. *JASA*, 85, 652-663.
- Freedman LS, Midthune D, Carroll RJ, and Kipnis V (2008). A comparison of regression calibration, moment reconstruction and imputation for adjusting for covariate measurement error in regression. *Stat. Med.* 27(25), 5195-5216.
- Guo*, Y., Harel, O. & Little, R.J. (2010). How well quantified is the limit of quantification? *Epidemiology*, 21, 4, S10-S16.
- Guo*, Y. & Little, R.J. (2011a). Regression analysis involving covariates with heteroscedastic measurement error. To appear in *Stat. Med.*
- Guo*, Y. & Little, R.J. (2011b). Multiple imputation for covariate measurement error based on summary statistics from an external calibration sample. In preparation.
- Spiegelman D, Carroll RJ, Kipnis V. (2001). Efficient regression calibration for logistic regression in main study/internal validation study designs with an imperfect reference instrument. *Stat. Med.* 20(1), 139-160.

and thanks to my recent students…

Hyonggin An, Qi Long, Ying Yuan, Guangyu Zhang, Xiaoxi Zhang, Di An, Yan Zhou, Rebecca Andridge, Qixuan Chen, Ying Guo, Chia-Ning Wang, Nanhua Zhang