eAppendix 1: Bias in the "Classical Calibration" estimator

Case 1. The classical measurement error

Under the classical measurement error model, $W = X + e$, where e is random unbiased error, $\beta_0 = 0,\ \beta_1 = 1$, and $\hat{X}_{CC} = (W - \hat{\beta}_0)/\hat{\beta}_1$, where β_0 and β_1 are the coefficients for the regression of W on X. Then, $\hat{X}_{CC} \xrightarrow{P} W$ since $\hat{\beta}_0 \xrightarrow{P} 0$ and $\hat{\beta}_1 \xrightarrow{P} 1$, and the estimator based on $\hat{X}_{CC}$ should converge to the naïve estimator. In the simulations investigating sensitivity to the multivariate normality assumptions intrinsic to Guo et al.'s method (Tables 2 and 3), the exact classical measurement error model was used. As expected by theory, results from the classical calibration estimator were virtually identical to those from the naïve.

Case 2: The linear measurement error

In the linear measurement error model, $W = \beta_0 + \beta_1 X + e$, where e is random unbiased error, the measurement error e has mean 0 and is uncorrelated with X. In addition, $\beta_1 = Cov(W, X)/Var(X) = \rho\sigma_W/\sigma_X$, and $\beta_0 = \mu_W - \rho\mu_X\sigma_W/\sigma_X$. Hence, the OLS estimator $\hat{\beta}_0 \xrightarrow{P} \mu_W - \rho\mu_X\sigma_W/\sigma_X$ and $\hat{\beta}_1 \xrightarrow{P} \rho\sigma_W/\sigma_X$. Similarly, $\hat{\alpha}_0 \xrightarrow{P} \mu_X - \rho\mu_W\sigma_X/\sigma_W$ and $\hat{\alpha}_1 \xrightarrow{P} \rho\sigma_X/\sigma_W$. Here, $E(X) = \mu_X, E(W) = \mu_W, Var(X) = \sigma_X^2, Var(W) = \sigma_W^2\ and\ Corr(X, W) = \rho$. So, it's clear that $1/\hat{\beta}_1$ will not be equal to $\hat{\alpha}_1$ unless the correlation between X and W equals 1. The same reasoning can be applied to $\hat{\beta}_0/\hat{\beta}_1$ and $\hat{\alpha}_0$. Hence, the "classical calibration" estimator based on $\hat{X}_{CC} = (W - \hat{\beta}_0)/\hat{\beta}_1$ in Guo et al.'s will be biased unless there is no measurement error.

eAppendix 2: Identifiability of Guo et al.'s method and their surrogacy condition.

When assuming $f(Y, X, Z|W)$ is multivariate normal with linear means and a constant variance-covariance matrix under the more restrictive surrogacy assumption used by the authors than is usually made in the measurement error literature, i.e. that $f(Y, Z|X, W) = f(Y, Z|X)$, the observed data likelihood is identifiable. It can be shown after some algebra that the assumption $f(Y, Z|X, W) = f(Y, Z|X\ )$ implies $f(Y|X, Z, W)/f(Y|X, Z) = f(W|X)/f(W|X, Z)$. Since under the standard surrogacy assumptions utilized throughout the measurement error literature, the left hand side equals 1, the condition utilized by Guo et al. is that, in addition to standard surrogacy, measurement error is independent of Z given X. This assumption is not required by regression calibration but as Guo et al. discuss, it may be reasonable in some settings such as the bioassay example motivating their work.