Measurement Errors

Graham Dunn, Biostatistics Group
School of Community Based Medicine

graham.dunn@manchester.ac.uk
General Aims of the Course

• To explain the concept and importance of the quantitative evaluation of measurement errors.
• To introduce the student to statistical models for measurements and, in particular, the estimation of reliability and the standard error of measurement (repeatability).
• To introduce statistical models and estimation techniques that allow for the presence of measurement errors.
Key Topics to be Covered

Four two-hour sessions:

• Measurement error and its implications for inferences concerning causal effects
• Measurement models and reliability estimation
• Linear errors-in-variables models
• Adjusting for measurement error in generalised linear models
Key Objectives 1

ATTITUDES
The acquisition of theoretical knowledge and associated technical skills should be motivated by the following changes in orientation of the student:

• To move away from the treatment of exposure and/or clinical measurements as if they were infallible. To realise that all assessments are fallible and that some are more fallible than others!

• To always question assumptions behind a given set of analyses. To ask how measurement errors might lead to misleading conclusions.
Key Objectives 2

KNOWLEDGE
After working through this course the learner should:

• Be able to explain the basic concepts of the classical measurement model (truth and error, true values, reliability, repeatability standard deviation/standard error of measurement, and so on).

• Be able to describe the various sources of measurement error and the design of reliability/validity studies to evaluate them.

• Know how to approach the problem of allowing for measurement errors in an analysis of exposure-response relationships.
Key Objectives 3

SKILLS
After the completion of working through this course the student should have acquired the following skills:

• Estimate, manipulate and use reliability coefficients in measurement error modelling.
• Demonstrate the impact of exposure measurement errors in simple exposure-response models.
• Use methods-of-moments methods to estimate parameters of simple linear measurement error models (as in those for method-comparison studies, for example).
• Be familiar with the use of instrumental variable methods, regression calibration, simulation extrapolation (SIMEX) and maximum likelihood to allow for exposure measurement errors in exposure-response modelling.
Assessment

100% by examination (25% of marks for the Advanced Biostatistics Module as a whole).

N.B. There will be no formally-marked in-course assessment.
Teaching Methods

We will have four two-hour sessions, each held in the computer lab.

Each two-hour session will be an informal mix of lecturing, discussion and computer-based exercises.
Resource Materials

• You will be given a hard copy of my Powerpoint notes at the beginning of each session.

• You will also receive a hard copy of my guide called “The Statistical Evaluation of Errors in Exposure Measurement”. You are encouraged to read through this material and to try the exercises yourself.

• A CD containing various papers, data files and other support materials for the use of Stata.
Web Resources

http://www.stata.com/merror/
Stata software for generalised linear measurement error models: qvf, rcal, simex and simexplot. Includes slides of a short course by Carroll & Ruppert (2003) - Highly recommended. (also available on the CD).

Supporting material from book by Carroll et al. (including software scripts & data)
Web Resources: gllamm
Generalized Linear and Latent Mixed Models

http://www.gllamm.org

For the gllamm cme wrapper:

http://www.gllamm.org/wrappers.html

Maximum likelihood for generalized linear models (and more) with errors in covariates.
ISO 5725-1:1994 (Free download)
Accuracy (trueness and precision) of measurement methods and results – 
Part 1: General principles and definitions
(there are several other parts of this Standard also available)
http://www.iso.org/iso/catalogue_detail.htm?csnumber=11833
The following textbooks are available for 48-hour loan from my PA, Wendy Lamb, in the Biostatistics Group (1st Floor, Jean McFarlane Building – ext. 55764 or wendy.lamb@manchester.ac.uk).


Clinical Measurements (examples)

- Body temperature
- Blood pressure
- Obesity
- Urine glucose concentration
- Blood cholesterol concentration
- Specific hormone or antibody levels
- Tumour volume (CAT scans or MRI)
- Regional brain blood flow (MRI)
- Severity of pain
- Depth of depression
- Quality of life
- Diagnosis
Exposure Measurements

Demographic, educational and social variables
Medical history
Reproductive and sexual history
Intake of illicit drugs, alcohol and tobacco
Diet and body size
Physical activity and fitness
Psychosocial variables (stress, depression etc.)
Occupation (coal miner; nuclear power worker; dentist?)
Radiation (environmental pollution, occupation, X-rays)
Environmental pollutants (pesticides, nitrates, carbon particles)
Distinction not always clear-cut

Hypertension (high blood pressure) or severe obesity could be seen as an indicators of a clinical conditions in their own right or they could be regarded as indicators of risk of, say, more serious conditions such as heart attack and stroke.
**Measurement Instruments**

*(White et al.)*

A *measurement instrument* is a procedure or set of procedures designed to measure one or several of the variables of interest in an epidemiological study.

**Examples:**

- Self-administered questionnaires
- Personal interviews
- Biochemical analysis of blood or other biological specimens
- Physical or chemical analysis of the environment
Measurement Error

The difference between true exposure and observed exposure (N.B. this difference is never known – but we can learn about its typical/average value and variability).

*Validity* refers to the capacity of an exposure variable to measure the true exposure in a population of interest.

*Precision* and *reliability* refer to the reproducibility of results. A precise/reliable instrument may not be valid (it could be measuring the wrong thing).
Measurement errors depend on context (what are we trying to measure?)

Consider intake of dietary fibre.

We may (or may not) be able to measure intake in a given 24 hour period with high precision (low error).

But if our aim is to measure average fibre intake over long period of time then day-to-day variability in intake will make the above a very imprecise measure of the characteristic we’re interested in.

Measurement error = Sampling error + Instrumental error
Differential measurement errors

Occur when measurement error differs according to the disease or outcome being studied.

Examples:

Recall bias in a case-control study – patients with a particular form of cancer, for example, may be more likely than controls to recall exposure to an exposure thought to be a potential risk factor.

Lack of blind outcome assessment in a randomised controlled trial of an antidepressant drug.
Notation:
An apology to the mathematicians

Usually, I will be a bit sloppy/informal.

I will frequently not bother with subscripts – assuming that the context will be obvious.

I will not explicitly distinguish between parameters and their estimated values – again, assuming that the context will make the distinction obvious.
A simple measurement model for a quantitative exposure

Observation $=\text{Truth + Bias + Random Error}$

$$W = X + B + U$$

(B is fixed; $X$ and $U$ are \textit{uncorrelated} random variables)

Total Error $= B + U = W - X$

$\text{Var}(X) = \sigma_x^2$; $\text{Var}(U) = \sigma_u^2$

It follows that

$$\sigma_w^2 = \sigma_x^2 + \sigma_u^2$$
Differential measurement errors

The following model now holds:

\[ W_D = X + B_D + U_D \]

\[ W_N = X + B_N + U_N \]

With either \( B_D \neq B_N \) and/or \( \sigma_{u_D}^2 \neq \sigma_{u_N}^2 \)
Linear regression: the effect of a quantitative exposure on a quantitative outcome

Let’s first assume that exposure, X, is measured without error. We are interested in predicting the outcome, Y, from a knowledge of exposure, i.e. the relationship

$$E(Y) = \alpha_x + \beta_x X$$

How do we estimate $\alpha_x$ and $\beta_x$?
Linear Regression

\[ E(Y) = \alpha_x + \beta_x X \]
\[ Y = \alpha_x + \beta_x X + V \]

\[ \text{Var}(X) = \sigma_x^2 \]
\[ \text{Var}(Y) = \sigma_y^2 = \beta_x^2 \sigma_x^2 + \sigma_v^2 \]
\[ \text{Cov}(Y, X) = \text{Cov}(Y, \alpha_x) + \text{Cov}(\beta_x X, X) + \text{Cov}(Y, V) \]
\[ = \text{Cov}(\beta_x X, X) = \beta_x \text{Var}(X) \]
\[ = \beta_x \sigma_x^2 \]

Therefore

\[ \beta_x = \frac{\text{Cov}(Y, X)}{\text{Var}(X)} \quad \text{(c.f. OLS)} \]
\[ \beta_x = \frac{\text{Cov}(Y,X)}{\text{Var}(X)} \]
\[ = \frac{\sigma_{xy}}{\sigma_x^2} \]

\[ \alpha = \mu_y - \beta_x \mu_x \]

\[ \sigma_e^2 = \sigma_y^2 - \beta_x^2 \sigma_x^2 \]
Introducing error into $X$

We are now interested in $E(Y) = \alpha_x + \beta_x X$

But have measured, say, $W = X + U$

We estimate $\beta_x$ from $\text{Cov}(Y, W)/\text{Var}(W)$ when, in fact, we should be using $\text{Cov}(Y, X)/\text{Var}(X)$
Introducing error into \( X \) (contd.)

The relative bias (\textit{attenuation}) in the estimator using \( X \) is

\[
\kappa = \frac{\text{Cov}(Y, W)}{\text{Var}(W)} / \frac{\text{Cov}(Y, X)}{\text{Var}(X)}
\]

(note \( \text{Cov}(Y, W) = \text{Cov}(Y, X) \))

\[
= \frac{\text{Var}(X)}{\text{Var}(W)}
\]

\[
= \frac{\text{Var}(X)}{\text{Var}(X) + \text{Var}(U)}
\]

This ratio has a range from 0 (very poor) to 1 (excellent) and is called the \textit{reliability ratio} or, simply, \textit{reliability}. 
Introducing error into X (contd.)

Effect on the slope

\[ \beta_w = \kappa \beta_x \]

Effect on the estimate of the intercept

\[ \alpha_w = \mu_y - \kappa \beta_x \mu_w \]
\[ = \alpha_x + (1-\kappa) \beta_x \mu_x \]

Effect on the estimate of the residual variance

\[ \sigma_{e*}^2 = \sigma_y^2 - \kappa^2 \beta_x^2 \sigma_w^2 \]
\[ = \sigma_e^2 + (1-\kappa) \beta_x^2 \sigma_x^2 \text{ (inflated)} \]
**ANCOVA with error in the Covariate**

We have two groups of subjects: Group A and Group B. Ideally, we are interested in the difference, $d_{AB|x}$, between the means for the two groups, after allowing for the linear effect of $X$ on $Y$:

$$d_{AB|X} = \mu_{yA} - \mu_{yB} - \beta_x(\mu_{xA} - \mu_{xB})$$

but, instead, we estimate

$$d_{AB|W} = \mu_{yA} - \mu_{yB} - \beta_w(\mu_{wA} - \mu_{wB})$$

where

$$\beta_w = \kappa \beta_x$$

Note that error in a covariate not only attenuates the regression coefficient for that covariate but also influences estimates of other effects. Here $d_{AB|W} > d_{AB|X}$ (i.e. the adjusted Group effect is increased, *not* attenuated).
Consider two groups: D and N, where \( P(D|X) + P(N|X) = 1 \).

\[
\ln\left[\frac{P(D|X)}{1 - P(D|X)}\right] = \ln\left[\frac{P(D|X)}{P(N|X)}\right] = \alpha_x + \beta_x X
\]

Let \( X \) be distributed as a Gaussian variate in each group with a common variance, \( \sigma_x^2 \). In addition, let the mean of \( X \) in Group D be \( \mu_{xD} \) and similarly, the mean of \( X \) in Group N be \( \mu_{zN} \). Let the relative proportions of subjects in Group D and Group N be \( \pi_D \) and \( \pi_N \), respectively, with \( \pi_D + \pi_N = 1 \).

Then it follows that

\[
P(D|X) = \frac{\pi_D}{(\sigma_x \sqrt{2\pi})}.\exp\left[-\frac{1}{2}(x - \mu_{xD})^2/\sigma_x^2\right]
\]
Logistic Regression (contd.)

Similarly,

\[ P(N|X) = \frac{\pi_N}{\sigma_x \sqrt{2\pi}} \cdot \exp\left[ -\frac{1}{2} (x - \mu_{xN})^2 / \sigma_x^2 \right] \]

It follows that

\[
\ln\left[ \frac{P(D|Z)}{P(N|Z)} \right] = \ln\left( \frac{\pi_D}{\pi_N} \right) + \frac{(\mu_{xD} - \mu_{xN})^2}{2\sigma_x^2} + x \cdot \frac{(\mu_{xN} - \mu_{xD})}{\sigma_x^2}
\]

which is equivalent to the initial logistic regression equation with

\[ \alpha = \ln\left( \frac{\pi_D}{\pi_N} \right) + \frac{(\mu_{xD} - \mu_{xN})^2}{2\sigma_x^2} \]

and

\[ \beta = \frac{(\mu_{xN} - \mu_{xD})}{\sigma_x^2} \]
Logistic regression with error in the covariate (1)

Now consider, $W = X + U$, with the same properties as before. Specifically, $\kappa = \sigma_x^2/(\sigma_x^2 + \sigma_u^2)$.

For the model
\[
\ln\left[\frac{P(D|W)}{1 - P(D|W)}\right] = \ln\left[\frac{P(D|W)}{P(N|W)}\right] = \alpha_w + \beta_w W
\]

We find that
\[
\beta_w = \frac{(\mu_{wN} - \mu_{wD})/\sigma_w^2}{\kappa}
\]

Since, $\mu_{wD} = \mu_{xD}$ and $\mu_{wN} = \mu_{xN}$, then it follows that
\[
\beta_w = \kappa \beta_x
\]
Logistic regression with error in the covariate (2) - X with differential measurement error

Now consider, \( W_D = B_D + X + U, \)

and \( W_N = B_N + X + U, \)

where \( B_D \) and \( B_N \) biases in the measurement processes for Groups D and N, respectively.

The Us are random measurement errors with the same properties as before. Specifically,

\[
\kappa = \frac{\sigma_x^2}{(\sigma_x^2 + \sigma_u^2)}.
\]
Logistic regression with differential errors (contd.)

For the model

\[
\ln\left(\frac{P(D|W)}{1 - P(D|W)}\right) = \ln\left(\frac{P(D|W)}{P(N|W)}\right) = \alpha_w + \beta_w W
\]

We find that

\[
\beta_w = \frac{(\mu_{wN} - \mu_{wD})}{\sigma_w^2} = \left[\frac{(\mu_{xN} - \mu_{xD}) + (B_N - B_D)}{(\sigma_x^2 + \sigma_u^2)}\right]
\]

and it follows that

\[
\beta_w = \left[ 1 + \frac{(B_N - B_D)}{(\mu_{xN} - \mu_{xD})}\right] \kappa \beta
\]
Logistic regression with differential errors (conclusion)

β_w is influenced by two factors:

(a) Its absolute size is reduced in proportion to the reliability, κ.

(b) (B_N - B_D) can be of any magnitude and either positive or negative!

...and when (B_N - B_D)=0 we get back to β_w=κβ_x.
Effects of non-differential measurement error on required sample size for given power 1-\( \beta \)

For an unmatched case-control study the number needed per group is given by

\[ N = 2\left( \frac{Z_{\alpha/2} + Z_\beta}{\sigma^2} \right)^2 \]

where \( d \) is the magnitude of the difference to be detected between the mean exposure of the cases and controls.

The increase in \( N \) due to measurement error will be \( \sigma_w^2/\sigma_x^2 \) (the reciprocal of the reliability ratio).
The same adjustment would apply to the use an error-prone outcome (Y) in say, a two-group randomised controlled trial.

Non-differential measurement error in this situation will not be a source of bias but will demand increases in sample size and corresponding costs.
Measurement error in a putative confounder

Let’s first assume that exposure is error-free but that the confounder (C) is subject to non-differential measurement error. Any adjustment for confounding will be incomplete. There will be residual confounding.

Adjustment for a poorly-measured confounder may be equivalent to no adjustment at all.
Measurement error in a putative confounder

Now let’s assume that we have an error-free confounder but exposure is subject to measurement error.

There will be attenuation of the effects of exposure on outcome and the attenuation can be greater than in an unadjusted analysis.
Measurement error in a putative confounder

If we have measurement errors in both exposure and confounder measurements then we have even more complex (and often unpredictable effects).

If we have several confounders, each subject to measurement error, then…
Summary and Concluding Remarks

Measurement error is practically universal.

It has implications for the estimation of causal effects – not just a matter of reducing power but an important source of bias.

It is important that we can quantify measurement errors, question their implications for the validity of our results and, if possible, rectify any biases that they might cause.
String lengths

Enter into Stata as:

id repeat string
1 1 7.6
2 2 8.2
3 1 13.6
4 2 11.1
. . . (i.e. 30 records)
Estimation of reliability

\textit{Stata command:}

\texttt{loneway string id}
### Results 1

One-way Analysis of Variance for string:

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Prob &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between id</td>
<td>852.08018</td>
<td>14</td>
<td>60.86287</td>
<td>18.04</td>
<td>0.0000</td>
</tr>
<tr>
<td>Within id</td>
<td>50.619316</td>
<td>15</td>
<td>3.374621</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>902.69949</td>
<td>29</td>
<td>31.127569</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Number of obs = 30
R-squared = 0.9439
### Results 2

<table>
<thead>
<tr>
<th>Intraclass correlation</th>
<th>Asy. S.E.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.89493</td>
<td>0.05232</td>
<td>0.79240 - 0.99747</td>
</tr>
</tbody>
</table>

- Estimated SD of id effect: 5.361355
- Estimated SD within id: 1.837014
- Est. reliability of a id mean: 0.94455
  (evaluated at n=2.00)
Reshaping the data

reshape wide string, i(id) j(repeat)

Produces:
id string1 string2

1  7.6  8.2
2  13.6 11.1

i.e. 15 records
Bland-Altman Plot

gen diff=string1-string2
gen ave=(string1+string2)/2
scatter diff ave

Could also have a look at:
scatter string1 string2
corr string1 string2