Evaluating bias and the associated uncertainty

Science
for a safer world

Overview

- Uncertainties associated with bias
- Estimating the effect of sample matrix on bias
- Including bias effects in the uncertainty budget

Evaluating bias
Issues to consider

- How to obtain a reasonable estimate of the bias
  - validation data (using CRMs or spiked samples)
  - EQA data?
- Is the bias significant?
  - statistically significant?
  - significant compared to the method precision?
- How to include bias in the uncertainty estimate

Obtaining an estimate of the precision is generally relatively straight forward. The true value for the material analysed to generate precision data doesn't need to be known - it is only the variation in results that is of interest. In addition, precision data generally provide 'Type A' uncertainty estimates (i.e. a standard deviation is calculated directly from experimental observations).

Evaluating bias is somewhat more complex. First, a suitable reference value is required. Second, we need to consider whether the observed bias is significant. Most bias estimates will be non-zero. However, when the precision of the measurement results and the uncertainty in the reference value are taken into account the observed bias may not be significantly different from zero. Finally we need to decide how to include any bias in the uncertainty estimate (e.g. how might we combine an estimate of precision with an estimate of bias?). It is this last point which is still open to debate. The ISO guide requires that results are corrected for all known significant biases. However, in many chemical analyses, results are not routinely corrected. Without the correction, a known bias is being ignored. If the uncorrected result is reported with its uncertainty, the range will not include the best estimate of the 'true value' – a simple report of the result and its uncertainty is therefore misleading.

Each of the above issues are discussed in the following slides.

Evaluating bias
Obtaining an estimate

- Need a reference value
  - certified reference material
  - in-house material prepared by spiking
  - reference method
- Bias scores from EQA
  - target values used to calculate scores often based on a consensus value obtained from participants results (and may be method specific)
  - gives an indication of bias compared to results produced by other laboratories (using the same method?)
- Empirical methods
  - method bias is zero (by definition)
  - need to asses laboratory bias

As mentioned previously, evaluation of bias requires a reference value – the different options are shown on the slide.

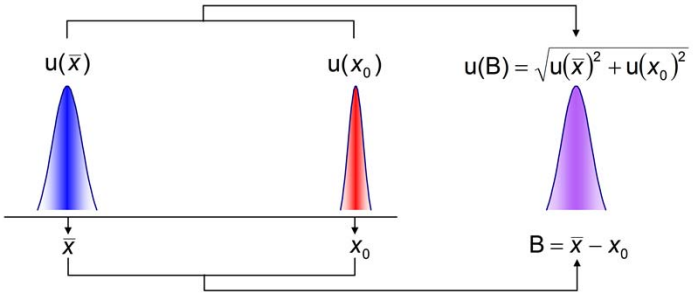
Empirical methods are those which define the measurand. Dietary fibre, extractable cadmium in ceramics, or standard methods for microbiological assay, are examples. In these cases, the method is considered to define the true value, so is, by definition, unbiased. The only uncertainty that applies is that associated with the laboratory's particular execution of the method.

In some cases, a reference material certified for use with the method may be available. Where this is so, a simple bias check against that reference material can be made.

Where no relevant reference material is available, it is not possible to estimate the uncertainty associated with the laboratory bias. There will still be uncertainties associated with bias, but they will be associated with possible bias in the temperatures, masses etc. used to define the method. It will normally be necessary to assess these individually where no relevant CRM exists.

Uncertainties associated with bias
CRM or spiked sample

- There will be an uncertainty associated with any bias estimate



- Uncertainty in the reference value $u(x_0)$
- Uncertainty in the average of measured results obtained fro the CRM/spiked sample
- The bias and associated uncertainty are usually expressed as relative values

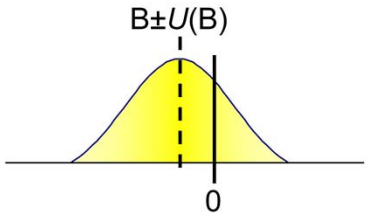
Bias is simply the difference between the mean of a number of observations and a reference value. As there will be uncertainties associated with both the mean of the observations and the reference value, there will also be an uncertainty associated with the bias estimate as shown on the slide.

For an unbiased method, the average bias, over a large number of measurements would be equal to zero. The approach to including bias in an uncertainty estimate will depend on whether the bias is significantly different from zero.

Estimating the uncertainties

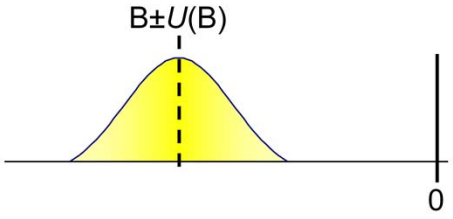
- Uncertainty in reference value
  - From CRM certificate – convert to standard uncertainty
  - Uncertainty in calculated concentration of spiked sample
  - Express as a relative value
- Uncertainty in average of results
  - Standard deviation of the mean of results for CRM or spike sample ($s/\sqrt{n}$)
  - Express as a relative value

Is there a significant bias?



The expanded uncertainty on the bias estimate includes zero

=> no significant bias



The expanded uncertainty on the bias estimate does not include zero

=> significant bias

The approach to handling bias in an uncertainty budget depends on whether or not it is significantly different from zero. In other words, is there really a significant bias, if the uncertainty in the estimate of the bias is taken into account.

To determine whether the bias is significant, the uncertainty in the bias estimate needs to be taken into account. To give the required level of confidence, the expanded uncertainty is used (typically at the 95% confidence level). If the range B±U(B) includes zero, the bias is considered not significant. A coverage factor of k=2 is generally used to calculate the expanded uncertainty when working at the 95% confidence level.

Bias calculation example
Creatinine in serum

Data

- CRM
  - Certified value (x_0): 449 µmol L^{-1}
  - Expanded uncertainty: 16 µmol L^{-1} (k=2.1, 95% confidence)
  - Standard uncertainty u(x_0): 16/2.1 = 7.62 µmol L^{-1}
- Experimental results
  - CRM analysed 10 times (n=10)
  - Mean ($\bar{x}$): 434 µmol L^{-1}
  - Standard deviation: 10.8 µmol L^{-1}
  - Standard deviation of the mean u($\bar{x}$): 10.8/$\sqrt{10}$ = 3.42 µmol L^{-1}

Bias calculation example
Creatinine in serum

Calculation

- Relative bias

$$\frac{434-449}{449} = -0.03$$

- Uncertainty – combine $u(x_0)$ and $u(\bar{x})$ as relative values

$$u(\mathrm{B})_{rel} = \sqrt{\left(\frac{3.42}{434}\right)^2 + \left(\frac{7.62}{449}\right)^2} = 0.0187$$

- Is the bias significant?

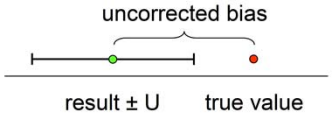
  -0.03±(2x0.0187)

  At 95% confidence, bias ranges from -0.067 to 0.0074

  Range includes zero => no significant bias

Including bias in uncertainty estimates (1)
Approaches in chemical analysis

- Insignificant bias
  - assume bias is zero with an uncertainty, *u*(B)
- Significant bias
  - develop method to remove/reduce bias
  - correct results for known significant bias (ISO Guide)
    - include *u*(B) in uncertainty estimate for corrected results
  - correction uncommon in clinical analysis
  - but, uncertainty is a range which includes the true value…..



  - …so significant bias should not be ignored
  - options: report bias and its uncertainty separately OR increase reported uncertainty to take account of the bias

Once the bias and its uncertainty have been estimated, the next step is to consider how to include bias in the uncertainty calculation. The issue of how to address bias when evaluating measurement uncertainty has been the subject of much (ongoing) debate. The current approaches used in analytical chemistry are outlined below.

If the bias is not significant, the bias is assumed to be equal to zero with an uncertainty *u*(B). The uncertainty associated with the bias estimate is included in the uncertainty estimate.

What if the bias is significant? The ISO guide requires that results are corrected for all known significant biases. If results are corrected for a significant bias, the uncertainty associated with the correction factor is simply included in the uncertainty estimate. However, in many chemical analyses, results are not routinely corrected. Without the correction, a known bias is being ignored. If the uncorrected result is reported with its uncertainty, the range will not include the best estimate of the 'true value' – a simple report of the result and its uncertainty is therefore misleading. Since the ISO guide does not allow for uncorrected bias, what is the most sensible approach? One acceptable option is to evaluate the uncertainty for uncorrected results (i.e. excluding any uncertainty associated with the bias) and to report the bias and its uncertainty separately. Using this approach users of the results are not being misled and can correct results themselves if necessary.

Including bias in uncertainty estimates (2)

- If a separate report of bias is not appropriate
  - increase reported uncertainty by including a bias uncertainty term
  - bias term combined with precision using 'root-sum-of-squares' rule
- Different approaches proposed for estimating bias term
  - root mean square (RMS) of bias estimates
  - mean bias
  - bias divided by coverage factor, *k*
- Further information in the literature

Where a separate report of the bias is not feasible, some authorities recommend increasing the reported uncertainty so that users of data are not misled. A number of different approaches have been proposed in the literature:

Measurement uncertainty revisited: Alternative approaches to uncertainty evaluation, Eurolab Technical Report 1/2007, 2007 (available at www.eurolab.org)

NORDTEST Report TR 537, Handbook for calculation of measurement uncertainty in environmental laboratories (available from www.nordicinnovation.net).

M. Thompson, S. L. R. Ellison, A. Fajgelj, P. Willetts, R. Wood, *Harmonised guidelines for the use of recovery information in analytical measurement*, Pure & Appl. Chem, 71, 337-348, 1999.

Two papers have reviewed the different approaches:

B. Magnusson, S. L. R. Ellison, *Treatment of uncorrected measurement bias in uncertainty estimation for chemical measurements*, Anal. Bioanal. Chem., 390, 201-213, 2008.

G. E. O'Donnell, D. Bryn Hibbert, *Treatment of bias in estimating measurement uncertainty*, Analyst, 130, 721-729, 2005.

Covering the method scope – matrix effects

- Estimating bias by analysing a single reference material or spiked sample gives a bias estimate for that sample
- May also need to evaluate bias for
  - different analyte concentrations
  - different sample matrices
- If there is significant variation in bias, include in the uncertainty budget
  - standard deviation of relative bias estimates

The results from the analysis of a single reference material or spiked material do not give any information about the bias for differing sample types. If the method scope covers a range of sample types, the bias for the determination of the analyte in a representative range of samples therefore needs to be studied. In the absence of suitable reference materials, this is generally done through spiking studies. A simple approach is to spike a representative range of sample matrices at a representative range of analyte concentrations. The spiked samples are then analysed and the bias calculated in each case. The average bias across the different sample types is assumed to be equal to zero. The standard deviation of the relative biases represents the uncertainty in the bias due to variations in sample composition. If this is significant, compared to the other sources of uncertainty for the method, then it should be included in the uncertainty budget.

Bias uncertainty from EQA

- Reference value is assigned value for EQA material
- Limitations
  - assigned value often based on consensus of participants results
  - bias assessed against peers rather than independent traceable value (CRM value)
- For well established measurements assigned value *may* provide a good estimate of the true value (but not always the case)
- Better than nothing if CRMs unavailable or spiking not possible…but use with caution

The majority of laboratories participate in external quality assessment schemes. A key feature of such schemes is that laboratories receive a 'performance score' for each round. The scoring systems used vary from scheme to scheme, but they generally involve comparing individual laboratory's results with a 'target value'. Results from individual rounds provide a 'snapshot' of method performance and are of limited use when trying to evaluate any underlying bias in measurement results. However, if scores are monitored over a number of rounds, patterns can emerge which may indicate the presence of a bias (e.g. if scores consistently have the same sign). There is an ongoing debate about how best to use EQA data in the evaluation of uncertainty. Care needs to be exercised when interpreting results from EQA. A bias estimate obtained from EQA data is calculated against the target value which is set by the scheme organisers for each EQA material. In many schemes, the target value is a consensus obtained from the participants' results (after processing to reduce the influence of extreme values). In addition, the consensus may be set on a method by method basis (and may be different for different methods). The consensus may not provide a reliable estimate of the true value for the material (i.e. the consensus value may be biased). Individual laboratories are therefore effectively comparing their results with those of their peers rather than with a traceable reference value. However, for many laboratories, EQA data may be the only data available to evaluate method bias.

Summary

- Bias is a measure of trueness
- Bias and its uncertainty can be determined from:
  - analysis of certified reference materials
  - spiking studies
  - interlaboratory studies (EQA)
- Even if the bias is not significant it will contribute to the uncertainty of measurement