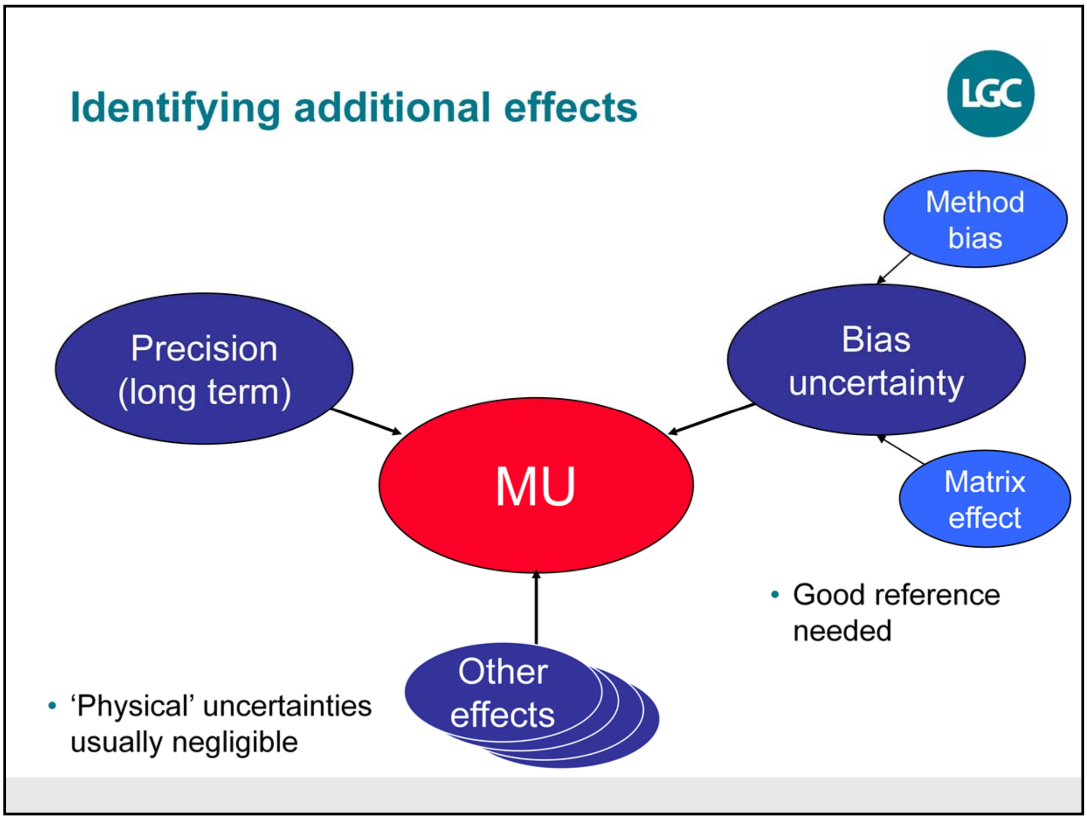
Completing the uncertainty budget

Overview

- Identifying additional effects
- Basis for considering additional effects
- Completing the uncertainty budget



The previous sessions have described the use of data from precision and bias studies in uncertainty calculations. The final step is to consider whether there are any other sources of uncertainty that need to be evaluated to complete the uncertainty budget.

What can be left out?

- Effects not covered by precision and bias data may need to be considered:
  - preparation of standards (step not varied during precision study)
  - balance/glassware (used controlled set)
- Criteria for inclusion/exclusion of additional effects in uncertainty estimation:
  - significance tests
  - calculations
  - experiments
  - professional judgement

In considering additional effects, the most pressing question is ‘how many effects need to be considered?’

This discussion will cover four common methods of deciding which effects need to be considered and which can reasonably be neglected.

Significance tests

- Objective tests to see if **inclusion** is justified on statistical grounds
- Null hypothesis assumed
- Studies* show that
  - including a statistically insignificant effect causes unreliable uncertainty estimates
  - most statistically insignificant effects lead to negligible uncertainty
- *Provided the significance test is good enough*
  - sufficient precision to detect significant effects

Significance tests begin with an assumption of no effect, and test the assumption. They are, accordingly, *tests of the validity of the measurement equation*. Information on significant testing can be found in statistical textbooks such as refs. 1 & 2 below.

From a purely statistical point of view, it is most unwise to add an effect into a model unless there is strong evidence to do so; it follows that an effect found to be statistically insignificant at, say, the 95% level of confidence, 'should' be ignored.

Studies[3] confirm that this is often a sensible approach, if the test is sensitive enough to pick up effects which might be significant. However, one caveat remains. While the equation may be tested and found complete within current knowledge, 'current knowledge' is usually limited. Most scientists will make some allowance for potential lack of knowledge in advising on scientific data, and metrologists are no exception. For that reason, it is not uncommon to find uncertainties retained despite lack of significant effects - for example because prior knowledge was not strong enough to justify a 'null hypothesis' assumption. Unfortunately, there is no clear point at which such decisions become objective; the judgement depends on confidence developed by accumulated scientific knowledge of technique, analyte and matrix, perhaps over long study in many institutions.

1. Statistics and Chemometrics for Analytical Chemistry, 6th edition, J Miller and J Miller, Prentice Hall, 2010, ISBN 978-0273730422
2. Practical Statistics for the Analytical Scientist, 2nd edition, S L R Ellison, V J Barwick, T J Duguid Farrant, RSC, 2009, ISBN 978-0-85404-131-2
3. S L R Ellison, D G Holcombe, M Burn (2001). Response surface modelling and kinetic studies for the experimental estimation of measurement uncertainty in derivatisation, *Analyst* **126** 199-210

Calculations

- Simple ‘worst case’ calculations can often show an effect is negligible
- Formal uncertainty calculations often show negligible components
- Typical criteria:
  – $u<u_{max}/3$ (<6% effect on combined uncertainty)
  – $u<u_{max}/5$ (<2% effect on combined uncertainty)

Calculations based on ‘worst case’ changes in an uncertainty source can show quickly whether an effect is worth pursuing. Full uncertainty budgets simply extend this principle to all sources.

A contribution is often considered negligible if it is below a third of the largest uncertainty, though some authorities prefer a factor of five.

Other considerations

- Experiments
  - usually combined with significance tests
  - need sufficient power to detect effects
- Professional judgement
  - always applied in identifying effects to check
  - always a matter of debate….

Experimental assessment is particularly common in analytical chemistry. It usually takes the form of either

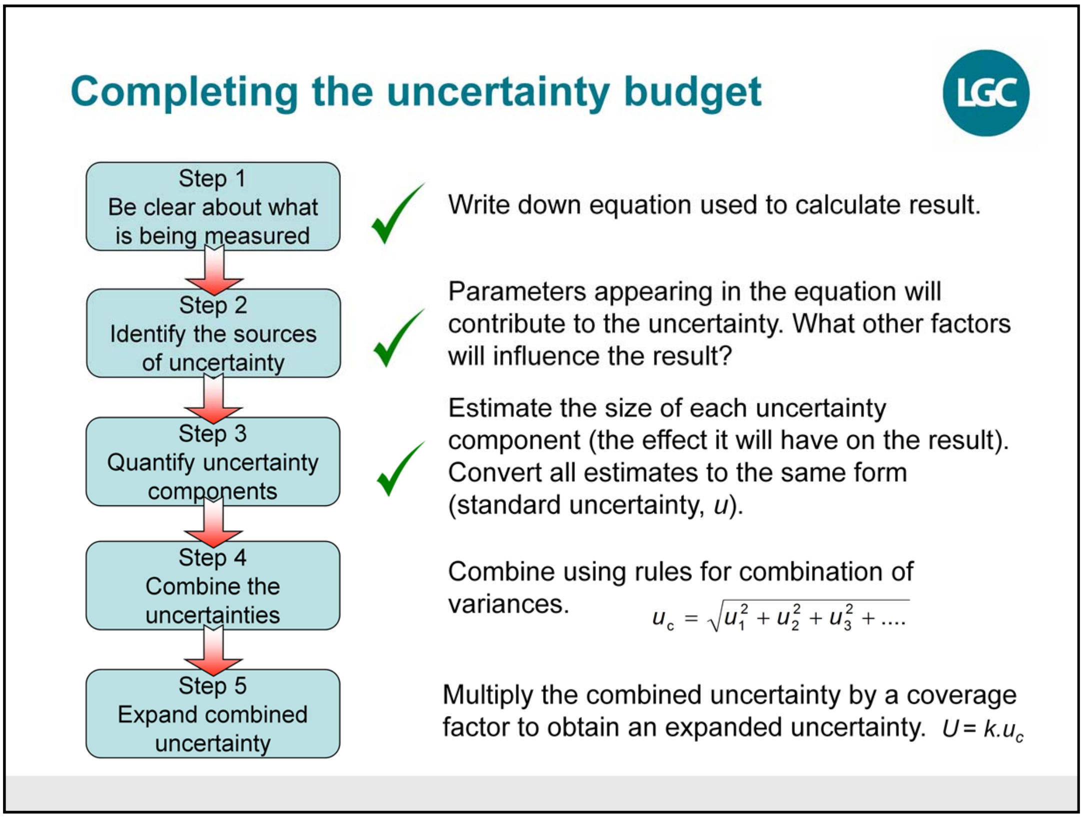
• simple experiments to determine the size of a particular effect

or

• multivariate 'screening' tests, like the common ruggedness test design.

These experiments are often backed up by significance tests to provide an objective decision as to whether the effects need further study or can be neglected.

Professional judgement is much more commonly applied to decisions about the effects to consider, than to estimates of the size of effects. Expert analysts should be well informed on the factors that can influence their results, and are usually aware of the relative importance of different effects. It is important, too, to realise that many years' experience in applying a particular measurement method, backed by intercomparisons and proficiency testing, does provide an ongoing test of the assumptions on which the method is based. Continued successful performance can be considered evidence that few additional effects are important. Professional judgement will always include a degree of subjectivity, however, and may well form grounds for debate.



The slide shows the general approach to evaluating measurement uncertainty.

For each stage of the method, we need to identify factors which could cause the result to change; these will be sources of uncertainty.

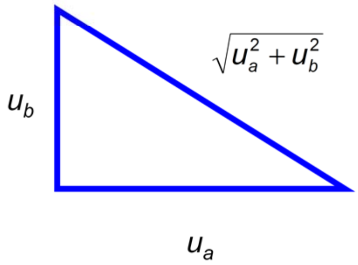
As we have seen during this workshop a ‘top-down’ approach, making use of method performance data, is a practical solution to obtaining a reasonable estimate of the measurement uncertainty.

Completing the uncertainty budget

Parameter	Relative uncertainty
Bias u(B)	
Precision u(P)	
Other contributions	
Combined (relative) standard uncertainty	
Expanded uncertainty (95% confidence)	

Combining uncertainties

- Combine using rules for combination of variances
  - Most common approach – express all uncertainty components as relative (or % relative) values

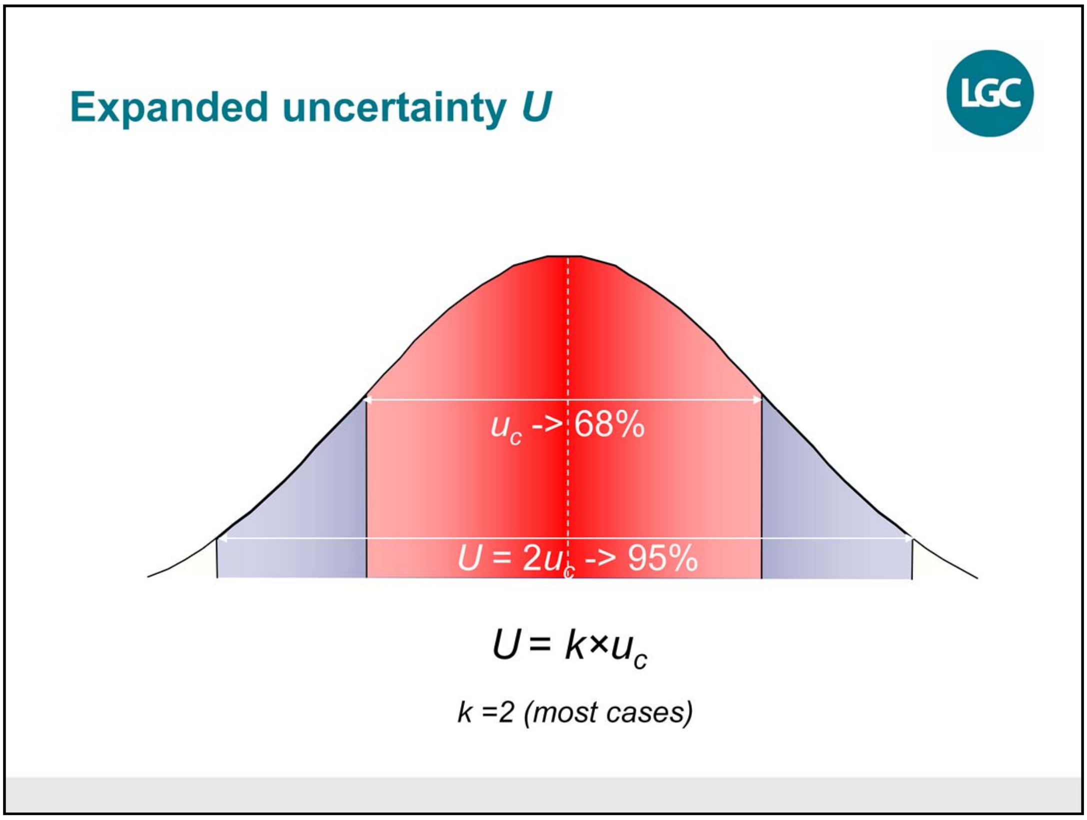


$$u_c = \sqrt{u_1^2 + u_2^2 + u_3^2 + ...}$$

To use the equation on the slide the different uncertainty components must be expressed in the same form. It is common the express all uncertainty components as relative (or %relative) values, for example %bias and %rsd.

Completing the uncertainty budget

Parameter	Relative uncertainty
Bias u(B)	
Precision u(P)	
Other contributions	
Combined (relative) standard uncertainty	$u_c = \sqrt{u(B)^2 + u(P)^2}$
Expanded uncertainty (95% confidence)	



For a normal distribution, a range of ±σ only includes about 68% of the possible values.

In ordinary statistics, the confidence interval is a range which includes a specified (usually 95%) fraction of the possible values. The confidence interval is typically calculated by multiplying the observed standard deviation *s* by Student's *t*. For 95% confidence and for large numbers of degrees of freedom, *t* is close to 2.

By analogy, the standard uncertainty *u* needs expanding to give a range with a high probability of including the right answer. The ISO Guide recommends that *u* be expanded by a coverage factor *k*. *k* is usually expected to be based on the Student *t* value corresponding to the number of degrees of freedom associated with the uncertainty estimate. In most cases, a reasonable number of degrees of freedom is assumed, so **for most practical purposes, *k* is set equal to 2**.

Completing the uncertainty budget

Parameter	Relative uncertainty
Bias $u(B)$	
Precision $u(P)$	
Other contributions	
Combined (relative) standard uncertainty	$u_c = \sqrt{u(B)^2 + u(P)^2}$
Expanded uncertainty (95% confidence)	$U = k \times u_c$

Summary

- Uncertainty in a result will arise from several contributions (random error and systematic error)
  - include precision and bias estimates plus any additional significant effects
- Uncertainty contributions are converted
  - uncertainty components may need to be converted to same form to allow combination => standard uncertainty (u)
- Standard uncertainties are then combined:
  - as relative or %relative values
  - use ‘root-sum-of-squares’ rule
- Combined uncertainty is then expanded to provide coverage at the 95% level of confidence