Using method validation and performance data for estimating measurement uncertainty

Ivo Leito
University of Tartu
ivo.leito@ut.ee

Overview

• The main question of uncertainty evaluation
• The different approaches
  – (Modelling approach)
  – Approach based on validation and QC data
• The role of performance data
  – Precision
  – (Trueness, bias)
The main question of uncertainty evaluation in an analytical lab:

- The uncertainty sources are more or less known
- There are different data available (control charts, PT results, parallel measurements …)

How to use these data to take these uncertainty sources into account?

Different approaches offer different solutions to this question

Uncertainty estimation approaches

- Definition of the measurand
- Single laboratory
  - Model-based?
    - Yes
    - No
    - Modelling
      - Component-by-component evaluation
      - ISO GUM
    - Single-lab validation
      - Within-lab reproducibility and bias
      - Nordtest TR537
- Interlaboratory
  - One procedure?
    - Yes
    - Interlaboratory validation
      - Reproducibility and bias
      - ISO 5725
      - ISO TS 21748
    - Proficiency testing (PT)
      - Between-lab variability
      - ISO Guide 43
      - ISO 13528

Uncertainty estimates by different approaches

- Modelling (classical ISO GUM)
  - Uncertainty of an individual result of a measurement can be obtained

- Single-lab validation
  - Typical uncertainty of results obtained using a procedure in the laboratory

- Interlaboratory validation
  - Uncertainty of results obtained using the same procedure in different laboratories

These uncertainties refer to different situations!

The Modelling Approach

Component by component evaluation
Validation parameters (procedure characteristics)

- identity, selectivity, specificity
- limit of detection
- limit of quantitation
- linear range
- accuracy, trueness (recovery), precision
- sensitivity
- ruggedness/robustness

Examples:
http://www.ut.ee/katsekoda/GUM_examples/
Examples:
http://www.ut.ee/katsekoda/GUM_examples/

Examples:
http://www.ut.ee/katsekoda/GUM_examples/
Examples:
http://www.ut.ee/katsekoda/GUM_examples/

Approach Based on Validation and Quality Control Data
on the example of the Nordtest approach

http://www.nordtest.info/index.php/technical-reports/category/chemistry.html
### Types of errors

- **Systematic error**
  - (total) error
  - Random error

### Performance characteristics

- **Trueness**
- **Accuracy**
- **Precision**

### Quantitative expression of performance characteristics

- **Bias**
- **Measurement uncertainty**
- **Standard deviation**
  - Repeatability/within lab reproducibility
  - Reproducibility

---

#### Single-laboratory validation approach

- The two groups of uncertainty contributions are quantified separately and then combined.
Single lab validation approach: in practice (1)

- The main equation:

\[ u_c = \sqrt{u(R_w)^2 + u(bias)^2} \]

- This and subsequent equations work with absolute and relative values

Validation parameters (procedure characteristics)

- identity, selectivity, specificity
- limit of detection
- limit of quantitation
- linear range
- accuracy, trueness (recovery), precision
- sensitivity
- ruggedness/robustness
**Precision component \( u(R_w) \)**

\[ u(R_w) = s_{R_w} \] is usually found from:

- the warning limits of X chart
  - using a stable control sample
- long term pooled standard deviation

Ideally: separately for different matrices and different concentration levels!

The control sample analysis has to cover the whole analytical process.

---

**How to determine precision?**

- Example:

An analyst analysed a food sample by HPLC. He carefully homogenized the sample in a blender and took a subsample. With the subsample he carried out sample preparation (consisting of extraction, precipitation and centrifugation). As a result he obtained a clear solution. He transferred it into a 50 ml volumetric flask and filled it up to the mark with the mobile phase. He analysed 10 aliquots of this solution during the same day and calculated the within-lab reproducibility as standard deviation of the results.

Did he do it right? If not, what should he do differently?

---

22.05.2012
Determining precision when sample is stable for a long time

Precision
Determination of fat content

<table>
<thead>
<tr>
<th>Date</th>
<th>Sample</th>
<th>Result (g/100g)</th>
<th>Mean: 23.1 g/100g</th>
<th>St Dev: 1.1 g/100g</th>
<th>DF: 12</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.02.2008</td>
<td>27</td>
<td>22.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16.02.2008</td>
<td>27</td>
<td>21.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26.02.2008</td>
<td>27</td>
<td>22.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.03.2008</td>
<td>27</td>
<td>23.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17.03.2008</td>
<td>27</td>
<td>23.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27.03.2008</td>
<td>27</td>
<td>23.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.04.2008</td>
<td>27</td>
<td>23.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16.04.2008</td>
<td>27</td>
<td>23.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26.04.2008</td>
<td>27</td>
<td>22.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.05.2008</td>
<td>27</td>
<td>25.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16.05.2008</td>
<td>27</td>
<td>22.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26.05.2008</td>
<td>27</td>
<td>23.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.06.2008</td>
<td>27</td>
<td>22.2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Within-lab reproducibility $s_{RW}$

Pooled Standard Deviation

- General formula:

$$s_{\text{pooled}} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \ldots + (n_k - 1)s_k^2}{n_1 + n_2 + \ldots + n_k - k}}$$

- Symbols:
  - $k$ number of groups (in this case samples)
  - $s_1$, $s_2$, etc are within group standard deviations
  - $n_1$, $n_2$, etc are numbers of measurements made with different samples
Determining precision when sample is not stable for a long time

**Precision**

### Determination of protein content

<table>
<thead>
<tr>
<th>Date</th>
<th>Sample</th>
<th>Result (g/100g)</th>
<th>Sample</th>
<th>St dev g/100g</th>
<th>Comp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.02.2008</td>
<td>1</td>
<td>10.2</td>
<td>1</td>
<td>0.354</td>
<td>0.125</td>
</tr>
<tr>
<td>10.02.2008</td>
<td>2</td>
<td>13.4</td>
<td>2</td>
<td>0.572</td>
<td>0.980</td>
</tr>
<tr>
<td>10.02.2008</td>
<td>3</td>
<td>17.6</td>
<td>3</td>
<td>0.473</td>
<td>0.447</td>
</tr>
<tr>
<td>13.02.2008</td>
<td>1</td>
<td>10.7</td>
<td>4</td>
<td>0.681</td>
<td>0.927</td>
</tr>
<tr>
<td>13.02.2008</td>
<td>2</td>
<td>14.2</td>
<td>5</td>
<td>0.707</td>
<td>0.500</td>
</tr>
<tr>
<td>13.02.2008</td>
<td>3</td>
<td>18.9</td>
<td>6</td>
<td>0.663</td>
<td>1.320</td>
</tr>
<tr>
<td>18.02.2008</td>
<td>2</td>
<td>12.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18.02.2008</td>
<td>3</td>
<td>16.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18.02.2008</td>
<td>6</td>
<td>12.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25.02.2008</td>
<td>6</td>
<td>13.5</td>
<td>1</td>
<td>0.598</td>
<td></td>
</tr>
<tr>
<td>25.02.2008</td>
<td>4</td>
<td>17.2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25.02.2008</td>
<td>5</td>
<td>19.2</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.03.2008</td>
<td>2</td>
<td>13.1</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.03.2008</td>
<td>6</td>
<td>12.9</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.03.2008</td>
<td>3</td>
<td>17.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.03.2008</td>
<td>4</td>
<td>18.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.03.2008</td>
<td>5</td>
<td>13.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.03.2008</td>
<td>6</td>
<td>13.5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Within-lab reproducibility \( s_{RW} \)**

\( s_{pooled} = 0.598 \) g/100g

**Trueness, bias**

- **bias** of lab’s results from the best estimate of true value is taken into account
- **\( u(bias) \)** can be found:
  - From analysis of the same samples with a reference procedure
  - From analysis of certified reference materials (CRMs)
  - From interlaboratory comparison measurements
  - From spiking experiments

**Ideally:** several reference materials, several spikings

(bias will in most cases vary with matrix and concentration range)

**Necessarily:** several replicate measurements for the same CRM

But less “long-term”

22.05.2012
This component accounts for the average bias of the laboratory results from the reference values $C_{ref}$

This component accounts for the average uncertainty of the reference values $C_{ref}$

\[ u(bias) = \sqrt{RMS_{bias}^2 + u(C_{ref})^2} \]

How to conduct a spiking experiment?

- Two analysts determined meropenem (an antibiotic) in blood plasma. Both needed to determine the bias of the procedure. They obtained blank plasma samples and did the following:

- **Analyst 1** took 500 µl of the blank plasma and added 400 µl of methanol. He separated the precipitated proteins by centrifugation and transferred the supernatant into an HPLC vial. He then added 100 µl of meropenem standard solution with suitable concentration to the supernatant and injected the resulting solution into the HPLC system for analysis.

- **Analyst 2** took 500 µl of the blank plasma and added 500 µl of methanol, which contained a suitable amount of meropenem. She separated the precipitated proteins by centrifugation and injected the resulting supernatant into the HPLC system for analysis.

Which analyst did it correctly? Why?

Analyte has to be added at as early stage as possible!
**Absolute vs relative uncertainties: Rules of Thumb**

- At low concentrations (near detection limit, trace level) use absolute uncertainties
  - Uncertainty is not much dependent on analyte level
- At medium and higher concentrations use relative uncertainties
  - Uncertainty is roughly proportional to analyte level
- In general: whichever is more constant

Available from: http://www.eurachem.org/

---

**Single lab validation approach: Determination of ammonium in water**

- According to EN/ISO 11732
- Concentration level 200 mg/L
- From the X chart: warning limits are set to ± 3.34%
  - Warning limits are set to 2s
  - Thus $u(R_w) = s_{RW} = 3.34\% / 2 = 1.67\%$

http://www.nordtest.info/index.php/technical-reports/category/chemistry.html
Single lab validation approach: Determination of ammonium in water

- From the interlaboratory comparison results bias over 3 years has been: +2.4%, +2.7%, +1.9%, +1.4%, +1.8% and +2.9%.
  - Thus $RMS_{bias} = 2.25\%$
  - Uncertainty of consensus values is estimated as $u(\text{Cref}) = 1.5\%$
  - Thus $u(\text{bias}) = 2.71\%$

$$u_{bias} = \sqrt{2.25^2 + 1.5^2} = 2.71\%$$

- Standard uncertainty:
  $$u_c = \sqrt{1.67^2 + 2.71^2} = 3.18\%$$

- Relative expanded uncertainty: $U = 6.4\%$ ($k = 2$)

- Absolute expanded uncertainty:
  $$U = 200 \text{ mg/l} \times 6.4\% / 100\% = 12.8 \text{ mg/l}$ ($k = 2$)

Choosing the approach

- If you have
  - Competence and time
  - Data on all important influencing quantities
    - Use the Modeling approach

- If you have
  - Quality control data and results of participation in ILC-s or CRM analysis
    - Use the Single-lab validation approach

- Interlab approaches are not generally recommended
Credits

• Parts of this presentation have been created in collaboration with Bertil Magnusson (SP, Sweden)

• The thoughts expressed yesterday by Steve Ellison, Ricardo da Silva and Wolfhard Wegscheider were very inspiring

• A part of this presentation has been used in the training materials

Thank you for your attention!

• The presentation is available from: http://www.ut.ee/ams/

• You are always welcome to contact me: ivo.leito@ut.ee
Validation, Traceability, Measurement Uncertainty: the Challenges for the 21st Century Analytical Science

On May 24-25, 2012, a meeting was held in Berlin to discuss the status of validation, traceability, and measurement uncertainty for the 21st Century Analytical Science. This meeting focused on the relationships between method validation, traceability, and measurement uncertainty.