Coming up to standards: Usefulness of standard solutions as test items

Ulrike Mülow-Stollin, Katrin Heider, Markus Fifeik, Katrin Kittler, Ferial Tadjine, Joachim Polzer
Proficiency tests (PTs) conducted by the EURL Berlin (residues of veterinary medicinal products)

PTs: Mandatory task of European Reference Laboratories (EURLs)

• National Reference Laboratories (NRLs) are required to participate
• Objective: harmonisation of lab performance, identification of areas of improvement

• EURL Berlin has long-standing experience in PT organisation (since 1996), has held ISO 17043 accreditation since 2016
• Focus on PTs using incurred samples

<table>
<thead>
<tr>
<th>substance group /matrix</th>
<th>anthelmintics</th>
<th>β-agonists</th>
<th>coccidiostats</th>
<th>nitroimidazoles</th>
<th>NSAIDs</th>
<th>multi</th>
</tr>
</thead>
<tbody>
<tr>
<td>egg</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>milk</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>liver</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>muscle</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>plasma</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>urine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>hair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>retina</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Substance group-matrix combinations previously investigated in EURL PTs
Why conduct a PT on standard solutions?

Picture this:

You are an analytical chemist participating in a proficiency test (PT). You just received your results.
How do you proceed?
Why conduct a PT on standard solutions?
PT Follow-up

Was it really a bad result?
• Critically evaluate the PT design
• Results reported correctly?

A bad result indeed:
• Check QC samples
• Check for anomalies in analytical series
• Check appropriateness of analytical method (with CRM)
• Inexperienced operator?
• Check the employed standard solutions!
Standard solutions – a common cause of errors

**Multiple standards**
Required standard **amount** for preparation of stock solution needs to be **corrected**

**Stability issues**

Stability studies are required for method validation, but cannot provide 100% certainty (freezing-thawing cycles)

**Comparative tests**

Comparison of freshly prepared stock solution with old solution is not mandatory

**CRM unavailable**

ISO 17034 certified mix solutions **not** readily **available** for all analytes of interest
• Participants: EU NRLs, NRLs from Third Countries, German official labs
  • 50 participants in total, 34-39 per substance group
• 3 samples of standard mixes containing anthelmintics, coccidiostats and NSAIDs
• Participants were asked to dilute samples as they wished (recommended 1/9), report results for undiluted sample
• All samples homogeneous and sufficiently stable
• Free choice of method
• Assigned value: consensus mean; target std. dev.: HORWITZ, reproducibility

<table>
<thead>
<tr>
<th>Sample</th>
<th>Solvent</th>
<th>Analytes</th>
<th>Σ</th>
<th>Concentration (ng/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANTH_B</td>
<td>EtOH</td>
<td>ABA, ALBSO2, CLOSAN, CLORSUL, DORA, EPRIN,</td>
<td>9</td>
<td>100-2000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FLUBZ, LEVA, TCBZSO2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COCC_B</td>
<td>EtOH</td>
<td>AMP, CLOL, DICLA, HLF, LAS, MON, ROB, TOLSO2</td>
<td>8</td>
<td>100-1000</td>
</tr>
<tr>
<td>NSAI_B</td>
<td>ACN/MeOH</td>
<td>MAA, CARPF, DICLO, FLU, MEFENA, MELXCM,</td>
<td>10</td>
<td>50-1500</td>
</tr>
<tr>
<td></td>
<td>9/1 v/v</td>
<td>NIFLUA, PBZ, TFAA, VEDAPRF</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Results: Consensus values vs. reference values

- **Good correlation** between consensus value and reference value under consideration of the respective uncertainties
- Indicates that all **relevant uncertainty** contributions were **considered**

Error bars represent $u(x_{pt})$ as estimated from $\sigma_R$.
Results: z-scores

>5 % of results with $|z'| > 2$

Possible reasons:

- In some cases use of HORWITZ instead of $\sigma_R$
- Lack of experience with quantification of standard solutions
- Gross errors (incorrect units, incorrect dilution factors)
Comparison with matrix PTs – same same but different

- **Comparison** of results from **standard solution** PT STRD0521 with results from **matrix** PTs
  - Matrices include milk, egg, liver, muscle
  - Several results from matrix PTs per analyte
  - Concentrations in the range 0.175-369 ppb
- **Robust statistics**
  - consensus mean: HAMPEL
  - repeatability, reproducibility: Q-method

≠
Comparison with matrix PTs – repeatability $\sigma_r$

Observations:
• $\sigma_r$ differs between analytes
• $\sigma_r$ often similar for same analyte in different matrices

• **Repeatability is higher in matrix samples**
  – Up to 7x higher in matrix samples,
  – **Median ratio: 2**
  – More factors influence analysis of matrix samples

Results of STRD0521 and matrix PTs matched by analytes
Comparison with matrix PTs – reproducibility $\sigma_R$

**Observations:**
- $\sigma_R$ differs between analytes
- Larger concentration for same analyte often comes with larger $\sigma_R$
- **Reproducibility is higher in matrix samples**
  - Up to 4x higher in matrix samples,
  - Median ratio: 1.5
  - More factors influence analysis of matrix samples

Usefulness of standard solutions as test items – Mülow-Stollin et al.
Observations:

- Ratio for solutions often higher
  - Up to 5x higher in solutions
  - **Median ratio: 0.84**
    - Ratio for solutions about 20% higher

- Lower $\sigma_r$ and same or higher $\sigma_R$ in solution
  - Labs are very good at analysis of their own solutions, but solutions might differ significantly between labs

Results of STRD0521 and matrix PTs matched by analytes

37% 63%
Observations:

- HORRAT dependant on analyte
  - HORRAT for AVER higher in solution
  - HORRAT for ANTH, COCC mostly higher in matrix
- Overall **good correlation** between HORRAT in solution and HORRAT in matrix
  - Median ratio: 0.9
  - If corrected for concentration, laboratory performance for matrix samples is comparable to performance for standard solutions
PTs on matrix samples and standard solutions differed
  - $\sigma_r$ and $\sigma_R$ for standard solutions often lower
  - Ratio $\sigma_R/\sigma_r$ for standard solutions often higher

External control of standard solutions is a valuable tool
  - Pitfalls of matrix sample analysis are eliminated
  - Immediate feedback on a common source for errors

PTs with standard solutions come with their own set of challenges
  - Participants’ methods are optimised for matrix
  - Participants are unfamiliar with analysis of standard solutions

An improvement of the quality of the employed standard solution leads to an overall improvement in performance of the laboratory network.
Thanks to unit 502, especially Katrin Heider who prepared all of the standard solutions.

The financial support of the European Commission is gratefully acknowledged.
Thank you for your attention and good luck with your upcoming proficiency tests!

Contact:
ulrike.muelow-stollin@bvl.bund.de
eurlvetdrug@bvl.bund.de