

EURL-CAMPYLOBACTER TRAINING COURSE

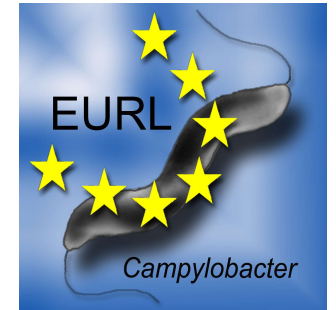
THE ORGANISATION OF

PROFICIENCY TESTS

Design and performance assessment

Helena Höök, DVM, PhD
9th October 2019

OUTLINE



- Planning
- Production of proficiency tests
- Statistical design
- Assessment of performance
- **ISO/IEC 17043:2010**: General requirements
 - **ISO 22117:2019**: Microbiological proficiency testing
 - **Guidance Document** for the organisation of Proficiency Tests by NRLs for national networks, including partial outsourcing

TO PLAN A PROFICIENCY TEST

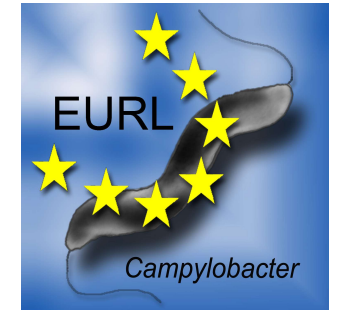


- PPF = project planning form
 - Purpose(s)
 - Objectives
 - Basic design
- Challenging or realistic approach?
- Interest of stakeholders (official laboratories)

PRODUCTION OF PROFICIENCY TESTS

- Microbiological material
 - Measurand/target organism(s) / background flora
 - Live cultures: spiking, how and when
 - Stressed bacteria?
 - Freeze-dried strains
 - Other forms of microbiological reference material
- Matrix/matrices
- How can homogeneity and stability of the tests be assured?

HOMOGENEITY AND STABILITY



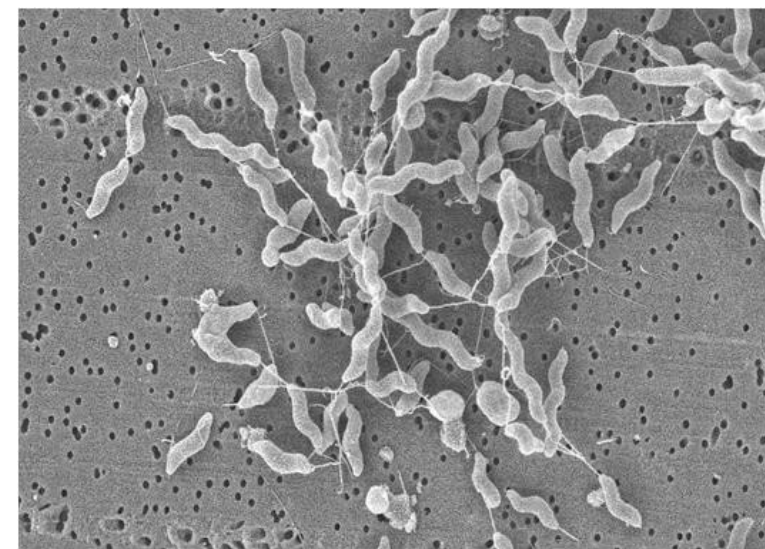
- Criteria for **suitable homogeneity and stability** ... shall be based on the **effect that inhomogeneity and instability** will have on the **evaluation of the participants' performance**

ISO/IEC 17043 4.4.3.1

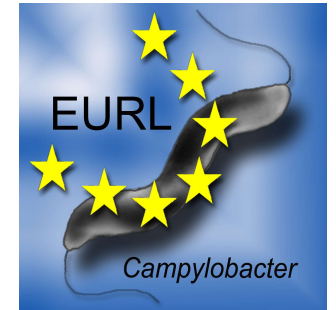
- *Note:* Materials not sufficiently homogeneous or stable can still be useful, provided that uncertainties of assigned values or evaluation are taken into account

SELECTION OF STRAINS

- Bacterial species
- Origin
- Specific characteristics
- Use of duplicates
- Mix of strains



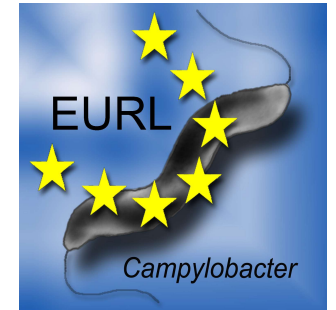
MATRIX MATTERS



- Relevant matrix?
- Realistic matrix?
- Target microorganism and background flora in relation to the matrix



DESIGN AND EVALUATION



- Design and evaluation: hand in hand
- Both are parts of the statistical design

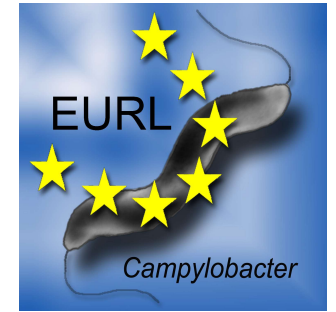
DESIGN ACCORDING TO ISO/IEC 17043:2010

- 4.4 Design of proficiency testing schemes: general strategies and specific design of quantitative tests (assigned values)
- 4.5 Choice of method or procedure
 - Normally expected to use test method of their choice
 - Should be consistent with their routine procedures
- 4.7 Data analysis and evaluation of proficiency testing scheme results
- Annex A: Types of proficiency testing schemes
- Annex B: Statistical methods for proficiency testing

DESIGN ACCORDING TO ISO 22117:2019

- 4 Scheme design and purpose
- 5 Technical requirements and guidance for sample design and content
- 6 Sample verification by the provider
- 7 Sample handling
- 8 Performance evaluations
- Annex C (informative) Methods of testing for variation between portions of test materials
- Annex E (informative) A practical method to assess long term performance of participants in PT schemes using enumeration methods

STATISTICAL DESIGN – CRITERIA (ISO/IEC 17043)

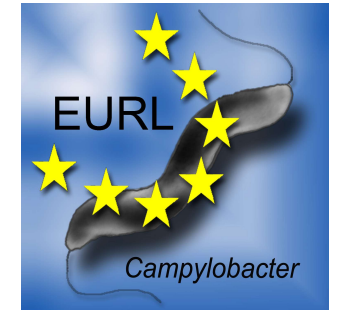


- Meet the objectives
- Based on
 - nature of the data: quantitative or qualitative (ordinal or categorical)
 - statistical assumptions
 - nature of errors
 - expected number of results

STATISTICAL METHODS FOR PROFICIENCY TESTING (ANNEX B ISO/IEC 17043)

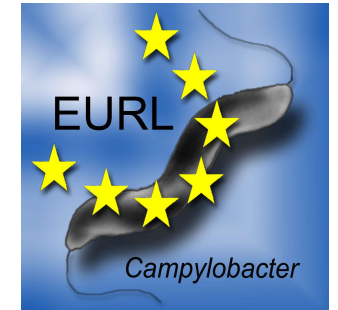
- Fundamental steps
 - a) determination of the assigned value,
 - b) calculation of performance statistics,
 - c) evaluation of performance, and
 - d) preliminary determination of proficiency test item homogeneity and stability

DESIGN OF QUANTITATIVE TESTS: ASSIGNING A VALUE



- The **assigned value**: value attributed to a particular property of a proficiency test item
- Alternative procedures according to Annex B in ISO/IEC 17043:
 - a) known values
 - b) certified reference values
 - c) reference values
 - d) consensus values from expert participants
 - e) consensus values from participants
- Uncertainty of the assigned value
- ISO 22117: mostly consensus values from participants

DESIGN OF QUANTITATIVE TESTS (GUIDANCE DOCUMENT)

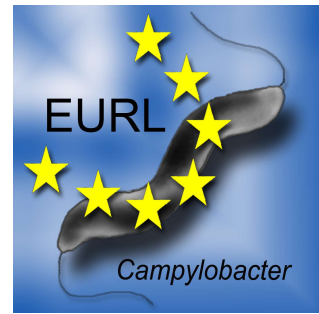


- 4 levels: negative, low, medium and high (1 replicate/level)
 - May be reduced to 3 levels
 - **Negative level:** to verify the absence of false positive results, possibly due to crosscontamination at a participating laboratory. The target bacterium has to be not detected, other background flora can be present.
 - **Low level:** close to the enumeration limit of the method, but avoid that participants may obtain a count on a plate of <10 cfu, corresponding to an estimated number according to EN ISO 7218.
 - **Medium level** (optional).
 - **High level.**
- The choice of the low, medium and high levels should take into account the regulatory limit(s) for the target microorganism and matrix.

ASSESSMENT OF QUANTITATIVE TESTS ACCORDING TO ISO 22117

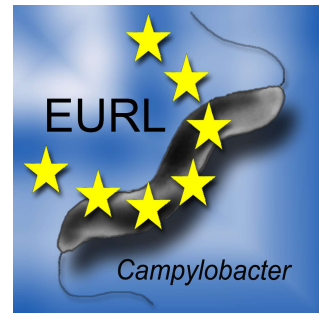
- Using z-scores
 - Using scaled median absolute deviation (MADe) from the median values
- Other methods
 - Using the $0.5\log_{10}$ rule
 - Using percentiles
 - Poisson 95% confidence interval

DESIGN OF QUALITATIVE TESTS



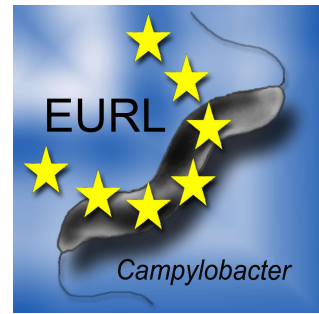
- **ISO 22117:2019:** 6 replicates at each of 3 levels (negative, low and high)
- **Alternative design with 10 samples (Guidance document):**
 - 3 levels: negative, low and high
- **Negative level** (2 samples per participant)
- **Low level** (6 samples per participant as a minimum):
 - goal: approach the regulatory limit or LOD_{50}
 - in practice: Poisson distribution → part of test portions may not be contaminated
 - ISO 22117: approach to deal with low level samples and interpret the results:
 - comparison of the proportion of positive replicates to the expected proportion according to the binomial distribution
 - may be difficult to achieve in practice – a level of up to 10 times acceptable (e.g. 1–10 cfu/test portion if the goal level is 1 cfu/test portion)
- **High level** (2 replicates per participant):
 - level giving 100% positive results for all replicates.
 - approx. 5–10 times the low level (e.g. 50–100 cfu/test portion if the low level is 10 cfu/test portion)

PERFORMANCE ASSESSMENT – QUALITATIVE TESTS



- Samples with low levels, where a fractional recovery can be expected, can be evaluated with a probabilistic approach
- Example: using the binomial distribution and the percentage of samples found positive at a 95% confidence level
- However, such low levels can be a both practical and pedagogical problem

PERFORMANCE ASSESSMENT – SCORING AND GRADES



- Scoring systems for more test items (e.g. 10 enumeration samples): how to set limits?
- Grades, how many and which levels should be used?

SUMMARY DESIGN AND ASSESSMENT

- Purpose and objectives
- Considerations in production of tests
 - Target
 - Matrix
 - Background
 - Quality control
- Design and evaluation hand in hand
 - Quantitative tests
 - Qualitative tests

