

Proficiency tests are widely used to assess the performance of laboratories. Reference standards for interlaboratory comparisons, ISO 5725 and ISO 13528, specify limits for computing the alerts, related to specified risks (usually 5% and 0,3%). Even if this is not written in the standards, these risks are of α -type (risk to trigger a warning that should not). There is actually another risk, usually called β -type (risk not to trigger a warning when it should). However, even if this question is of main importance, this β -type risk is quite hard to compute, and for this reason, is almost always just ignored, including in the reference standards. Everybody knows that an enough number of participants is necessary to ensure the efficiency of the PT, but nobody has any clear view of what should be "an enough number". On the other hand, test methods for which there are very few potential participants to a PT are quite numerous. There is then no opportunity for them to get the usual advantages of the organization of a PT. This paper proposes to overcome the difficulty of computing the β -risk by using a Monte-Carlo method and to provide a beginning of answer to the question: does it make sense or not to organize PTs with 5 or 8 or 12 participants, especially when there are few potential participants?

Results:

- ↓ A ratio $\lambda = \sigma_r/(\sigma_L \times \sqrt{N_r})$ (where σ_r is the inner participant standard deviation, σ_L is the interlaboratory standard deviation and N_r is the number of results provided by each participant) is of main importance to control the efficiency of a PT scheme, even more than the number of participants (see Figure 1). The PT providers should then care Nr, number of test results per participant that they request;
- Even in adverse conditions, the α-risk is always very low (less than 0,7%, see Figures 1 and 2);
- Robust algorithms improve the efficiency of the PT program (i.e., β-risk) at a slight expense on α-risk (which always remain very low, see Figures 2 and 3. This comes from a significantly better estimation of the standard deviation of reference when an outlier occurs;
- A number of 6 participants is large enough to detect a strongly outlying participant provided that good PT conditions (i.e. λ value close to or lower than 0,17) are present, see Figures and 4:
- PT with a low number of participants is (almost) always better than no PT at all.

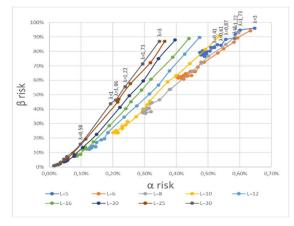


Figure 1. α and β -risks for participants without outlier in function of the number of participants (L is the number of participants)

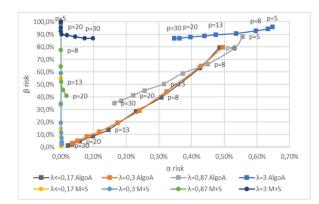
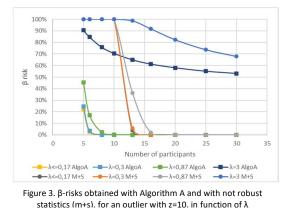
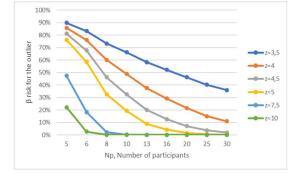


Figure 2. Comparison of α and β -risks obtained with Algorithm A and with not robust statistics (m+s), for participants without any outlier in function of λ (p is the number of participants)





Conclusions:

ISO 5725-2 and ISO 13528 recommend not to organise an ILC with less than 12 participants. This makes sense for ISO 5725, which goal is to determine the performance of a test method. It makes less sense for ISO 13528, which goal is to check the performance of a lab. Obviously, when no PT is organised, β -risk is 100%: any lab having a problem can never at all realise it! Consequently, for test methods that are performed by a little number of labs, it is obviously better to organise PT with 6 participants than nothing. In those cases, the PT provider should specially care the number of results per participant that it requests, to ensure a proper λ value and consequently assure an efficiency as good as possible.

Figure 4. β -risks obtained with Algorithm A and λ =0,17 for an outlier in function of its z-score.