Talanta 88 (2012) 769–771

Contents lists available at SciVerse ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Letter to the Editor

Comments on "Uncertainty profiles for the validation of analytical methods" by Saffaj and Ihssane

Saffaj et al., recently proposed an uncertainty profile for evaluating the validity of analytical methods using the statistical methodology of γ-confidence β-content tolerance intervals [1]. This uncertainty profile estimates the measurement uncertainty of analytical methods using method validation data. Therefore this profile assesses the validity of the method by comparing the method measurement uncertainty to a pre-defined acceptance limit stating the maximum uncertainty suitable for the method under study. Several years earlier as stated by these authors a SFSTP (Société Française des Sciences et Techniques Pharmaceutique) commission has developed a similar profile called accuracy profile used to assess the validity of analytical methods [2–5]. This accuracy profile also uses the methodology of statistical tolerance intervals, but β-expectation tolerance intervals. The uncertainty profile of Saffaj et al. and the accuracy profile of the SFSTP commission are both fulfilling the same final purpose. The core question is finally what statistical tolerance interval to use? The β-expectation tolerance interval can be described by the following general formula when the mean μ and the standard deviation σ are known [6]:

$$E(P[L \leq x_i \leq U]) = \beta \quad (1)$$

where E is the expression for statistical expectation and P stands for probability.

This interval is also known as the "mean coverage tolerance interval" and is equivalent to the "prediction intervals" for a single observation. If $\beta = 0.95$, this means that each future result has a probability of 0.95 to fall within the computed interval $[L; U]$ [7,8].

The second type of tolerance interval is called the "β-content, γ-confidence tolerance interval" (also "guaranteed-coverage tolerance interval"). For μ and σ known, this type of interval is defined by [6]:

$$P(P[L \leq x_i \leq U] \geq \beta) = \gamma \quad (2)$$

where β is the proportion of the individual observations of the population covered by the interval and γ is the confidence level to claim this proportion will be achieved. It is an interval that one can claim to contain a specified proportion β of the population with a specified degree of confidence γ. For instance for $\beta = 0.95$ and $\gamma = 0.90$ the β-content, γ-confidence tolerance interval expresses that there is a probability (P) of 0.90 that 95% of the individual observations (*e.g.* results) of the population are included in the interval $[L; U]$ [7], assuming a rather large number of future observations is envisaged.

The key difference is thus to distinguish proportions and probabilities. While it is generally assumed in Frequentist statistics that both are always similar it is only the case asymptotically, *i.e.* when the sample size tends to infinity, because the proportion, or frequency, is used to estimate the probability. This is illustrated in Fig. 1 using the binomial distribution $Bi(n, 0.95)$ where n is the total number of trials and 0.95 is the probability of success of a single trial. As can be seen in Fig. 1, when the total number of trials n increases (x-axis), the proportion of successful trials (depicted on the y-axis) gets closer to the probability value. This figure also shows that when the probability that an event has a probability of success of 0.95 there is high confidence that a high proportion of events will be successful. This implies that for analytical methods it should be enough to demonstrate that *each* future result has a high probability (*e.g.* 0.8, 0.9, 0.95 and so on) to fall inside the specifications to declare a method as valid. Indeed, when releasing a result to an end customer, what matters beyond everything is to provide assurance that the result is likely to be within specifications. Giving a confidence that most of results are within specifications has limited interest for the customer. For example if a laboratory in a hospital produce results about a disease status, the physician that needs to make decision for a patient, would like to know how much he can trust the very result of that patient, *i.e.* the probability of this result being in specifications. Knowing that 95% of the many results produced during that run (day) has no interest for the physician that make this decision; it only helps the analyst to track and ensure that his run (of many results) can be accepted. The β-expectation tolerance interval correctly assesses the former and not the quality of runs. For instance, Boulanger et al. [9] have shown that a probability of 0.80 should be used when assessing the validity of bioanalytical methods with β-expectation tolerance intervals in order to correctly accept routine runs in high proportion (90% of the runs) when following the FDA in-study validation rule [10].

This distinction between proportion and probability has in fact strong practical consequences. In this same work [9], these authors also emphasized that the two risks of making erroneous decisions (false compliance and false non-compliance risks) were adequately balanced using β-expectation tolerance intervals (*i.e.* when making decision about method validity by focusing on the probability of having results within the acceptance limits) by opposition to β-content, γ-confidence tolerance intervals. Although this last type of tolerance interval highly controls the false compliance risk it highly overlooks the false non-compliance risk leading to excessively rejecting valid methods. By opposite, the β-expectation tolerance interval adequately balances both risks.

Another practical consequence is that trying to correctly declare a method as valid using β-content, γ-confidence tolerance intervals such as proposed by Saffaj et al. requires a greater sample size than β-expectation tolerance intervals as can be seen in Fig. 2. This figure shows that for similar sample sizes, the β-expectation tolerance intervals ($\beta = 0.95$) correctly declares analytical methods as compliant more frequently than β-content, γ-confidence tolerance

0039-9140/$ – see front matter
doi:10.1016/j.talanta.2011.11.011

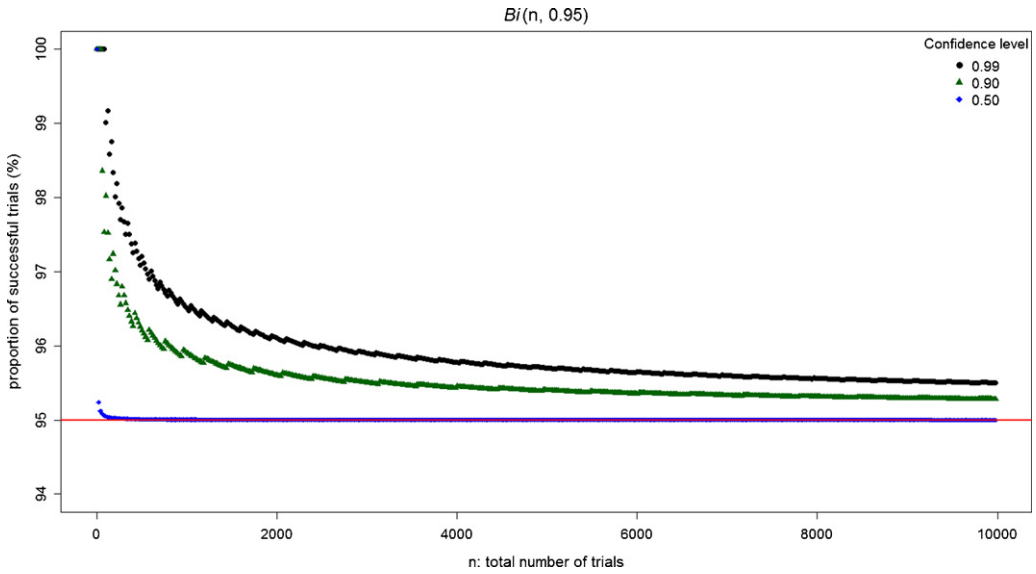

Fig. 1. Proportion of successful trials versus the number of total trials (n) for the binomial distribution $Bi(n, 0.95)$ for three different confidence levels: 0.50, 0.90, and 0.99.

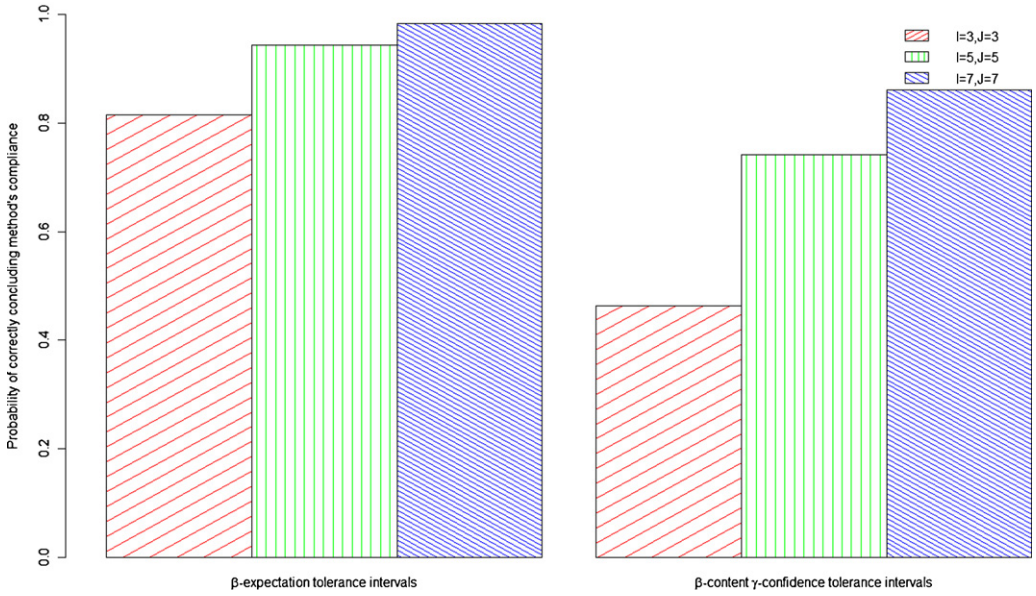

Fig. 2. Probability of correctly declaring analytical methods as compliant (*i.e.* valid methods) when increasing the sample size used in method validation study for the β-expectation tolerance interval methodology ($\beta = 0.95$) and for the β-content, γ-confidence tolerance interval one ($\beta = 0.95$, $\gamma = 0.95$). I: number of series, J: number of repetitions per series.

intervals ($\beta = 0.95$, $\gamma = 0.95$). This implies thus a higher sample size when using β-content, γ-confidence tolerance intervals (*i.e.* either an increase in the number of series, of repetitions or of both) to have the same probability to correctly declare as valid a method than with β-expectation tolerance intervals.

At least for these reasons the methodology to decide about the validity of analytical methods should rely on the estimation of the predictive probability to produce a result that will fall within pre-specified acceptance limits rather than measuring the proportion of the results population falling within these limits. The former is the very objective of an analytical method, while the later is more related to the quality check process performed after each run. The use of β-expectation tolerance intervals is one possibility to achieve this, but other solutions directly estimating this probability are possible and available [11–14].

Acknowledgement

A research grant from the Belgium National Fund for Scientific Research (FRS-FNRS) to E. Rozet is gratefully acknowledged.

References

[1] T. Saffaj, B. Ihssane, Talanta 85 (2011) 1535–1542.
[2] Ph. Hubert, J.-J. Nguyen-Huu, B. Boulanger, E. Chapuzet, P. Chiap, N. Cohen, P.-A. Compagnon, W. Dewe, M. Feinberg, M. Lallier, M. Laurentie, N. Mercier, G. Muzard, C. Nivet, L. Valat, J. Pharm. Biomed. Anal. 36 (2004) 579–589, http://hdl.handle.net/2268/6169.
[3] Ph. Hubert, J.J. Nguyen-Huu, B. Boulanger, E. Chapuzet, P. Chiap, N. Cohen, P.A. Compagnon, W. Dewé, M. Feinberg, M. Lallier, M. Laurentie, N. Mercier, G. Muzard, C. Nivet, L. Valat, E. Rozet, J. Pharm. Biomed. Anal. 45 (2007) 70–81, http://hdl.handle.net/2268/6170.

[4] Ph. Hubert, J.J. Nguyen-Huu, B. Boulanger, E. Chapuzet, N. Cohen, P.-A. Compagnon, W. Dewé, M. Feinberg, M. Laurentie, N. Mercier, G. Muzard, L. Valat, E. Rozet, J. Pharm. Biomed. Anal. 45 (2007) 82–96, http://hdl.handle.net/2268/41022.
[5] Ph. Hubert, J.-J. Nguyen-Huu, B. Boulanger, E. Chapuzet, N. Cohen, P.-A. Compagnon, W. Dewé, M. Feinberg, M. Laurentie, N. Mercier, G. Muzard, L. Valat, E. Rozet, J. Pharm. Biomed. Anal. 48 (2008) 760–771, http://hdl.handle.net/2268/6171.
[6] I. Guttman, Statistical Tolerance Regions: Classical and Bayesian, Hafner, Darian, 1969, 147 pp.
[7] G.J. Hahn, W.Q. Meeker, Statistical Intervals: A Guide for Practitioners, Wiley Edition, 1991, 392 pp.
[8] S.B. Vardeman, Am. Stat. 46 (1992) 193.
[9] B. Boulanger, E. Rozet, F. Moonen, S. Rudaz, Ph. Hubert, J. Chromatogr. B 877 (2009) 2235–2243, http://hdl.handle.net/2268/18777.
[10] Guidance for industry: Bioanalytical Method Validation, US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Rockville, May 2001.
[11] W. Dewé, B. Govaerts, B. Boulanger, E. Rozet, P. Chiap, Ph. Hubert, Chemometr. Intell. Lab. Syst. 85 (2007) 262–268, http://hdl.handle.net/2268/4372.
[12] B. Boulanger, W. Dewé, A. Gilbert, B. Govaerts, M. Maumy, Chemometr. Intell. Lab. Syst. 86 (2007) 198–207.
[13] B. Govaerts, W. Dewé, M. Maumy, B. Boulanger, Qual. Reliab. Eng. Int. 24 (2008) 667–680.
[14] E. Rozet, B. Govaerts, P. Lebrun, K. Michail, E. Ziemons, R. Wintersteiger, S. Rudaz, B. Boulanger, Ph. Hubert, Anal. Chim. Acta 705 (2011) 193–206, http://hdl.handle.net/2268/91261.

E. Rozet [*,1]
E. Ziemons
R.D. Marini
Laboratory of Analytical Chemistry, CIRM, Institute of Pharmacy, University of Liège, CHU, B36, Tour 4, Avenue de l'Hôpital 1, B-4000 Liège, Belgium

B. Boulanger
Arlenda SA, Avenue de l'Hôpital 1, B-4000 Liège, Belgium

Ph. Hubert
Laboratory of Analytical Chemistry, CIRM, Institute of Pharmacy, University of Liège, CHU, B36, Tour 4, Avenue de l'Hôpital 1, B-4000 Liège, Belgium

[*] Corresponding author. Tel.: +32 4366 4320; fax: +32 4366 4317.
E-mail address: eric.rozet@ulg.ac.be (E. Rozet)

[1] FRS-FNRS Post-Doctoral Researcher (Belgium).

10 October 2011
Available online 6 November 2011