

# **A New Interpretation To The Equivalence Test (TOST)**

## **By (Bayesian) Success Probabilities.**

Bernard G Francq<sup>1</sup>, Ron S. Kenett<sup>2</sup>

<sup>1</sup> Technical R&D - CMC Statistical Sciences, GSK, Belgium, \_\_\_\_\_

<sup>2</sup> The KPA group and the Samuel Neaman Institute, Technion, Israel

In medical and pharma research, statistical significance is often based on confidence intervals (CIs) for means or mean difference and p-values, the reporting of which is included in publications in most top-level medical journals. However, recent years have seen ongoing debates on the usefulness of these inferential tools. Misinterpretations of CIs for means and p-values can lead to misleading conclusions and nonreproducible claims. On the other hand, the two one sided tests (TOST) approach is usually applied in pharma industry for equivalence testing, robustness study or stability analysis. Yet, the TOST is also commonly based on CIs for mean difference or p-value.

Here, we propose a unified framework based on success probability (SP), which has a wider definition based on the tolerance interval's methodology. The SP allows a straightforward and identical interpretation between both frequentist and Bayesian paradigms. The SP extends also the concept of 'probability of agreement' and (Bayesian) 'comparative probability metrics' (CPM). While the CPM is calculated from the posterior distributions, we show that the confidence bound of such probabilities is crucial but rarely applied in practice. The confidence bound for the SP is indeed a one-to-one function of the p-value with enhanced interpretability properties and has a default cut-off value of 50% whatever the type I error.

Performance of our methodology will be evaluated by simulations and applications to case studies within CMC statistics and vaccines development. We argue that success probabilities should be preferred by researchers in pharma industry.