

# The harmonic mean $\chi^2$ test to substantiate scientific findings

Leonhard Held  
Epidemiology, Biostatistics and Prevention Institute (EBPI)  
and Center for Reproducible Science (CRS)  
University of Zurich

Statistical methodology plays a crucial role in drug regulation. Decisions by the FDA or EMA are typically made based on multiple primary studies testing the same medical product, where the two-trials rule is the standard requirement, despite a number of shortcomings. A new approach is proposed for this task based on the (weighted) harmonic mean of the squared study-specific test statistics. Appropriate scaling ensures that, for any number of independent studies, the null distribution is a  $\chi^2$ -distribution with one degree of freedom. This gives rise to a new method for combining one-sided p-values and calculating confidence intervals for the overall treatment effect. Further properties are discussed and a comparison with the two-trials rule is made, as well as with alternative research synthesis methods. An attractive feature of the new approach is that a claim of success requires each study to be convincing on its own to a certain degree depending on the overall significance level and the number of studies. The new approach is motivated by and applied to data from five clinical trials investigating the effect of Carvedilol for the treatment of patients with moderate to severe heart failure.