Letter to the Editor

Letter: biological drugs for inducing remission in patients with ulcerative colitis – determining statistical equivalence according to evidence-based methods

D. Maratea*, V. Fadda*, S. Trippoli*, R. Gatto*, M. De Rosa†, C. Marinai† & A. Messori*

*HTA Unit, ESTAV Toscana Centro, Regional Health Service, Florence, Italy.
†SIFACT, Italian Society for Clinical Pharmacy and Therapeutics, Milan, Italy.
‡Department of Pharmaceutical Logistics, ESTAV Toscana Centro, Regional Health Service, Florence, Italy.
E-mail: andrea.messori.it@gmail.com

doi:10.1111/apt.12690

Sirs, When nonsignificant results are found in trials or meta-analyses, differentiating between no proof of difference (an ‘inconclusive result’) and proof of no difference (equivalence, a ‘conclusive’ result) is increasingly recognised to be a crucial step for a correct interpretation.2–4

We have reanalysed the trials examined by Stidham and co-workers5 for the end-point of induction of remission. Firstly, the meta-analysis results were re-expressed using risk difference (RD) rather than relative risk.6 Then, the pooled RDs for direct comparisons of biologics

Figure 1 | Head-to-head indirect comparisons of three biological treatments for inducing remission in ulcerative colitis: network meta-analysis (a) and equivalence testing based on a Forest plot (b). The outcome measure for each of these indirect pair-wise comparisons was the achievement of remission (expressed as a percentage). The meta-analytic values of RD (with 95% CIs) were extracted from reference 2. (a) This type of graph (simplified figure according to Fadda et al.5) summarises the results, but does not allow us to differentiate between “no proof of difference” and “proof of no difference”. Statistical calculations according to Bucher.10 Symbols: +, more effective at statistical level of P < 0.05; −, less effective at statistical level of P < 0.05; =, no difference; t, indicates which treatment is favoured by a trend in cases of no difference. RD, risk difference; CI, confidence interval. (a) The equivalence test is based on the area comprised between the two vertical dashed lines that reflect the pre-determined equivalence margins (from −12% to +12%). Each horizontal bar indicates the two-sided 95% CI for the RD (solid square). The criterion for demonstrating equivalence is when both extremes of the 95% CI remain within the two vertical dashed lines. Comparisons: [1] infliximab vs. adalimumab (in green): RD = 7% (95% CI: −6.7% to 20.7%); [2] infliximab vs. golimumab (in brown): RD = 2.3% (95% CI: −11.8% to 16.4%); [3] adalimumab vs. golimumab (in blue): RD = −4.7% (95% CI: −12% to 2.6%).

AP&T invited commentary and correspondence columns are restricted to letters discussing papers that have been published in the journal. A letter must have a maximum of 300 words, may contain one table or figure, and should have no more than 10 references. It should be submitted electronically to the Editors via http://mc.manuscriptcentral.com/apt.
Letter to the Editor

vs placebo were subjected to network meta-analysis. In this way, the pooled values of RD were estimated for the three indirect head-to-head comparisons between individual biologics.

The results of our net-work meta-analysis revealed nonsignificant differences in all indirect comparisons (Figure 1a). Then, we extended our analysis by performing an equivalence test among these three biologics. Testing equivalence requires that a margin is pre-specified to separate clinically relevant improvements in the outcome from clinically irrelevant ones. Margins can be retrieved from the statistical power sections of original trials.

According to this procedure, we adopted the margin of ±12% employed by Reinisch et al. and we finally combined, in a Forest plot, this margin with the RD values for indirect comparisons. Equivalence testing frequently relies on these Forest plots.

Based on our equivalence testing, the comparisons of infliximab vs. adalimumab or golimumab (Figure 1b) showed no proof of difference, but failed to demonstrate proof of no difference (i.e. equivalence). So, these two comparisons remained inconclusive. More interestingly, the indirect comparison between the two subcutaneous agents (adalimumab vs. golimumab) showed proof of no difference (i.e. a conclusive result).

In conclusion, this latter finding indicates therapeutic equivalence of these two agents at least for this therapeutic indication. One limitation of this analytical approach is that margins intrinsically have a certain degree of arbitrariness.

ACKNOWLEDGEMENT

Declaration of personal and funding interests: None.

REFERENCES

8. Norman G, Monteiro S, Salama S. Sample size calculations: should the emperor’s clothes be off the peg or made to measure? BMJ 2012; 344: e5278. doi:10.1136/bmj.e5278.