Biological drugs for inducing remission in patients with Crohn's disease: determining statistical equivalence according to evidence-based methods

Andrea Messori, Valeria Fadda, Dario Maratea, Sabrina Trippoli, Claudio Marinai

Regional Health System, ESTAV, HTA Unit, Firenze, Italy

Corresponding author:
Andrea Messori PharmD
Regional Health System
ESTAV, HTA Unit
via San Salvi 12
50100 Firenze, Italy
Phone: 393389513583
Fax: 390574701319
E-mail:
andrea.messori.it@gmail.com

In studying the effectiveness of infliximab, adalimumab and certolizumab for inducing remission in Crohn's disease, the article by Kawalec *et al.* [1] included a comprehensive series of pair-wise meta-analyses that compared individual biologic agents versus placebo (direct comparisons), but did not study the indirect comparisons of biologics with one another, an issue that can be managed by application of network meta-analysis. To better address the clinical relevance of differences between these biologics, equivalence testing is another point that can be worthwhile to investigate based on these data. In fact, differentiating between no proof of difference (an "inconclusive result" [2]) and proof of no difference (equivalence, a "conclusive" result [2]) is increasingly recognised to be a crucial step for a correct interpretation of both meta-analyses and clinical trials [3, 4].

We have reanalyzed the trials examined by Kawalec *et al.* [1] for the end-point of induction of remission. Firstly, the meta-analysis results were re-expressed using risk difference (RD) rather than relative risk [5]. Then, the pooled RDs for direct comparisons of biologics 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 network meta-analysis revealed non-significant differences for the three indirect comparisons (Figure 1 A). Then, we extended our analysis by performing an equivalence test [2–4] 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 [2]. Margins can be retrieved from the statistical power sections of original trials.

According to this procedure, we adopted the margin of ±15% employed by Sandborn *et al.* [6] 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 [2–4].

Based on our equivalence testing, the comparisons of infliximab versus adalimumab or certolizumab showed no proof of difference (Figure 1 B), 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. certolizumab) showed proof of no difference, i.e. a conclusive result.
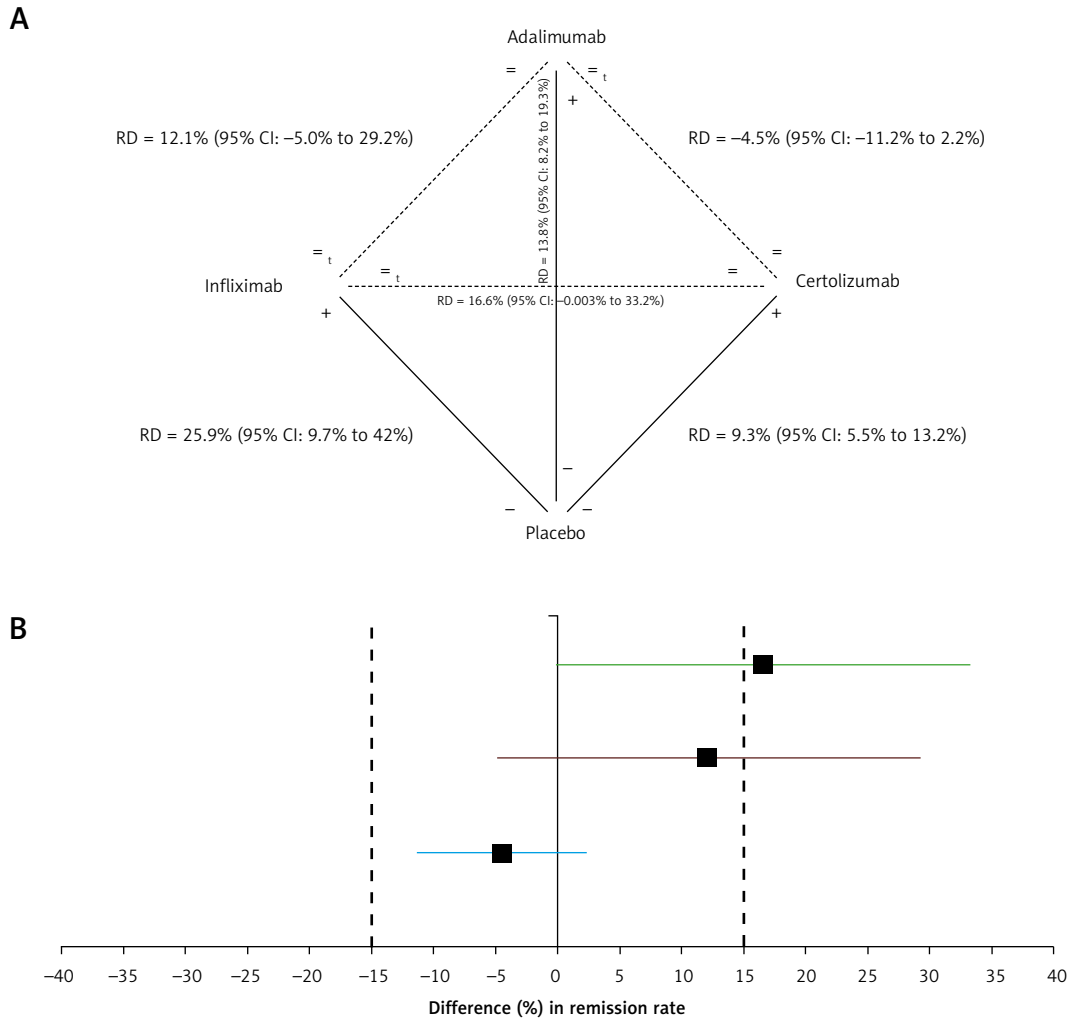
AMS

A

Adalimumab

RD = 12.1% (95% CI: –5.0% to 29.2%)   RD = –4.5% (95% CI: –11.2% to 2.2%)

RD = 13.8% (95% CI: 8.2% to 19.3%)

Infliximab   Certolizumab

RD = 16.6% (95% CI: –0.003% to 33.2%)

RD = 25.9% (95% CI: 9.7% to 42%)   RD = 9.3% (95% CI: 5.5% to 13.2%)

Placebo

B

Difference (%) in remission rate

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 [6]. A – This type of graph (simplified figure according to Fadda *et al.* [7]) 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's method [7]. 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. B – The equivalence test is based on the area comprised between the two vertical dashed lines, that reflect the pre-determined equivalence margins (from –15% to +15%). 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. certolizumab (in green): RD = 16.6% (95% CI: 0.15% to 33.0%); [2] infliximab vs adalimumab (in brown): RD = 12.1% (95% CI: –5.0% to 20.2%); [3] certolizumab vs adalimumab (in blue): RD = –4.5% (95% CI: –10.9% to 1.9%)

RD – risk difference, CI – confidence interval.

In summary, our results (Figure 1) indicate that these two subcutaneous agents are therapeutically equivalent, at least for this indication. There are, of course, some limitations in the present study. One controversial point concerns these mixed analytical approaches based on meta-analysis plus pre-specified margins, because margins are known to possess a certain component of arbitrariness [8]. Furthermore, there can be some controversy as to whether the end-point of remission

achievement can fully account for all the main effects of these agents, because other outcomes can be relevant as well [9, 10].

In conclusion, this network meta-analysis showed the equivalence between the two subcutaneous biologics. Despite this evidence, our findings confirm that there is still an unmet need for large, well-designed, controlled trials providing reliable comparisons between anti-TNF therapies in the treatment of Crohn's disease.

Andrea Messori, Valeria Fadda, Dario Maratea, Sabrina Trippoli, Claudio Marinai

Conflict of interest

The authors declare no conflict of interest.

References

1. Kawalec P, Mikrut A, Wiśniewska N, Pilc A. Tumor necrosis factor-alpha antibodies (infliximab, adalimumab and certolizumab) in Crohn's disease: systematic review and meta-analysis. Arch Med Sci 2013; 9: 765-79.
2. Ahn S, Park SH, Lee KH. How to demonstrate similarity by using noninferiority and equivalence statistical testing in radiology research. Radiology 2013; 267: 328-38.
3. Messori A, Fadda V, Maratea D, Trippoli S. Outcomes with short-term versus long-term antiplatelet dual therapy after drug-eluting stenting: quantifying the equivalence margins. Int J Cardiol 2014; 172: 469-70.
4. Messori A, Maratea D, Fadda V, Trippoli S. Risk of intracranial haemorrhage in patients with atrial fibrillation treated with novel oral anticoagulants: testing the equivalence margins between dabigratran, rivaroxaban and apixaban. Eur J Clin Pharmacol 2014; 70: 505-6.
5. Messori A, Fadda V, Gatto R, Maratea D, Trippoli S. PubMed Commons (comment), 15 February 2014, available at http://www.ncbi.nlm.nih.gov/pubmed/24273556.
6. Sandborn WJ, Rutgeerts P, Enns R, et al. Adalimumab induction therapy for Crohn disease previously treated with infliximab: a randomized trial. Ann Intern Med 2007; 146: 829-38.
7. Fadda V, Maratea D, Trippoli S, Messori A. Network meta-analysis. Results can be summarised in a simple figure. BMJ 2011; 342: d1555.
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; 345: e5278.
9. Lis K, Kuzawińska O, Bałkowiec-Iskra E. Tumor necrosis factor inhibitors – state of knowledge. Arch Med Sci 2014; 10: 1175-85.
10. Mozaffari S, Nikfar S, Abdolghaffari AH, Abdollahi M. New biologic therapeutics for ulcerative colitis and Crohn's disease. Expert Opin Biol Ther 2014; 14: 583-600.