

Tocilizumab for reduction of mortality in severe COVID-19 patients: how should we GRADE it?

Vladimir Trkulja¹

¹Department of Pharmacology, School of Medicine, University of Zagreb

December 7, 2021

Tocilizumab for reduction of mortality in severe COVID-19 patients: how should we GRADE it?

Vladimir Trkulja

Vladimir Trkulja, MD, PhD

Department of Pharmacology

Zagreb University School of Medicine

Šalata 11

10000 Zagreb, Croatia

e-mail: vladimir.trkulja@mef.hr

Number of words: 799

Number of figures/tables: 1

To the Editor,

A recent systematic review/meta-analysis¹ of randomized trials (RCTs) of tocilizumab (plus standard of care [SoC] vs. SoC w/wo placebo) in severe COVID-19 patients was a pleasure to read owing to a clear presentation of a thorough approach to data (e.g., sensitivity analyses, accounting for corticosteroid use, need for mechanical ventilation [MV] at baseline). Authors assigned high quality (certainty) GRADE levels to the evidence of efficacy in reduction of mortality overall (10 RCTs) and in patients without MV at baseline (data from 9 RCTs), and reduction of incident MV (10 RCTs). The grading was based on fixed-effect pooling, likely owing to low inconsistency index (I^2) and closely similar fixed-effect and random-effects estimates¹. It is this point that deserves a few comments. Conceptually, fixed-effect meta-analysis of RCTs in medicine is rarely justified, since the underlying assumption is practically inevitably violated due to variety of elements contributing to clinical heterogeneity². The authors¹ presented a range of differences in trial designs (e.g., one or repeated tocilizumab dose, more or less use of concomitant corticosteroids, differences in proportion of subjects on MV). When variance across trials is low, fixed and random-effects estimates are numerically close/identical, but the conceptual differences remain. Again, conceptually, the random-effects method is a preferred approach² (regardless of numerical closeness of fixed/random estimates) and the choice (fixed/random) should not be based on the heterogeneity estimates². At this point, the issue of the choice of the variance (τ^2) estimator should be mentioned. A number of estimators have been explored: performance depends on the nature of the outcome, may vary across trial sizes, depends on the differences in size of included trials, and is problematic when the number of studies is low^{e.g.,2-5}. Variance reflects on the assigned trial weights and measures of uncertainty about the pooled estimate. While no τ^2 estimator is ideal²⁻⁵, it has

been suggested that the Paule-Mandel (PM) estimator performs better than the common DerSimonian-Laird estimator for binary outcomes³. Another point to consider is the method to calculate confidence intervals (CIs) around the pooled estimate. While not without certain limitations⁶, the Hartung-Knapp-Sidik-Jonkman (HKSJ) method has been repeatedly shown (under variety of scenarios) to result in more adequate coverage probability than the standard method^{4,7}. Figure 1A re-creates meta-analysis (data presented by the authors¹) on mortality across the 10 RCTs (all subjects) – it is only that it uses PM variance estimator and HKSJ correction: random-effects estimate suggests that the mean of the distribution of the effects is 0.88 (as reported¹), but the CIs extend to 1.04, suggesting that it includes also effects that are somewhat above unity. It also provides prediction intervals (wider) - the best illustration of heterogeneity^{2,8}. When viewed from the present standpoint, data indicate a non-trivial level of imprecision and heterogeneity. The authors themselves reported apparent differences (mortality reduction vs. no reduction) between estimates based on RCTs with a high proportion vs. low proportion of patients concomitantly treated with corticosteroids¹ (or those generated accounting only for corticosteroid-treated vs. not treated patients, but such data were very scarce¹): so, there is apparent inconsistency of the estimates across clinical settings. As re-created in Figure 1B-C, there was a tendency of reduced mortality in trials with a high proportion of patients co-treated with corticosteroids (corticosteroid treatment regimen likely variable), but with quite some imprecision and heterogeneity; and no such tendency with “low corticosteroid use”. Similarly, in patients not on MV at baseline, there was a consistent reduction in mortality risk across trials with a high proportion of steroid co-treated patients, but not in trials with a low proportion of co-treated patients (Figure 1D-E). There was also a consistent reduction of risk of incident MV in trials with a high proportion of corticosteroid co-treated patients (Figure 1F), whereas the estimate in trials with “low steroid use” is burdened with heterogeneity and imprecision (Figure 1G).

Considering the above, if one were to assign a GRADE level⁹ to evidence of benefit of tocilizumab in severe COVID-19 patients based on the 10 RCTs addressed in the published meta-analysis¹, then the following seems reasonable: a) considering (indiscriminately) all 10 RCTs (and all patients), certainty about reduced mortality is closer to “low/moderate” than to “high” due to imprecision (CIs 0.75-1.04) and heterogeneity/inconsistency; b) data on the effect of tocilizumab+corticosteroid combination that could be extracted from the 10 RCTs are scarce. Trials with high vs. low concomitant use of corticosteroids could be perceived as a proxy, but this is indirect, suggestive and not conclusive evidence. Therefore, while the effects of tocilizumab on the risk of incident MV and mortality in patients not on MV at baseline in trials with a high proportion of corticosteroid co-treated patients were consistent and reasonably precisely estimated, certainty about the benefit of tocilizumab (on top of corticosteroids; regimen?) in this setting is at best moderate/low.

References

1. Vela D, Vela-Gaxha Z, Rexhepi M, Olloni R, Hyseni V, nallbani R. Efficacy and safety of tocilizumab versus standard of care/placebo in patients with COVID-19; a systematic review and meta-analysis of randomized controlled trials. *Br J Clin Pharmacol* . 2021; doi: 10.1111/bcp.15124.
2. Higgins JPT, Thomson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *J R Statist Soc A* . 2009; 172(Pt1):137-159.
3. Veroniki AA, Jackson D, Viechtbauer W, Bender R, Bowden J, Knapp G, Kuss O, Higgins JPT, Langan D, Salanti G. Methods to estimate the between-study variance and its uncertainty in meta-analysis. *Res Synth Methods* . 2016;7(1): 55-79.
4. Langan D, Higgins JPT, Jackson D, Bowden J, Veroniki AA, Kontopantelis E, Viechtbauer W, Simmonds M. A comparison of heterogeneity variance estimators in simulated random-effects meta-analyses. *Res Synth Methods*. 2019; 10(1):83-98.
5. IntHout J, Ioannidis JPA, Borm GF, Goeman JJ. Small studies are more heterogeneous than large ones: a meta-meta-analysis. *J Clin Epidemiol* . 2015; 68(8):860-869.
6. Jackson D, Law M, Rucker G, Schwarzer G. The Hartung-Knapp modification for random-effects meta-analysis: a useful refinement but are there any residual concerns? *Stat Med* . 2017; 36(25):3923-3934.
7. IntHout J, Ioannidis JPA, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random ef-

- fects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Med Res Methodol* . 2014; 14:25 doi:10.1186/1471-2288-14-25.
8. IntHout J, Ioannidis JPA, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open* . 2016; 6:e010247 doi: 10.1136/bmjopen-2015-010247
 9. Guyatt GH, Oxman, AD, Vist GE, Kurz R, Falck-Ytter Y, Schunemann HJ. GRADE: what is “quality of evidence” and why is it important to clinicians. *BMJ* . 2008;336(7651):995-998.
 10. Balduzzi S, Rucker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health* . 2019; 22(4):153-160.

Figure 1 . Re-creation of the published meta-analysis¹ using data provided in the published figures: the difference is in that the present estimates are generated using the Paule-Mandel variance estimator (Q-profile method for variance estimate confidence intervals) instead of the DerSimonian-Laired method available in the RevMan software used by the authors¹, and Hartung Knapp Sidik Jonkman correction for random effects (see text for explanation). Panel A corresponds to published¹Figure 1, panels B and C correspond to published¹supplemental Figure S4. Published meta-analysis¹ does not include figures that would correspond to panels D-G. Panels E and G are reduced to summaries for brevity. Note that in all meta-analyses point-estimates of I^2 and τ^2 were low, but the upper limits of their confidence intervals were rather high, particularly when only 4 RCTs were included (except in panel F with highly consistent results across trials). “High%” or “low %” steroid use refers to trials (as presented in the published meta-analysis¹) in which >50% or <50% of the patients were co-treated with corticosteroids. Meta-analyses were performed using `packagemeta`¹⁰ in R.

MV – mechanical ventilation; RCT – randomized controlled trial; SoC – standard of care

Hosted file

Trkulja-Figure01-Tocilizumab.docx available at <https://authorea.com/users/437234/articles/548463-tocilizumab-for-reduction-of-mortality-in-severe-covid-19-patients-how-should-we-grade-it>