Intensive Care Medicine

Editorial Un-edited accepted proof

# Multistakeholder social media peer review and fixing slips

Daniele Poole[1], Audrey De Jong[2], Giuseppe Citerio[3]

1. Operative Unit of Anesthesia and Intensive Care, S. Martino Hospital, Belluno, Italy.
2. Department of Anesthesia and Intensive Care unit, St-Eloi Hospital. PhyMedExp, Montpellier University, INSERM, CNRS, CHU de Montpellier, France
3. School of Medicine and Surgery, University of Milano-Bicocca, Monza, Italy.

**DOI: 10.1007/s00134-022-06747-5**

A trial on the antibiotic duration for Pseudomonas aeruginosa ventilator-associated pneumonia in adults was released a couple of days ago [1] during the LIVES 40 ESICM congress. The manuscript had two peer revisions, including an external statistical revision. In the hours following the release, a discussion on social media began highlighting potential methodological flaws and misinterpretation of the results.

ICM took all these multistakeholder public evaluations of the published manuscripts seriously and immediately reevaluated the situation. Two in-house statisticians reviewed the methods and the manuscript. We concluded that the methods were sound and correctly applied. However, the interpretation by the authors of this non-inferiority trial seemed to be at least ambiguous, if not overtly wrong.

The trial was aimed at demonstrating that a short duration of antibiotics (8 days) was non-inferior to prolonged antibiotic therapy (15 days) in ventilator-associated pneumonia due to Pseudomonas aeruginosa (PA-VAP) in terms of a composite outcome including death at 90-days and VAP recurrence. The prespecified margin of non-inferiority was 10%, consistently with the European Medicine Agency indication for studies dealing with hospital-acquired pneumonia.

In theory, the non-inferiority margin is established based on at least one previous trial comparing the active control of the non-inferiority trial with a placebo, demonstrating superiority. The lower confidence interval boundary of the outcome difference in this trial will then become the non-inferiority margin (figure 1, trial 7) [2]. The non-inferiority margin is usually kept lower than this to prevent, in the most extreme cases of recognized non-inferiority, conditions (Figure 1) where the new treatment effect is equivalent to or close to the impact of the placebo [3]. In non-inferiority trials, experimental and standard treatments are compared. The latter is known to be superior to the placebo, and the differences in outcomes between the two study arms are measured. Concerning the non-inferiority query, we can have three different scenarios (see figure 1). First, the higher boundary of the outcome rate difference confidence interval is lower than the non-inferiority margin. In this case, we reject the null hypothesis of inferiority and declare non-inferiority. Second, the confidence interval encompasses the non-inferiority margin. In this case, we have an inconclusive result, and we cannot either rule out neither inferiority or non-inferiority. Third, the lower confidence interval boundary is higher than the non-inferiority margin. This corresponds to evidence of inferiority [3]. The authors report that in the intention-

to-treat (ITT) population, 25 (25.5%) patients in the prolonged antibiotic therapy group and 31 (35.2%) patients in the short antibiotic therapy group had the primary outcome (difference of 9.7%, 90% CI − 2.4 to 21.9%). The per-protocol analysis, the first-choice analysis in a non-inferiority trial [4], showed similar results (difference 12.8%, 90% CI 0–25.6%). Both analyses provided inconclusive findings since the null hypothesis of inferiority could not be rejected since the upper limit of the confidence interval was higher than the 10% value of the non-inferiority margin. Although, in the results section, the authors correctly state that non-inferiority could not be demonstrated, in the abstract, in the discussion, and in the takeaway message, they incorrectly point out, upon one reviewer's suggestion, that no differences could be found between the study arms. Indeed, in general, this would be the objective of a superiority trial, but not for a non-inferiority trial that, by design, admits that the new treatment may be worse by a certain amount (that should be clinically neglectable) the known effective treatment. Moreover, in this specific case, the results were inconclusive, not surprising since the study was prematurely interrupted after about one-quarter of the planned number of patients was randomized because of difficulties in recruiting patients.

Thus, the trial of Bouglé et al. failed to provide any information about the non-inferiority of short compared to prolonged antibiotic therapy in PA-VAP. These results concordance with the secondary outcomes, 90-day overall survival rate and recurrence of PA-VAP, showing trends in favor of prolonged antibiotic treatment. Even the post-hoc Bayesian analysis, with priors that could be discussed [5], confirmed low probability rates of non-inferiority (40%).

For clinicians and ICM readers, this study results do not suggest that a short duration of antibiotic therapy in PA-VAP is equivalent to a longer one.

We recognize that peer review is not perfect [6] and that we, as humans, are prone to make errors. Knowing that authors, reviewers, and editors do make errors and that peer review is not a fail-proof system, we heard the voice of the online researchers to correct the research record. However, if "*errare humanum est, perseverare autem diabolicum*". Once we recognized the flaws, we speedily started to repair them, contacting the Authors immediately for a correction and a reply. Moreover, the authors were not aware of the discussion on social media, and we, as authors and clinicians, are not obliged to follow all the discussions on the internet. ICM would like to keep and promote the discussion of the manuscripts as letters to the journal. We are a pretty fast journal, and the delay in publication of a debate is very short.

As we wrote in an editorial during the early phase of the pandemic [7], "*we hope that scientific journals, with their imperfect but rigorous and fair review processes, will have reached even a higher standard for their readers, proving to be the right place to find, in a reasonable amount of time, appropriate clinical science for patients*."

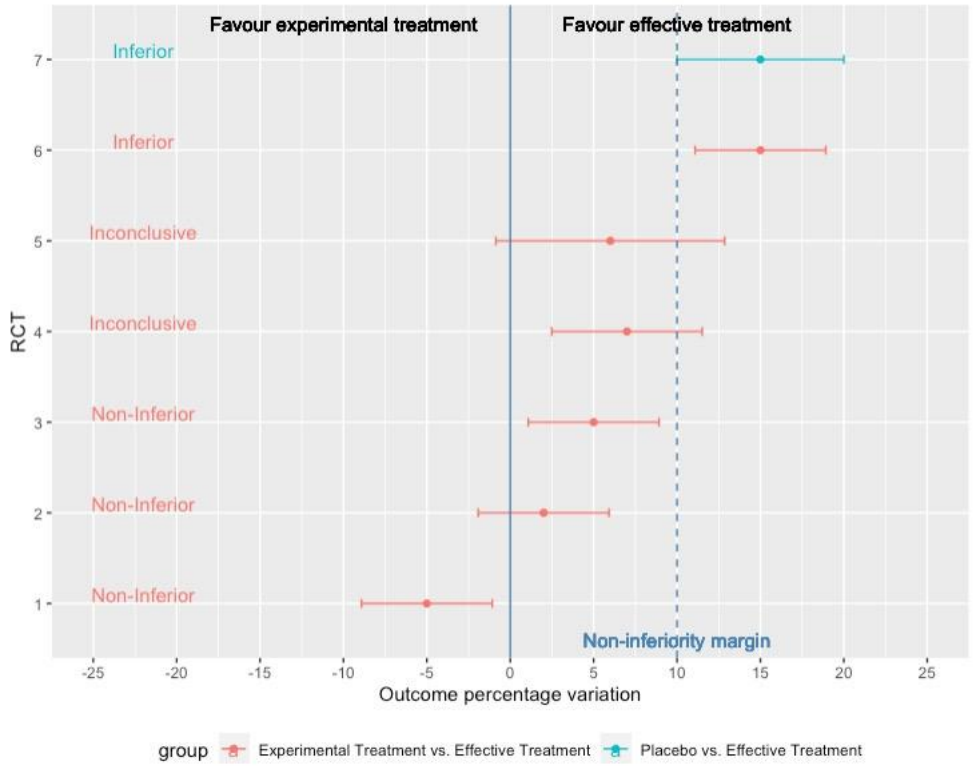


Figure 1: simulation of one superiority trial including the active control vs. placebo (RCT 7) and six non-inferiority RCTs (from 1 to 6) comparing the experimental treatment with the active control. The lower confidence interval boundary of the superiority trial defines the non-inferiority margin for the non-inferiority RCTs (dashed blue line). Three different scenarios can be identified concerning the non-inferiority query: non-inferiority, inconclusiveness, and inferiority. Besides RCT 7, in this example, we have not considered superiority issues in relation to the zero threshold (continuous blue line). Indeed, in trial 1 we also have superiority of the experimental treatment compared to the active control, while in trials 3, 4, and 6, it is inferior. RCT 3 is a particular, and mostly theoretical, case of non-inferiority in the non-inferiority perspective, and inferiority when adopting a superiority design (using the zero-difference threshold).

## Conflicts of interest:

The authors declare no conflicts of interest.

References

1. Bougle A, Tuffet S, Federici L, Leone M, Monsel A, Dessalle T, Amour J, Dahyot-Fizelier C, Barbier F, Luyt CE, Langeron O, Cholley B, Pottecher J, Hissem T, Lefrant JY, Veber B, Legrand M, Demoule A, Kalfon P, Constantin JM, Rousseau A, Simon T, Foucrier A, i DTI (2022) Comparison of 8 versus 15 days of antibiotic therapy for Pseudomonas aeruginosa ventilator-associated pneumonia in adults: a randomized, controlled, open-label trial. Intensive Care Med. https://doi.org/10.1007/s00134-022-06690-5
2. Kaul S, Diamond GA (2006) Good enough: a primer on the analysis and interpretation of non-inferiority trials. Ann Intern Med 145: 62-69
3. Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG, Group C (2012) Reporting of non-inferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. JAMA 308: 2594-2604
4. Mo Y, Lim C, Watson JA, White NJ, Cooper BS (2020) Non-adherence in non-inferiority trials: pitfalls and recommendations. BMJ 370: m2215
5. Granholm A, Munch MW, Møller MH, Lange T, Perner A (2022) Choice of priors: how much scepticism is appropriate? Intensive Care Medicine 48: 372-373
6. Csiszar A (2016) Peer review: Troubled from the start. Nature 532: 306-308
7. Citerio G, Bakker J, Brochard L, Buchman TG, Jaber S, Mazzone PJ, Teboul JL, Vincent JL, Azoulay E (2020) Critical care journals during the COVID-19 pandemic: challenges and responsibilities. Intensive Care Med 46: 1521-1523