Metering the METEOR in methotrexate failure: is propensity score a falling star?

The debate on the preferred treatment of rheumatoid arthritis (RA) after methotrexate (MTX) failure has always been fascinating. Sytske Anne Bergstra and associates have analysed data from patients in the METEOR cohort who had failed MTX monotherapy. It was absorbing to see how they used multiple propensity score (PS) to control for bias in non-randomised data collected from an international registry (‘METEOR’). They have pragmatically outlined how the PS was analysed.

The analysis has demonstrated clear superiority of biological disease-modifying anti-rheumatoid drugs (bDMARDs) over conventional synthetic DMARDs (cs-DMARDs) in both unadjusted and PS-adjusted results at 6 months and at 1 year. There is already plenty of evidence that addition of a bDMARD to MTX (on failure of monotherapy) is better in the short term than addition of a cs-DMARD. A network meta-analysis of 33 trials has shown the superiority of bDMARDs at 6 months in achieving higher disease control, that is, in the American College of Rheumatology 70% response criteria (ACR70). Similarly, a meta-analysis of eight such trials demonstrated similar benefits with bDMARDs at 6 months, but the differences disappeared at 24 months. The European League Against Rheumatism (EULAR) already recommends adding a bDMARD (over adding a csDMARD) in MTX failure if ‘unfavourable factors’ are present. At the present date, the evidence for benefits of this approach is only for the short term, and not beyond 2 years.

The contrast in this analysis is that Bergstra et al have used rheumatoid arthritis disease activity score (DAS) as the primary outcome measure. Clinical trials usually use the ACR20 criteria. Even in the RACAT trial that showed non-inferiority of triple drug therapy versus addition of etanercept in MTX failure, DAS28 reduction was more in the etanercept group at 6 months. Nevertheless, ACR20 and ACR50 were similar at 6 months (and at 2 years). Differences in DAS28 and ACR70 vanished by the end of 2 years. Again, difference in DAS (includes 44 swollen joint count and 78 tender joint count) may be numerically different than difference in DAS28.

I would have liked to know what proportion of patients in the three groups of the study met ACR20 at the two time points. And, hopefully, the authors will publish the results of the follow-up of the METEOR MTX monotherapy failure cohort at 24 months as well. It is interesting that the multiple PS adjustments did not change the differences in DAS by much. Improvement in DAS was almost double in the bDMARD group compared with the others. The bDMARD group had higher number of swollen joints and thus seemed likely to show greater difference on adjustment by the PS. But this was not so.

PS basically combines the effect of all known confounders into one covariate. It is good for studies with rare events and multiple confounders. As the authors have mentioned, the PS cannot control bias due to factors not included in the analyses. One such factor can be the placebo effect of biologicals on patients. Again use of the Ritchie articular index (RAI) and the DAS (that includes the RAI) allows for variability in the subjective evaluation. This might unconsciously augment a bias in the evaluator who is aware that the patient is on biologicals. Another potential confounder was the MTX dose. Patients in the csDMARD +steroids group were getting almost half the MTX dose compared with the other two groups. If this difference was maintained throughout the observation period, that itself would lead to differences in the DAS. Also, I did not find information on steroid use in the bDMARD group.

This paper has reiterated the short-term superiority of bDMARDs over csDMARD after MTX failure. In a comparison of cost-effectiveness and quality of life, results depended on the methods used. We are yet to know what is best for the patient in the long term.

The highlight of the paper seems to be use of the PS. But PS-adjusted analysis did not differ much from the unadjusted analysis. Maybe because it has good validity. Maybe its use is superfluous until we understand all the confounders in RA. Either way, it is for the reader to decide if PS is a guiding star or a falling star!

Sakir Ahmed

Correspondence to Dr Sakir Ahmed, Clinical Immunology and Rheumatology, Kalinga Institute of Medical Sciences, Bhubaneswar, Odisha 751024, India; sakir005@gmail.com

Handling editor Josef S Smolen

Contributors The sole author has completely drafted and finalised the manuscript.

Competing interests None declared

Patient consent Not required.

Provenance and peer review Not commissioned; internally peer reviewed.

© Author(s) (or their employer(s)) 2019. No commercial re-use. See rights and permissions. Published by BMJ.

To cite Ahmed S. Ann Rheum Dis 2019;78:e131.

Received 19 October 2018
Accepted 26 October 2018
Published Online First 8 November 2018

http://dx.doi.org/10.1136/annrheumdis-2018-214612


REFERENCES


