

EDITORIAL

Efficacy assessed in follow-ups of clinical trials: methodological conundrum

Robert BM Landewé*

See related research by Rantalaiho *et al.*, <http://arthritis-research.com/content/12/3/R122>

Abstract

Increasingly, we see papers describing the long-term follow-up results of randomised clinical trials. Sometimes, like the article by Rantalaiho and colleagues in the previous issue of *Arthritis Research & Therapy*, the follow-up extends to more than 10 years. It is not uncommon that authors of such articles describe their results as a comparison of the original treatment groups in the original randomised clinical trial. Methodologically, such a comparison is fallible for several reasons. In this editorial, two important sources of bias that may jeopardise the results of such follow-up studies are discussed: confounding by indication and confounding by trial completion.

Long-term follow-ups of randomised clinical trials are a *contradictio in terminis*.

With this rather bold statement I do not mean that such studies are impossible to conduct. Rantalaiho and colleagues have proven with the publication of the 11-year follow up of their world-famous Fin-RACo trial that dedicated investigators and patients who believe in the goals of the study can create a dataset that is insurmountable in terms of wealth, from which we can learn a lot about the long-term fate of patients with rheumatoid arthritis (RA) [1]. The authors have carefully analysed the available radiographic data, they have investigated important long-term outcomes such as mortality and joint-replacement surgery, and they have appropriately modelled longitudinal data. Their conclusion that early aggressive therapy with combinations of conventional disease-modifying antirheumatic drugs including corticosteroids pays off in terms of long-term radiographic and clinical benefits is credible. And their

argument that 'treat to target' is the best way to exploit those benefits is convincing [1].

What concerns me most in Rantalaiho and colleagues' interpretation – and admittedly in similar exercises in which I took part myself [2,3] – is the implicit assumption that two groups of patients formed a decade ago by a stochastic process that we call randomisation can be compared 11 years later under the same premise of prognostic similarity.

Groups in randomised clinical trials (RCTs) may violate prognostic similarity even at baseline. Chance theory tells us that if we were to perform the procedure of randomisation 1,000 times, we may face a number of attempts with a number of imbalances, sometimes even in prognostically relevant variables. We usually ignore such baseline differences, assuming that imbalances may occur in either direction, and their combined net effect on the outcome of interest is probably negligible. The important consideration is that these baseline differences are completely by chance (random), which means 'not driven by any tangible or impressionable process'.

I need this piece of theory to convince you that Rantalaiho and colleagues' 11-year-old RCT follow-up has suffered from many influences that may have jeopardised prognostic similarity. Let us look through the spectacles of the trial methodologist and play devil's advocate by working out two important biases: confounding by indication and confounding by trial completion.

The Fin-RACo trial had a protocol for only 2 years [4], implying that any treatment choice thereafter was up to the discretion of the doctor and the patient. Undoubtedly, the physician wanted the best for the patient, thus prioritising the patient's wellbeing over the fate of the study. A consequence of good clinical practice, however, is that – as confirmed by Rantalaiho and colleagues – the worst patients may have received the most intensive (effective, costly) treatment, which may in turn have unquantifiable influences on the outcome of interest. If such events occur in an unbalanced fashion, we speak about confounding by indication. I think in RA, with its many effective treatments to choose and its inextricable relationship between disease activity (determinant) and

*Correspondence: r.landewe@mumc.nl
Maastricht University Medical Center, Department of Internal Medicine/
Rheumatology, P.O. Box 5800, 6202AZ Maastricht, The Netherlands

radiographic progression (outcome measure) [5], confounding by indication should be a number-one reason to refrain from statistical between-group comparisons in long-term follow-ups of RCTs.

The second issue is related to the first, but is slightly different in nature: confounding by trial completion. Obviously, the investigators have done their best in obtaining the outcome of interest in as many patients as possible. Expectedly, they have not been able to assess outcome in every patient. What is important from a methodological point of view is whether this loss to follow-up was completely random. Usually it is impossible to determine the exact reasons for patients not showing up at a control visit or an end-of-study assessment. Usually, therefore, it is impossible to conclude that a no-show (or missing) had nothing to do with the severity and activity of the RA. What follows is that you cannot be sure that such events are distributed evenly across trial groups, and therefore every between-group comparison under the assumption of prognostic similarity is meaningless. Rantalaiho and colleagues have done their best to collect as many radiographs from as many patients as possible, but – not unexpectedly – more than 30% of the patients miss their 11-year radiographic assessment. The investigators may, like many authors do, provide inferential arguments that drop-out is not relevant in their study, but unfortunately one cannot judge.

These two biases mean I am rather reluctant to accept firm conclusions from follow-ups of RCTs that have been analysed a decade after the randomisation procedure, however credible they may seem. Many events may have occurred in every individual patient in the trial that may have broken prognostic similarity. I therefore do not truly believe in the explanation of differences after 10 years of intangibly trying to influence patients' fates.

Does this make Rantalaiho and colleagues' results useless? Absolutely not. We welcome cohorts of patients that have been followed for years in order to find out what eventually determines the disease course. Ideally such cohorts include patients with severe and less severe disease, with more and less active RA, with more and less aggressive initial treatment. We should know a lot more about these patients' fates; their baseline values and their

baseline biomaterials are extremely important in defining new prognostic biomarkers. Such carefully conducted studies may give insight into what is really important in determining an individual patient's prognosis in a world full of treatment choices that differ in efficacy, effectiveness and cost.

Explained in terms of *contradictio in terminis*, the contradiction is in the recognition that the randomised part of a RCT is not necessarily a licence for harmlessly comparing treatment effects after a decade of follow-up of that trial.

Abbreviations

Fin-RACo, Finnish Rheumatoid Arthritis Combination Therapy; RA, rheumatoid arthritis; RCT, randomised clinical trial.

Competing interests

The author declares that he has no competing interests.

Published: 30 July 2010

References

1. Rantalaiho V, Korpela M, Laasonen L, Kautiainen H, Järvenpää S, Hannonen P, Leirisalo-Repo M, Blåfield H, Puolakka K, Karjalainen A, Möttönen T; FIN-RACo Trial Group: Early combination disease-modifying antirheumatic drug therapy and tight disease control improve long-term radiologic outcome in patients with early rheumatoid arthritis: the 11-year results of the Finnish Rheumatoid Arthritis Combination Therapy trial. *Arthritis Res Ther* 2010, **12**:R122.
2. Landewé RB, Boers M, Verhoeven AC, Westhovens R, van de Laar MA, Markusse HM, Jacobs P, Boonen A, van der Heijde DM, van der Linden S: COBRA combination therapy in patients with early rheumatoid arthritis: long-term structural benefits of a brief intervention. *Arthritis Rheum* 2002, **46**:347-356.
3. van Tuyl LH, Boers M, Lems WF, Landewe RB, Han H, van der Linden S, Peeters AJ, Jacobs P, Huizinga TW, van de Brink H, Dijkmans BA, Voskuyl AE: Survival, comorbidities and joint damage 11 years after the COBRA combination therapy trial in early rheumatoid arthritis. *Annals Rheum Dis* 2009, **69**:807-812.
4. Mottonen T, Hannonen P, Leirisalo-Repo M, Nissila M, Kautiainen H, Korpela M, Laasonen L, Julkunen H, Luukkainen R, Vuori K, Paimela L, Blåfield H, Hakala M, Ilva K, Yli-Kerttula U, Puolakka K, Jarvinen P, Hakola M, Piirainen H, Ahonen J, Palvimaki I, Forsberg S, Koota K, Friman C: Comparison of combination therapy with single-drug therapy in early rheumatoid arthritis: a randomised trial. FIN-RACo trial group. *Lancet* 1999, **353**:1568-1573.
5. Welsing PM, Landewe RB, van Riel PL, Boers M, van Gestel AM, van der Linden S, Swinkels HL, van der Heijde DM: The relationship between disease activity and radiologic progression in patients with rheumatoid arthritis: a longitudinal analysis. *Arthritis Rheum* 2004, **50**:2082-2093.

doi:10.1186/ar3080

Cite this article as: Landewé RBM: Efficacy assessed in follow-ups of clinical trials: methodological conundrum. *Arthritis Research & Therapy* 2010, **12**:132.