Rheumatology 2000;39:1–6

## Editorials

# Potential bias in Kaplan–Meier survival analysis applied to rheumatology drug studies

Survival analysis using methods due to Kaplan and Meier [1] is the recommended statistical technique for use in cancer trials [2]. It is applied by analysing the distribution of patient survival times following their recruitment to a study. The analysis expresses these in terms of the proportion of patients still alive up to a given time following recruitment. In graphical terms, a plot of the proportion of patients surviving against time has a characteristic decline (often exponential), the steepness of the curve indicating the efficacy of the treatment being investigated. The more shallow the survival curve, the more effective the treatment. Kaplan–Meier analysis can be used to test the statistical significance of differences between the survival curves associated with two different treatments.

In recent years, such survival analysis has found widespread application in the field of rheumatology [3–7], being used to examine how long patients continue to tolerate a particular drug before either side-effects or lack of efficacy cause them to be switched to some other therapy. Although intuitively this seems analogous to oncology trial analysis, little research seems to have been done to confirm that survival analysis is a valid method to examine drug survivorship. In spite of this, survival analysis is becoming not only accepted within rheumatology, but dogmatically insisted upon. Indeed the authors have experienced harsh criticism from referees for not using it.

Of course, as a well-developed and investigated statistical method, the mathematical logic underlying Kaplan–Meier analysis is impeccable and we certainly would not dispute its value when applied in appropriate circumstances. However, just because a statistical technique has been shown to be mathematically sound and applicable to one clinical study does not make it universally applicable to all clinical studies.

Frustrated by numerous referees insisting on Kaplan–Meier analysis and concerned about the applicability of the method to drug survival, we have carried out some simple tests of the assumptions underlying the method. Sadly, it does not fare well.

In order to test the method, we used data for 606 patients recruited as part of a follow-up study investigating the long-term progression of patients initially recruited with early rheumatoid arthritis (RA). This forms part of a programme of research known as the Early Rheumatoid Arthritis Study (ERAS). Our tests of Kaplan–Meier methods concern a subgroup of the full ERAS patient group comprising 606 patients who completed at least 2 yr follow-up and were prescribed sulphasalazine as their first disease-modifying antirheumatic drug (DMARD). We have focused attention on patients whose first DMARD was sulphasalazine

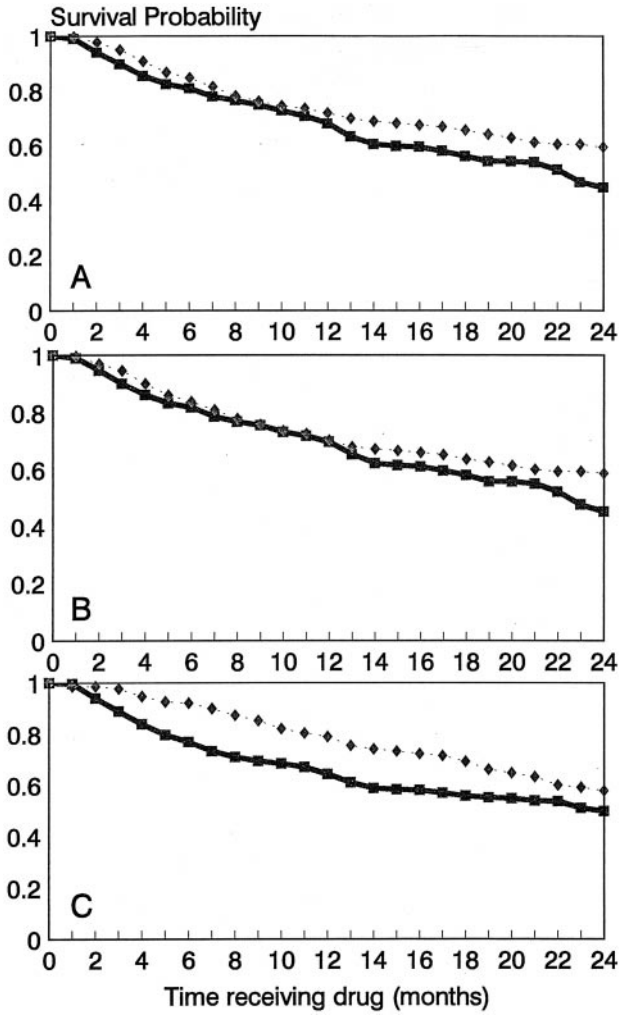


FIG. 1. Drug survival curves for patients subdivided in different ways: (A) two groups of 303 patients defined by recruitment date being before (top curve) or after (bottom curve) November 1991; (B) two groups of 303 patients defined by DMARD start date being before (top curve) or after (bottom curve) February 1992; and (C) two groups defined by the lag between recruitment and DMARD start being more than 3 months (233 patients, top curve) or within 3 months (373 patients, bottom curve).

since it was the drug most commonly prescribed as a first DMARD within the ERAS cohort.

We have investigated three assumptions that are fundamental to the validity of Kaplan–Meier survival analysis in relation to drug survivorship:

(A) patients recruited early in the study should have the same drug survival characteristics as those recruited later;
(B) patients receiving their first DMARD later in the study should have the same drug survival characteristics as those receiving it earlier;
(C) drug survival characteristics should be independent of the time that a patient has been in the study before first receiving a DMARD.

In order to examine if these assumptions are valid, we divided data for our cohort into two groups, according to the criterion being examined. Thus, to examine whether drug survival depends on when patients were recruited to the study (assumption A), the 606 patients were divided into two groups of 303 patients each, the former group having earlier recruitment dates than the latter. Survival curves were plotted to compare the drug survival characteristics of sulphasalazine between the two groups. Similar methods were used to examine assumptions (B) and (C). The results of this process are summarized in Fig. 1. In each case, the drug survival characteristics were significantly different between the two groups ($P < 0.01$, log rank test for each case). It should be noted that these tests are not wholly independent. However, correction for multiple testing would have little effect on our conclusions.

It is not clear why there should be these differences in drug survivorship and one can only speculate about the reasons. Indeed to do more than speculate would be unscientific, since the ERAS study was not set up specifically to examine such issues. However, whether or not we can explain the differences is irrelevant to the issue under discussion. The fact is that there are differences and, for the cohort studied, the assumptions required by Kaplan–Meier analysis are not valid.

This, of course, should be disturbing news for advocates of Kaplan–Meier drug survival analysis. Of course we cannot infer whether the same is true for other follow-up studies, which may be some solace. But against this, the onus should now be on those who use such drug survival analysis to show that it is valid, given the evidence that there is one case where it is not.

M. Utley, S. Gallivan, A. Young[1], N. Cox[2], P. Davies[3], J. Dixey[4], P. Emery[5], A. Gough[6], D. James[7], P. Prouse[8], P. Williams[9] and J. Winfield[10]

*Clinical Operational Research Unit, University College London, [1]St Albans City Hospital, St Albans, [2]Royal Hampshire County Hospital, Winchester, [3]Broomfield Hospital, Chelmsford, [4]Robert Jones and Agnes Hunt Orthopaedic Hospital, Oswestry, [5]Rheumatology Research Unit, University of Leeds, Leeds, [6]Harrogate Hospital, Harrogate, [7]Grimsby Hospital, Grimsby, [8]Basingstoke District Hospital, Basingstoke, [9]Medway Hospital, Gillingham and [10]Royal Hallamshire Hospital, Sheffield, UK*

Correspondence to: S. Gallivan, Clinical Operational Research Unit, 4 Taviton Street, London WC1H 0BT, UK.

## References

1. Kaplan EL, Meier P. Non parametric estimation from incomplete observations. J Am Stat Assoc 1958;53:457–81.
2. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV *et al.* Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. Br J Cancer 1976; 34:585–612.
3. Wijnands M, van Riel P, van't Hoff M, Gribnau F, van de Putte L. Longterm treatment with nonsteroidal anti-inflammatory drugs in rheumatoid arthritis: a prospective drug survival study. J Rheumatol 1991;18:184–7.
4. Buchbinder R, Hall S, Sambrook PN, Champion GD, Harkness A, Lewis D *et al.* Methotrexate therapy in rheumatoid arthritis: a life table review of 587 patients treated in community practice. J Rheumatol 1993; 20:639–44.
5. Jones E, Jones JV, Woodbury JF. Response to sulphasalazine in rheumatoid arthritis: life table analysis of a 5-year follow-up. J Rheumatol 1991;18:195–8.
6. De La Mata J, Blanco FJ, Gomez-Reino JJ. Survival analysis of disease modifying antirheumatic drugs in Spanish rheumatoid arthritis patients. Ann Rheum Dis 1995;54:881–5.
7. Saag KG, Koehnke R, Caldwell JR, Brasington R, Burmeister LF, Zimmerman B *et al.* Low dose long-term corticosteroid therapy in rheumatoid arthritis: an analysis of serious adverse events. Am J Med 1994;96:115–23.