Editorial

Surviving survival analysis

RM Szydlo

Department of Haematology, Hammersmith Hospital, London, UK

The two-part paper entitled 'Statistical Methods for the Analysis and Presentation of the Results of Bone Marrow Transplants' summarises current methods for appropriate analysis of data generated in a bone marrow, or more correctly, a stem cell transplant programme. The first part, which appears in this issue of *Bone Marrow Transplantation*, covers unadjusted analyses and addresses the importance of understanding the relationship between outcome and time, thereby providing a basis for estimating the probability for an event occurring. The second part, to be published in a forthcoming issue of the journal, contrasts regression with matched pair analysis and addresses basic concepts underlying the proportional hazards (Cox) regression model. This section concludes with specific recommendations for analysing transplant data as preferred by a group of acknowledged experts in the field.

Part 1 outlines two methods for calculating a probability curve. Kaplan–Meier analysis is appropriate where the outcome is readily defined (eg death or treatment failure) and occurs at a random time. Cumulative incidence is appropriate where there is the possibility of competing risks (eg death in remission will preclude relapse, and so is treated as a competing risk). Where a competing risk is present, use of the Kaplan–Meier method will overestimate the outcome in question and yield a quantity that is uninterpretable.

Three methods for making unadjusted comparisons of outcomes between groups are outlined. (a) Outcomes that occur early after SCT (eg transplant-related deaths within 100 days post SCT) may utilise the standard chi-squared test. However, patients who are not followed up for the entire time period are not evaluable and cannot therefore be included. (b) If survival curves are calculated using the Kaplan–Meier estimator, then a log rank test can be used to compare the entire survival experience between groups. However, this test compares outcomes over the whole time interval, and may not detect differences between the groups that occur early or late in the time interval, and so a more appropriate weighted log rank test should be used. (c) At specific time-points, the survival probabilities can be compared by calculating a Z-score based on the difference in probabilities and their variance estimates.

The advent of therapies such as donor lymphocyte infusions to successfully induce a second durable remission after relapse of original disease post SCT, has prompted

the development of a new summary outcome called current leukaemia-free survival (CLFS). Unlike the conventional leukaemia-free survival curve, the CLFS curve can increase, reflecting patients attaining a second remission.

Part 2 describes how proportional hazards regression is best utilised to study outcomes associated with stem cell transplantation. This regression method is a model for a patient's hazard rate, defined as the chance of the event of interest occurring in the next instant of time for a patient yet to experience the event. It can be used to investigate relationships for patients with different risk factors or covariates. Before any analysis can be carried out, covariates have to be appropriately coded in order to make interpretation of the results of the regression analysis meaningful. Continuous data can be treated as such, but as illustrated in the paper, interpretation may be difficult, thus dichotomisation with a choice cut-point may be preferable. The merits of these approaches are presented.

In addition to the classification issues described above, covariates also have to be described in terms of their relationship with time. Thus variables such as age at the time of transplant or patient sex are called time-fixed covariates, whilst those such as engraftment or the first occurrence of acute graft-versus-host disease (AGVHD) are time-dependent covariates. It is important to make this distinction and not to misinterpret time-dependent variables as time-fixed. The authors present results of a study with survival after diagnosis being the outcome, and therapy (BMT vs chemotherapy) being the covariate of interest. They find that the effect of therapy changed over time, thus defining therapy as a time-varying fixed effect.

Once an analysis has been carried out and a model has been found it is imperative to check that the underlying assumptions of the model are correct. The most important of these is that the effects of the risk factors are constant over the follow-up time period. This can be tested by adding a time-dependent covariate to the model and looking at its regression coefficient (this should be zero). Alternative graphical approaches can also be utilised.

In addition, this detailed paper includes a commentary on the use of matched pairs analysis, and sections on how to deal with missing data, how to determine if there is an interaction between two factors in their effect on outcome, and how best to present results in a publication.

The two papers include a comprehensive reference list, as well as a glossary of statistical terms that should satisfy the most enthusiastic of readers in their quest for coming to terms with an important technology that is all too often misapplied.