Liver and biliary

Variceal bleeding: does variable delay in trial entry invalidate comparison between trials?

D BULLIMORE
From St James's Hospital, Leeds

Summary Variation in time of patient admission to studies of prognosis after variceal haemorrhage has been proposed as a major factor in the wide range of reported results. A study of 144 unselected subjects with a low initial mortality (3% at two days) suggests that the effect has been overemphasised and studies in which time of entry is later than the date of bleeding may be usefully compared. Reanalysis of previous work suggests that even in populations with a high initial mortality limited comparisons between studies can still be usefully made. It is stressed that survival analysis must be started from the date of study entry and not from the date of the index bleed which involves incorporating a retrospective period of survivorship.

The mortality after variceal haemorrhage remains high although variable from series to series at from 15% to over 50%. Part of the variability is because of the differences in the populations studied and part is to the selection of subjects accepted for study. In their recent series Graham and Smith emphasised the importance of the time delay from the onset of bleeding to the entry of subjects into a study as a major contributor to prognosis. Thus of 85 consecutive subjects with variceal bleeding a delay of as little as two days before inclusion of subjects in a study would have resulted in greatly differing survival curves. One consequence of this is that comparison of published series where entry occurred several days after the initial bleed has become very difficult. Because the very high mortality of 36% by two days reported by Graham and Smith did not appear typical of experience outside the USA a similar analysis on 144 subjects with bleeding oesophageal varices was undertaken to determine whether a comparable dramatic effect of time on survival would be apparent.

Methods

Subjects
Data on all subjects with oesophageal varices diagnosed in the endoscopy unit at St James's Hospital, Leeds are stored on a purpose designed computer system with associated statistical software. Data on all 144 different subjects with bleeding oesophageal varices were analysed. The index bleed was the first episode of variceal bleeding identified within the unit and the date of entry was the date of admission with that bleed. Endoscopic criteria for the diagnosis of variceal bleeding were identical to those used by Graham and Smith and functional hepatic reserve was assessed using Pugh's modification of Child's grade. Life table analysis and the log-rank test was used to compare survivorship of different populations. Associated $\chi^2$ values are all for one degree of freedom. Survivorship was compared for the whole of the follow up period. Sixty two, 36, 17, and eight subjects were followed up beyond one, two, three, and five years respectively.

Results

Fifty eight per cent of subjects were men and 28% had alcoholic liver disease. Other diagnoses included idiopathic cirrhosis (17%), primary biliary cirrhosis (13%), chronic active hepatitis (13%), extra-hepatic obstruction (12%, including that secondary to malignancy), and non-cirrhotic portal hypertension (5%). Twelve subjects (8%) were unclassified and 12% had a variety of other diagnoses—for example, alpha one antitrypsin deficiency.
subjects had more than one diagnosis—for example, chronic active hepatitis plus extrahepatic obstruction. A Child's grade was available in 140 cases (97%). The distribution of Child's grade was A 24%, B 40%, and C 36%. Using life table analysis survival of alcoholic and non-alcoholic subjects was similar (p>0.1).

**EARLY MORTALITY**

Of 144 subjects at entry 140 survived to two days, 134 to seven days, and 121 to 14 days. Beyond 14 days numbers available for analysis at 30 days, six weeks, two months and three months were 114, 109, 100, and 93.

**SURVIVAL CURVES**

The effect of changing the zero time for calculating survival of the remaining population by between two days and three months is shown in Figure 1. The entry and two day curves are not shown separately as they differ by only 2% (p>0.75). There is no statistically significant difference between the entry and two week curves (p>0.1) nor the two week and three month curves (p>0.1). Entry and three month curves are significantly different (p<0.025).

**Discussion**

Both this series and the series of Graham and Smith include all subjects with bleeding oesophageal varices. Their series included subjects with alcoholic and non-alcoholic cirrhosis, benign extrahepatic obstruction and extrahepatic obstruction secondary to malignancy. Ninety one per cent of their series had alcoholic cirrhosis, a figure typical of series from the United States of America. Series from the United Kingdom, similarly not restricted to cirrhotic subjects, have fewer patients with alcoholic cirrhosis, frequently in the range 30–40%.

The purpose of this paper is to concentrate on major differences in the overall survival curves presented in this study and that of Graham and Smith and on the consequences of those differences. In particular whether comparison of trials whose entry date is variable, but later than the date of the variceal haemorrhage, may be undertaken without that variability invalidating the comparison. The size of this study is sufficient so that the conclusions reached are applicable to the published studies available for comparison with each other. The survival curve from the current study is shown in Figure 1 and the survival curve from their study is reproduced, with permission, in Figure 2.

Mortality in their study was 36% (31/85) by day two compared with 3% (4/144) by day two in this study. (The entry and day two curves have not been plotted separately on Figure 1 as they differ by only
Variceal bleeding: does variable delay in trial entry invalidate comparison between trials?

2%, \( \chi^2 = 0.024, p > 0.75 \). Therefore two studies one of which included all subjects from day two and one of which included all subjects from the date of the index bleed could be usefully compared, given other factors were equal, in a UK population or other population with a similar low initial mortality.

The curves remain comparable at seven days (p>0.5) and at 14 days (p>0.1) although the difference between them is increasing. Thus studies which admitted subjects within one to two weeks of the initial bleeding period may also be compared without major bias because of the variation in the time of entry to the trial. The survival curves of the populations show only modest differences. This contrasts with the survival curves shown in Figure 2 where 48% of the population had died by two weeks.

Similarly late entry studies in which therapy has been considerably delayed may to some extent be compared. Thus survival curves of the subpopulations surviving two weeks, four weeks, six weeks, and two and three months do not differ significantly in a statistical sense (p>0.1 for the two week v three month curve).

While it is desirable to include subjects in a study from as close to the date of the index bleed as is possible, to minimise the risk of bias, in populations with low initial and early mortality, appropriate comparison of studies in which the time of entry is later than the index date can be undertaken. The very different two day (3% v 36%) and two week (16% v 48%) mortality largely accounts for the difference between the conclusions of this study and the previous study.2

Are there any other lessons to be learnt from the survival curves shown in Figure 2?

First, the form of the curves requires comment. Succeeding curves represent subpopulations of the entry population. Beyond three months a death affecting one population will affect them all. As the '100% population' of successive curves from entry to three months is decreasing, a death must cause an increasing percentage reduction in survival from entry to three month curve. That is the curves must converge, they can not diverge or run parallel but decreasing. For example, if an initial population of 100 subjects was reduced to 50 by two weeks and 25 by three months then a death occurring at four months would reduce the entry curve by 1%, the two week curve by 2% and the three month curve by 4%. The form of the plotted curves cannot therefore be correct (Fig. 2).

Second, Graham and Smith noted that most surgical series accept the zero time for survival as the date of the operation. Their figure was said to illustrate 'the effect of changing the zero time for the survival curve from the index bleed to two days, two weeks and three months after the index bleed'.2 Figure 2 does not show any alteration in 'zero time'. 'Zero time' for all curves remains the date of the index bleed. What is in fact altered is the time after the index bleed (two days, two weeks or three months) at which the remaining population is considered to be '100%'. Hence the plateau between zero and three months on the three month curve, absent from the entry curve, which shows the retrospective survival of this subpopulation to the date of the index bleed. The dotted curve shows the three month curve moved three months to the left: a survival curve with the zero time for survival truly shifted from entry to three months. It is evident that the two day, two week, and three month curves are not statistically significantly different. A subpopulation admitted to a study two weeks or three months after the index bleed with survival calculated from time of entry to the study, as is the normal convention with such studies, would on this basis have similar survival curves. Beyond the first few days of very high mortality their data do not tend to support the statement that 'the longer a patient survives the better is his prognosis'.

Differences between survival curves can be accentuated both by using a relatively short follow up period and by using retrospective survival plateaux of the form discussed above. The two week and three month survival curves of Figure 1 were compared statistically using the whole of the patient follow up period, as were the data of Graham and Smith, although for comparability only the first year is plotted. Curtailing the life table analysis for the two week v three month curve to one year increases the \( \chi^2 \) value from 1.6 to 3.6, significant at the 10% level. Incorporating a retrospective survival plateau of 14 and 90 days respectively in addition increases the \( \chi^2 \) value to 4.35, a difference significant at the 5% level. Such plateaux may be generated in surgical series when, for example, 30 day postoperative mortality is excluded but survivorship in survivors is calculated from the day of operation.

In conclusion variation in the time of entry of subjects to studies following variceal bleeding, in populations with low two day and two week mortalities, causes relatively little bias when such studies are compared provided differences do not become extreme. Even in populations with a high two day mortality the difficulties of comparison of late entry trials—for example, two weeks to three months after the index bleed, may have been overemphasised. Survivorship must, however, be calculated from the time of trial entry with no retrospective survivorship to the date of the index bleed.
References


Forthcoming Scientific Meetings of the British Society of Gastroenterology

AUTUMN 1987 — 15–18 September
Jubilee Meeting at the University of London Institute of Education, London

SPRING 1988 — 23–25 March
University of Leicester

AUTUMN 1988 — 13–16 September
University of Sheffield

For details of registration please contact: The Administrative Secretary, BSG, 3 St Andrew's Place, Regent’s Park, London NW1 4LB. Tel: 01–387 3534 (International 44–1–387 3534)