



The logrank test

J Martin Bland and Douglas G Altman

BMJ 2004;328:1073-
doi:10.1136/bmj.328.7447.1073

Updated information and services can be found at:
<http://bmj.com/cgi/content/full/328/7447/1073>

These include:

References

This article cites 3 articles, 2 of which can be accessed free at:
<http://bmj.com/cgi/content/full/328/7447/1073#BIBL>

9 online articles that cite this article can be accessed at:
<http://bmj.com/cgi/content/full/328/7447/1073#otherarticles>

Rapid responses

5 rapid responses have been posted to this article, which you can access for free at:
<http://bmj.com/cgi/content/full/328/7447/1073#responses>

You can respond to this article at:
<http://bmj.com/cgi/eletter-submit/328/7447/1073>

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top left of the article

Topic collections

Articles on similar topics can be found in the following collections
[Other Statistics and Research Methods: descriptions](#) (604 articles)

Correction

A correction has been published for this article. The contents of the correction have been appended to the original article in this reprint. The correction is available online at:
<http://bmj.com/cgi/content/full/328/7453/1412>

Notes

To order reprints follow the "Request Permissions" link in the navigation box

To subscribe to *BMJ* go to:
<http://resources.bmj.com/bmj/subscribers>

Statistics Notes

The logrank test

J Martin Bland, Douglas G Altman

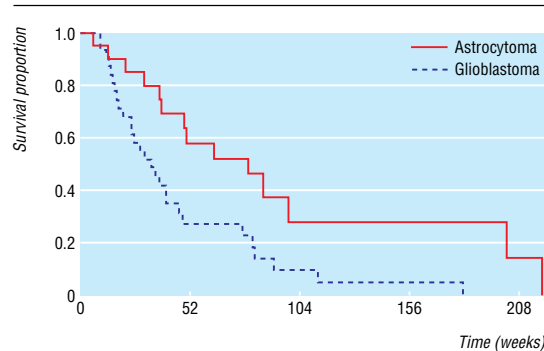
We often wish to compare the survival experience of two (or more) groups of individuals. For example, the table shows survival times of 51 adult patients with recurrent malignant gliomas¹ tabulated by type of tumour and indicating whether the patient had died or was still alive at analysis—that is, their survival time was censored.² As the figure shows, the survival curves differ, but is this sufficient to conclude that in the population patients with anaplastic astrocytoma have worse survival than patients with glioblastoma?

We could compute survival curves³ for each group and compare the proportions surviving at any specific time. The weakness of this approach is that it does not provide a comparison of the total survival experience of the two groups, but rather gives a comparison at some arbitrary time point(s). In the figure the difference in survival is greater at some times than others and eventually becomes zero. We describe here the logrank test, the most popular method of comparing the survival of groups, which takes the whole follow up period into account. It has the considerable advantage that it does not require us to know anything about the shape of the survival curve or the distribution of survival times.

The logrank test is used to test the null hypothesis that there is no difference between the populations in the probability of an event (here a death) at any time point. The analysis is based on the times of events (here deaths). For each such time we calculate the observed number of deaths in each group and the number expected if there were in reality no difference between the groups. The first death was in week 6, when one patient in group 1 died. At the start of this week, there were 51 subjects alive in total, so the risk of death in this week was $1/51$. There were 20 patients in group 1, so, if the null hypothesis were true, the expected number of deaths in group 1 is $20 \times 1/51 = 0.39$. Likewise, in group 2 the expected number of deaths is $31 \times 1/51 = 0.61$. The second event occurred in week 10, when there were two deaths. There were now 19 and 31 patients at risk (alive) in the two groups, one having died in week 6, so the probability of death in week 10 was $2/50$. The expected numbers of deaths were $19 \times 2/50 = 0.76$ and $31 \times 2/50 = 1.24$ respectively.

The same calculations are performed each time an event occurs. If a survival time is censored, that individual is considered to be at risk of dying in the week of the censoring but not in subsequent weeks. This way of handling censored observations is the same as for the Kaplan-Meier survival curve.³

From the calculations for each time of death, the total numbers of expected deaths were 22.48 in group 1 and 19.52 in group 2, and the observed numbers of deaths were 14 and 28. We can now use a χ^2 test of the null hypothesis. The test statistic is the sum of $(O - E)^2/E$ for each group, where O and E are the totals of the observed and expected events. Here $(14 - 22.48)^2 / 22.48 + (28 - 19.52)^2 / 19.52 = 6.88$. The degrees of freedom are the number of groups minus one, i.e.



Survival curves for women with glioma by diagnosis

$2 - 1 = 1$. From a table of the χ^2 distribution we get $P < 0.01$, so that the difference between the groups is statistically significant. There is a different method of calculating the test statistic,⁴ but we prefer this approach as it extends easily to several groups. It is also possible to test for a trend in survival across ordered groups.⁴ Although we have shown how the calculation is made, we strongly recommend the use of statistical software.

The logrank test is based on the same assumptions as the Kaplan Meier survival curve³—namely, that censoring is unrelated to prognosis, the survival probabilities are the same for subjects recruited early and late in the study, and the events happened at the times specified. Deviations from these assumptions matter most if they are satisfied differently in the groups being compared, for example if censoring is more likely in one group than another.

The logrank test is most likely to detect a difference between groups when the risk of an event is consistently greater for one group than another. It is unlikely to detect a difference when survival curves cross, as can happen when comparing a medical with a surgical intervention. When analysing survival data, the survival curves should always be plotted.

Because the logrank test is purely a test of significance it cannot provide an estimate of the size of the difference between the groups or a confidence interval. For these we must make some assumptions about the data. Common methods use the hazard ratio, including the Cox proportional hazards model, which we shall describe in a future *Statistics Note*.

Competing interests: None declared.

1 Rostomily RC, Spence AM, Duong D, McCormick K, Bland M, Berger MS. Multimodality management of recurrent adult malignant gliomas: results of a phase II multiagent chemotherapy study and analysis of cytoreductive surgery. *Neurosurgery* 1994;35:378-8.

2 Altman DG, Bland JM. Time to event (survival) data. *BMJ* 1998;317:468-9.

3 Bland JM, Altman DG. Survival probabilities. The Kaplan-Meier method. *BMJ* 1998;317:1572.

4 Altman DG. *Practical statistics for medical research*. London: Chapman & Hall, 1991: 371-5.

Department of Health Sciences, University of York, York YO10 5DD

J Martin Bland
professor of health statistics

Cancer Research UK/NHS Centre for Statistics in Medicine, Institute of Health Sciences, Oxford OX3 7LF

Douglas G Altman
professor of statistics in medicine

Correspondence to: Professor Bland

BMJ 2004;328:1073

Weeks to death or censoring in 51 adults with recurrent gliomas¹ (A=astrocytoma, G=glioblastoma)

| A | G |
|------|-----|
| 6 | 10 |
| 13 | 10 |
| 21 | 12 |
| 30 | 13 |
| 31* | 14 |
| 37 | 15 |
| 38 | 16 |
| 47* | 17 |
| 49 | 18 |
| 50 | 20 |
| 63 | 24 |
| 79 | 24 |
| 80* | 25 |
| 82* | 28 |
| 82* | 30 |
| 86 | 33 |
| 98 | 34* |
| 149* | 35 |
| 202 | 37 |
| 219 | 40 |
| | 40 |
| | 40* |
| | 46 |
| | 48 |
| | 70* |
| | 76 |
| | 81 |
| | 82 |
| | 91 |
| | 112 |
| | 181 |

*Censored survival time (still alive at follow up).

What is already known on this topic

Epidural analgesia during labour is effective but has been associated with increased rates of instrumental delivery

Studies have included women of mixed parity and high concentrations of epidural anaesthetic

What this study adds

Epidural infusions with low concentrations of local anaesthetic are unlikely to increase the risk of caesarean section in nulliparous women

Although epidural analgesia is associated with an increased risk of instrumental vaginal delivery, operator bias cannot be excluded

Epidural analgesia is associated with a longer second stage labour and increased oxytocin requirement, but the importance of these is unclear as maternal analgesia and neonatal outcome may be better with epidural analgesia

Differences in protocols for management of labour could have contributed to the differences in rates of instrumental vaginal delivery.

Another limitation was the large number of women who changed from parenteral opioids to epidural analgesia. Our intention to treat approach would likely render any estimation of the effects of epidural analgesia more conservative, but it was necessary to prevent selection bias. Comparing a policy of offering epidural analgesia with one of offering parenteral opioids reflects real life. As the definitions of stages of labour varied between trials, we were unable to determine if epidural analgesia prolonged the first stage, and the actual duration can only be estimated. Unlike previous reviews, we focused on nulliparous women because the indications for, and risks of, caesarean section differ with parity.

We limited our analysis to trials that used infusions of bupivacaine with concentrations of 0.125% or less, to reflect current practice. In a randomised controlled trial, low concentrations have been shown to reduce the rate of instrumental delivery.⁷

Epidural analgesia may increase the risk of instrumental delivery by several mechanisms. Reduction of serum oxytocin levels can result in a weakening of uterine activity.⁹⁻¹⁰ This may be due in part to intravenous fluid infusions being given before epidural analgesia, reducing oxytocin secretion.¹¹ The increased use of oxytocin after starting epidural analgesia may indicate attempts at speeding up labour. Maternal efforts at expulsion can also be impaired, causing fetal malposition during descent.¹² Previously, the association of neonatal morbidity and mortality with longer labour (second stage longer than two hours) had justified expediting delivery, leading to increased rates of instrumental delivery.¹³

Delaying pushing until the fetus's head is visible or until one hour after reaching full cervical dilation may reduce the incidence of instrumental delivery and its attendant morbidity.¹³ Although women receiving epidural analgesia had a longer second stage labour,

this was not associated with poorer neonatal outcome in our analysis.

We thank Christopher R Palmer (University of Cambridge) for guiding us and Shiv Sharma (University of Texas Southwestern Medical Centre) for helping with the data.

Contributors: See bmj.com

Funding: None.

Competing interests: None declared.

Ethical approval: Not required.

- Hawkins JL, Koonin LM, Palmer SK, Gibbs CP. Anesthesia-related deaths during obstetric delivery in the United States, 1979-1990. *Anesthesiology* 1997;86:277-84.
- Chestnut DH. Anesthesia and maternal mortality. *Anesthesiology* 1997;86:273-6.
- Roberts CL, Algert CS, Douglas I, Tracy SK, Peat B. Trends in labour and birth interventions among low-risk women in New South Wales. *Aust NZ J Obstet Gynaecol* 2002;42:176-81.
- Halpern SH, Leighton BL, Ohlsson A, Barrett JF, Rice A. Effect of epidural vs parenteral opioid analgesia on the progress of labor: a meta-analysis. *JAMA* 1998;280:2105-10.
- Zhang J, Klebanoff MA, DerSimonian R. Epidural analgesia in association with duration of labor and mode of delivery: a quantitative review. *Am J Obstet Gynecol* 1999;180:970-7.
- Howell CJ. Epidural versus non-epidural analgesia for pain relief in labour. *Cochrane Database Syst Rev* 2000;CD000331.
- Comparative Obstetric Mobile Epidural Trial (COMET) Study Group UK. Effect of low-dose mobile versus traditional epidural techniques on mode of delivery: a randomised controlled trial. *Lancet* 2001;358:19-23.
- Scottish Intercollegiate Guidelines Network. Methodology checklist 2: randomised controlled trials. www.sign.ac.uk/guidelines/fulltext/50/checklist2.html (accessed 7 Mar 2004).
- Goodfellow CF, Hull MG, Swaab DF, Dogterom J, Buijs RM. Oxytocin deficiency at delivery with epidural analgesia. *Br J Obstet Gynaecol* 1983;90:214-9.
- Bates RG, Helm CW, Duncan A, Edmonds DK. Uterine activity in the second stage of labour and the effect of epidural analgesia. *Br J Obstet Gynaecol* 1985;92:1246-50.
- Cheek TG, Samuels P, Miller F, Tobin M, Gutsche BB. Normal saline i.v. fluid load decreases uterine activity in active labour. *Br J Anaesth* 1996;77:632-5.
- Newton ER, Schroeder BC, Knappe KG, Bennett BL. Epidural analgesia and uterine function. *Obstet Gynecol* 1995;85:749-55.
- Miller AC. The effects of epidural analgesia on uterine activity and labor. *Int J Obstet Anesth* 1997;6:2-18.

(Accepted 18 March 2004)

doi 10.1136/bmj.38097.590810.7C

Corrections and clarifications*The logrank test*

An error in this Statistics Notes article by J Martin Bland and Douglas G Altman persisted to publication (1 May, p 1073). The third sentence—relating to the values in a table of survival times—should have ended: “but is this sufficient to conclude that in the population, patients with anaplastic astrocytoma have better [not ‘worse’] survival than patients with glioblastoma?”

Minerva

In the seventh item in the 17 April issue (p 964), Minerva admits to not being religious. Maybe this is why she reported that in the Bible Daniel has 21 chapters. It doesn't; it has only 12.

Exploring perspectives on death

In this News article by Caroline White (17 April, p 911), we referred to average life expectancy as running to 394 000 minutes—a rather brief life span, some might think, given that there are only 525 600 minutes in one non-leap year. Assuming that the average life expectancy referred to in the article was intended to relate to UK males (that is, 75 years), some straightforward maths reveals that 75 years (even discounting the extra length of leap years) in fact contain 39 420 000 minutes. Quite a lot of minutes were lost somewhere in our publication process.