Statistically speaking: methodological madness

Comparing times in clinical studies with a finite ending

Dear Ms Method Matters,

I am planning a study to investigate the effect of perioperative intravenous lidocaine versus saline, the primary outcome being time to first request of analgesics. I thought that the best way to compare the lidocaine and saline group would be to investigate the mean time to request of analgesics and simply perform a t-test or a non-parametric equivalent, but I have seen a similar study on different doses of intravenous dexamethasone in Anaesthesia [1], which displayed results using the Kaplan-Meier time-to-event curve and a 'time-dependent interaction term' to assess response to the drug, because the 'proportional hazards assumption was invalid'. Why can I not simply use the t-test to compare the times between the two groups, or logistic regression to investigate the relationship between the treatment and my primary outcome, and what are interaction terms?

Crossed-lines (Gravely Hill, Birmingham, UK)

Dear Crossed-lines,

At first glance at your experimental design, it would appear that you could possibly compare the mean times to first analgesic request between the two groups. We wouldn't really consider this method because there might be patients in either group, known as 'censored cases' who never request analgesia (at least, not request analgesia right up to the end of your observation time) and so it would be difficult to ascertain what your denominator would be, if you were to calculate the mean time. Therefore, it is common in analysis of 'time-to-event' to employ the Kaplan-Meier distribution and Cox proportional hazards regression. The Kaplan-Meier method will estimate the probability of not experiencing the event (in this case, not requesting analgesia) past that particular time point taking into

consideration the presence of censored cases. In other words, the Kaplan-Meier method estimates the "experience" of not requesting analgesia over a period of time.

The Kaplan-Meier with log-rank test is useful to graph and compared the time-to-event distributions in two or more groups [2]. The log-rank test is used to test the null hypothesis that there is no difference between the saline and lidocaine groups in the probability of an event (request for analgesics) at any time point. The analysis is based on the times of events, and for each time-point, the number of people who have asked for analgesics in each group and the expected number of people who would have asked for analgesics in each group is calculated if there were really no difference between the groups. The log-rank test is a test of significance and it cannot provide an estimate of the size of the difference between the groups or a confidence interval, in addition you may be interested in analysing the effect of several risk factors on your primary outcome. For this, other common methods of regression must be employed.

In the study you cited and several recent studies in Anaesthesia, [3-5] the Cox proportional hazards regression method was used to investigate variables such as age, sex, type of surgery and body mass index on a time-to-event primary outcome. The Cox proportional regression model assumes that the effects of the predictor variables are constant over time (proportional hazards assumption), in the example you have given, it means that lidocaine will change the overall time to first analgesic request, rather than change the pattern of analgesic request over time. A quick method to assess whether the effect of the predictor variables are constant over time is to look at your time-to-event distribution. If at any point the two lines cross (figure 1), then the proportional hazards assumption has been violated, and that there is an ‘interaction effect’. An interaction effect exists when the impact of one variable (lidocaine or saline) on time to first analgesic request is not the same at all levels of a second explanatory variable (let’s say, for example, age).

Desmet et al [1] used age, sex, type of operation and body mass index as other explanatory variables, and so the investigators therefore had to assess the effect of all these on time to first analgesic request (detailed in Box 1). They found that post-operative analgesia was given sooner after rotator cuff repairs compared to subacromial decompression, but post-operative analgesia was given later in older participants, and that there was no interaction of analgesia duration with body mass index and sex.

The output you would expect for a Cox regression analysis is a coefficient, (shown as Exp(b) in most statistical software). For example, if the Exp(b) for the covariate age is 2.14, this is interpreted as for an increase in one unit of time, the hazard ratio for 'request to analgesic' increases by a factor of 2.14, for two units of time, the hazard ratio for request to analgesic increases by a factor of 2.14^2.

Since the request for analgesic is a binary response, you may consider analysing your data using logistic regression, as long as there are no 'censored' cases. But survival analysis, such as Cox regression differ from logistic regression in that Cox regression models the incidence (otherwise known as hazard) rate, that is the number of new cases of event (ie, request for analgesia) per population per unit time. The hazard function is the probability that if a person is pain-free to time ***t***, they will ask for analgesia in the next instant. Logistic regression considers the proportion of new cases that develop in a given time period, in other words, logistic regression estimates the cumulative incidence given as the odds ratio, whilst Cox regression estimates the hazard ratio.

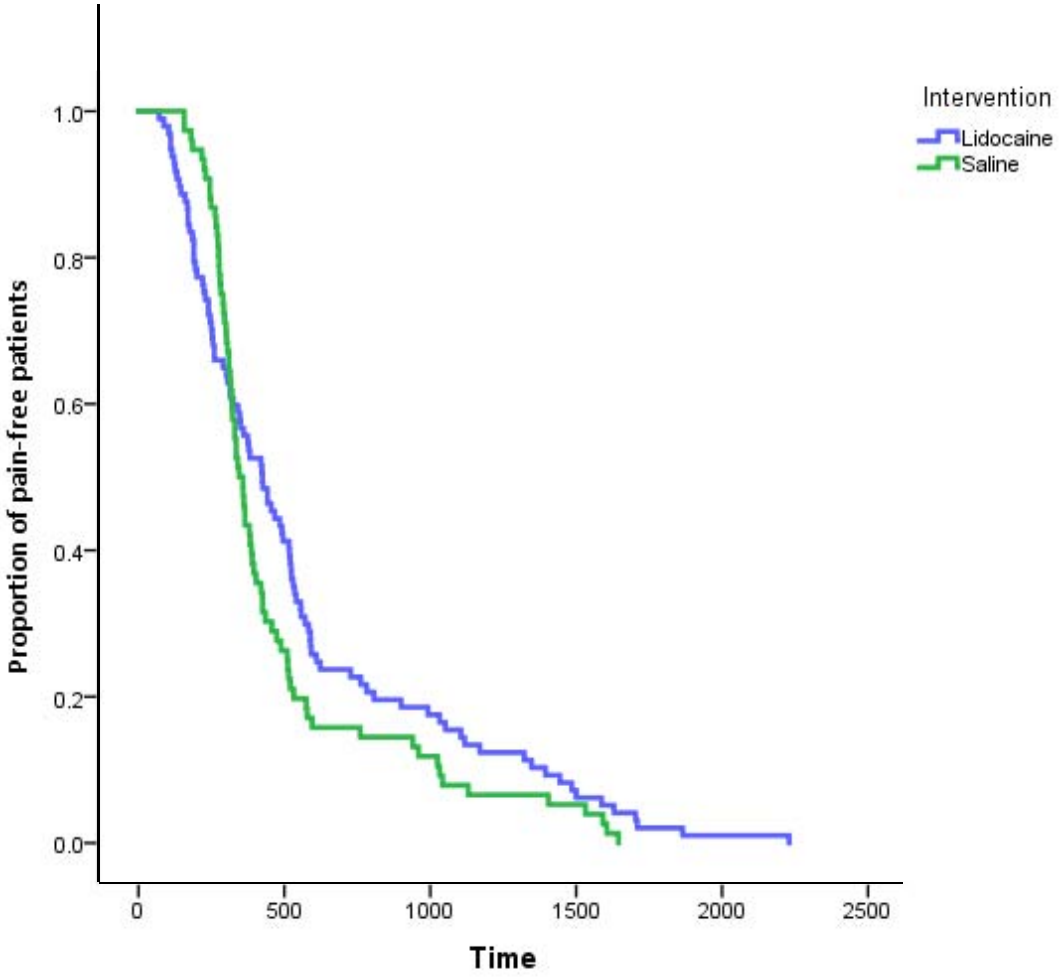
If you have any questions on methodology, please direct them to Ms Method Matters at

msmethodmatters@gmail.com

Box 1. Possible interaction terms

Age * sex
Age * type of surgery
Age * body mass index
Sex * type of surgery
Sex * body mass index
Age * sex * type of surgery
Age * sex * body mass index
Sex * type of surgery * body mass index

Fig 1


Intervention
Lidocaine
Saline
Proportion of pain-free patients
1.0
0.8
0.6
0.4
0.2
0.0
0
500
1000
1500
2000
2500
Time

References

1. Desmet M, Vanneste B, Reynvoet M, et al. A randomised controlled trial of intravenous dexamethasone combined with interscalene brachial plexus blockade for shoulder surgery. *Anaesthesia* 2015; **70:** 1180-5.
2. Bland JM, Altman DG The logrank test. *BMJ* 2004; **328:** 1073.
3. Wong DJN Cox regression models for frailty in vascular surgery patients. *Anaesthesia* 2016; **71:** 1117-8.
4. Coulson TG, Bailey M, Reid CM, et al. The association between peri-operative acute risk change (ARC) and long-term survival after cardiac surgery. *Anaesthesia***:** n/a-n/a.
5. O'Neill BR, Batterham AM, Hollingsworth AC, Durrand JW, Danjoux GR Do first impressions count? Frailty judged by initial clinical impression predicts medium-term mortality in vascular surgical patients. *Anaesthesia* 2016; **71:** 684-91.

Acknowledgements

SC is statistical advisor to Anaesthesia. No other competing interests declared.

S. W. Choi

Assistant Research Officer

Department of Anaesthesiology,

The University of Hong Kong,

Hong Kong, HKSAR

Email: htswchoi@hku.hk

D. M. H. Lam

Resident

Department of Anaesthesiology,

Queen Mary Hospital,

Hong Kong, HKSAR