JAMA Guide to Statistics and Methods

Time-to-Event Analysis

Juliana Tolles, MD, MHS; Roger J. Lewis, MD, PhD

Time-to-event analysis, also called survival analysis, was used in the study by Nissen et al[1] published in this issue of *JAMA* to compare the risk of major adverse cardiovascular events (MACE) in a noninferiority trial of a combination of naltrexone and bupropion vs placebo for overweight or obese patients with cardiovascular risk factors. The authors used a type of time-to-event analysis called Cox proportional hazards modeling to compare the risk of MACE in the 2 groups, concluding that the use of naltrexone-bupropion increased the risk of MACE per unit time by no more than a factor of 2.

Related article page 990

Use of the Method

Why Is Time-to-Event Analysis Used?

One way to evaluate how a medical treatment affects patients' risk of an adverse outcome is to analyze the time intervals between the initiation of treatment and the occurrence of such events. That information can be used to calculate the hazard for each treatment group in a clinical trial. The hazard is the probability that the adverse event will occur in a defined time interval. For example, Nissen et al[1] could measure the number of patients who experience MACE while taking naltrexone-bupropion during week 8 of the study and calculate the risk that an individual patient will experience MACE during week 8, assuming that the patient has not had MACE before week 8. This concept of a discrete hazard rate can be extended to a hazard function, which is generally a continuous curve that describes how the hazard changes over time. The hazard function shows the risk at each point in time and is expressed as a rate or number of events per unit of time.[2]

Calculating the hazard function using time-to-event observations is challenging because the event of interest is usually not observed in all patients. Thus, the time to the event occurrence for some patients is invisible—or censored—and there is no way to know if the event will occur in the near future, the distant future, or never. Censoring may occur because the patient is lost to follow-up or did not experience the event of interest before the end of the study period. In Nissen et al,[1] only 243 patients experienced MACE before the termination of the study, resulting in 8662 censored observations, meaning there were 8662 patients for whom it is not known when they experienced MACE, if ever. Common nonparametric statistical tests, such as the Wilcoxon rank sum test, could be used to compare the time intervals seen in the 2 groups if the analysis was limited to only the 243 patients who had observed events; however, when censored data are excluded from analysis, the information contained in the experience of the other 8662 patients is lost. While it is unknown when in the future, if ever, these patients will experience an event, the knowledge that these patients did not experience MACE during their participation in the trial is informative. The information contained in censored observations varies: patients whose data are censored early, such as a patient who is lost to follow-up in the first weeks of a study, contribute less information than those who are observed for a long time before censoring. However, all observations provide some information, and to avoid bias, methods of analysis that can accommodate censoring are used for time-to-event studies.

Kaplan-Meier plots and the Cox proportional hazards model are examples of methods for analyzing time-to-event data that account for censored observations. A Kaplan-Meier curve plots the fraction of "surviving" patients (those who have not experienced an event) against time for each treatment group. The height of the Kaplan-Meier curve at the end of each time interval is determined by taking the fraction or proportion of patients who remained event-free at the end of the prior time interval and multiplying that proportion by the fraction of patients who survive the current time interval without experiencing an event. The value of the Kaplan-Meier curve at the end of the current time interval then becomes the starting value for the next time interval. This iterative and cumulative multiplication process begins with the first time interval and continues in a stepwise manner along the Kaplan-Meier curve; the Kaplan-Meier curve is thus sometimes called the "product limit estimate" of the survival curve. Censoring is properly taken into account because only patients still being followed up at the beginning of each time interval are considered in determining the fraction "surviving" at the end of that time interval.[3] Figures 2A and 2B in Nissen et al[1] plot the cumulative incidence of MACE in each group vs time, an "upside-down" version of Kaplan-Meier, which provides similar information.

While a Kaplan-Meier plot elegantly represents differences between different groups' survival curves over time, it gives little indication of their statistical significance. The statistical significance of observed differences can be tested with a log-rank test.[3] This test, however, does not account for confounding variables, such as differences in patient demographics between groups.

The Cox proportional hazards model both addresses the problem of censoring and allows adjustment for multiple prognostic independent variables, or confounders such as age and sex. The model assumes a "baseline" hazard function exists for individuals whose independent predictor variables are all equal to their reference value. The baseline hazard function is not explicitly defined but is allowed to take any shape. The output of a Cox proportional hazards model is a hazard ratio for each independent predictor variable, which defines how much the hazard is multiplied for each unit change in the variable of interest as compared with the baseline hazard function. Hazard ratios can be calculated for all independent variables, both confounders and intervention variables.

What Are the Limitations of the Proportional Hazards Model?

The Cox proportional hazards model relies on 2 important assumptions. The first is that data censoring is independent of outcome of interest. If the placebo patients in the trial by Nissen et al[1] were both less likely to experience MACE and less likely to follow up with trial investigators because they did not experience weight loss, the probability of censoring and the risk of MACE would be correlated,

threatening the validity of the analysis. The second assumption is that the hazard functions, representing the risk of an event over time, are proportional to each other for all patient groups. In other words, the hazard functions all have the same shape and differ only in overall magnitude; the effect of each independent predictor or confounder is on the overall magnitude of the hazard function. In this trial, it is reasonable to assume that the baseline hazard function for MACE in patients taking placebo looks like a line with a positive slope: age likely increases the hazard of a cardiovascular event. The assumption of proportional hazards means that the hazard function of MACE in patients taking naltrexone-bupropion is assumed to be the baseline hazard multiplied by an unknown, constant value. This assumption would be violated if, for example, patients taking the drug experience an early increase in risk of MACE after initiating treatment as a result of adverse effects but then experience decreased risk over the long-term as they lose weight. In that scenario, the treatment group hazard function would be shaped like a peak with a long tail and would not be proportional to the baseline hazard function.

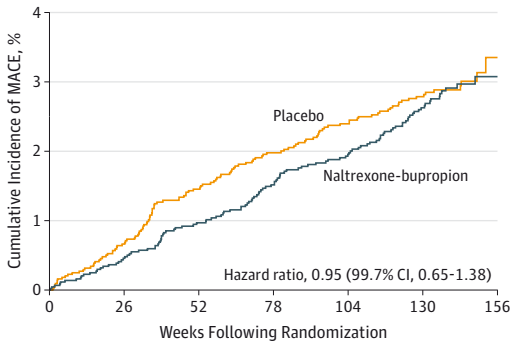
How Should Time-to-Event Findings Be Interpreted in This Particular Study?

The trial was designed as a noninferiority study and statistically powered to assess the null hypothesis that the hazard ratio of naltrexone-bupropion to placebo for MACE is greater than 2.0 at the 25% interim analysis point. Using a Cox proportional hazard model with randomized treatment as a predictor, the estimated hazard ratio was 0.59 (95% CI, 0.39-0.90). It can therefore be concluded that the hazard ratio of MACE associated with the active treatment was less than 2.0. Although it might be tempting to conclude that the hazard ratio is smaller (eg, less than 1.0), the hypothesis testing structure of the noninferiority trial only allowed a rigorous conclusion to be drawn about the hypothesis that the hazard ratio was less than 2.0.

Caveats to Consider When Looking at Results From a Time-to-Event Analysis

Nissen et al[1] used a Cox proportional hazards model to estimate the hazard ratio associated with naltrexone-bupropion compared with placebo for MACE in overweight or obese patients with cardiovascular risk factors. This trial likely meets the assumptions of the Cox proportional hazards model: the censoring is likely to be independent of hazard, and the hazard functions for all groups are likely to be roughly proportional. Readers should interpret with caution any time-to-event analysis in which the probability of being lost to follow-up or the duration of observation is likely to be correlated with the risk of experiencing an event. Readers should also be cautious in accepting Cox proportional hazards models in which the hazard function of a treatment group is unlikely to be proportional to the baseline hazard. If 2 survival curves cross at any point, such as seen in the far right of Figure 2B in the article by Nissen et al,[1] this might suggest that the hazard ratio between the 2 groups has reversed and the proportionality assumption has been violated (**Figure**). There are also several diagnostic tests that researchers can use to verify the proportionality assumption, including using Kaplan-Meier curves, testing the significance of time-dependent covariates, and plotting Schoenfeld residuals.[4] Selection of an appropriate verification method depends on the types of covariates used in the Cox proportional hazards model.

Figure. Time to MACE in the Final End-of-Study Analysis



No. at risk							
Placebo	4450	4289	4183	4053	3886	3333	90
Naltrexone-bupropion	4455	4317	4228	4092	3951	3403	102

The survival curves cross in this figure from Nissen et al,[1] suggesting that the proportionality assumption may have been violated. MACE indicates major adverse cardiovascular events.

ARTICLE INFORMATION

Author Affiliations: Department of Emergency Medicine, Harbor-UCLA Medical Center, Torrance, California (Tolles, Lewis); Los Angeles Biomedical Research Institute, Torrance, California (Tolles, Lewis); David Geffen School of Medicine at UCLA, Los Angeles, California (Tolles, Lewis); Berry Consultants LLC, Austin, Texas (Lewis).

Corresponding Author: Roger J. Lewis, MD, PhD, Department of Emergency Medicine, Harbor-UCLA Medical Center, Bldg D9, 1000 W Carson St, Torrance, CA 90509 (roger@emedharbor.edu).

Section Editors: Roger J. Lewis, MD, PhD, Department of Emergency Medicine, Harbor-UCLA Medical Center and David Geffen School of Medicine at UCLA; and Edward H. Livingston, MD, Deputy Editor, *JAMA*.

Conflict of Interest Disclosures: Both authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

REFERENCES

1. Nissen SE, Wolski KE, Prcela L, et al. Effect of naltrexone-bupropion on major adverse cardiovascular events in overweight and obese patients with cardiovascular risk factors: a randomized clinical trial. *JAMA*. doi:10.1001/jama.2016.1558.

2. Lee ET. *Statistical Methods for Survival Analysis*. 2nd ed. New York, NY: John Wiley & Sons; 1992.

3. Young KD, Menegazzi JJ, Lewis RJ. Statistical methodology: IX, Survival analysis. *Acad Emerg Med*. 1999;6(3):244-249.

4. Hess KR. Graphical methods for assessing violations of the proportional hazards assumption in Cox regression. *Stat Med*. 1995;14(15):1707-1723.