PERSPECTIVES

Statistical Analysis of Clinical COVID-19 Data: A Concise Overview of Lessons Learned, Common Errors and How to Avoid Them

This article was published in the following Dove Press journal:
Clinical Epidemiology

Martin Wolkewitz[1]
Jerome Lambert[1]
Maja von Cube [ID][1]
Lars Bugiera[1]
Marlon Grodd[1]
Derek Hazard[1]
Nicole White [ID][2]
Adrian Barnett [ID][2]
Klaus Kaier [ID][1]

[1]Institute of Medical Biometry and Statistics, Medical Center - University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany; [2]School of Public Health and Social Work, Queensland University of Technology, Brisbane, QLD, Australia

Abstract: By definition, in-hospital patient data are restricted to the time between hospital admission and discharge (alive or dead). For hospitalised cases of COVID-19, a number of events during hospitalization are of interest regarding the influence of risk factors on the likelihood of experiencing these events. The same is true for predicting times from hospital admission of COVID-19 patients to intensive care or from start of ventilation (invasive or non-invasive) to extubation. This logical restriction of the data to the period of hospitalisation is associated with a substantial risk that inappropriate methods are used for analysis. Here, we briefly discuss the most common types of bias which can occur when analysing in-hospital COVID-19 data.

Keywords: competing risk bias, immortal-time bias, competing events, time-dependent bias, time-varying exposure, time-to-event analysis

Introduction

During the COVID-19 pandemic, the analysis of data from COVID-19 patients is essential to understand the clinical progress, potential treatment effects and expected outcomes. Numerous methodological articles have shown that adequate statistical models are necessary to acknowledge the time-dependent complexity in particular of hospital data. It has also been shown that in several applications, standard models such as Kaplan–Meier survival plots, Cox or logistic regression models can easily lead to biased estimates and wrong conclusions. In this article, we summarize some of the most important pitfalls, provide intuitive clinical examples (mostly related to other types of infections) and refer to methodological articles which explain the mathematical framework. We also describe common errors which can occur in the presence of competing events and time-dependent covariates, or when conditioning on future events.

Methods and Results
Inappropriate Use of Kaplan–Meier Survival Plots

One important outcome of admitted COVID-19 patients are the combined endpoint admission to intensive care, invasive ventilation or hospital mortality. In this case, standard Kaplan–Meier plots are often used for time-dependent events such as admission to intensive care unit, invasive ventilation or hospital mortality, if only hospital data is available.[1–6] However, this implies that patients discharged without

Correspondence: Klaus Kaier
Institute of Medical Biometry and Statistics, Medical Center - University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany
Email kaier@imbi.uni-freiburg.de

the event are censored. This is not a meaningful model assumption, since discharged patients are usually in a better health condition than hospitalized patients. Cumulative incidence functions should be used instead, accounting for discharge as a competing event.

It is also common to use standard Kaplan–Meier plots to display effects of time-dependent covariates such as infectious disease consultation or medical treatment.[7,8] This is misleading and does more harm than good. One way to address this issue is to use landmark models where covariate values are updated at each landmark.[9]

Inappropriate Use of Standard Cox Regression Models

Competing events are often ignored in time-to-event analysis.[8,10-12] In case of events other than death (eg, time-to-extubation) being analysed, death should be considered a competing event. The same is true when death is the event of interest in a time-to-event analysis of in-hospital data: here, discharge alive should be considered a competing event and patients should not be censored at the time of discharge. There are two reasons why ignoring competing events should be avoided. First, it is incomplete[13] and second, these events are often very informative and the key to understanding your research question.

For example, Wolkewitz et al[2] do not detect an effect of Oseltamivir on the death hazard of patients hospitalized with confirmed pandemic influenza A/H1N1 2009 infection. However, the discharge hazard is significantly decreased, resulting in a prolonged length of stay for the treated patients. The prolonged length of stay results in an increased risk of death in the hospital. Cause-specific Cox regression for both in-hospital mortality and discharge alive avoids competing risk bias and provides a detailed understanding of potential treatment effects.

Handling time-dependent predictors (eg, time-dependent anti-viral treatment, time to ventilation or time to oxygen saturation below 90%) as time-fixed leads to immortal-time bias.[2,4,14-16] A mathematical formula shows that the hazard ratio for the predictor's effect on death is always underestimated, meaning that predictors that increase the risk of death may be missed or even appear to be beneficial. In contrast, this bias always overestimates the effect for additional length of hospital stay.

For example, in the primary analysis of Rosenberg et al[8] on patients hospitalized with COVID-19, the hazard ratio of Hydroxychloroquine + azithromycin is 1.35 (0.67–2.40). The hazard ratio corrected for the immortal time bias is presented in the appendix of the paper by Rosenberg et al (eTable 6).[8] This sensitivity analysis that properly accounts for the time-dynamics results in a clearly higher hazard ratio: 1.83 (1.02–3.28). The same is applied for Hydroxychloroquine and Azithromycin alone.

This is one of the severe types of bias in epidemiology and very common in the hospital setting. More advanced Cox models can account for time-dependent covariates.[17]

The increased mortality in the treatment groups of Rosenberg et al may be explained by time-dependent confounding, which is generally present if treatment is administered over time. Time-dependent propensity scores in the Cox model (known as marginal structural model) adequately take the time-dynamic confounder-treatment feedback into account. For example, Arabi et al discussed the influence of immortal time bias and time-dependent confounding in their excellent retrospective cohort study of patients infected with Middle East Respiratory Syndrome Coronavirus (MERS-CoV).[18] In their study, they find a differential effect of immortal-time bias and time-dependent confounding on the magnitude of the hazard ratio: the hazard ratio of Ribavirin/Recombinant Interferon for 90-day mortality from the marginal structural Cox model was 67% lower than from a time-dependent Cox model adjusted only for baseline confounding (1.52 (1.13–1.06) versus 1.03 (0.73–1.44)).

Inappropriate Use of Standard Logistic Regression

Standard logistic regression is a popular tool for examining associations between risk factors and the event of interest if patient data are available until discharge. However, during a pandemic situation data contain mainly active COVID-19 cases (ie, patients are still in hospital). For instance, Zhou et al[19] discarded infected patients who were still alive on 31 January 2020. The reported numbers are 813 hospitalized and 613 excluded because they were still hospitalized or not confirmed as infected; the numbers for each of the two exclusion reasons are not given. Hence a large proportion of the cohort of interest was excluded, which has the potential to create selection bias which could impact the key risk estimates. In this case, time-to-event models should be used over logistic regression for examining associations between patient-level predictors and the event of interest (eg, death). Time-to-event models allow censoring of patients who are still hospitalised at the last follow-up date. Excluding these patients when using

logistic regression introduces selection bias, which can be strong enough to nullify the effect of important predictors. The bias depends on the difference between patients who have long and short lengths of stay. For example, if longer-staying patients are generally older, then the effect of age may be underestimated and the generalisability of the results questionable.

Inappropriate Use of Advanced Regression Models

Sub-distribution hazard models, such as the Fine & Gray model,[20] have been developed to evaluate effects of time-fixed covariates on competing events. However, be mindful of time-dependent covariates and competing endpoints.[21] The risk of finding artefacts is high and results need to be carefully considered before reaching any conclusions about whether covariates truly influence the risk of events.

Inappropriate Conditioning on Future Events

Conditioning on future events, for example, excluding patients who experience events that are not of primary interest,[22–24] violates one of the main principles in time-to-event analysis[25] by introducing selection bias. A recent COVID-19 example is the retrospective study by Mehra et al,[26] where

> patients for whom one of the treatments of interest was initiated more than 48 h after diagnosis or while they were on mechanical ventilation, as well as patients who received remdesivir, were excluded.

In general, it is best to include the whole cohort over time, as excluding patients always has the potential to introduce serious bias. Picture being in a hospital and collecting patient data prospectively. All the patients in your cohort are at risk of the key events, and you have no way of knowing what patients to exclude based on future events.

Discussion

These issues have been known for many years from studies of which factors influence hospital morbidity and mortality. The potential biases in these studies have been explained using mathematical formulae, simulation studies and real-world analyses. Furthermore, most methods to reduce bias are easy to apply and available in standard statistical software. Unfortunately, statistical methods that are better suited to these data have yet to be widely adopted, and too many researchers continue to use simple survival analysis

techniques that are not suited to complex hospital data with variables that change over time and multiple outcomes.[27] The rush to analyse in-hospital COVID-19 data to inform clinical practice has reduced the time for detailed peer review.[28] We are concerned that many current results that examine the in-hospital treatment of COVID-19 patients are untrustworthy. Greater involvement of statisticians with experience in appropriate methods will improve the quality of analysis undertaken and their interpretation before being shared with the wider research community.

Disclosure

The authors report no conflicts of interest for this work.

References

1. Wolkewitz M, Sommer H. Evaluating mortality in an intensive care unit requires extended survival models. *Crit Care*. 2014;18(2):415. doi:10.1186/cc13757
2. Wolkewitz M, Schumacher M. Survival biases lead to flawed conclusions in observational treatment studies of influenza patients. *J Clin Epidemiol*. 2017;84:121–129. doi:10.1016/j.jclinepi.2017.01.008
3. Wolkewitz M, Cooper BS, Bonten MJ, Barnett AG, Schumacher M. Interpreting and comparing risks in the presence of competing events. *BMJ*. 2014;349(aug21 5):g5060. doi:10.1136/bmj.g5060
4. Schumacher M, Allignol A, Beyersmann J, Binder N, Wolkewitz M. Hospital-acquired infections: appropriate statistical treatment is urgently needed. *Int J Epidemiol*. 2013;42:1502–1508.
5. Wolkewitz M, Harbarth S, Beyersmann J. Daily Chlorhexidine Bathing and Hospital-Acquired Infection. *N Engl J Med*. 2013;368:2330–2332.
6. Wu C, Chen X, Cai Y, et al. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA Intern Med*. 2020;180(7):934. doi:10.1001/jamainternmed.2020.0994
7. Lambert J, von Cube M, Biard L, Rieg S, Wolkewitz M. Infectious disease consultation for candidaemia. *Lancet Infect Dis*. 2020;20 (2):164–165. doi:10.1016/S1473-3099(19)30745-5
8. Rosenberg ES, Dufort EM, Udo T, et al. Association of treatment with hydroxychloroquine or azithromycin with in-hospital mortality in patients with COVID-19 in New York State. *JAMA*. 2020;323 (24):2493. doi:10.1001/jama.2020.8630
9. von Cube M, Schumacher M, Putter H, Timsit J-F, van de Velde C, Wolkewitz M. The population-attributable fraction for time-dependent exposures using dynamic prediction and landmarking. *Biometrical J*. 2020;62(3):583–597. doi:10.1002/bimj.201800252
10. Weber S, von Cube M, Sommer H, Wolkewitz M. Necessity of a competing risk approach in risk factor analysis of central line–associated bloodstream infection. *Infect Control Hosp Epidemiol*. 2016;37(10):1255–1257. doi:10.1017/ice.2016.166

11. Wolkewitz M. Estimating the risk of ventilator-associated pneumonia as a function of time. *Am J Respir Crit Care Med*. 2015;192 (5):640–641. doi:10.1164/rccm.201502-0282LE

12. Wolkewitz M, Cooper B, Barnett AG, Binder N, Schumacher M. Treatment of hospital-acquired pneumonia. *Lancet Infect Dis*. 2011;11(10):731–732. doi:10.1016/S1473-3099(11)70262-6

13. Latouche A, Allignol A, Beyersmann J, Labopin M, Fine JP. A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions. *J Clin Epidemiol*. 2013;66(6):648–653. doi:10.1016/j.jclinepi.2012.09.017

14. Suissa S. Immeasurable time bias in observational studies of drug effects on mortality. *Am J Epidemiol*. 2008;168:329–335. doi:10.1093/aje/kwn135

15. Wolkewitz M. Avoidable statistical pitfalls in analyzing length of stay in intensive care units or hospitals. *Crit Care*. 2014;18(1):408. doi:10.1186/cc13735

16. Wolkewitz M, Allignol A, Harbarth S, de Angelis G, Schumacher M, Beyersmann J. Time-dependent study entries and exposures in cohort studies can easily be sources of different and avoidable types of bias. *J Clin Epidemiol*. 2012;65(11):1171–1180. doi:10.1016/j.jclinepi.2012.04.008

17. Beyersmann J, Allignol A, Schumacher M. *Competing Risks and Multistate Models with R*. Springer Science & Business Media; 2011.

18. Arabi YM, Shalhoub S, Mandourah Y, et al. Ribavirin and interferon therapy for critically Ill patients with middle east respiratory syndrome: a multicenter observational study. *Clin Infect Dis*. 2020;70 (9):1837–1844. doi:10.1093/cid/ciz544

19. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*. 2020;395(10229):1054–1062. doi:10.1016/S0140-6736(20)30566-3

20. Fine J, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc*. 1999;94(446):496–509. doi:10.1080/01621459.1999.10474144

21. Poguntke I, Schumacher M, Beyersmann J, Wolkewitz M. Simulation shows undesirable results for competing risks analysis with time-dependent covariates for clinical outcomes. *BMC Med Res Methodol*. 2018;18(1). doi:10.1186/s12874-018-0535-5

22. Sommer H, Timsit J-F, Wolkewitz M; COMBACTE-NET Consortium. Bezlotoxumab and Recurrent Clostridium difficile Infection. *N Engl J Med*. 2017;376(16):1594. doi:10.1056/NEJMc1702531

23. Heister T, Wolkewitz M, Kaier K. Determining the attributable costs of clostridium difficile infections when exposure time is lacking: be Wary of "Conditioning on the future". *Infect Control Hosp Epidemiol*. 2018;39(6):759–760. doi:10.1017/ice.2018.42

24. Heister T, Wolkewitz M, Kaier K. Estimating the additional costs of surgical site infections: length bias, time-dependent bias, and conditioning on the future. *J Hosp Infect*. 2018;99(1):103–104. doi:10.1016/j.jhin.2018.02.012

25. Andersen PK, Keiding N. Interpretability and importance of functionals in competing risks and multistate models: interpretability and importance of functionals in competing risks and multistate models. *Stat Med*. 2012;31(11–12):1074–1088. doi:10.1002/sim.4385

26. Mehra MR, Desai SS, Ruschitzka F, Patel AN. Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. *Lancet*. 2020. doi:10.1016/S0140-6736(20)31180-6

27. Wynants L, Van Calster B, Bonten MM, et al. Systematic review and critical appraisal of prediction models for diagnosis and prognosis of COVID-19 infection. *medRxiv*. 2020.

28. Glasziou PP, Sanders S, Hoffmann T. *Waste in Covid-19 Research*. British Medical Journal Publishing Group; 2020.