Discussion of 'Why test for proportional hazards?' by Stensrud & Hernán

Paul W Dickman

12 January 2023

http://pauldickman.com/talk/

JAMA Guide to Statistics and Methods

March 13, 2020

Why Test for Proportional Hazards?

Mats J. Stensrud, MD, DrPhilos[1,2]; Miguel A. Hernán, MD, DrPH[1,3,4]

» Author Affiliations | Article Information

JAMA. 2020;323(14):1401-1402. doi:10.1001/jama.2020.1267

Overview of my thoughts on the paper [1]

- Nice paper; I agree with essentially everything.
- The statement that 'statistical tests for proportional hazards are unnecessary' is potentially controversial, but I agree.
- I am concerned that the statement may be (mis)interpreted by some as 'assessing proportional hazards is unnecessary'.
- Researchers should understand the concept of proportional hazards, to which this paper makes a valuable contribution.
- Researchers should consider the time-varying nature of hazard ratios in the design and reporting of their studies and should assess the proportional hazards assumption in the analysis.
- Do formal tests have any value in assessing PH?
- Does the 'tests are unnecessary' claim apply to all effect modifiers and to other models?

Why Are Hazards Usually Not Proportional?

Quotes from Stensrud & Hernán [1]

1. Hazards are not proportional when the treatment effect changes over time.
2. Hazards may also not be proportional because disease susceptibility varies between individuals [2].

- (1) is just the familiar assumption of constancy of effect, often called no interaction or no effect modification, where the potential effect modifier in this case is time.
- (1) applies to other covariates in the Cox model and to other regression models whereas (2) is specific to time.
- Does this mean we should never perform statistical tests for effect modification?

How Should Hazard Ratios Be Interpreted?

Quote from Stensrud & Hernán [1]

As a weighted average of the time-varying hazard ratios, the hazard ratio estimate from a Cox proportional hazards model is often used as a convenient summary of the treatment effect during the follow-up. However, a hazard ratio from a Cox model needs to be interpreted as a weighted average of the true hazard ratios over the entire follow-up period.

- I agree with the intepretation (second sentence) but I'm not sure I understand the distinction between what they claim is often done (first sentence) and what should be done.

'Statistical tests for PH are unnecessary'

Because it is expected that the hazard ratio will vary over the follow-up period, tests of proportional hazards yielding high P values are probably underpowered.

- I agree, but am concerned that the 'tests are unnecessary' statement may be interpreted by some as 'assessing PH is unnecessary' or 'it's fine to just report the HR from a PH model'.
- Researchers should consider the time-varying nature of hazard ratios in the design and reporting of their studies and should assess the proportional hazards assumption in the analysis.
- Another issue is that there is no omnibus test of PH.
- Arguably the most common test, based on scaled Schoenfeld residuals, tests the null of PH against the alternative that the HR changes as a linear or log-linear function of time.

Alternative measures

Quote from Stensrud & Hernán [1]

Reports of hazard ratios should be supplemented with reports of effect measures directly calculated from absolute risks, such as the survival differences or the restricted mean survival difference, at times prespecified in the study protocol. These measures are arguably more helpful for clinical decision-making and more easily understood by patients.

- I very much agree.

Estimating the HR from a PH model

Quote from Stensrud & Hernán [1]

Another limitation is that the magnitude of the Cox HR depends on the distribution of losses to follow-up (censoring), even if the losses occur at random. This limitation can be overcome by estimating an inverse probability-weighted hazard ratio.

- The statement is indisputably true, but how much difference does it make in practice?
- The authors show using simulations (see next slide taken from supplementary material) that differences can be considerable.
- Those three scenarios, however, concern large departures from PH and I would not consider reporting the HR from a PH model.
- How large is the 'bias' when a PH model is reasonable?

Table from supplementary material

Table. Simulated trials under the 3 scenarios described in the Figure in the main text. Each trial included 50,000 individuals and was analyzed first including all individuals and then after randomly censoring individuals such that about 20% of the events were unmeasured. The magnitude of the Cox hazard ratio depends on the censoring proportion even though the survival difference does not change.

Scenario	Censoring	Hazard ratio (95% CI), Cox proportional hazards model	3-year survival difference, % (95% CI), Kaplan-Meier estimator
1	No	0.69 (0.66 to 0.72)	3.2 (2.6 to 3.8)
	Yes	0.71 (0.67 to 0.74)	3.1 (2.5 to 3.8)
2	No	0.51 (0.48 to 0.54)	3.6 (3.1 to 4.1)
	Yes	0.62 (0.58 to 0.66)	3.6 (3.0 to 4.1)
3	No	1.27 (1.22 to 1.32)	–5.2 (–5.8 to –4.5)
	Yes	1.34 (1.28 to 1.40)	–5.2 (–5.9 to –4.5)

Estimating the HR from a PH model

Quote from Stensrud & Hernán [1]

One limitation of using Cox regression models when the hazard ratio is not constant during the follow-up period is reporting an incorrect standard variance estimator when the statistical model includes covariates other than the treatment group indicator [3]. This limitation can be overcome, and valid 95% confidence intervals can be estimated, by using bootstrapping methods.

- The statement is indisputably true, but how much difference does it make in practice?
- How many of you do this?

Annals of Internal Medicine ORIGINAL RESEARCH

Risk for Arterial and Venous Thrombosis in Patients With Myeloproliferative Neoplasms

A Population-Based Cohort Study

Malin Hultcrantz, MD, PhD; Magnus Björkholm, MD, PhD; Paul W. Dickman, MSc, PhD; Ola Landgren, MD, PhD; Åsa R. Derolf, MD, PhD; Sigurdur Y. Kristinsson, MD, PhD*; and Therese M.L. Andersson, MSc, PhD*

Background: Patients with myeloproliferative neoplasms (MPNs) are reported to be at increased risk for thrombotic events. However, no population-based study has estimated this excess risk compared with matched control participants.

Objective: To assess risk for arterial and venous thrombosis in patients with MPNs compared with matched control participants.

Design: Matched cohort study.

Setting: Population-based setting in Sweden from 1987 to 2009, with follow-up to 2010.

Patients: 9429 patients with MPNs and 35 820 matched control participants.

Measurements: The primary outcomes were rates of arterial and venous thrombosis. Flexible parametric models were used to calculate hazard ratios (HRs) and cumulative incidence with 95% CIs.

Results: The HRs for arterial thrombosis among patients with MPNs compared with control participants at 3 months, 1 year, and 5 years were 3.0 (95% CI, 2.7 to 3.4), 2.0 (CI, 1.8 to 2.2), and 1.5 (CI, 1.4 to 1.6), respectively. The corresponding HRs for venous thrombosis were 9.7 (CI, 7.8 to 12.0), 4.7 (CI, 4.0 to 5.4), and 3.2 (CI, 2.9 to 3.6). The rate was significantly elevated across all age groups and was similar among MPN subtypes. The 5-year cumulative incidence of thrombosis in patients with MPNs showed an initial rapid increase followed by gentler increases during follow-up. The HR for venous thrombosis decreased during more recent calendar periods.

Limitation: No information on individual laboratory results or treatment.

Conclusion: Patients with MPNs across all age groups have a significantly increased rate of arterial and venous thrombosis compared with matched control participants, with the highest rates at and shortly after diagnosis. Decreases in the rate of venous thrombosis over time likely reflect advances in clinical management.

Primary Funding Source: The Cancer Research Foundations of Radiumhemmet, Blodcancerfonden, the Swedish Research Council, the regional agreement on medical training and clinical research between Stockholm County Council and Karolinska Institutet, the Adolf H. Lundin Charitable Foundation, and Memorial Sloan Kettering Cancer Center.

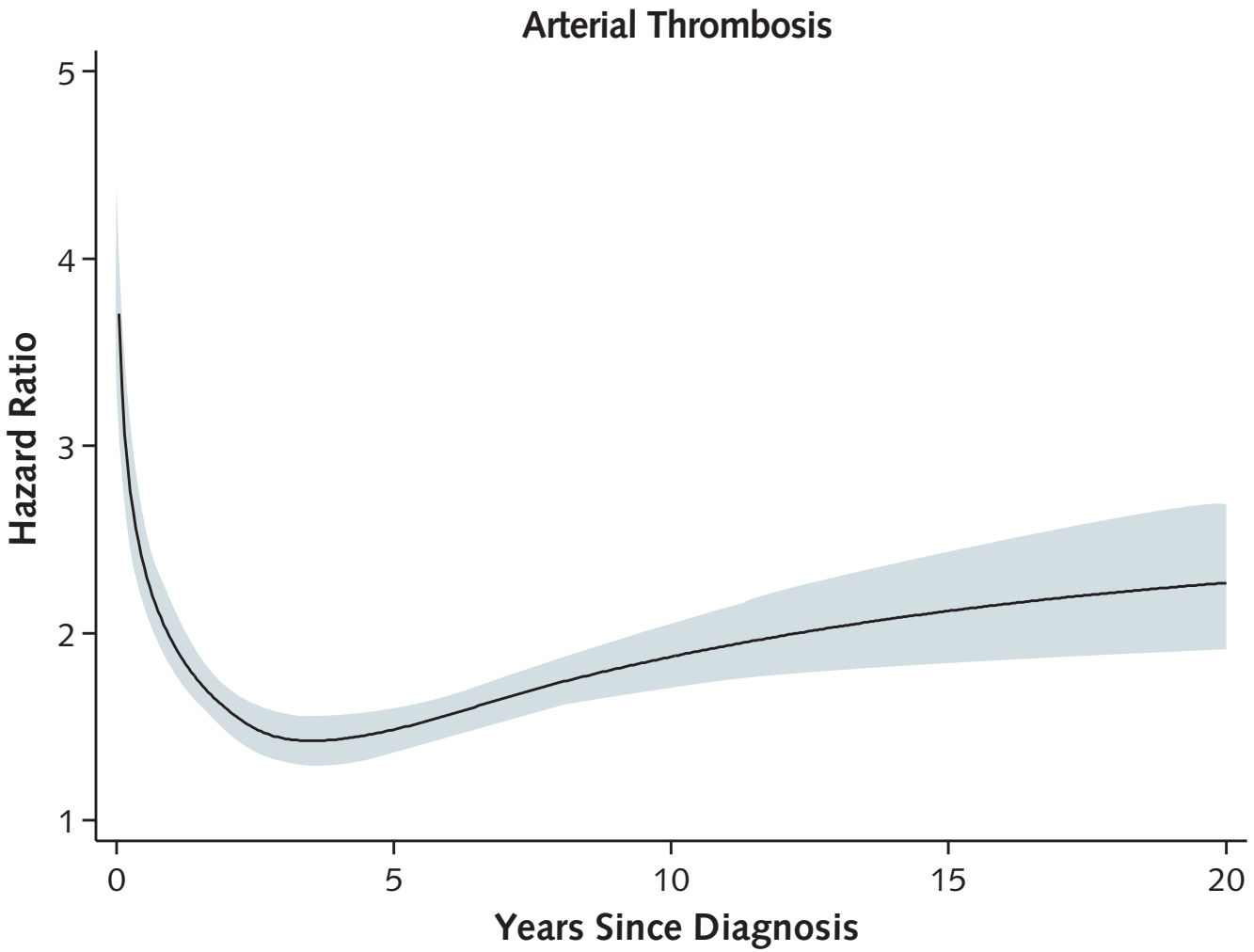
Ann Intern Med. 2018;168:317-325. doi:10.7326/M17-0028 **Annals.org**

For author affiliations, see end of text.

This article was published at Annals.org on 16 January 2018.

* Drs. Kristinsson and Andersson contributed equally to this work.

Figure 1. Arterial (*top*) and venous (*bottom*) thrombosis during follow-up in patients with MPNs versus matched control participants.



References

[1] Stensrud MJ, Hernán MA. Why Test for Proportional Hazards? *JAMA* 2020; **323**:1401–1402.

[2] Hernán MA. The hazards of hazard ratios. *Epidemiology* 2010;**21**:13–15.

[3] DiRienzo AG, Lagakos SW. Effects of model misspecification on tests of no randomized treatment effect arising from Cox's proportional hazards model. *Journal of the Royal Statistical Society Series B Statistical Methodology* 2001;**63**:745–757.