BJOG An International Journal of Obstetrics and Gynaecology

DOI: 10.1111/1471-0528.13884
www.bjog.org

Commentary

Cumulative risk of stillbirth in the presence of competing events

Al Naimi,[a] N Auger[b]

[a] Department of Epidemiology, University of Pittsburgh, Pittsburgh, PA, USA [b] Institut national de santé publique du Québec, Montreal, QC, Canada

Correspondence: Al Naimi, Department of Epidemiology, University of Pittsburgh, 130 DeSoto Street, 503 Parran Hall, Pittsburgh, PA 15261, USA. Email ashley.naimi@pitt.edu

Accepted 1 December 2015. Published Online 29 February 2016.

Please cite this paper as: Naimi AI, Auger N. Cumulative risk of stillbirth in the presence of competing events. BJOG 2016;123:1071–1074.

Linked article This article is commented on by GCS Smith and IR White, p. 1075 in this issue. To view this article mini commentary visit http://dx.doi.org/10.1111/1471-0528.13943.

Evidence suggests a sharp increase in stillbirth at the end of gestation. Between 35 and 38 weeks of gestation, cumulative risk of stillbirth in twin pregnancies increases by 60%, from 30 to 50 deliveries per 1000.[1] A similar increase occurs after 37 weeks in singleton pregnancies.[2] Although this evidence has implications for clinical practice, it is misleading. Increases in the cumulative risk of stillbirth at the end of gestation occur because researchers use the Kaplan–Meier method, which fails to account for live births as competing events.

In studies of stillbirth, women can have three possible outcomes: stillbirth, live birth, or right-censoring. Right-censoring occurs when a woman is no longer followed in a study (because of dropout or loss-to-follow-up), but is still at risk for the event. When a live birth occurs, the woman is no longer followed in a study because she is no longer at risk for the event. For this reason, live births (whether preterm or term) are competing events because they eliminate the possibility of stillbirth.[3] Competing events are recognised in many areas of research.[4,5] In obstetric studies of stillbirth, competing events are arguably more important because of the natural increase in live births near term. However, to our knowledge, only one letter in the obstetrics literature has raised concern over inappropriate treatment of competing events.[6]

The purpose of this commentary is to clarify why the Kaplan–Meier approach should not be used when competing events are present. To illustrate, we calculate the incorrect risk of stillbirth in a cohort of roughly two million pregnancies in Quebec, Canada. We describe the degree of error, explain why the error occurs, and show how to obtain correct estimates of cumulative risk when competing events are present. We conclude with a mention of the broader debate on how best to define the cumulative risk of stillbirth.

Kaplan–Meier versus cumulative incidence

There are several ways to compute cumulative risk. The Kaplan–Meier method[7] is the most common. The Kaplan–Meier method correctly estimates cumulative risk when there is right-censoring.[8] Women who are right-censored are still at risk of stillbirth even though they are lost to follow up, but the exact timing and outcome of pregnancy is not known.

The Kaplan–Meier approach is incorrect when competing events are present.[3–6,8–11] To censor live births, the Kaplan–Meier method 'redistributes' the potential stillbirth risk for a given pregnancy to subsequent stillbirths from other pregnancies. In doing so, the Kaplan–Meier method requires the assumption that women with live births are still at risk of stillbirth, which evidently is impossible. As a result, when there is no loss to follow up but competing events are present, the Kaplan–Meier stillbirth risk estimated after all pregnancies have delivered is not identical to the overall (empirical) risk of stillbirth in the cohort.

An alternative approach is known as the cumulative incidence method, which can estimate cumulative stillbirth risk without right-censoring live births.[12] Use of the cumulative incidence method does not require the assumption that women with live births remain at risk of stillbirth. When there is no loss to follow up but competing events are present, the cumulative incidence stillbirth risk after all pregnancies have delivered is identical to the overall (empirical) risk of stillbirth in the cohort.

Although straightforward, the cumulative incidence method is rarely used in obstetrics. We show how three software packages can be used to calculate cumulative still-

birth risk in the presence of competing events in an accompanying Appendix).

Data example

We extracted information on 2 521 597 live births and 10 551 stillbirths from Quebec birth registration certificates between 1981 and 2010. There were 59 641 multiple and 2 472 507 singleton pregnancies. Complete follow up was available for all individuals: pregnant women had either a stillbirth (event) or a live birth (competing event), and no women were lost to follow up. In Quebec, stillbirths must have a birthweight ≥500 g to be registered.

We estimated the cumulative risk of stillbirth for women with multiple and singleton pregnancies using (i) the Kaplan–Meier method, which incorrectly right-censors live births, and (ii) the cumulative incidence method, which correctly treats live births as competing events. We used completed weeks of gestation as the time scale. To assist researchers with calculating cumulative risk, we include an Appendix (Appendix S1) with annotated SAS (SAS Institute, Cary, NC, USA), STATA (StataCorps, College Station, TX, USA) and R (R Core Team, Vienna, Austria) code and a detailed explanation of the procedures.

Results

The overall (empirical) risk of stillbirth in the cohort was 4.2 per 1000. There were 765 stillbirths among 59 641 births from multiple pregnancies, yielding a stillbirth risk of 12.8 per 1000. There were 9786 stillbirths among 2 472 507 singleton pregnancies, yielding a risk of 4.0 per 1000.

Figure 1 displays the cumulative risk of stillbirth in women with multiple and singleton pregnancies using the Kaplan–Meier and cumulative incidence methods. Regardless of method, risk of stillbirth in multiple pregnancies increased to 10.5 per 1000 by 35 weeks of gestation. After 35 weeks, risk of stillbirth increased rapidly to 60.5 per 1000 with the Kaplan–Meier method, but to only 12.8 per 1000 with the cumulative incidence method. For singleton pregnancies, risk of stillbirth increased to 3.7 per 1000 by 39 weeks of gestation. After 39 weeks, risk of stillbirth increased rapidly to 10.1 per 1000 with the Kaplan–Meier method, but to only 4.0 per 1000 with the cumulative incidence method.

The Kaplan–Meier method incorrectly estimated a five-fold greater cumulative risk for multiple births, and a 2.5-fold greater cumulative risk for singleton births. The cumulative incidence method, on the other hand, yielded cumulative risks that were identical to the actual risk of stillbirth in the cohort. We next explain why the Kaplan–Meier approach is incorrect.

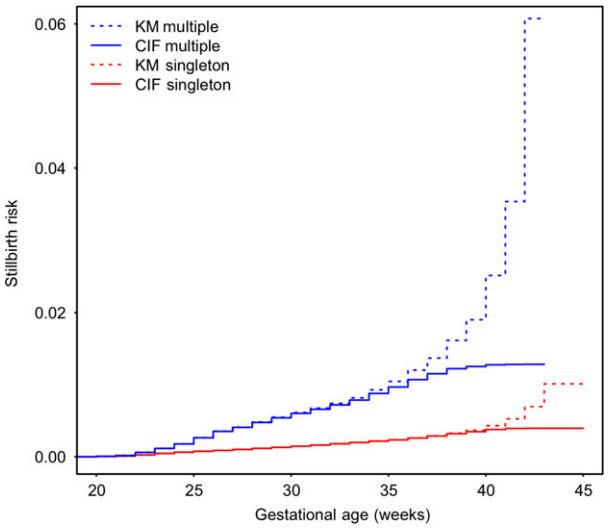

Figure 1. Cumulative risk of singleton and multiple stillbirth, Kaplan–Meier versus cumulative incidence function, Quebec, Canada, 1981–2011. KM Multiple, Kaplan–Meier method, multiple pregnancies; CIF Multiple, Cumulative incidence function method, multiple pregnancies; KM Singleton: Kaplan–Meier method, singleton pregnancies; CIF Singleton: Cumulative incidence function method, singleton pregnancies.

Discussion

We showed that, using appropriate methods, the cumulative risk of stillbirth in multiple and singleton births does not increase sharply at term in a population-wide cohort of women under routine care in Quebec. In contrast, we found a sharp increase in the cumulative risk of stillbirth at term using inappropriate methods that right-censored live births. This sharp increase became substantial when live births occurred more frequently. Furthermore, the increase is similar to previous studies that used the Kaplan–Meier method to right-censor live births.[1,2,13] Using the Kaplan–Meier approach, we showed that the risk of stillbirth in multiple pregnancies increased considerably after 35 weeks of gestation, with a five-fold greater overall risk compared with the cumulative incidence method. Similarly, the cumulative risk for singleton pregnancies increased sharply after 40 weeks, yielding a 2.5-fold greater overall risk. The artificial increase in risk after 35 and 40 weeks of gestation with the Kaplan–Meier method is a consequence of treating competing events (live births) as right-censored observations, and does not reflect an actual increase in stillbirth risk late in gestation.

To show why the Kaplan–Meier method is incorrect, consider the increase in the cumulative risk of stillbirth at 38 weeks among singleton pregnancies in our cohort. Consider further that there were 2 472 507 singleton pregnancies at risk, such that with every stillbirth, the cumulative

risk increases by 1/2 472 507. At 38 weeks there were 702 stillbirths. Hence, we can manually calculate the increase in cumulative risk at 38 weeks as 702/2 472 507 = 0.284 per 1000. Although the cumulative incidence method correctly estimated the same cumulative risk increase of 0.284 per 1000, the Kaplan–Meier method incorrectly estimated a cumulative risk increase of 0.397 per 1000 at 38 weeks.

Why is the Kaplan–Meier estimate higher? The Kaplan–Meier increase in risk at 38 weeks is a combination of the actual risk increase (0.284 per 1000), plus the risk increase that would have been observed had some of the live births at or before 38 weeks been stillbirths (0.397−0.284 per 1000). This unrealised stillbirth risk at ≤38 weeks is redistributed to all remaining stillbirths.[14] The difference (0.397−0.284) is proportional to the number of live births that occurred before 38 weeks. This addition to the risk at 38 weeks is the problem because it is based on the untenable assumption that some live births are succumbing to stillbirth after delivery. Redistribution of risk to later stillbirths makes sense when there is right-censoring, but not when there are competing events such as live birth.[4]

With true right-censoring, the Kaplan–Meier method correctly estimates the cumulative risk by redistributing the potential risk of right-censored pregnancies to remaining pregnancies at risk.[14] Consider a hypothetical study of ten pregnant women over 42 weeks of gestation. Each woman potentially contributes one-tenth to the cumulative risk (i.e. if they have the event, the risk increases by one-tenth). If a woman is lost to follow up at 35 weeks, then her potential contribution to the cumulative risk is lost. The Kaplan–Meier method accounts for this loss by redistributing (or 'spreading') the potential risk of the censored woman to the remaining women, slightly increasing each remaining woman's contribution to the cumulative risk.[3] This redistribution of risk is the reason that the Kaplan–Meier method estimated a cumulative risk at 38 weeks of 0.397 per 1000 instead of 0.284 per 1000.

Redistribution of potential risk assumes that censored women could have had a stillbirth in the study had they not been censored. With live birth, this assumption is violated. Considering our previous example, suppose a woman has a live birth at 35 weeks, and is censored even though she is not lost to follow up. The Kaplan–Meier method assumes she can still have a stillbirth, and redistributes her potential risk to the remaining women at risk.[3] It is technically impossible, however, for a woman who had a live birth to be at risk of stillbirth during the same pregnancy. Hence, the Kaplan–Meier method is incorrect and should not be used to calculate cumulative stillbirth risks. The cumulative incidence method can account for both competing events and right-censoring, and is a more appropriate alternative.

In this paper, we showed that the divergence between the Kaplan–Meier and cumulative incidence methods can be sharp. Importantly, what we have shown here is well understood and accepted in biostatistics and other areas of clinical medicine. Indeed, mis-application of the Kaplan–Meier method has received attention in many areas of clinical research,[5,8,9,11,15] but less so in reproductive, perinatal, and obstetrics research, despite the greater likelihood of severe errors. These errors can be seriously misleading. One study, for example, reported that a 'second stillbirth epoch' after 36 weeks of gestation may be mediated through impaired placental function measured by Doppler ultrasound.[2] However, the study used Kaplan–Meier estimates of cumulative risk. It is most probable that the entirety of the 'second stillbirth epoch' was due to inappropriate censoring of live births.

Defining and estimating stillbirth and infant mortality risk across gestation has long been of interest in obstetrics research. Much of this research has been geared towards methods for conveying instantaneous risk. These methods include use of incidence-based measures,[16] and proportional hazards regression,[17] which resolve several puzzling artefacts, including the problem of intersecting mortality curves. However, the issue we discuss in this commentary is separate from the debate over measurement of instantaneous risks. When the Kaplan–Meier method is used to estimate cumulative stillbirth risk, one must assume that a given pregnancy is still at risk of stillbirth after actually ending in live birth. This is an untenable assumption. The Kaplan–Meier method should not be used to estimate cumulative risk when the outcome is subject to competing events.

Disclosure of interests

None declared. Completed disclosure of interests form available to view online as supporting information.

Contribution to authorship

AIN conceived of the study and design, and was primarily responsible for the interpretation of findings, execution of statistical analyses, and the devising of all manuscript drafts; NA provided important intellectual contributions to the study design and analysis, interpretation of findings, and manuscript drafting.

Details of ethics approval

Ethical review was waived by the University of Montreal Hospital Centre Internal Review Board, as data were anonymised and conformed to requirements for research involving humans in Canada.

Funding

None.

Acknowledgements
We thank Michael S. Kramer and Jay S. Kaufman for their helpful comments.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Supplementary material illustrating the implementation of the cumulative incidence method for estimating cumulative risk in three software platforms. ■

References

1 Danon D, Sekar R, Hack KEA, Fisk NM. Increased stillbirth in uncomplicated monochorionic twin pregnancies: a systematic review and meta-analysis. *Obstet Gynecol* 2013;121:1318–26.
2 Singh T, Leslie K, Bhide A, D'Antonio F, Thilaganathan B. Role of second-trimester uterine artery Doppler in assessing stillbirth risk. *Obstet Gynecol* 2012;119:256–61.
3 Tableman M, Kim JS. *Survival Analysis Using S: Analysis of Time-to-Event Data*. Boca Raton, FL: CRC Press, 2003.
4 Gooley TA, Leisenring W, Crowley J, Storer BE. Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. *Stat Med* 1999;18:695–706.
5 Pepe MS, Mori M. Kaplan-Meier, marginal or conditional probability curves in summarizing competing risks failure time data? *Stat Med* 1993;12:737–51.
6 Fong F, Thangaratinam S, Zamora J. Increased Stillbirth in Uncomplicated Monochorionic Twin Pregnancies: A Systematic Review and Meta-analysis. *Obstet Gynecol* 2013;122:1302.
7 Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958;53:457–81.
8 Grunkemeier GL, Jin R, Eijkemans MJC, Takkenberg JJM. Actual and actuarial probabilities of competing risks: apples and lemons. *Ann Thorac Surg* 2007;83:1586–92.
9 Southern DA, Faris PD, Brant R, Galbraith PD, Norris CM, Knudtson ML, et al. Kaplan-Meier methods yielded misleading results in competing risk scenarios. *J Clin Epidemiol* 2006;59:1110–4.
10 Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. *Am J Epidemiol* 2009;170:244–56.
11 Koller MT, Raatz H, Steyerberg EW, Wolbers M. Competing risks and the clinical community: irrelevance or ignorance? *Stat Med* 2012;31:1089–97.
12 Kalbfleisch JD, Prentice RL. *The Statistical Analysis of Failure Time Data*. Hoboken, NJ: John Wiley & Sons, 2011.
13 August EM, Salihu HM, Weldeselasse H, Biroscak BJ, Mbah AK, Alio AP. Infant mortality and subsequent risk of stillbirth: a retrospective cohort study. *BJOG* 2011;118:1636–45.
14 Efron B. The two sample problem with censored data. In: Le Cam LM Neyman J, editors. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley, CA: University of California Press, 1967:831–53.
15 Verduijn M, Grootendorst DC, Dekker FW, Jager KJ, le Cessie S. The analysis of competing events like cause-specific mortality–beware of the Kaplan-Meier method. *Nephrol Dial Transplant* 2011;26:56–61.
16 Joseph KS. Incidence-based measures of birth, growth restriction, and death can free perinatal epidemiology from erroneous concepts of risk. *J Clin Epidemiol* 2004;57:889–97.
17 Platt RW, Joseph KS, Ananth CV, Grondines J, Abrahamowicz M, Kramer MS. A Proportional hazards model with time-dependent covariates and time-varying effects for analysis of fetal and infant death. *Am J Epidemiol* 2004;160:199–206.