

Comment

Dealing with the Problem of Monotone Likelihood in the Inflation of Estimated Effects in Clinical Studies. Comment on Hasegawa et al. Impact of Blood Type O on Mortality of Sepsis Patients: A Multicenter Retrospective Observational Study. *Diagnostics* 2020, 10, 826

I-Shiang Tzeng ^{1,2,3} 

¹ Department of Research, Taipei Tzu Chi Hospital, Buddhist Tzu Chi Medical Foundation, New Taipei City 23142, Taiwan; istzeng@gmail.com

² Department of Statistic, National Taipei University, Taipei 10478, Taiwan

³ Institute of Education, Tzu Chi University, Hualien 97004, Taiwan



Citation: Tzeng, I.-S. Dealing with the Problem of Monotone Likelihood in the Inflation of Estimated Effects in Clinical Studies. Comment on Hasegawa et al. Impact of Blood Type O on Mortality of Sepsis Patients: A Multicenter Retrospective Observational Study. *Diagnostics* 2020, 10, 826. *Diagnostics* 2022, 12, 2295. <https://doi.org/10.3390/diagnostics12102295>

Academic Editor: Eric Deconinck

Received: 22 July 2022

Accepted: 19 September 2022

Published: 23 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

I read the press article from the Fujita Health University School of Medicine titled “Impact of Blood Type O on Mortality of Sepsis Patients: A Multicenter Retrospective Observational Study”. It was based on an article by Hasegawa et al. [1]. This study aimed to explore the association between the risk of mortality and blood type for sepsis patients based on a retrospective cohort study. The authors found that sepsis patients with blood type O are associated with a significantly higher risk of 90-day mortality from septic shock. Indeed, this research has provided important contributions to the field; however, there are some biostatistics issues that should be clarified.

Firstly, the authors compared the risk of 90-day mortality in patients with septic shock between blood type groups and found that the adjusted odds ratio (OR) of mortality increased 3.26 times in patients with blood type O compared to that in the blood type non-O reference group (blood type O patients with sepsis shock (death number/total number = 21/40), blood type non-O patients with sepsis shock (death number/total number = 21/76); adjusted OR = 3.26; 95% confidence interval (CI: 1.34–7.9)). The massive effect estimation that may result from specificity in the data set, called “monotonic likelihood” [2,3], is of concern to scientists. The true OR cannot be confirmed if the CI of the OR is too wide. We suggest that the researchers reconstruct the interval estimation based on profile penalized log likelihood (PPL) owing to the wide CI with a massive effect estimate [3].

In addition to the longitudinal cohort study, a similar phenomenon was observed for the odds ratio (OR) in a cross-sectional cohort study. As a demonstration, a lung cancer data set [4] was analyzed with the package logistic [5], which applies the Firth regression in the software R (version 4.0.0, Vienna, Austria). A total of 649 male patients with cancer were included (cases), 647 of whom were reported to be smokers. An equal number of participants (649 men without cancer) were recruited as the control group, 622 of whom were reported to be smokers. The OR of lung cancer in smokers compared with that of non-smokers can be calculated as $(647 \times 27)/(2 \times 622) = 14.04$ —that is, the odds of contracting lung cancer in smokers was estimated to be 14 times the odds of contracting lung cancer in non-smokers. We would like to determine the reliability of this estimate based on the width of the 95% CI. The 95% CI for this OR is between 3.33 and 59.3. After modification based on PPL, the OR and 95% CI of lung cancer for smokers compared with that of non-smokers decreased (modified OR = 11.44; 95% CI: 3.74–56.42). Figure 1 shows that the range of the 95% CI of smoking status, which was estimated based on a PPL, is shorter than that of the 95% CI before modification.

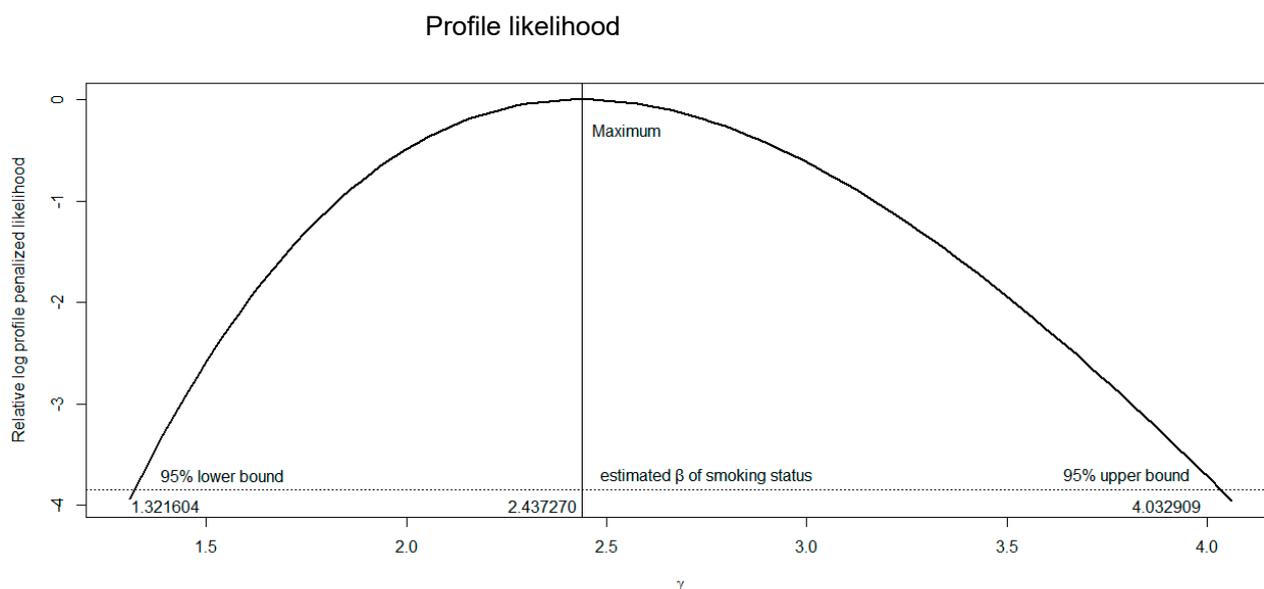


Figure 1. Plot of the profile penalized log-likelihood function for variable smoking status.

As shown in Table 3 of the study conducted by Hasegawa et al. [1], only 21 cases were present in blood type O of patients with septic shock, which means that their data were subject to sparse events, which may have increased the probability of monotone likelihood. Several methods have been proposed to solve the problem of monotonic likelihood through data penalization, and augmentation has been suggested as an efficient method [2]. After adjusting for age, sex, and Sequential (Sepsis-related) Organ Failure Assessment score in Table 4 [1], we found that the effect estimate was eliminated and the CI was significantly increased (adjusted OR = 3.26; 95% CI: 1.34–7.9 in contrast to crude OR = $(21 \times 76)/(21 \times 40) = 1.9$, 95% CI: 0.93–3.89 calculated by a website was available at https://www.medcalc.org/calc/odds_ratio.php, accessed on 1 June 2022). We speculate that this problem is not limited to univariate models alone [2]. When considering multiple regressions, the sample size, event number, dimension of association with the result, size of the balance, a linear combination of variables [6,7], and count of binary covariates affect the generality of monotonic likelihood. This is a comment article aiming to evaluate whether the estimate of the 90-day mortality risk of blood type O (i.e., inherent risk factor) patients with septic shock may have had a bias. As this was a retrospective observational study [1] with several potential sources of error in the measurement of the sepsis event and independent variables, the point estimates must be seen as approximate and certainly the accuracy would have been increased by the PPL modification suggested by the author of the letter. However, we also must not overemphasize the magnitude of those non-sampling errors. Therefore, we commented on a study [1] that may be a potentially useful note for researchers involved the problem of the monotonic likelihood of the univariate and multivariate models in clinical studies.

Funding: This research received no external funding.

Data Availability Statement: Data was available at <https://www.mdpi.com/2075-4418/10/10/826>.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Hasegawa, D.; Nishida, K.; Kawaji, T.; Hara, Y.; Shimomura, Y.; Moriyama, K.; Niimi, D.; Kuriyama, N.; Shintani, A.; Komura, H.; et al. Impact of Blood Type O on Mortality of Sepsis Patients: A Multicenter Retrospective Observational Study. *Diagnostics* **2020**, *10*, 826. [[CrossRef](#)] [[PubMed](#)]
2. Heinze, G.; Schemper, M. A solution to the problem of monotone likelihood in Cox regression. *Biometrics* **2001**, *57*, 114–119. [[CrossRef](#)] [[PubMed](#)]

3. Heinze, G.; Ploner, M. SAS and SPLUS programs to perform Cox regression without convergence problems. *Comput. Methods Programs Biomed.* **2002**, *67*, 217–223. [[CrossRef](#)]
4. Doll, R.; Hill, A.B. Smoking and carcinoma of the lung; preliminary report. *Br. Med. J.* **1950**, *2*, 739–748. [[CrossRef](#)] [[PubMed](#)]
5. Heinze, G.; Ploner, M.; Jiricka, L. *Logistf: Firth's Bias-Reduced Logistic Regression*; R Package Version 1.24; CRAN: Vienna, Austria, 2020.
6. Tsiatis, A.A. A large sample study of Cox's regression model. *Ann. Stat.* **1981**, *9*, 93–108. [[CrossRef](#)]
7. Albert, A.; Anderson, J. On the existence of maximum likelihood estimates in logistic regression models. *Biometrika* **1984**, *71*, 1–10. [[CrossRef](#)]