Letter to the Editor

Necessary sample size for solid conclusion

Dear Editor,

We read the paper, "Infection after operative fixation of tibia plateau fractures. A risk factor analysis," [1] recently published online, with great honor and interest. In this research the authors claimed that excessive alcohol consumption is the only independent predictor of infection after fixation, while smoking, diabetes, and/or obesity are not. These results are drawn from clinical data of patients with tibia plateau fractures, from multiple trauma centers. We congratulate the authors on raising an interesting clinical question and void in knowledge. However, there are some statistical concerns of this research worth noting:

1) A multiple logistic regression was used with final 8 variables included, however, the reported sample size (13 infection cases and 162 non-infection cases) is not adequately sized for such an analysis. There is a wide acceptance of a minimal 10 events per variable (EPV) rule for medical research [2], or preferably more [3], which means the smaller group should contain samples more than ten times the number of variables. In this research, 8 variables are included in final analysis; therefore, each group should have more than 80 patients, while the number in the infected group (13 patients) is far more than enough. The small EPV (13/8=1.6) of this multiple logistic regression analysis may affect the accuracy and precision of regression coefficients for independent variables. To prove the reliability of its conclusion, it is better to report the quality test methods and results (such as Hosmer and Lemeshow test) of this regression model.

2) The description and explanation of BMI is conflicting. In the methods section, the BMI was described as a dichotomous variable since original BMI data was converted into elevated or non-elevated. However, in the discussion section the BMI was treated as a continuous variable with no relationship to post-operative infection. Since the meaning of result from dichotomous and continuous variable in logistic regression is different, this inconsistency is confusing.

Availability of patients' data from multiple centers motivates clinicians with ambition to identify risk factors related to vexing clinical troubles, which help to shape medical practice worldwide. However, the analysis of such data needs appropriate statistical methods as well as sufficient samples. The interpreting of the results, especially from retrospective research, should also be with utmost caution to avoid misleading or answers to clinical question.

Declaration of Competing Interest

No potential and real conflicts of interest.

Acknowledgements

References

[1] Chan G, Iliopoulos E, Jain A, Turki M, Trompeter A. Infection after operative fixation of tibia plateau fractures. A risk factor analysis. Injury 2019, doi:http://dx.doi.org/10.1016/J.INJURY.2019.06.022.

[2] Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49:1373–9, doi:http://dx.doi.org/10.1016/S0895-4356(96)00236-3.

[3] van Smeden M, de Groot JAH, Moons KGM, Collins GS, Altman DG, Eijkemans MJC, et al. No rationale for 1 variable per 10 events criterion for binary logistic regression analysis. BMC Med Res Methodol 2016;16:163, doi:http://dx.doi.org/10.1186/s12874-016-0267-3.

Zhen Zhang
Guanglin Wang*
Department of Orthopedics, West China Hospital, Sichuan University, Chengdu, China

* Corresponding Author at: No. 37, Guoxue Lane, Wuhou District, Chengdu, Sichuan Province, 610041, China.
E-mail address: wangglfrank@hotmail.com (G. Wang).