CORRESPONDENCE

Reply to "Limitations of multivariate survival analysis"

Modern Pathology (2022) 35:994–995; https://doi.org/10.1038/s41379-021-00988-3

TO THE EDITOR:

We recently reported a detailed clinicopathologic study of 116 cases of anal squamous cell carcinoma (SqCC) that demonstrated the utility of HPV and *TP53* mutation status in prognostication and described corresponding patterns of p16 and p53 immunohistochemical staining that are clinically meaningful[1]. As part of our study, we performed Cox proportional hazards multivariate analysis, which showed that HPV status, tumor T stage, lymph node involvement, and age were statistically associated with survival while controlling for the other predictors in the model. A recent correspondence by Ramia de Cap and Kaul questioned the validity of this analysis, labeling it "likely… inaccurate," a dismissive characterization that is unsubstantiated.

Ramia de Cap and Kaul argue that an events-per-variable (EPV) ratio of >10:1 is "required" for Cox proportional hazards analysis. In fact, a 10:1 EPV ratio is not, as the letter authors suggest, an inviolable law whose breach will "mislead the medical field" but rather a rough guideline for sample size that is based on simulation studies with design limitations[2]. Many studies have challenged the 10:1 EPV rule[3–5], and some epidemiological researchers have concluded that "there is no single rule based on EPP [events per predictor] that would guarantee an accurate estimation of logistic regression parameters[5]." Others have concluded that the current evidence underlying the 10:1 EPV rule as a minimum sample size criterion for logistic regression is weak[6]. Similarly, it has been established that in many cases an EPV threshold of 10:1 can be relaxed without significantly impacting the accuracy of results[3], particularly in uncommon disorders with a low disease prevalence, such as anal SqCC[4]. Indeed, when our analysis is limited to the four variables initially found to be statistically significant (thereby increasing the EPV to 7.8), all remain significantly associated with survival (HPV status: $p = 0.0092$; age at diagnosis: $p = 0.0014$; T stage: $p = 0.0223$; lymph node involvement: $p = 0.0324$). Similarly, when this analysis was repeated with Firth's correction, a method for increasing the efficiency of the estimators in logistic regression with small samples[6], all variables remained statistically significant. When age is dropped and the analysis is limited to three variables (HPV status, T stage, and lymph node involvement), thereby increasing the EPV to >10, all variables similarly remain statistically significant. These findings confirm that our results are consistent and robust.

To their credit, Ramia de Cap and Kaul have emphasized a limitation inherent to many studies of uncommon subtypes of tumors, in which sample size and the number of events are intrinsically low, precluding an EPV of >10:1. In such studies on the clinical significance of uncommon variants of neoplasia, strict adherence to the 10:1 EPV recommendation would sever important lines of research[7]. Higher case numbers, resulting in higher EPV, would be preferable in nearly every study in the field of pathology, but this is not always possible due to disease rarity. While we acknowledge that our statistical analysis is limited by the sample size and number of events in our cohort, we note that our analysis is larger than previous studies with detailed clinical, pathologic, and molecular analysis of anal carcinoma[8,9].

Tumor T stage and lymph node involvement are well-established as markers of prognosis in anal SqCC. These variables form the basis for the TNM staging of anal SqCC and are routinely used to place patients into clinically relevant risk groups that are associated with outcome. Similarly, it is established that HPV status is correlated with survival, with HPV-negative status portending a poor prognosis[8–10]. Our findings fall squarely in line with these established associations, supporting their validity.

Finally, we strongly encourage the letter authors to abstain from heedlessly attributing perceived flaws in the work of their peers to ignorance, and instead to consider, in good faith, that their peers' analyses are deliberate and informed of limitations, which no study or statistical analysis lacks.

Xiaoqin Zhu[1], Jian Zou[2], Mark C. Mochel[3] and Jacob R. Bledsoe[4]✉

[1]*Department of Pathology, University of Massachusetts, UMass Memorial Health Care, Worcester, MA, USA.* [2]*Department of Mathematical Sciences, Worcester Polytechnic Institute, Worcester, MA, USA.* [3]*Departments of Pathology and Dermatology, Virginia Commonwealth University School of Medicine, Richmond, VA, USA.* [4]*Department of Pathology, Boston Children's Hospital, Boston, MA, USA.* ✉*email: Jacob.Bledsoe@childrens.harvard.edu*

REFERENCES

1. Zhu, X. et al. Molecular and immunophenotypic characterization of anal squamous cell carcinoma reveals distinct clinicopathologic groups associated with HPV and TP53 mutation status. *Mod. Pathol.* **34**, 1017–1030 (2021).
2. Peduzzi, P. et al. A simulation study of the number of events per variable in logistic regression analysis. *J. Clin. Epidemiol.* **49**, 1373–1379 (1996).
3. Vittinghoff, E. & McCulloch, C. E. Relaxing the rule of ten events per variable in logistic and cox regression. *Am. J. Epidemiol.* **165**, 710–718 (2007).
4. Riley, R. D. et al. Minimum sample size for developing a multivariable prediction model: PART II—binary and time-to-event outcomes. *Stat. Med.* **38**, 1276–1296 (2019).
5. Courvoisier, D. S. et al. Performance of logistic regression modeling: beyond the number of events per variable, the role of data structure. *J. Clin. Epidemiol.* **64**, 993–1000 (2011).
6. Smeden, M. et al. No rationale for 1 variable per 10 events criterion for binary logistic regression analysis. *BMC Med. Res. Methodol* **16**, 163 (2016).
7. Talhouk A., McAlpine J. N. Reply to Survival analysis and treatment effects in patients with endometrial cancer and POLE mutations. *Cancer* 2021. https://doi.org/10.1002/cncr.33808.
8. Meulendijks, D. et al. HPV-negative squamous cell carcinoma of the anal canal is unresponsive to standard treatment and frequently carries disruptive mutations in TP53. *Brit. J. Cancer* **112**, 1358–1366 (2015).
9. Soares, P. C. et al. HPV positive, wild type TP53, and p16 overexpression correlate with the absence of residual tumors after chemoradiotherapy in anal squamous cell carcinoma. *BMC Gastroenterol.* **18**, 30 (2018).
10. Parwaiz, I. et al. A systematic review and meta-analysis of prognostic biomarkers in anal squamous cell carcinoma treated with primary chemoradiotherapy. *Clin. Oncol.* **31**, e1–e13 (2019).

Received: 23 November 2021 Revised: 23 November 2021 Accepted: 29 November 2021
Published online: 28 December 2021

COMPETING INTERESTS
The authors declare no competing interests.

ADDITIONAL INFORMATION
Correspondence and requests for materials should be addressed to Jacob R. Bledsoe.