**Point-by-point review response and revision summary of Manuscript BA2012-009**

“Bayesian inference for Cox Proportional Hazard Models with Partial likelihoods, Semi-parametric Covariate Effects, and Correlated Observations”

Dear Dr Guindani (Editor-in-Chief) and Editors,

We wish to thank you for giving us this opportunity to revise our manuscript, and thank the associate editor and the referee for their detailed and helpful comments. We have provided point-by-point responses to the comments below. The main revisions we made can be summarized as the followings:

* In response to helpful comments from the Editor, we reformatted our section 3 on methodology, to better emphasize the methodological innovations we made in order to extend the method of (Stringer et al., 2020) to work in our case. Furthermore, on top of the proposed method shown in the original manuscript, we have made some additional methodological improvements in this revision, which are explained in details in the point-by-point responses.
* In response to helpful comments from the Associate Editor, we expanded the simulation and example sections in our revised manuscript. In the revised version, we not only implemented our proposed Laplace-approximation based method for the inference, but also the MCMC method for the same model based on partial likelihood, with appropriate comparisons between the two approaches.
* In response to helpful comments from the reviewer, we have included more thorough comparisons between the proposed method and the existing method, and illustrated the accuracy of the proposed posterior approximation in terms of Mean Square Error and posterior coverage probability through independent replications.

We believe this revision is a significant improvement compared to our original manuscript, and we hope it is now suitable for publication in *Bayesian Analysis*. Thank you for considering our work.

Sincerely,

Ziang Zhang Alex Stringer Jamie Stafford Patrick Brown

PhD student PhD Candidate Professor Professor

**Point-by-point review response**

**Comments from the Editor:**

*1. As explained in the reports to the authors, despite being a solid piece of work, both reviewers have raised concerns about the methodological innovation of the manuscript, especially if compared with a paper (Stringer et al., 2020) recently published by three of the four authors of this manuscript.*

*A resubmission will have to address the points raised by the referees, in particular with regard to the methodological novelty and the practical utility of their proposed approach.*

**Response:** We understand the concern from the editor on the methodological innovation of the manuscript compared to the paper of (Stringer et al., 2020), but we believe this concern is mostly due to the way we present our proposed approach in the previous version of the manuscript. The method from the paper of (Stringer et al., 2020) only works with a simplest special case of partial likelihood, in which the Hessian matrix of the log (partial) likelihood will be sparse. However, in most applications of Cox Proportional Hazard (Cox PH) Model with partial likelihood, the Hessian matrix of log likelihood if directly adopted the method of Stringer et al (2020), will be completely dense and having dimension growing with the sample size. Through the use of Adaptive Gaussian Quadrature and quasi-Newton optimization method, the proposed method extends the approach of Stringer et al (2020) to work for general partial likelihood, and accommodate a more flexible class of models.

To better emphasize the methodological innovation of the proposed method over the method of Stringer et al (2020), we significantly change the presentation of the section 3 on Methodology in the revised manuscript (page 6). At the same time, we also give a more detailed review of the method of Stringer et al (2020) in the third paragraph of section 2.1 (page 3), and in the last two paragraphs of section 2.3 (page 5&6), to better show the flexibility of the proposed method over the method of Stringer et al (2020).

Furthermore, we have made additional improvements to the proposed method in the previous manuscript, to better accommodate the computation of the dense Hessian matrix especially when the sample size is relatively large. These improvements include the followings:

1. We removed the additional Gaussian noise in the additive linear predictor, which made the dense Hessian matrix required in the inference to have a constant dimension that will no longer grow with sample size, and hence reduced the computational load to compute, store or factorize the Hessian matrix. More details of this methodological improvement can be found in section 3.1 (page 6-8).

2. We adopted the automatic differentiation method from the R package TMB (Kristensen et al., 2021) for the computation of both the Laplace approximation and its derivatives. This reduced the number of computations required to compute the mode of the Laplace approximation, and hence improved on the computational efficiency of the proposed method. More details of this methodological improvement can be found in the last paragraph of section 3.2 (page 8-9).

**Comments from the Associate Editor:**

1. *I recommend expanding your paper so as to include appropriate conceptual and numerical comparisons between MCMC based inference for the Cox model and the proposed methodology, to better inform applied statisticians.*

**Response**: We would like to thank the associate editor for this constructive suggestion. In our revised manuscript, we have included comparisons between the inference from the proposed method and the inference from MCMC method, both in the two simulation examples (section 4.1, page 9-12), and in the real data analyses (section 4.2, page 12-13).

**Comments from the Referee:**

1. *The model specification presented in Equation 1 is general as it accounts for predictors whose association with the log-hazard is modelled as linear or semi-parametrically, and for a frailty term. It is not immediately clear though if the computational challenges provided by the use of the Cox PH model with partial likelihood have anything to do with the specific model that was considered or the same would arise regardless of the model adopted for the effect of the predictors. If this is the case, in my opinion, the presentation of the method might be more effective if introduced, at first, for the simplest model, for example the one only counting only the covariates x\_{ij}. The more complex model could be introduced at a later stage as the one implemented in the examples of Section 4.*

**Response:** Although we see the point suggested by the reviewer, we decided to keep the current presentation for the following reason. The model form presented in Equation (1) aims to show the flexibility of the proposed method to include different kinds of covariate effects and frailties in the analysis. In the revised manuscript, we have made this part clearer by providing several references to the model forms accommodated by the existing Bayesian inference methods and contrasted them with the flexible model form we considered in this work (page 3, the first paragraph of section 2.1).

2. *The starting point of the paper is that existing approaches based on INLA cannot be applied to Cox PH models with partial likelihood. It should be clarified in the introduction if other Bayesian methods, either exact or approximate, have been used in the literature. This aspect should be clarified.*

*In addition, I think it might be worth mentioning another class of nonparametric models for the baseline hazard function, such as the one of Dykstra & Laud, (1981), where the baseline hazard function is modelled by means of a gamma process.*

**Response:** We would like to thank the referee for this helpful comment. In the revised manuscript, we have added a more complete literature overview on the existing Bayesian methods for this model, including the nonparametric method mentioned in the reviewer’s comment (page 2, the second paragraph of section 1).

3. *I find the simulation study of Section 4 not very compelling as, in my opinion, it fails at satisfactorily address two questions that I think are relevant when adopting a new and approximate method for posterior computation. Namely, 1) how good is the approximation? 2) How is the method compared with alternative strategies for posterior computations, both in terms of accuracy and computational efficiency?*

*I will list some points which I believe would help making the simulation study more compelling. These suggestions are nothing but a possible option, other ideas might be equally valid. As for point 1, it would be interesting -for example- to explore the coverage of the approximate posterior estimates on a set of several replicates. Another option could be to compare posterior estimates of the proposed approximate method with those obtained with standard MCMC (when the parametric form of the baseline hazard which generated the data is known). As for point 2, right now the only comparisons are rather qualitative and made only on the analysis of the real data sets of Sections 4.2 and 4.3. A comparison could be carried out also in the case of simulated data: one option that seems in line with how the material is presented would be to compare posterior inference obtained with the proposed method, with the one produced by the approach of Martino et al., (2011). Such comparison could be done under various settings, e.g. settings where the smoothness assumptions required by Martino and co-authors are met by the data generating process, and settings where the same requirements are not met.*

**Response:** To address the first question mentioned by the referee, we followed the referee’s suggestions and examined the posterior coverage of the proposed approximate method (as well as the Mean Square Error) on a set of independent replications. At the same time, we also compared the posterior estimates of the proposed method with those obtained with an equivalent MCMC method, both in the two simulation studies and in the real data analyses.

As for the second question of the referee, we provided two specific examples on which the proposed approximate inference method is shown to perform better than the existing approximate inference method of Martino et al., (2011) based on full likelihood. This is shown to occur when the frailties are sparse, and when the true baseline hazard function is rapidly varying. Also, as the referee has mentioned, we have now provided the simulation results under various settings (e.g. settings where the smoothness of baseline hazard is at different levels, settings where the sparsity of frailties is at different levels). The advantages of the proposed method are shown both in the smaller MSE and in the posterior coverage rates closer to the nominal levels. These changes are reflected in the revised section 4 (page 9-14).