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 to extend the method of (Stringer et al., 2020) to work in our case. At the same time, we have made major improvements on the proposed method in the original manuscript which significantly enhance the flexibility and the computational efficiency of the proposed method.
- 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 and the existing Laplace-approximation based method INLA for the inference, but also the MCMC method for the same model based on partial likelihood, with appropriate comparisons between these 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	Professor	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), and we have made significant changes both in the proposed methodology and in its presentation in the revised manuscript.

In the revised manuscript, we 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 major methodological improvements to the proposed method in the previous manuscript, which significantly improve its flexibility and computational efficiency, and hence enhances the practical utility of the proposed approach. These improvements include the followings:
1. Removing the Gaussian noise in the additive linear predictor to reduce the computational load to compute, store or factorize the Hessian matrix (section 3.1, page 6-8).
2. Adopting the automatic differentiation method for the computation of both the Laplace approximation and its derivatives required in the use of Adaptive Gauss-Hermite Quadrature (section 3.2, page 8-9).

We also further illustrated the practical utility of the proposed approach compared to existing methods, in the real data example where the proposed approach accommodates the inclusion of spatial variations observed at all sampled locations with a fully dense precision matrix (section 4.3, page 13-15). For real data example, we compared the computational efficiency of the improved approach with the previous approach, demonstrating the necessity of our methodological improvements in the revised manuscript (last paragraph in section 4.3, page 15).

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: We understand the point mentioned by the reviewer, and thanks the reviewer for the suggestion. 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 (first paragraph in section 2.1, page 3).

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 as well as similar methods based on MCMC (second paragraph in section 1, page 2).

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 (section 4.1, page 9-12). 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 (section 4.1, page 9-12) and in the real data analyses (section 4.2, page 12-13).

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 (section 4.1, page 9-12). Also, as the referee has mentioned, we have now provided the simulation results under various settings: settings where the smoothness of baseline hazard is at different levels (section 4.1.2, page 11-12), and settings where the sparsity of frailties is at different levels (section 4.1.1, page 9-10). The advantages of the proposed method are shown both in the smaller MSE and in the posterior coverage rates closer to the nominal levels.