Author’s response to reviews

Title: Dynamic prediction of repeated events data based on landmarking model: application to colorectal liver metastases data

Authors:

Isao Yokota (isao-yokota@umin.ac.jp)
Yutaka Matsuyama (matuyama@epistat.m.u-tokyo.ac.jp)

Version: 2 Date: 13 Dec 2018

Author’s response to reviews:

Line 119: Clarify what "terminal event issue" is. This paragraph should have the following clearly stated:

(a) What exactly is the issue that DPOs addresses
(b) What is the motivation of using them in the repeated events context and how do they address the issue described
(c) What benefit do they provide over existing methods

We appreciate your constructive suggestion. Although we have written the content which you stated (a) and (b) in the third paragraph of introduction session, the ambiguous expression “provides a solution to the terminal event issue” made this content difficult. So we modified the expression. Following that, the content that you stated in (c) was added.

Old: So our proposal is regarded as an extension of DPOs for semi-competing risks settings, and that provides a solution to the terminal event issue.

Revised: So our proposal is regarded as an extension of DPOs for semi-competing risks settings, and that provides a dynamic prediction method in the presence of a terminal event. It is unnecessary to specify the correlation between repeated events and a terminal event, or model hazard function addressed by existing methods mainly.

Line 399: The phrase "predicted probabilities wiggly" is awkward.
We apologize for the strange grammar. We modified them as below;

, whereas the supermodel of the marginal Cox modeling approach returns wiggly function of predicted probabilities against landmark time ***

>Line 401: Is it possible to provide an example of when bias and efficiency are superior? Otherwise it is not convincing that the "wiggly" behavior is a disadvantage and that the smooth behavior of the DPOs is preferable.

You have raised an important insight. We re-considered that the bias in a prediction model was nonsense. In a context of the prediction model, it was difficult to specify the correct model, which meant a pair of covariates and these functional form was accurate. Rather a well-calibrated model for each patient or an accurate prediction for future patients is important. So we delete this sentence. (But the bias of DPOs should be evaluated because the regression model was not established if DPO had a significant bias.) On the other hand, there is a reason that wiggly curve against time was undesirable. It is hard to suppose and interpret that the predicted probabilities repeat increasing and decreasing within a short-time span. (Of course, predicted probabilities might repeat increasing and decreasing over time within a long-time span.) So we should get the curve of predicted probabilities smoothly.

Old: Predicted probabilities from supermodel of DPO is smooth function against time, however, whether bias and efficiency are superior depends on the situation.

Revised: Since it is hard to think that the predicted probabilities repeat increasing and decreasing over time within a short time-span, it would be better to get the smooth curves for interpretation.

>Line 404-405: Clarify "details of hazards" and "model misspecifications". Are you referring to deviations from the proportional hazards assumption as mentioned in the beginning of the paragraph or are there other situations? Was model misspecification considered in the simulation study?

Thank you for your comments to clarify our manuscript. We would like to express that our methods do not restrict any links in GLM, so we modified this sentence as below;
Old: Because details of hazards within \((s, s + w)\) are not required for predictions at only one-time point \(s + w\), our methods should predict more accurately and more robustly for model misspecifications.

Revised: Although a prediction model based hazard function is an analog of generalized linear model linked with complementary log-log function, our methods do not restrict any link function such as log, logit or probit.

>Line 420-421: Bias-variance trade-off is not clear here. It seems possible that with the smoothers the model is actually borrowing information from other time points.

Thank you for your comments. We revised them clearly.

Old: Such a bias-variance trade-off is worth investigating, ***

Revised: The number of landmarking time and the interval of landmarking on a supermodel are worth investigating,