`casebase`: **An Alternative Framework For Survival Analysis and Comparison of Event Rates** by Sahir Rai Bhatnagar*, Maxime Turgeon*, Jesse Islam, James A Hanley and Olli Saarela. (*joint co-authors)

Response to the reviewers[1]

We thank the reviewers for their additional comments. In this document, we reproduce the reviewers comments, and provide our response to each of them below. For a quick reference, here are the main changes to the manuscript:

- We added a paragraph after Figure 2 discussing how the results for the time-dependent hazard ratio suggest a more parsimonious model, and how such a model could be fitted within the framework of case-base sampling.
- We clarified the notation in two places as suggested by one of the reviewers.

Reviewer 1

Overall, the abstract and manuscript is much improved and is now focused on the casebase package. The authors have addressed many of my concerns, and the addition of the confidence band in Figure 4 is a good addition. They have done a better job describing what they are showing, so although I'm not enthusiastic about the approach and would not use it, it probably meetings some need somewhere. The comments below more illustrate where I'm still struggling.

Reviewer Point P 1.1 — P2.3 and Figure R2. The smoothed HR could be informative, however I'm not sure if I believe (or could sell to investigators) that the HR really changes from 1.25 at year 1 to 0.75 at year 2, back to 1.0 at year 6, then decreasing thereafter. Plotting cox.zph results, it also shows that the HR decreases starting at about 6-7 years. It seems like you are modifying 2 things at once (effect of the covariate over time and change in baseline hazard over time), making it perhaps more difficult to tease out what is going on.

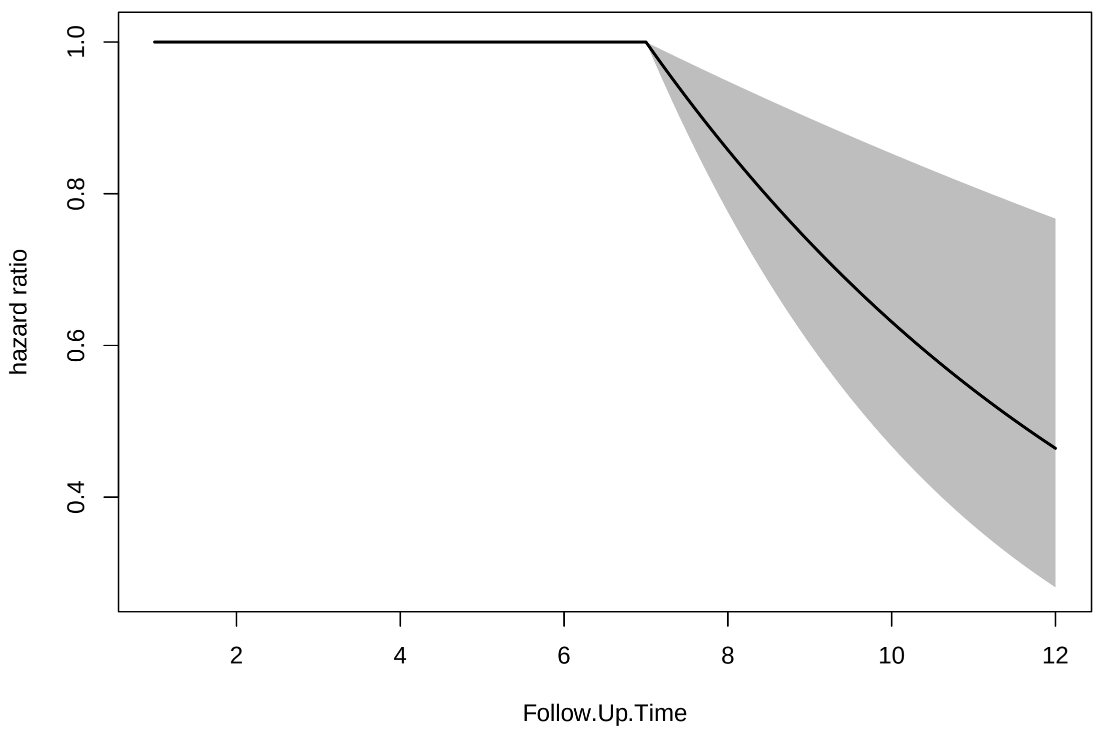
Reply: We wholeheartedly agree with the reviewer that a time-dependent hazard ratio with so much variability is hard to believe. Indeed, based on the results shown in Figure R2 (both the point-wise estimates and the confidence bands) and the principle of parsimony, we would be more inclined to investigate a model where the hazard ratio is constant until year $t_0 = 7$ (approximately) and then decreases. We could even estimate the breakpoint within the framework of case-base sampling using segmented regression. We want to emphasize that such a model would still be more biologically meaningful (and perhaps easier to sell to investigators) than a model with a piecewise constant hazard ratio. The purpose of Figure R2 (and Figure 2 in the manuscript) was to illustrate the flexibility of the `casebase` package, rather than to provide a complete and definitive analysis of the ERSPC dataset. In the manuscript, we added a paragraph discussing this more parsimonious approach along with the following code to fit such the fixed-breakpoint model using out `casebase` package:

[1] July 2, 2022

```
fit <- fitSmoothHazard(DeadOfPrCa ~ as.numeric(Follow.Up.Time >= t0) :
                                    ScrArm : I(Follow.Up.Time - t0),
                      data = ERSPC)
```

The resulting hazard ratio over time can be easily visualized with the plot method:

```
new_time <- seq(1, 12, by  = 0.1)
new_data <- data.frame(ScrArm = factor("Control group",
                                   levels = c("Control group","Screening group")),
                   Follow.Up.Time = new_time)
plot(fit, type = "hr", newdata = new_data,
     var = "ScrArm", xvar = "Follow.Up.Time", ci = TRUE)
```



Reviewer Point P 1.2 — Figure 1 - I find panels C and D in Figure 1 to be confusing. It isn't clear to me why you would want to randomly assign values in that way, and for the naïve user I think this could cause interpretation issues. However, they have at least described what they did. If the goal is to look at censoring, then you could also do the following by group.

```
sfit2 <- survfit(Surv(Follow.Up.Time,DeadOfPrCa==0) ~ ScrArm, data=ERSPC)
plot(sfit2, col=1:2, lty=1:2)
```

Reply: As mentioned in the caption to Figure 1, the case series is randomly redistributed vertically to provide a visualization of the incidence density. A uniform distribution of the case series over the entire study base would be consistent with a constant hazard. Moreover, by visually comparing the incidence densities across strata, we can assess whether the hazards are equal.

Figure 1 does not directly provide a visualization of the censoring process, since the boundary of the grey polygon corresponds to both censored observations and events of interest.

Reviewer 2

Reviewer Point P 2.1 — "First, let $N_i(t) \in \{0, 1\}$"

this notation is inconsistent with the statement further on, where you state that 'the counting process jumps are less than or equal to one'. This suggests that $N_i(t) \in [0, 1]$ or $(0, 1]$. Which is it?

Reply: The notation is not inconsistent, but the clarity of our explanation could be improved. What we mean is that the jumps $dN_i(t)$ of our counting process (which we assume is absolutely continuous) is either zero or one, and so cannot be a larger positive integer. But since we are restricting ourselves to the setting $N_i(t) \in \{0, 1\}$ (i.e. no recurrence of the event of interest), this clarification is redundant: once the process jumps once, it cannot jump again. Therefore, we removed the corresponding statement from the manuscript.

Reviewer Point P 2.2 — please explain the unusual notation `\stackrel{\theta}{\propto}`

Reply: We thank the reviewer for bringing to our attention that this notation is not standard and perhaps unusual. We clarified in the manuscript that the two expressions on each side of the symbol $\propto$ can be considered equal up to (multiplicative) factors that do not depend on θ. As a consequence, both expressions attain their maximum at the same value of θ.