This paper proposes comparing estimators of discriminative performance of time-to-event models. The validity of the tests is performed through numerical demonstrations. Overall, I think the paper is well-written and I enjoyed reading it. I believe the proposed methods are valuable additions to existing literature and I would recommend if for publication after addressing my few questions/suggestions.

1. Time dependent AUC is used to measure the discriminations. Did you consider other measures such as Brier scores or integrated AUC?
2. Section 2 and Section 3 can be integrated into a single section to make the paper more concrete and readable.
3. In the simulation setting, did you consider other censoring rates
4. Some papers are not cited in the paper but I can see them in the reference lists. Could you confirm it again?
5. Did you also consider Accelerated Failure Time models? It would be interesting if you also consider semiparametric AFT models too.
6. Could you specify why the behavior of nonparametric estimators is consistent with the expectations, but semiparametric estimators show the opposite in Section 4.2.2?
7. What are the criteria for Table 1? More specifically, what are the criteria for unbiased/slightly biased/biased upwards? Variability and out-of-sample behavior too.