Yale SCHOOL OF PUBLIC HEALTH
Biostatistics

DANIEL J. ECK, PHD
Postdoctoral Associate
Department of Biostatistics
Yale School of Public Health
60 College Street
New Haven, CT 06510
daniel.eck@yale.edu
http://campuspress.yale.edu/danieleck/

Reviewer 1 checklist

Comment: Summary: The manuscript approaches this from a different angle than I have seen in the past. The binomial application in Section 4 seems sound and is easy to understand. The part about weighting and sensitivity analysis and generally why this might be the right thing to do, is less clear. The manuscript has several interesting observations. The critiques are interesting though opinionated. I like the concept and it seems potentially appropriate for Chance. The manuscript needs minor revisions and clarifications including possibly shortening or excluding some sections.

Response: Thank you for your interest in this work. The manuscript is shortened, Section 6.4 is removed, and clarity has been added to the weighting and sensitivity analysis. The critiques are still present, though the tone is softened. Additional justification for the weighting has been added.

Required revisions:

Comment: It should be clarified what demographic subgroups are included/excluded at each point in time (from Table 1). The first paragraph in the data section does not discuss race/ethnicity at all, while later paragraphs seem to suggest but are not crystal clear on whether and what adjustments are being made. I found that discussion in general to be long-winded and not necessary.

Response: Thank you for your feedback. The first paragraph is now clearer on race/ethnicity. The entire discussion has been shortened.

Comment: The weighting and sensitivity analysis should likely have further justification as to why this is a reasonable thing to do. Having a half-page appendix on the weighting regime seems inappropriate. Probably incorporate that into the section on this part of the analysis.

Response: The weighting material is now incorporated into the text. Further justification as to why the weighting and sensitivity analysis is a reasonable thing to do has been added.

Comment: I dont see what Section 6.4 really adds over previous sections. Also seems out of place.

Response: Thank you for pointing this out, Section 6.4 is now removed.

Recommendations / questions:

Comment: Page 4, Table 1 – not sure how the cumulative population proportions are relevant. I found myself more interested to know the noncumulative proportion over the timeframe.

Response: We define baseball players from earlier eras to be those that started their MLB careers in the 1950 season or before. Therefore the cumulative proportion in row 8 of Table 1 is one of the most important quantities in the paper.

Comment: Why did the authors stop at Top 25? It might be interesting to add top 100. Does the "trend" continue past the top 25?

Response: We included top 10 and top 25 lists because those lists are easily digestible by fans and these lists are abundant online. I am interested in this query as well. The trend continues and is worse for fWAR, bWAR, and Ranker. The trend is still apparent for the ESPN rankings of the top 100 players, but it is dampened. The chance of extreme event in the ESPN top 100 list is about 1 in 125, which is between chances corresponding to ESPN's top 10 and top 25 lists.