

March 15, 2022

From: Dr. Holmes Finch, Pstat
To: Dr. Seth Leopold
Re: Reanalysis of data from Poon, et al., (2020)

Dear Dr. Leopold,

I am writing this letter to report the findings from my reanalysis of the data used in the study by Dr. Selina Poon and colleagues, and published in the manuscript “Race, but not gender, is associated with admissions into orthopaedic residency programs” in *Clinical Orthopaedics and Related Research* in 2020.

Executive Summary:

- The goals of the analysis were (1) to replicate the initial analysis from the paper of Poon et al. so as to verify the integrity of the data set with which I was provided, (2) to analyze those data within a framework that I believed was most methodologically appropriate, and (3) to compare the main findings from my analysis with the main results reported in the original publication by Poon et al. My analysis was performed using SPSS, version 27.

- Based upon my reading of the manuscript and consideration of the data structure, I believe the most appropriate analytic strategy was a multilevel modeling framework, with individual students at level-1 nested within schools at level-2. This is the strategy that I used to produce the results that I report here.

- The intraclass correlation coefficient associated with medical school in my null model was 0.057, which means that 5.7% of the variance in admission to an orthopedic program was associated with the medical school that the students attended.

- In light of this finding, results from the most-definitive multilevel model I fit (a model that included all of the variables initially explored in the paper by Poon et al., in addition to medical school) did not differ substantively from that reported in the original manuscript; that is, the effect direction did not change and the coefficient sizes did not change by very much. Although the multilevel model is the more-correct analytic approach in this context, the findings as initially reported do not differ substantially from those based upon what I believe to be a more correct statistical model.

- Below, I report the analytic approach and findings in more detail.

Null Model

I first fit a null model with a random intercept term in which there were no independent variables and the dependent variable was admission to an orthopedic residency program. Of the 8,966 individuals used in this analysis, 6,218 were admitted to such a residency. The random effect estimates for the null model appear in Table 1.

Table 1: Random effect estimates for the null model

Effect	Estimate	Standard error	Z	P
Intercept	0.01217	0.002152	5.65	<0.001
Residual	0.2027	0.003055	66.37	<0.001

The intraclass correlation (ICC) for this model was calculated as:

$$ICC = \frac{0.01217}{0.01217+0.2027} = 0.057$$

Therefore, approximately 5.7% of the variance in admission to an orthopedic program was associated with the school that students attended.

Models including independent variables

After fitting the null model and obtaining the ICC, I then fit a series of models to the data including sets of independent variables. These models can be divided into two families: (1) Replication of the analyses in Poon, et al. (2020) which was done in a single level fashion and (2) Replication of the *form* of the two Poon, et al. models including the same independent variables but also including a level-2 structure accounting for medical school. More specifically, model 1 included the variables AOA, STEP1, STEP2, Publications, Volunteer, Work, and Research. Model 2 included each of these variables as well as Woman, Asian, Black, Hispanic, and Other.

The odds ratios and their confidence intervals for these models appear in Table 2. Three sets of results are reported: (1) my replication of Dr. Poon's results, (2) Dr. Poon's results as reported in the 2020 manuscript, and (3) my results based on the multilevel models.

These results demonstrate that I was able to replicate Dr. Poon's results using a single level model. In addition, the results of the multilevel model were very close to those from Dr. Poon's manuscript. Therefore, it appears that the results reported in Poon, et al. (2020) would not have been substantively different from those obtained using were what I believe to be the more statistically correct model been used. This is likely due to the fact that the ICC was relatively low (0.057).

Sincerely,

Holmes Finch, Pstat

Table 2: Odds ratios and confidence intervals by model

Model	AOA	STEP1	STEP2	Publications	Volunteer	Work	Research	Woman	Asian	Black	Hispanic	Other
Odds ratios												
1%	2.12	1.04	1.01	1.04	1.03	0.98	1.02					
P1*	2.12	1.04	1.01	1.04	1.03	0.98	1.02					
2%	2.07	1.04	1.01	1.04	1.03	0.98	1.03	1.02	0.81	0.66	0.44	0.63
P2*	2.07	1.04	1.01	1.04	1.03	0.98	1.03	1.02	0.78	0.63	0.48	0.65
1 ML+	2.57	1.04	1.01	1.04	1.03	0.98	1.00					
2 ML+	2.46	1.04	1.01	1.04	1.03	0.98	1.00	1.05	0.79	0.67	0.53	0.62
Confidence intervals for odds ratios												
1	1.80,2.50	1.03,1.04	1.01,1.02	1.03,1.05	1.01,1.04	0.96,1.01	0.99,1.06					
P1	1.80,2.50	1.03,1.04	1.01,1.02	1.03,1.05	1.01,1.04	0.96,1.01	0.99,1.06					
2	1.75,2.44	1.03,1.04	1.01,1.01	1.03,1.05	1.01,1.04	0.96,1.00	1.00,1.07	0.88,1.18	0.69,0.94	0.53,0.81	0.30,0.65	0.50,0.79
P2	1.75,2.44	1.03,1.04	1.01,1.01	1.03,1.05	1.01,1.04	0.96,1.00	1.00,1.07	0.89,1.18	0.67,0.92	0.51,0.77	0.36,0.65	0.55,0.77
1 ML	2.06,3.20	1.03,1.04	1.01,1.02	1.03,1.06	1.02,1.04	0.96,1.01	0.97,1.04					
2 ML	1.95,3.01	1.03,1.04	1.01,1.02	1.03,1.06	1.02,1.04	0.96,1.00	0.98,1.04	0.93,1.24	0.68,0.92	0.55,0.81	0.30,0.63	0.50,0.77

%Finch replication of Dr. Poon's results

*Dr. Poon's results as reported in the 2020 manuscript

+Multilevel model results