Draft of Statistical Methods Section

Kipp Johnson

July 6th, 2014

*Software*

All statistical analysis was performed in the R statistical software environment (R Foundation for Statistical Computing, Vienna, Austria)[1]. The R graphics package “ggplot2” was used to construct figures [2], the function used to compute Cronbach’s alpha was taken from the R package “psych,” [3] and the R package “dunn.test” was used for computing Dunn’s test [4].

*Question Construction*

The Evidence-Based Practice Attitude Scale (EBPAS) was adapted from [???] in order to assess attitudes towards personalized genomics in a student population. The modified EBPAS scale includes questions intended to measure (1) “openness” to genomic testing concepts and (2) “divergence” from typical genomic testing best practices and attitudes. Responses for each question were scored 1-5 depending upon the degree of agreement with each EBPAS statement (“Not at all”, “To a slight extent”, “To a moderate extent”, “To a great extent”, “To a very great extent”). Responses from the divergence category of questions were reverse-scored.

*Questions Assessing Perception of Genomics Education (?)*

*Questions Assessing Comfort in Genomic Testing Ability and Knowledge (?)*

Eight questions assessing students’ comfort in their genomic testing conceptual knowledge and ability to interpret genomic testing results were added to the survey, with responses coded 1-5, ranging from “Not at all comfortable” to “Very comfortable.”

*Statistical Analysis*

Cronbach’s alpha, a measure of whether purportedly related questions are actually assessing the same concepts, was computed for the two subsections. When Cronbach’s alpha for a scale exceeds 0.7, the scale is considered to consistently assess the same underlying concept. Cronbach’s alpha was computed for the two EBPAS subsections, the four education questions, and the genomic testing comfort and ability questions separately.

Answers for questions 13-25 on the survey were collapsed into binary categories and tested for an association with EBPAS score. Multiple testing was corrected for with the Benjamini-Hochberg false discovery rate (BH-FDR) method at a 5% rate. The sample distribution of EBPAS scores was approximately normal. Each of the four educational questions was separately tested for an association with questions 13-25 in the same fashion. All results reported as significantly significant are p<0.05 after correction for multiple testing.

Each of the four educational questions was tested for an association with sample covariates of age group, gender, medical student year, dual degree program status, and interest in a career involving research using the Kruskal-Wallis (KW) test, a non-parametric test similar to one-way ANOVA. Significant KW test results were further examined with Dunn’s post-hoc test of rank sums, controlling for false discovery rate at the 5% level.

Because scores for comfort in ability to interpret genomic testing and genomic testing knowledge were approximately normally distributed, we elected to use one-way analysis of variance (ANOVA) with post-hoc Tukey honest significant difference tests to examine significant associations between genomic testing ability and genomic testing knowledge with our collected sample covariates. ANOVA p-values were controlled for multiple testing with the BH-FDR method.

1. R Core Team (2015). R: A lang
2. uage and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
3. H. Wickham. ggplot2: elegant graphics for data analysis. Springer New York, 2009.
4. Revelle, W. (2015) psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, http://CRAN.R-project.org/package=psych Version = 1.5.4.
5. Dinno, A. (2015). dunn.test: Dunn's Test of Multiple Comparisons Using Rank Sums. R package version 1.2.4. http://CRAN.R-project.org/package=dunn.test