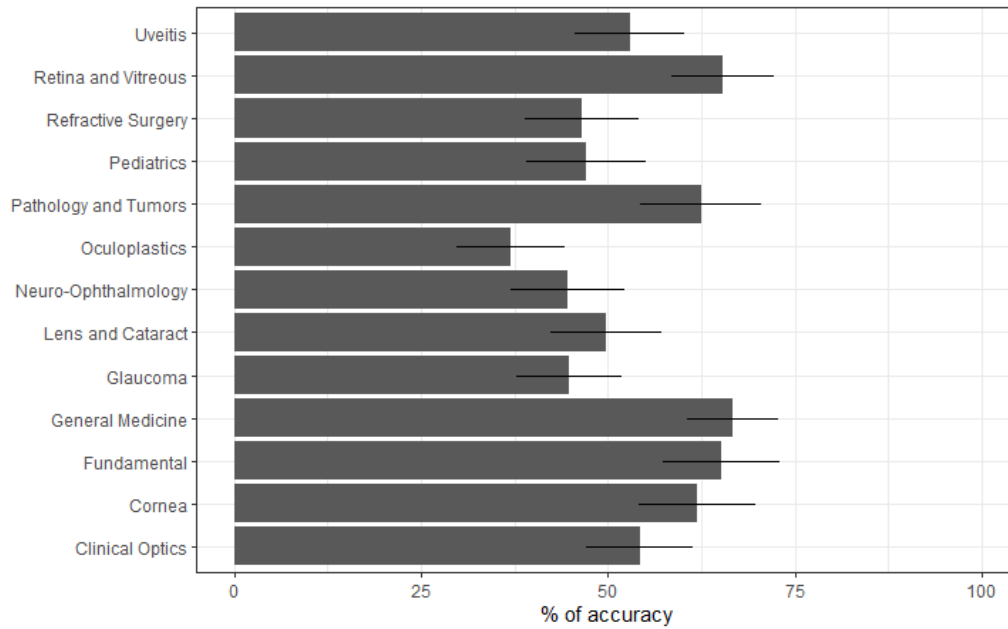


Supplemental Figure 6: Post hoc analysis using Tukey's test to isolate the effect of exam section and difficulty index using all questions for ChatGPT Plus. (A) Bar plot of the percentage of accuracy by exam section. There were no differences in performance across the exam sections. **(B)** Predicted percentage of accuracy by difficulty index. Accuracy increased with increasing difficulty index (easier questions).

(A)



(B)

