Clinical Reading Questions #6

HST 190: Introduction to Biostatistics

These discussion questions are based on Perlis et al. (2026).

1. What were the major goals of the study? Describe the design, the sampling procedure, and the period over which the survey was administered. What is the outcome, how is it scored, and what does a score of 10 or greater indicate?

2. The authors describe this as a *nonprobability* survey and then apply post-stratification survey weights. What problem are the weights meant to solve? Does weighting fully substitute for probability sampling? Which limitation outlined in the discussion speaks directly to this?

3. The exposure — frequency of generative AI use — was recorded on a seven-level ordinal scale but dichotomized to "at least daily" for several analyses. What is gained and what is lost by this dichotomization? Where in the paper do the authors retain the full ordinal scale, and what do you see there that the dichotomy would have concealed?

4. From the Table, 32.0% of participants using AI at least daily had a PHQ-9 score of 10 or greater, versus 25.9% of those using it less often. Does this comparison adjust for any other factors? Using the same Table, name two characteristics that differ sharply between the two groups and explain how each could plausibly distort the presented contrast.

5. In Figure 2, daily use carries an association of $\beta = 1.08$ (95% CI: [0.55–1.62]) and use multiple times a day of $\beta = 0.86$ (95% CI: [0.35–1.37]). State what the units of β are and what a value of 1.08 means.

6. Figure 3 presents three nested regression models: unadjusted, adjusted for sociodemographic variables, and further adjusted for social media use. Compare the estimates across the three. What does the similarity of regression models 2 and 3, together with the reported polychoric correlation of $\rho = 0.00$ between AI use and social media frequency, say about whether social media use is a confounder of the association of interest?

7. In Figure 4, personal AI use is associated with depressive symptoms ($\beta = 0.31$, $p = 0.004$) while work and school use are not. How much weight would you place on this contrast? Consider how many hypothesis tests are reported across Figures 1–4 (and the supplement); is any multiplicity adjustment described?

8. The authors report no interaction between AI use and gender but a significant interaction with age group, then present stratified estimates ($\beta = 1.22$ for ages 25–44, $\beta = 1.38$ for 45–64, and null elsewhere) alongside an odds ratio of 1.29 overall. What does an interaction term mean here? The paper's conclusions call for work "to understand the potential for a causal

relationship"—given the design, list the specific reasons a causal reading may be unwarranted, and describe a study that could support one.

References

Perlis, Roy H, Faith M Gunning, Ata A Uslu, et al. 2026. "Generative AI Use and Depressive Symptoms Among US Adults." *JAMA Network Open* 9 (1): e2554820. https://doi.org/10.1001/jamanetworkopen.2025.54820.