Consistency of GPT–4 annotations by HPO level $\chi^2_{\text{Pearson}}(13) = 22.17, \, \rho = 0.05, \, \widehat{V}_{\text{Cramer}} = 0.03, \, \text{Cl}_{95\%} \, [0.00, \, 0.05], \, n_{\text{obs}} = 7,680$

 $\chi_{\text{Pearson}}(13) = 22.17, p = 0.05, V_{\text{Cramer}} = 0.03, Cl_{95\%} [0.00, 0.05], n_{\text{obs}} = 7,68$

