

Table 1 (P6): Results of reliability and validity testing for the blinding questions.

Question	Reliability (n = 233)			Validity (n = 107)
	Kappa (4 categories)	Weighted kappa (4 categories)	Kappa (2 categories)	Kappa (2 categories)
Blinding of patients	0.71	0.88	0.90	0.96
Blinding of providers	0.68	0.83	0.90	1.0
Blinding of data collectors	0.65	0.76	0.82	0.93
Blinding of outcome adjudicators	0.52	0.60	0.66	0.85
Blinding of data analysts	0.42	0.50	0.43	0.42

For the collapsed 2 categories' analyses, we merged "probably yes" with "definitely yes", and "probably no" with "definitely no". Weighted kappa is based on partial agreement using quadratic weights.