

**Supplementary material, Table 1. Quality Assessment (MERSQI Scale) and Inter-rater Reliability**

	MERSQI scale items		Mean score/ maximum score	K (SE)	CI <i>P</i>	Agreement
Study design	1	Study design	2.2/3.0	1.000 (0.000)	1.000-1.000 < .001	Almost perfect
Sampling	2	No. of institutions studied	0.7/1.5	0.892 (0.132)	0.607-1.000 < .001	Almost perfect
	3	Response rate, %	1.4/1.5	0.534 (0.250)	0.000-1.000 < .001	Moderate
Type of data	4	Type of data	3.0/3.0	1.000 (0.000)	1.000-1.000 < .001	Almost perfect
Validity of evaluation instruments	5	Internal structure	0.4/1.0	0.851 (0.103)	0.624-1.000 < .001	Almost perfect
	6	Content	0.7/1.0	0.822 (0.124)	0.512-1.000 <.001	Almost perfect
	7	Relationship to other variables	0.1/1.0	0.362 (0.293)	0.000-0.926 .015	Fair
Data analysis	8	Appropriateness of analysis	1.0/1.0	1.000 (0.000)	1.000-1.000 < .001	Almost perfect
	9	Complexity of analysis	1.9/2.0	0.362 (0.277)	0.000-0.910 .015	Fair
Outcomes	10	Outcomes	1.5/3.0	1.000 (0.000)	1.000-1.000 < .001	Almost perfect
		<b>Mean score</b>	13.0/18.0	0.782		Substantial

K, Cohen's kappa coefficient.