

Table S5 – Sensitivity analyses for each eligible outcome.

Outcome measure	Main findings			Sensitivity analysis			
	k, n	I ² , P value	SMD (CI 95%)	k, n	I ² , P value	SMD (CI 95%)	Excluded studies (reasons)
Exam scores	40, 4263	94%, p<0.001	2.37 (1.25 to 3.49)	29, 3301	95%, p<0.001	3.06 (1.54 to 4.58)	<i>Control group (CBL vs DL)</i> : Diwan et al. (2017); Panja et al. (2013) (1, 2); Patil et al. (2016) [Tutorial method]; Fortun et al. (2017) [Independent lectures]; Montaldo et al. (2013) [DL + Tutorial with patients]; Willis et al. (2020) [unspecified]; Palappallil et al. (2019) (2) [TBL]; Jamkar et al. (2007) [PBL]
				6, 473	78%, p<0.001	0.90 (0.14 to 1.66)	<i>Control group (CBL vs Tutorial method)</i> : Adiga et al. (2011) (1, 2); Cendan et al. (2011); Chengyi et al. (2017) (1, 2, 3, 4); Ciraj et al. (2010) (1, 2); Grover et al. (2020); Joshi et al. (2014); Kamat et al. (2012) (1,2); Kaur et al. (2020); Latif et al. (2014); Ma et al. (2016); Maas et al. (2018); Nair et al. (2013); Palappallil et al. (2019) (2); Sudhakar et al. (2017) (1, 2); Surapaneni et al. (2010) (1, 2, 3, 4, 5); Tathe et al. (2014); Yang et al. (2021); Zhao et al. (2020) [Didactic lectures]; Fortun et al. (2017) (1, 2) [Independent readings]; Jamkar et al. (2007) [PBL]; Palappallil et al. (2019) (1) [TBL]; Willis et al. (2020) [unspecified]
				31, 2973	91%, p<0.001	1.44 (0.67 to 2.21)	<i>Experimental group</i> : Chengyi et al. (2017) [CBL+PBL]; Maas et al. (2018) [TB+CBL]; Grover et al. (2020) [CBL+whatsapp]; Willis et al. (2020) [CBL]; Yang et al. (2021) [CBL + FC]; Zhao et al. (2020) [CBL+PBL]
				29, 2971	95%, p<0.001	3.07 (1.55 to 4.60)	<i>Type of study: (crossover)</i> : Adiga et al. (2011); Cendan et al. (2011); Diwan et al. (2017); Grover et al. (2020); Kaur et al. (2020); Ma et al. (2016); Sudhakar et al. (2017); Tathe et al. (2014)
				31, 3439	95%, p<0.001	1.99 (0.95 to 3.04)	<i>Without global academic performance</i> : Kamat et al. (2012); Ciraj et al. (2010); Sudakar et al. (2017); Chengyi et al. (2017)

				15, 1591	39%, p=0.06	1.16 (0.99 to 1.32)	<i>Outliers:</i> Surapaneni et al. (2010) (1, 2, 3, 4, 5); Cendan et al. (2011); Jamkar et al. (2007); Chengyi et al. (2017) (1, 2, 3, 4); Ciraj et al. (2010) (1, 2); Diwan et al. (2017) (1); Zhao et al. (2020); Fortun et al. (2017) (2); Kamat et al. (2012) (2); Kaur et al. (2020); Palappallil et al. (2019) (1, 2); Latif et al. (2014); Panja et al. (2013) (1); Ma et al. (2016); Maas et al. (2018); Yang et al. (2021)
Self-learning	4, 654	97%, p<0.001	0.87 (-0.49 to 2.23)	2, 240	0%, p=0.58	0.20 (-0.73 to 1.12)	<i>Experimental group:</i> Grover et al. (2020) [CBL+ Whatsapp]; Zhao et al. (2020) [CBL + PBL]
				2, 414	97%, p<0.001	1.17 (-9.61 to 11.95)	<i>Type of study (crossover):</i> Ma et al. (2016); Grover et al. (2020)
Critical thinking	5, 390	85%, p<0.001	0.96 (-0.33 to 2.26)	3, 476	96%, p<0.001	2.76 (-2.43 to 7.94)	<i>Control group (CBL vs DL):</i> Yang et al. (2021) [CBL + FC]; Chengyi et al. (2017) [CBL + PBL]
Time allotted for classes	4, 565	61%, p=0.05	0.15 (-0.26 to 0.57)	3, 493	73%, p=0.03	0.17 (-0.56 to 0.90)	<i>Control group (CBL vs DL):</i> Palappallil et al. (2019) (2) [TBL]
				3, 395	39%, p=0.20	0.26 (-0.35 to 0.87)	<i>Type of study (crossover):</i> Ma et al. (2016)
Interest and motivation	6, 841	90%, p<0.001	0.79 (0.13 to 1.44)	3, 363	84%, p=0.002	0.36 (-1.01 to 1.72)	<i>Experimental group:</i> Yang et al. (2021) [CBL + FC]; Zhao et al. (2020) [CBL+PBL]; Palappadi et al. (2) [TBL]
				4, 549	37%, p=0.19	1.16 (0.69 to 1.62)	<i>Type of study (crossover):</i> Ma et al. (2016); Hansen et al. (2005)
				5, 769	93%, p<0.001	0.77 (-0.09 to 1.63)	<i>Control group (CBL vs DL):</i> Palappadi et al. (2) [TBL]
Inter-peer interaction in team work	5, 716	96%, p<0.001	1.45 (-0.34 to 3.25)	2, 240	92%, p=0.005	0.38 (-6.13 to 6.88)	<i>Experimental group:</i> Grover et al. (2020) [CBL + Whatsapp]; Hashim et al. (2015) [CBL + Whatsapp]; Yang et al. (2021) [CBL + FC]; Zhao et al. [PBL + CBL]
				3, 476	94%, p<0.001	2.17 (-1.35 to 5.69)	<i>Type of study (crossover):</i> Ma et al. (2016); Grover et al. (2020)

Legend: OSCE – objective structured clinical examination score; k – number of comparisons (each study corresponds to more than one comparison when (i) exists more than 3 groups of intervention or (ii) one group has more than one evaluation group with overlapping); n – total number of individuals analyzed in each group study; I² – heterogeneity; SMD – standardized mean difference values; CI – 95% confidence interval at 95%.

Bold values – statistically significant values.