Table s10. GRADE Evidence Profile of Test Accuracy Results for Prevalence/Pre-Test Probability of 40% and 80% for URT vs LRT Sampling (7 studies)

	Upper respiratory tract sample	**Lower respiratory tract sampling**
Sensitivity	0.57 (95% CI: 0.42 to 0.71)	0.81 (95% CI: 0.73 to 0.86)
Specificity	1.00 (95% CI: 0.99 to 1.00)	1.00 (95% CI: 0.99 to 1.00)

Outcome	**№ of studies (№ of patients)**	**Study design**	**Factors that may decrease certainty of evidence**					**Effect per 1,000 patients tested**				**Test accuracy CoE**
								pre-test probability of 40% [g]		pre-test probability of 80% [h]		
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	URT sampling	LRT sampling	URT sampling	LRT sampling	
True positives (patients with COVID-19)	7 studies 1244 patients	cohort & case-control type studies	serious [a,b]	serious [c]	serious [d]	serious [e]	none	228 (168 to 284)	324 (292 to 344)	456 (336 to 568)	648 (584 to 688)	⊕◯◯◯ VERY LOW
								96 fewer TP in URT sampling		**192 fewer TP in URT sampling**		
False negatives (patients incorrectly classified as not having COVID-19)								172 (116 to 232)	76 (56 to 108)	344 (232 to 464)	152 (112 to 216)	
								96 more FN in URT sampling		**192 more FN in URT sampling**		
True negatives (patients without COVID-19)	1 study 8 patients	cross-sectional (cohort type accuracy study)	not serious	serious [c]	not serious	very serious [f]	none	600 (594 to 600)	600 (594 to 600)	200 (198 to 200)	200 (198 to 200)	⊕◯◯◯ VERY LOW
								0 fewer TN in URT sampling		**0 fewer TN in URT sampling**		
False positives (patients incorrectly								0 (0 to 6)	0 (0 to 6)	0 (0 to 2)	0 (0 to 2)	

classified as having COVID-19)								**0 fewer FP in URT sampling**	**0 fewer FP in URT sampling**	

Explanations: This table is based on applying the sensitivity and specificity estimates to calculate True and false positives and negatives in a hypothetical population of 1000 individuals

a. The case-control design leads to a serious study population bias.
b. The Fengting Y. (2020) study showed results as the number of samples and not the number of patients.
c. There was no direct evidence comparing the accuracy of a strategy with starting with upper sample and then conducting a lower sample if the upper sample is negative. Additionally, studies reported test accuracy results but did not report on patient-important and population-important outcomes based on the results.
d. There is serious unexplained heterogeneity.
e. Considering the upper vs lower limits of the sensitivity's confidence interval would lead to different clinical decisions.
f. A very low number of patients.
g. Typically seen in patients meeting clinical definition for COVID-19 who were hospitalized.
h. Typically seen in patients meeting clinical definition for COVID-19 who were admitted to intensive care units.