

**Table S1.** Relevant ratio based on keywords-based search.

TEXT LEVEL	RELEVANT	TOTAL	EST. RELEVANT RATIO
TITLE	447	529,397	0.1%
ABSTRACTS	1,691	385,194	0.5%
FULL TEXTS	3,427	186,530	2%

**Table S2.** Input processing approach optimization result.

Datasets	Category	Abstract	Summary	Sentences	Sentence (Consensus)
Training	Sample Size	2,561	2,561	20,244	
(Before 2020-12-31)	10-fold CV (ACC)	<b>0.972</b>	0.939	0.890	
Testing	Sample Size	819	819	7,328	819
(After 2020-12-31)	Testing (ACC)	<b>0.940</b>	0.902	0.819	0.921