

**Supplementary Table 3. Subgroup Analysis of RQS according to Publish Date**

	Publications before 2019 (n = 18)	Publications after 2019 (n = 8)	<i>P</i> *
Total 16 items (ideal score 36)	0.8 ± 5.2	8.1 ± 6.3	0.006
Domain 1-protocol quality and stability in image and segmentation (0 to 5 points)	3 (2-3)	3 (2-3)	0.461
Protocol quality (2 points)	2 (1-2)	2 (2-2)	0.567
Test-retest (1 point)	0 (0-0)	0 (0-0)	0.644
Phantom study (1 point)	0 (0-0)	0 (0-0)	0.849
Multiple segmentation (1 point)	1 (0-1)	1 (1-1)	0.338
Domain 2-feature selection and validation (-8 to 8 points)	-2 (-8- -1)	6 (-1-6)	0.019
Feature reduction or adjustment of multiple testing (-3 or 3 points)	-3 (-3-3)	3 (-3-3)	0.238
Validation (-5, 2, 3, 4, or 5 points)	-5 (-5-0)	3 (0-3)	0.035
Domain 3-biologic/clinical validation and utility (0 to 6 points)	0 (0-0)	0 (0-1)	0.367
Non-radiomics features (1 point)	0 (0-0)	0 (0-1)	0.978
Biologic correlations (1 point)	0 (0-0)	0 (0-0)	0.644
Comparison to "gold standard" (2 points)	0 (0-0)	0 (0-1)	0.461
Potential clinical utility (2 points)	0 (0-0)	0 (0-0)	NA
Domain 4-model performance index (0 to 5 points)	1 (0-1)	2 (1-2)	0.062
Cut-off analysis (1 point)	0 (0-0)	0 (0-0)	NA
Discrimination statistics (2 points)	1 (0-1)	2 (1-2)	0.062
Calibration statistics (2 points)	0 (0-0)	0 (0-0)	NA
Domain 5-high level of evidence (0 to 8 points)	0 (0-0)	0 (0-0)	NA
Prospective study (7 points)	0 (0-0)	0 (0-0)	NA
Cost-effectiveness analysis (1 point)	0 (0-0)	0 (0-0)	NA
Domain 6-open science and data (0 to 4 points)	1 (0-1)	1 (1-1)	0.724

\*Calculated from Student *t* test or Mann-Whitney's test according to normality tests.