Supplementary Materials

Table S1. Radiomic quality score criteria

Criteria		Scoring criteria
1	Image protocol	+ 1 (if protocols are well-documented) + 1 (if public protocol is used)
2	Multiple segmentations	+ 1
3	Phantom study on all scanners	+ 1
4	Imaging at multiple time	+ 1
5	Feature reduction or adjustment for multiple testing	– 3 (if neither measure is implemented) + 3 (if either measure is implemented)
6	Multivariable analysis with non-radiomics features	+ 1
7	Detect and discuss biological correlates	+ 1
8	Cut-off analyses	+ 1
9	Discrimination statistics	+ 1 (if a discrimination statistic and its statistical significance are reported) + 1 (if a resampling method technique is also applied)
10	Calibration statistics	+ 1 (if a calibration statistic and its statistical significance are reported) + 1 (if a resampling method technique is also applied)
11	Prospective study	+ 7 (for prospective validation of a radiomics signature in an appropriate trial)
12	Validation	– 5 (if validation is missing) + 2 (if validation is based on a dataset from the same institute) + 3 (if validation is based on a dataset from another institute) + 4 (if validation is based on two datasets from two distinct institutes) + 4 (if the study validates a previously published signature) + 5 (if validation is based on three or more datasets from distinct institutes)
		*Datasets should be of comparable size and should have at least 10 events per model feature
13	Comparison to 'gold standard'	+ 2
14	Potential clinical utility	+ 2
15	Cost-effectiveness analysis	+ 1
16	Open science and data	+ 1 (if scans are open source) + 1 (if region of interest segmentations are open source) + 1 (if code is open source) + 1 (if radiomics features are calculated on a set of representative ROIs and the calculated features and representative ROIs are

ROI, region of interest.