

**Table S5.** Methods used to quantify the performance of the models reported in the 6 articles with development and validation studies and the 25 articles with an external validation

	Development & Validation studies N=6	Stand-alone external validation studies N=25
Discrimination ( <i>ability to differentiate between patients who do or do not experience the event</i> )		
AUC ROC curve with 95%CI	4	16
AUC ROC curve without 95%CI	2	6
Sensitivity & specificity	1	8
Calibration ( <i>agreement between predictions from the model and observed outcomes</i> )		
Hosmer–Lemeshow test	3	9
GiViTI calibration belt		2
Contingency table		1
Calibration histogram	1	4
Calibration curve & statistical tests		4
Global performance ( <i>simultaneous evaluation of calibration and discrimination</i> )		
Accuracy	1	1
Brier score		1
R <sup>2</sup>	1	3

AUC ROC: Area Under the Receiving Operative Curve; 95%CI: 95% Confident Interval