Table 5. Calibration outcome (Calibration slope at 3 months)

Author	Model	N of Variables	N	Death (n)	Calibration slope (95% CI)	Performance	Certainty of the Evidence (Quality of evidence)	Conclusions
Bartels, 2016	Bartels 2011	5	110	90	Until >10 months: 0.64 (0.34-0.94)	Not sufficient	**Moderate**[1]	The Bartels 2011 model likely results in insufficient calibration when predicting time-to-event survival among patients eligible for surgery and/or radiotherapy to treat spinal metastasis.
Bindels, 2025	Bollen 2014	3	953	225/360/474§	At 3 months: 0.76 (0.73-0.80)	Not sufficient	**Moderate**[2]	The Bollen model likely results in insufficient calibration when predicting 3 month survival (yes/no) in patients eligible for treatment of spinal metastasis.
	Mizumoto 2008	7	953	225/360/474§	At 3 months: 0.69 (0.66-0.73)	Not sufficient	**Moderate**[2]	The Mizumoto model likely results in insufficient calibration when predicting 3 months survival (yes/no) among patients eligible for treatment of spinal metastasis.
	Modified Bauer 2008	3	953	225/360/474§	At 3 months: 0.70 (0.66-0.74)	Not sufficient	**Moderate**[2]	The Modified Bauer model likely results in insufficient calibration when predicting 3 month survival (yes/no) among patients eligible for treatment of spinal metastasis.
	NESMS 2015	5	953	225/360/474§	At 3 months: 0.59 (0.55-0.63)	Not sufficient	**Moderate**[2]	The NESMS model likely results in insufficient calibration when predicting 3 months survival (yes/no) among patients eligible for treatment of spinal metastasis.
	Original Bauer 1995	4	953	225/360/474§	At 3 months: 0.66 (0.63-0.70)	Not sufficient	**Moderate**[2]	The Original Bauer model likely results in insufficient calibration when predicting 3 months survival (yes/no) among patients eligible for treatment of spinal metastasis.
	OSRI 2013	2	953	225/360/474§	At 3 months: 0.75 (0.72-0.79)	Not sufficient	**Moderate**[2]	The OSRI model likely results in insufficient calibration when predicting 3 months survival (yes/no) among patients eligible for treatment of spinal metastasis.

PathFx 2020	11	953	225/360/474§	At 3 months: 0.66 (0.62-0.70)	Not sufficient	**Moderate**[2]	The PathFx model likely results in insufficient calibration when predicting 3 months survival (yes/no) among patients eligible for treatment of spinal metastasis.
Revised Katagiri (ten categories) 2014	11	953	225/360/474§	At 3 months: 0.79 (0.76-0.82)	Not sufficient	**Moderate**[2]	The Revised Katagiri (ten categories) model likely results in insufficient calibration when predicting 3 months survival (yes/no) among patients eligible for treatment of spinal metastasis.
Revised Katagiri (3 categories) 2014	11	953	225/360/474§	At 3 months: 0.75 (0.72-0.77)	Not sufficient	**Moderate**[2]	The Revised Katagiri (3 categories) model likely results in insufficient calibration when predicting 3 months survival (yes/no) among patients eligible for treatment of spinal metastasis.
Revised Tokuhashi 2005	6	953	225/360/474§	At 3 months: 0.70 (0.67-0.73)	Not sufficient	**Moderate**[2]	The Revised Tokuhashi model likely results in insufficient calibration when predicting 3 months survival (yes/no) among patients eligible for treatment of spinal metastasis.
SORG-MLA 2019	18	953	225/360/474§	At 3 months: 0.68 (0.63-0.72)	Not sufficient	**Moderate**[2]	The SORG-MLA model likely results in insufficient calibration when predicting 3 months survival (yes/no) among patients eligible for treatment of spinal metastasis.
Tomita 2001	3	953	225/360/474§	At 3 months: 0.64 (0.60-0.68)	Not sufficient	**Moderate**[2]	The Tomita model likely results in insufficient calibration when predicting 3 months survival (yes/no) among patients eligible for treatment of spinal metastasis.
Van der Linden 2005	3	953	225/360/474§	At 3 months: 0.68 (0.65-0.71)	Not sufficient	**Moderate**[2]	The Van der Linden model likely results in insufficient calibration when predicting 3 months survival (yes/no) among patients eligible for treatment of spinal metastasis.