

GRADE tables for review question: How effective are radiological imaging techniques in the diagnosis of spinal metastases, direct malignant infiltration of the spine or associated spinal cord compression?

Table 4: Evidence profile for screening spinal MRI in people at high risk of metastatic spinal cord compression

Quality assessment							No. of patients		Effect		Quality	Importance
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Screening spinal MRI	Control	Relative (95% CI)	Absolute		
Overall survival (event is death from any cause; maximum follow-up 36 months in survivors)												
Dearnaley 2022	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious ¹	none	172/210 (84%)	174/210 (56.9%)	Adj. HR 0.98 (0.79 to 1.21) ³	not estimable	MODERATE	CRITICAL
Neurological and functional status - clinical spinal cord compression (follow-up 24 months)												
Dearnaley 2022	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious ¹	none	19/210 (9.2%)	26/210 (12.6%)	Adj. HR 0.61 (0.35 to 1.08) ³	not estimable	MODERATE	CRITICAL
Neurological and functional status - persistent neurological functional deficit (Frankel score A-D; follow-up 24 months)												
Dearnaley 2022	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ²	none	15/210	23/210	RR 0.73 (0.42 to 1.28)	30 fewer per 1000 (from 64 fewer to 31 more)	LOW	CRITICAL
Quality of life - EQ-5D-5L – health state today (range 0 to 100, higher scores are better; change from baseline to 12 months)												
Dearnaley 2022	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	112	121	not estimable	-1.5 (-5.7 to 2.7)	HIGH	CRITICAL
Pain - Brief Pain Index – severity (range 0 to 10, lower scores are better; change from baseline to 12 months)												
Dearnaley 2022	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	107	111	not estimable	0.4 (-0.2 to 0.9)	HIGH	CRITICAL

Adj: adjusted; CI: confidence interval; HR: hazard ratio; RT: radiotherapy.

1 95% CI crosses 1 MID

2 95% CI crosses 2 MIDs

3 Adjusted for time since development of castration-resistant prostate cancer, time since start of continuous hormone treatment, ECOG performance status (0, 1, and 2), and natural logarithm of PSA concentration.

Table 5: Evidence profile for early MRI referral in people with suspected metastatic spinal cord compression

Quality assessment							No. of patients		Effect		Quality	Importance
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Early MRI referral (MSCC hotline)	Usual care	Relative (95% CI)	Absolute		
Neurological and functional status - ambulant at MSCC diagnosis												
Allan 2009	Cohort study	very serious ¹	no serious inconsistency	no serious indirectness	serious ²	none	34/44 (77%)	175/324 (54%)	RR 1.43 (1.18 to 1.73)	232 more per 1,000 (from 97 more to 394 more)	VERY LOW	CRITICAL
Time to treatment - time from referral to diagnosis of MSCC, days, median (IQR)												
Allan 2009	Cohort study	very serious ¹	no serious inconsistency	no serious indirectness	no serious imprecision	none	N=44 (median 1 day [0 to 21])	N=324 (median 15 days [3 to 66])	Not estimable	Median 14 fewer days (P<0.002)	LOW	IMPORTANT

CI: confidence interval; MSCC: malignant spinal cord compression; RR: risk ratio

1 Very serious risk of bias as per ROBINS-I

2. 95% CI crosses 1 MID

Table 6: Evidence profile: tests for differential diagnosis of malignant and non-malignant vertebral bone marrow lesions

No. of studies	Study design	Sample size	Prevalence of malignant BML (%)	Sensitivity (95% CI)	Specificity (95% CI)	Likelihood ratios (95% CI)	Predictive values (95% CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality	Importance
Diagnostic accuracy of Chemical Shift MRI													
10 ¹	Cohort studies	788	Range 21 to 78	0.90 [0.86–0.94]	0.85 [0.75–0.91]	LR+ 6.22 [3.63–10.30]	PPV 90% (78% to 96%)	Not serious	Not serious	Not serious	Serious ²	MODERATE	CRITICAL
						LR- 0.12 [0.08–0.17]	NPV 89% (83% to 93%)						

BML: bone marrow lesions; CI, confidence interval; LR+, positive likelihood ratio; LR-, negative likelihood ratio; NPV, negative predictive value; PPV, positive predictive value

1 Douis 2016, El-Samie 2015, Kim 2014, Maeder 2018, Perry 2018, Rathore 2017, Schmeel 2021, Shi 2017, Tadros 2016, Taheri 2017

2 95% CI of LR+ crosses 1 default MID (2, 5)

Table 7: Evidence profile: tests for differential diagnosis of malignant and non-malignant vertebral compression fractures

No. of studies	Study design	Sample size	Prevalence of malignant VCF (%)	Sensitivity (95% CI)	Specificity (95% CI)	Likelihood ratios (95% CI)	Predictive values (95% CI)	Risk of bias	Inconsistency	Indirectness	Imprecision ⁹	Quality	Importance
Diagnostic accuracy – FDG-PET or FDG-PET CT													
5 ¹	Cohort studies	274	Range 34 to 71	0.96 [0.82–0.99]	0.77 [0.56–0.89]	LR+ 4.1 [2.1–8.0]	not estimable	Not serious	Serious ²	Not serious	Serious ³	LOW	CRITICAL
						LR- 0.05 [0.01–0.23]	not estimable						
Diagnostic accuracy – chemical Shift MRI													
12 ⁵	Cohort studies	690	Range 28 to 55	0.89 [0.80–0.94]	0.86 [0.81–0.89]	LR+ 6.28 [7.83–26.88]	PPV 84% (78% to 88%)	Not serious	Not serious	Not serious	Not serious	HIGH	CRITICAL
						LR- 0.14 [0.07–0.23]	NPV 92% (85% to 95%)						
Diagnostic accuracy –conventional MRI sequences + contrast enhanced MRI													
4 ⁶	Cohort studies	231	Range 33 to 60	0.89 [0.66–0.97]	0.89 [0.79–0.95]	LR+ 8.85 [3.9–17.70]	PPV 88% (73% to 95%)	Not serious	Not serious	Not serious	Serious ³	MODERATE	CRITICAL
						LR- 0.15 [0.03–0.39]	NPV 92% (76% to 98%)						
Diagnostic accuracy –conventional MRI sequences + diffusion weighted imaging													
11 ⁷	Cohort studies	782	Range 11 to 70	0.84 [0.75–0.90]	0.88 [0.81–0.92]	LR+ 6.85 [4.49–10.20]	PPV 86% (77% to 91%)	Not serious	Not serious	Not serious	Serious ³	MODERATE	CRITICAL
						LR- 0.19 [0.11–0.28]	NPV 89% (83% to 94%)						
Diagnostic accuracy – conventional MRI sequences													
3 ⁷	Cohort studies	221	Range 50 to 77	0.93 [0.87–0.96]	0.88 [0.45–0.99]	LR+ 13.4 [1.69–60.00]	PPV 92% (82% to 97%)	Not serious	Serious ²	Not serious	Very serious ⁸	VERY LOW	CRITICAL
						LR- 0.09 [0.04–0.19]	NPV 89% (73% to 96%)						

CI, confidence interval; LR+, positive likelihood ratio; LR-, negative likelihood ratio; NPV, negative predictive value; PPV, positive predictive value; VCF: vertebral compression fractures

1 Aggarwal 2013, Bredella 2008, Cho 2011, He 2018, Shin 2008 (reported in Kim 2020 systematic review – the results are taken directly from Kim 2020 and were not updated, so there is no Forest plot)

2 Serious heterogeneity unexplained by subgroup analysis

- 3 95% CI of LR+ crosses 1 default MID (2, 5)
- 4 95% CI of LR- crosses 1 default MID (0.2, 0.5)
- 5 Bacher 2021, Eryly 2006, Geith 2012, Kim 2017, Mittal 2016, Ogura 2012, Ovali 2017, Ragab 2009, Schmeel 2018, Zampa 2002, Zidan 2014
- 6 Arvelo-Perez 2015, Geith 2013, Jung 2003, Pongorsop 2009
- 7 Bhugaloo 2006, Biffar 2010, Biffar 2011, Geith 2014, Mubarak 2011, Oztekin 2009, Pozzi 2012, Razek 2019, Sung 2014, Wonglaksanapimon 2012, Zafar 2020
- 7 Kato 2015, Tokuda 2011, Zou 2016
- 8 95% CI of LR+ crosses 2 default MIDs (2, 5)
- 9 Precision estimates based separately on LR+ and LR-

Table 8: Evidence profile: tests for diagnosis of metastatic spinal cord compression

No. of studies	Study design	Sample size	Prevalence of MSCC (%)	Sensitivity (95% CI)	Specificity (95% CI)	Likelihood ratios (95% CI)	Predictive values (95% CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality	Importance
Diagnostic accuracy – plain radiograph plus neurological examination													
Husband 2001	Cohort study	280	72	0.44 [0.37–0.51]	0.98 [0.91–1.00]	LR+ 19.70 [4.96–78.24]	PPV 98% (92% to 99%)	Not serious	NA	Serious ¹	Serious ²	LOW	CRITICAL
						LR- 0.57 [0.50–0.66]	NPV 44% (41% to 47%)				Serious ³	LOW	
Diagnostic accuracy – T1-weighted sagittal MRI images													
Kim 2000	Cohort study	57	23	0.71 [0.59–0.79]	0.97 [0.94–0.98]	LR+ 20.29 [10.87–37.86]	Not estimable	Very serious ²	NA	Not serious	Not serious	LOW	CRITICAL
						LR- 0.31 [0.22–0.42]	Not estimable				Not serious	LOW	

CI: confidence interval; LR+, positive likelihood ratio; LR-, negative likelihood ratio; MSCC: metastatic spinal cord compression; NPV, negative predictive value; PPV, positive predictive value

1 Index test is seriously indirect – composite of X-ray and neurological examination

2 95% CI of LR+ crosses 1 default MID (2, 5)

3 95% CI of LR- crosses 1 default MID (0.2, 0.5)

2 Very serious risk of bias per QUADAS-2

Table 9: Evidence profile: CT guided biopsy of suspected malignant spinal lesions

No. of studies	Study design	Sample size	Sensitivity (95% CI)	Specificity (95% CI)	Diagnostic yield	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality	Importance
Test failure/success: diagnostic yield of CT-guided biopsy (proportion of biopsies providing sufficient material to make diagnosis)											

3 ¹	Cohort study	150	Not reported	Not reported	Median 89% (range 81% to 99%)	Serious ²	Very serious ³	Not serious	Serious ⁴	VERY LOW	IMPORTANT
----------------	--------------	-----	--------------	--------------	-------------------------------	----------------------	---------------------------	-------------	----------------------	----------	-----------

CI: confidence interval; CT: computed tomography

1 Laufer 2009, Phadke 2001, Spinnato 2018

2 Serious risk of bias per QUADAS-2

3 Very serious heterogeneity unexplained by subgroup analysis

4 Sample size < 300