



Development of a whole spinal MRI-based tumor burden scoring method in participants with multiple myeloma: a pilot study of prognostic significance

Sha Cui^{1,2} · Yinnan Guo³ · Jianting Li² · Wenjin Bian¹ · Wenqi Wu² · Wenjia Zhang² · Qian Zheng² · Haonan Guan⁴ · Jun Wang² · Jinliang Niu²

Received: 1 November 2023 / Accepted: 20 January 2024 / Published online: 7 February 2024
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

Abstract

The aim of the study was to develop a new whole spinal MRI-based tumor burden scoring method in participants with newly diagnosed multiple myeloma (MM) and to explore its prognostic significance. We prospectively recruited participants with newly diagnosed MM; performed whole spinal MRI (sagittal FSE T₁WI, sagittal IDEAL T₂WI, and axial FLAIR T₂WI) on them; and collected their clinical data, early treatment response, progression-free survival (PFS), and overall survival (OS). We developed a new tumor burden scoring method according to the extent of bone marrow infiltration in five MRI patterns. All participants were divided into good response and poor response groups after four treatment cycles. Univariate, multivariate analyses, and ROC were used to determine the performance of independent predictors. Thresholds for PFS and OS were calculated using X-tile, and their prognostic significance were assessed by Kaplan–Meier. The Kruskal–Wallis *H* test was used to compare the differences of tumor burden score between the revised International Staging System (R-ISS) stages. The new tumor burden scoring method was used in 62 participants (median score, 12; range, 0–18). The tumor burden score (OR 1.266, *p* = 0.002) was an independent predictor of poor response and the AUC was 0.838. Higher tumor burden scores were associated with shorter PFS (*p* = 0.002) and OS (*p* = 0.011). The tumor burden score was higher in R-ISS-III than in R-ISS-I and R-ISS-II (*p* = 0.016 and *p* = 0.006, respectively). The tumor burden score was an excellent predictor of prognosis and may serve as a supplemental marker for R-ISS.

Keywords Multiple myeloma · Tumor burden · Magnetic resonance imaging · Prognosis · R-ISS

Introduction

Multiple myeloma (MM) is a monoclonal plasma cell proliferative disorder and causes bone marrow infiltration or bone destruction, which is the most prominent feature of

MM, occurring in approximately two-thirds of patients at diagnosis and in nearly all patients during their disease [1, 2]. Quantifying the extent of bone marrow infiltration or bone destruction plays a key role in assessing tumor burden, guiding treatment, and evaluating prognosis [3].

Conventional radiography and computed tomography (CT) can visualize the number and size of bone destruction, but its sensitivity is limited because it cannot show bone marrow infiltration [4]. The limitation is now often complemented by fluorodeoxyglucose (FDG) positron emission tomography (PET)/CT, which contains both tumor morphology and metabolism information. However, FDG PET/CT is expensive, radiative, insensitive to bone marrow infiltration and bone destruction located in the skull or ribs, and has a high false-positive rate and false-negative rate [5–7]. Magnetic resonance imaging (MRI) is highly sensitive for detecting bone marrow infiltration because of the excellent soft-tissue contrast [8]. Moreover, whole-body (WB) MRI

Sha Cui and Yinnan Guo contributed equally to this work.

✉ Jinliang Niu
sxlscjy@163.com

¹ Department of Medical Imaging, Shanxi Medical University, Taiyuan, Shanxi, China

² Department of Radiology, Second Hospital of Shanxi Medical University, 382 Wuyi Road, Taiyuan, Shanxi, China

³ Department of Pain, Fifth Hospital of Shanxi Medical University, Taiyuan, Shanxi, China

⁴ GE Healthcare, MR Research China, Beijing, China

has been proved to have greater sensitivity and specificity in detecting bone marrow infiltration or bone destruction than FDG PET/CT [9]. There are five MRI patterns of bone marrow infiltration in MM: normal, focal, diffuse, combined diffuse and focal, and salt-and-pepper [10]. Previous studies have shown that the tumor burden and prognosis differ among five MRI patterns [11–14]. Subsequently, semi-quantitative tumor burden scoring methods based on MRI pattern began to emerge in MM. However, previous scoring methods had some controversies: (I) they did not cover all five MRI patterns (such as normal and salt-and-pepper patterns) [15], (II) the studies had variational scoring weight for the number and size of focal lesions [15, 16], and (III) scoring weight for diffuse and combined diffuse and focal patterns were not proper and did not correspond to their tumor burden [1, 15].

WB MRI has the disadvantages of long scanning time, high requirements for technology and equipment, and difficult observation of humeral lesions due to limited field of view (FOV). MM lesions are mainly located in the axial skeleton, and the whole spine is the most affected area [17]. The whole spine scan is quick and convenient, and is widely used in clinical practice as an alternative to the WB MRI.

In our study, we try to develop a new, easy-to-implement scoring method for all five MRI patterns on whole spine scanning. We explored the prognostic significance of the new tumor burden score by evaluating its role in predicting the early treatment response and its association with progression-free survival (PFS), overall survival (OS), and the revised International Staging System (R-ISS) stage.

Materials and methods

Participant cohort

We prospectively recruited participants with newly diagnosed MM who were determined by the International Myeloma Working Group (IMWG) criteria from August 2020 to October 2022, had not received any treatment, and could tolerate MRI examination [18]. Exclusion criteria were participants who had any other diseases in the whole spine and had not

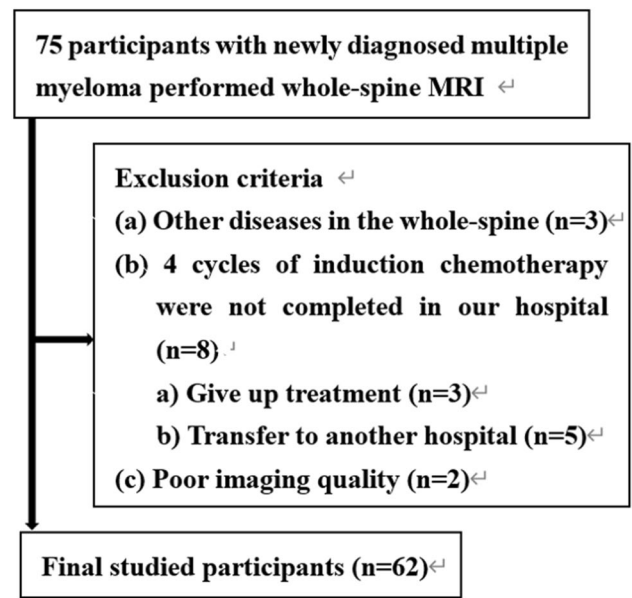


Fig. 1 Participant inclusion and exclusion flowchart

completed the four cycles of induction chemotherapy in our hospital or inferior quality of MRI images (Fig. 1). Clinical data and laboratory test results including sex, age, hemoglobin, platelet, serum albumin, serum β_2 -microglobulin, serum lactate dehydrogenase (LDH), serum calcium, serum c-reactive protein (CRP), bone marrow plasma cell (BMPC) percentage, and flow cytometry of bone marrow cells were collected at the first diagnosis. Before treatment, all participants were divided into three groups based on the R-ISS stage, and then all participants were treated with one of the following first-line induction regimens: bortezomib, lenalidomide, dexamethasone ($n = 43$); bortezomib, dexamethasone ($n = 9$); bortezomib, thalidomide, dexamethasone ($n = 6$); bortezomib, cyclophosphamide, dexamethasone ($n = 4$). Treatment response was evaluated according to the IMWG response criteria after the completion of four cycles of induction therapy [19]. The treatment response categories include stringent complete remission (sCR), complete response (CR), very good partial response (VGPR), partial response (PR),

Table 1 Bone marrow infiltration patterns and scoring method in each anatomic site

Bone marrow infiltration patterns	Score for number of lesions			Score for size of the largest lesion		
	0 (0)	1 (1)	2 (2–9)	3 (≥ 10)	2 (5–15 mm)	3 (> 15 mm)
Focal				3 (≥ 10)		
Salt and pepper				3 (≥ 10)	1 (<5 mm)	
Diffuse				3 (≥ 10) *		3 (> 15 mm) *
Combined diffuse and focal				3 (≥ 10) *		3 (> 15 mm) *
Normal	0 (0) ‡					

*For the diffuse and combined diffuse and focal infiltration patterns, a score of 6 for each anatomic site was equal to the focal infiltration pattern with ≥ 10 lesions and the largest lesion size > 15 mm

‡For the normal pattern, each anatomic site received a score of 0 due to the lack of macroscopic lesions

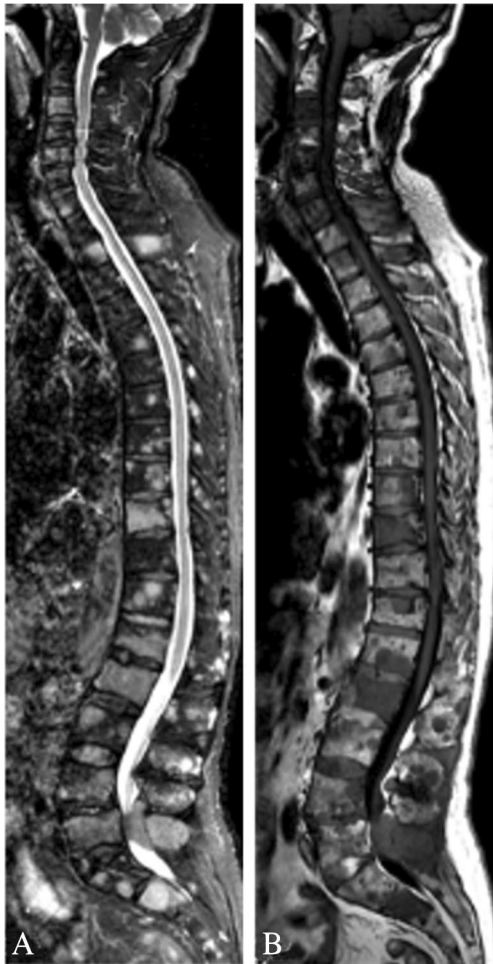


Fig. 2 Exemplary images of MR focal pattern. **A** Whole spine sagittal T₂WI IDEAL and **B** whole spine sagittal T₁WI. IDEAL, iterative decomposition of asymmetric echoes

stable disease (SD), and progressive disease (PD). All participants were categorized into two groups: good response group (sCR, CR, and VGPR status) and poor response group (PR, SD, and PD status). The PFS and OS were also recorded, with PFS defined as the time from MRI scan to progression or death or December 2023, and OS defined as the time from MRI scan to death or December 2023. The median follow-up was 24 months (range, 1–37 months).

MRI examination

All whole-spine scanning were performed on a 3.0-T MRI scanner (Signa Pioneer, GE Healthcare, Milwaukee, WI, USA) with a 16-channel head-and-neck coil, 64-channel spine coil, and 16-channel phased-array body coil before treatment. The whole-spine sequences included cervical, thoracic, and lumbar sagittal fast spin-echo (FSE) T₁WI, sagittal iterative decomposition of asymmetric echoes



Fig. 3 Exemplary images of MR focal pattern. **A** Whole spine sagittal T₂WI IDEAL and **B** whole spine sagittal T₁WI. IDEAL, iterative decomposition of asymmetric echoes

(IDEAL) T₂WI, and axial fluid-attenuated inversion recovery (FLAIR) T₂WI.

Two radiologists (J.N. and S.C, with 13 and 7 years of experience in musculoskeletal imaging, respectively) were responsible for the evaluation of tumor burden score of the whole spine. The average value from the two observers was used for statistical analysis.

Evaluation of tumor burden score

The whole spine consists of three anatomic sites: cervical, thoracic, and lumbar spine. The scoring method of each anatomic site was developed based on the pattern and extent of bone marrow infiltration. The total tumor burden score was the sum of three anatomic sites. The diagnostic

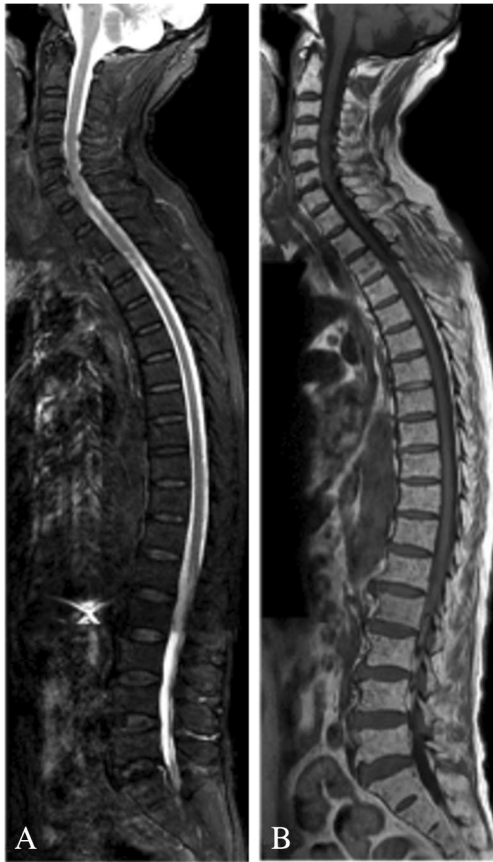


Fig. 4 Exemplary images of MR salt-and-pepper pattern. **A** Whole spine sagittal T₂WI IDEAL and **B** whole spine sagittal T₁WI. IDEAL, iterative decomposition of asymmetric echoes

criteria for myeloma lesions were that the signal intensity was lower than that of intervertebral disc or muscle on T₁WI and higher than that of muscle on T₂WI [10].

The scoring method for each anatomic site in five MRI patterns was as follows (Table 1): (1) focal: circumscriptive lesions with diameter at least 5 mm [10]. The number of lesions ≥ 10 was scored as 3, 2–9 was scored as 2, 1 was scored as 1, and 0 was scored as 0 (see exemplary MRI in Figs. 2 and 3) [15]. The size of the largest lesion > 15 mm was scored as 3 and 5–15 mm was scored as 2 (see exemplary MRI in Figs. 2 and 3) [15]. (2) Salt-and-pepper: widespread tiny nodular infiltrates, diameter of less than 5 mm, with preserved normal bone marrow between them [10, 20]. The size of the largest lesion < 5 mm was scored as 1 [15]. We thought the number of lesions of this pattern was at least 10, with a score of 3, and a score of 1 for the largest lesion size, so each anatomic site received a score of 4 (see exemplary MRI in Fig. 4). (3) Diffuse, homogenous infiltration [10]: we considered the tumor burden of this pattern to be equivalent to that of the focal pattern when the

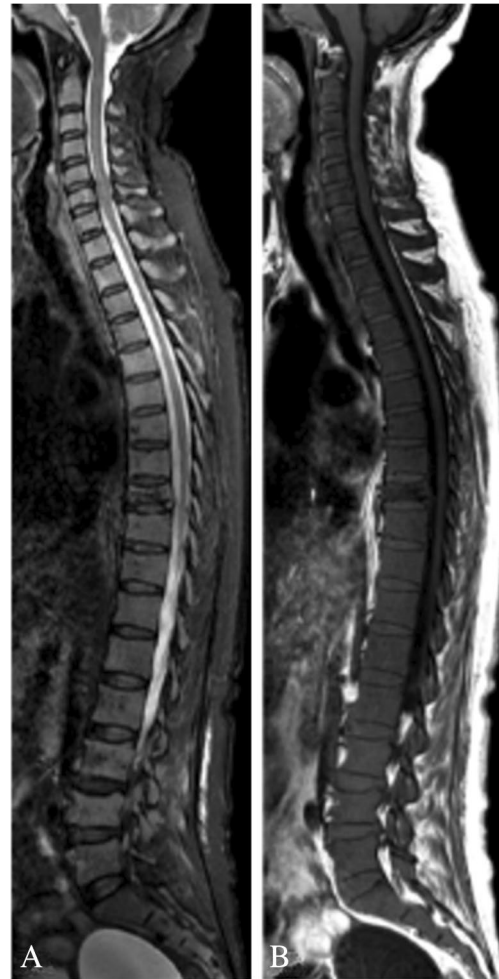


Fig. 5 Exemplary images of MR diffuse pattern. **A** Whole spine sagittal T₂WI IDEAL and **B** whole spine sagittal T₁WI. IDEAL, iterative decomposition of asymmetric echoes

number of lesions ≥ 10 and the largest lesion size > 15 mm, so the score for each anatomic site was 6 (see exemplary MRI in Fig. 5). (4) Combined diffuse and focal [10]: the score of this pattern was same as the diffuse pattern (see exemplary MRI in Fig. 6). (5) Normal, similar in appearance to normal adult bone marrow [10]: we considered this pattern to be no macroscopic lesions, so the score for each anatomic site was 0 (see exemplary MRI in Fig. 7).

Statistical analysis

For univariate analysis, the clinical and MRI variables were compared between two groups using the Mann–Whitney *U* test for continuous variables and chi-square tests or Fisher's exact tests for categorical variables. Variables with a *P*-value < 0.05 in univariate analyses were included as covariates in multivariate (logistic regression) analyses to identify independent predictors, and their predictive



Fig. 6 Exemplary images of MR combined diffuse and focal and diffuse pattern. **A** Whole spine sagittal T₂WI IDEAL and **B** whole spine sagittal T₁WI. IDEAL, iterative decomposition of asymmetric echoes

performance were assessed by receiver operating characteristics (ROC) curve analyses. Thresholds for PFS and OS for independent predictors were calculated using X-tile, and their prognostic significance were assessed using Kaplan–Meier analyses. The differences of tumor burden score at three R-ISS stages were compared by Kruskal–Wallis *H* test, and an adjusted *P*-value by Bonferroni's correction was used for post-hoc tests. Statistical significance was defined as *P* < 0.05. All statistical analyses were performed by using software (SPSS version 22.0; SPSS, Chicago, IL).

Results

Participant cohort

A total of 75 participants were initially included in this study. Three participants had other diseases in the spine,



Fig. 7 Exemplary images of MR normal pattern. **A** Whole spine sagittal T₂WI IDEAL and **B** whole spine sagittal T₁WI. IDEAL, iterative decomposition of asymmetric echoes

eight participants had not completed the four cycles of induction chemotherapy in our hospital, and two participants had poor imaging quality. Ultimately, we collected 62 participants. Among them 37 participants received good response and 25 received poor response after four cycles of induction chemotherapy (Table 2).

There were 29 participants with focal MRI pattern (median score, 10; range, 3–18), 7 participants with diffuse pattern (median score, 18; range, 12–18), 8 participants with salt-and-pepper pattern (median score, 8; range, 4–12), 9 participants with normal pattern (median score, 0; range, 0–0), and 9 participants with combined focal and diffuse pattern (median score, 18; range, 12–18).

Comparison of clinical and MRI characteristics between two groups

Baseline characteristics for the two groups are shown in Table 2. Poor response group and good response group

Table 2 Participant characteristics and group differences

Variables	Participants with VGPR or better (<i>n</i> = 37)	Participants with less than VGPR (<i>n</i> = 25)	<i>p</i> -value
Sex			
Male	25	17	0.971 ^a
Female	12	8	
Age (year)	63 (47,85)	64 (50,78)	0.275 ^b
Hemoglobin (g/L)	99 (42,153)	90 (42,141)	0.139 ^b
Platelet counts (10 ⁹ /L)	172.0 (98.5,262.5)	156.0 (81.5,217.5)	0.532 ^b
β ₂ -microglobulin (mg/L)	3.41 (2.87,7.02)	6.91 (4.50,11.24)	0.004 ^b
Albumin (g/L)	33.6 (29.2,40.5)	30.9 (27.6,36.9)	0.199 ^b
LDH (U/L)	182.0 (135.5,233.5)	180.0 (136.0,229.5)	0.920 ^b
Calcium (mmol/L)	2.32 (2.18,2.42)	2.28 (2.17,2.79)	0.509 ^b
Creatinine (umol/L)	73.0 (58.0,94.5)	114.0 (67.5,281.5)	0.012 ^b
BMPC (%)	35.0 (20.0,60.5)	55.0 (34.7,67.8)	0.066 ^b
CRP (mg/L)	3.13 (3.13,6.22)	4.18 (3.13,10.82)	0.168 ^b
Tumor burden score	6.5 (3,12)	16 (14,18)	<0.001 ^b
Induction regimen			0.860 ^c
VRD	25	18	
PD	5	4	
VTD	4	2	
BCD	3	1	
R-ISS stage			0.010 ^a
I	9	4	
II	25	11	
III	3	10	

Data expressed as median (interquartile range). ^aThe statistical method used chi-square test. ^bThe statistical method used Mann–Whitney *U* tests. ^cThe statistical method used Fisher's exact tests

VGPR, very good partial response; LDH, lactate dehydrogenase; BMPC, bone marrow plasma cell; CRP, C-reactive protein; VRD, bortezomib, lenalidomide, dexamethasone; PD, bortezomib, dexamethasone; VTD, bortezomib, thalidomide, dexamethasone; BCD, bortezomib, cyclophosphamide, dexamethasone; R-ISS, revised International Staging System

had similar clinical characteristics at diagnosis, except for higher β₂-microglobulin in poor response group (6.91 mg/L vs. 3.41 mg/L; *p* = 0.004), higher creatinine in poor response group (114.0 umol/L vs. 73.0 umol/L; *p* = 0.012), and higher tumor burden score in poor response group (16 vs. 6.5; *p* < 0.001). The R-ISS stages (*p* = 0.010) were different between the two response groups (Table 2).

Regression analyses for the prediction of poor response

Logistic regression analyses indicated that tumor burden score (odds ratio, 1.266; 95% CI, 1.094–1.464; *p* = 0.002) (Table 3) was correlated with poor response, but β₂-microglobulin, creatinine, and R-ISS stage had no correlation with the poor response (Table 3).

Diagnostic performance for prediction of early treatment response

In ROC analysis, the tumor burden score had sensitivity of 96.0%, specificity of 78.0%, and AUC of 0.838 (Fig. 8).

Tumor burden score for predicting prognosis of MM

By using X-tile, the cut-offs of tumor burden score for PFS and OS were both 12. Kaplan–Meier survival curves showed that higher tumor burden scores were associated with shorter PFS (*p* = 0.002, Fig. 9A) and OS (*p* = 0.011, Fig. 9B).

Table 3 OR of poor response

Variables	OR (95%CI)	<i>p</i> -value
β ₂ -microglobulin	0.987 (0.789,1.235)	0.909
Creatinine	1.005 (0.998,1.013)	0.178
Tumor burden score	1.266 (1.094,1.464)	0.002
R-ISS (II vs. I)	0.390 (0.031,4.964)	0.468
R-ISS (III vs. I)	0.258 (0.032,2.097)	0.205

OR, odds ratio; CI, confidence interval

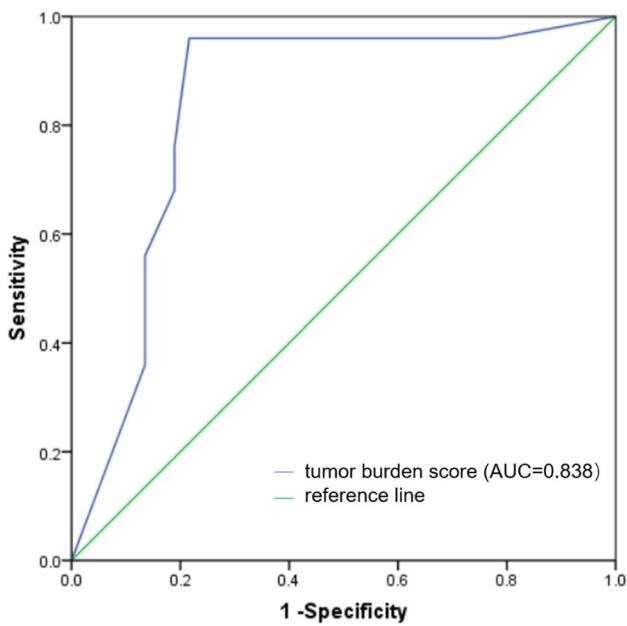


Fig. 8 The ROC curve of tumor burden score for the diagnosis of poor early treatment response, and the AUC was 0.838. ROC, receiver operating characteristic; AUC, area under the curve

Comparison of tumor burden score between different R-ISS stages

The results showed that the tumor burden score was significantly different among R-ISS stages, and after pairwise comparisons within the subgroups of R-ISS stage, the tumor burden scores in

stage I and II were lower than that in stage III. However, there was no significant difference between stage I and stage II (Table 4).

Discussion

The tumor burden based on bone marrow infiltration plays a key role in evaluating prognosis in MM [3]. We developed a new tumor burden scoring method according to the extent of bone marrow infiltration from whole spine. Our results showed that the tumor burden score was an independent predictor of early treatment response to newly diagnosed MM with a good diagnostic performance and associated with PFS and OS. In addition, the tumor burden scores in R-ISS-I and R-ISS-II stage were lower than in R-ISS-III stage.

Bone marrow infiltration is the feature that reflects the pathogenesis of MM. Earlier studies had described a variety of semi-quantitative scoring methods to reflect the extent of bone marrow infiltration in MM, but there was no uniform standard. For example, the scoring method of Dong H. et al., firstly, only focused on the focal, diffuse, and combined diffuse and focal MRI patterns. Secondly, a score of 1 per anatomic site for diffuse pattern is considered to underestimate the tumor burden in this pattern, as in this pattern the total amount of plasma cells is large, and this pattern has also been shown to be associated with poor cytogenetics, advanced disease, and worse prognosis [21, 22]. Thirdly, for the combined diffuse and focal pattern, we thought the combined scores of focal and diffuse infiltrations overestimate

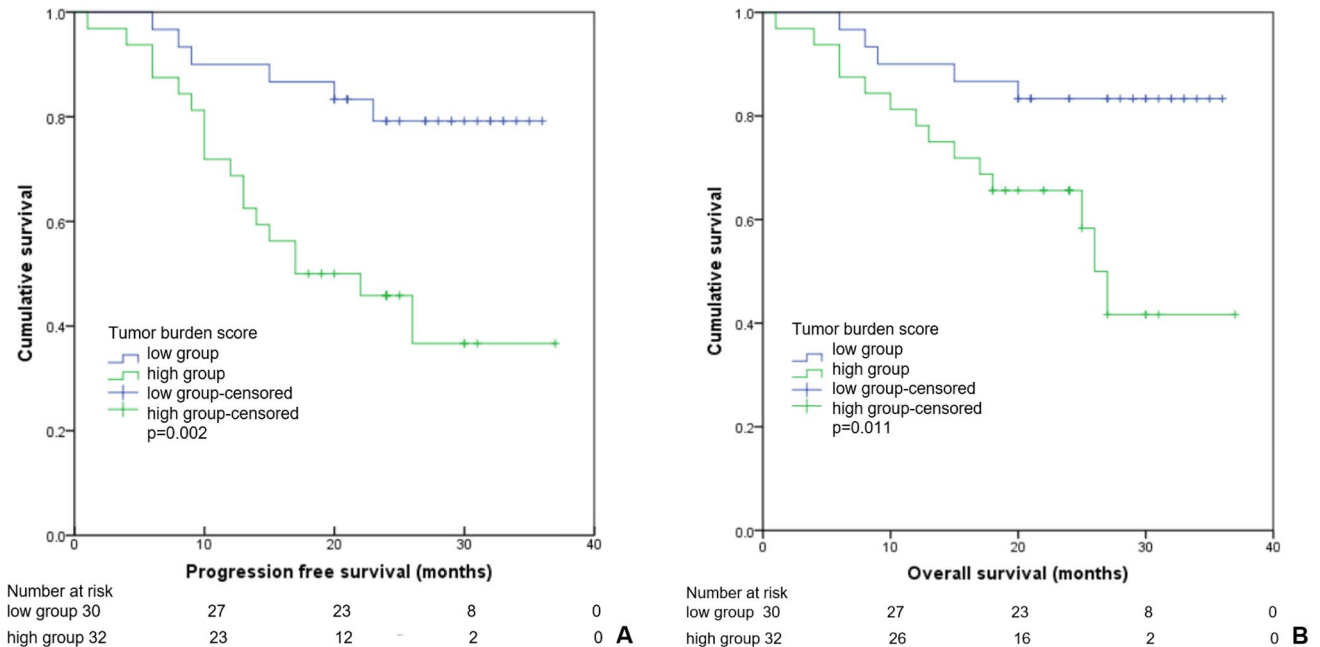


Fig. 9 **A** Kaplan–Meier curves illustrating progression free survival differences between participants with low and high tumor burden score. **B** Kaplan–Meier curves illustrating overall survival differences between participants with low and high tumor burden scores

Table 4 Tumor burden score in different R-ISS stages

	R-ISS-I	R-ISS-II	R-ISS-III	p^*	$p^\#$	p^\ddagger
Tumor burden score	8 (5,13)	11 (3,15)	18 (12,18)	0.864	0.016	0.006

Data expressed as median (interquartile range). *Pairwise comparison of R-ISS-I and R-ISS-II. #Pairwise comparison of R-ISS-I and R-ISS-III. ‡Pairwise comparison of R-ISS-II and R-ISS-III. An adjusted $p < 0.017$ regarded as significant

R-ISS, revised International Staging System

the tumor burden in this pattern. In our study, we developed scoring methods for normal and salt-and-pepper patterns and adjusted the scoring methods for diffuse and combined diffuse and focal patterns. Therefore, our tumor burden score, which developed on the basis of previous studies and considered the characteristics of five MRI patterns, was a potential semi-quantitative tool to fully reflect tumor burden.

To further evaluate the prognostic values of tumor burden score in MM participants, we analyzed the correlation between tumor burden score and R-ISS stages. The R-ISS combines the International Staging System (ISS) with high-risk cytogenetic abnormalities [deletion (17p), translocation t (4;14), or t (14;16)] and serum lactate dehydrogenase (LDH) level [23, 24]. The ISS is based on the serum β_2 -microglobulin and albumin levels, and identifies three different prognosis patients: median OS of 62 months in ISS stage I, 44 months in ISS stage II, and 29 months in ISS stage III groups [23]. The high-risk cytogenetic abnormalities and LDH denote increased disease aggressiveness and a high rate of tumor proliferation. The R-ISS was validated in independent cohorts to be more potent than the ISS [25–29]. Our results showed that with the progression of R-ISS stage, the tumor burden score tended to increase, and the tumor burden score of stage I and stage II was significantly lower than that of stage III, indicating that the tumor burden score could identify the patients with poor prognosis. However, there was no significant difference in tumor burden scores between stage I and stage II, which may be related to the small sample size. We speculate that the tumor burden score, which reflects the radiological tumor burden, may serve as a supplemental marker for R-ISS.

There were several limitations in the current study. First, it was a small sample size study. Studies with large samples are needed to confirm these findings. Second, due to the long acquisition time of WB MRI, our study only applied it to the whole spine. In the future, the WB MRI can be performed by using deep learning-based image reconstruction of under sampled MRI data to shorten scan times. Deep learning image reconstruction of under-sampled MRI data is a technique that interpolates the missing k-space data points directly with the help of training data sets or finds out a mapping function between the under-sampled and fully sampled MR images to estimate the reconstructed image with better accuracy [30, 31].

In conclusion, the new tumor burden scoring method was applicable to five MRI patterns. The tumor burden score was an excellent predictor of prognosis in newly diagnosed MM and may serve as a supplemental marker for R-ISS.

Author contribution Conception and design of study: Sha Cui, Yinnan Guo, and Jinliang Niu; collection and assembly of data: Sha Cui, Jianting Li, Wenjin Bian, and Wenqi Wu; follow-up of participants: Sha Cui and Yinnan Guo; statistical analysis: Sha Cui and Wenjia Zhang; analysis and interpretation of data: Sha Cui, Qian Zheng, and Haonan Guan; manuscript writing: Sha Cui and Yinnan Guo; critical review: Jinliang Niu and Jun Wang.

Funding This study has received funding by National Natural Science Foundation of China (NSFC 82071898).

Data availability The data that support the findings of this study are available on request from the corresponding author.

Declarations

Ethics approval This study received approval from the ethics committee of the Second Hospital of Shanxi Medical University (protocol number 2020–040).

Consent to participate Informed consent was obtained from all individual participants included in the study.

Conflict of interest The authors declare no competing interests.

References

- Ormond Filho AG, Carneiro BC, Pastore D et al (2019) Whole-body imaging of multiple myeloma: diagnostic criteria. *Radiographics* 39(4):1077–1097. <https://doi.org/10.1148/rg.2019180096>
- Terpos E, Berenson J, Raje N, Roodman GD (2014) Management of bone disease in multiple myeloma. *Expert Rev Hematol* 7(1):113–125. <https://doi.org/10.1586/17474086.2013.874943>
- Dall'Olio FG, Marabelle A, Caramella C et al (2022) Tumor burden and efficacy of immune-checkpoint inhibitors. *Nat Rev Clin Oncol* 19(2):75–90. <https://doi.org/10.1038/s41571-021-00564-3>
- Giles SL, deSouza NM, Collins DJ et al (2015) Assessing myeloma bone disease with whole-body diffusion-weighted imaging: comparison with x-ray skeletal survey by region and relationship with laboratory estimates of disease burden. *Clin Radiol* 70(6):614–621. <https://doi.org/10.1016/j.crad.2015.02.013>
- Shortt CP, Gleeson TG, Breen KA et al (2009) Whole-body MRI versus PET in assessment of multiple myeloma disease activity. *AJR Am J Roentgenol* 192(4):980–986. <https://doi.org/10.2214/AJR.08.1633>

6. Regelink JC, Minnema MC, Terpos E et al (2013) Comparison of modern and conventional imaging techniques in establishing multiple myeloma-related bone disease: a systematic review. *Br J Haematol* 162(1):50–61. <https://doi.org/10.1111/bjh.12346>
7. Cavo M, Terpos E, Nanni C et al (2017) Role of 18F-FDG PET/CT in the diagnosis and management of multiple myeloma and other plasma cell disorders: a consensus statement by the International Myeloma Working Group. *Lancet Oncol* 18(4):e206–e217. [https://doi.org/10.1016/S1470-2045\(17\)30189-4](https://doi.org/10.1016/S1470-2045(17)30189-4)
8. Hillengass J, Landgren O (2013) Challenges and opportunities of novel imaging techniques in monoclonal plasma cell disorders: imaging “early myeloma.” *Leuk Lymphoma* 54(7):1355–1363. <https://doi.org/10.3109/10428194.2012.740559>
9. Lecouvet FE, Boyadzhiev D, Collette L et al (2020) MRI versus 18F-FDG-PET/CT for detecting bone marrow involvement in multiple myeloma: diagnostic performance and clinical relevance. *Eur Radiol* 30(4):1927–1937. <https://doi.org/10.1007/s00330-019-06469-1>
10. Baur-Melnyk A, Buhmann S, Dürr HR, Reiser M (2005) Role of MRI for the diagnosis and prognosis of multiple myeloma. *Eur J Radiol* 55(1):56–63. <https://doi.org/10.1016/j.ejrad.2005.01.017>
11. Mouloupoulos LA, Gika D, Anagnostopoulos A et al (2005) Prognostic significance of magnetic resonance imaging of bone marrow in previously untreated patients with multiple myeloma. *Ann Oncol* 16(11):1824–1828. <https://doi.org/10.1093/annonc/mdi362>
12. Dimopoulos MA, Hillengass J, Usmani S et al (2015) Role of magnetic resonance imaging in the management of patients with multiple myeloma: a consensus statement. *J Clin Oncol* 33(6):657–664. <https://doi.org/10.1200/JCO.2014.57.9961>
13. Dutoit JC, Verstraete KL (2016) MRI in multiple myeloma: a pictorial review of diagnostic and post-treatment findings. *Insights Imaging* 7(4):553–569. <https://doi.org/10.1007/s13244-016-0492-7>
14. Mai EK, Hielscher T, Kloth JK et al (2015) A magnetic resonance imaging-based prognostic scoring system to predict outcome in transplant-eligible patients with multiple myeloma. *Haematologica* 100(6):818–825. <https://doi.org/10.3324/haematol.2015.124115>
15. Dong H, Huang W, Ji X et al (2022) Prediction of early treatment response in multiple myeloma using MY-RADS total burden score, ADC, and fat fraction from whole-body MRI: impact of anemia on predictive performance. *AJR Am J Roentgenol* 218(2):310–319. <https://doi.org/10.2214/AJR.21.26534>
16. Lai AYT, Riddell A, Barwick T et al (2020) Interobserver agreement of whole-body magnetic resonance imaging is superior to whole-body computed tomography for assessing disease burden in patients with multiple myeloma. *Eur Radiol* 30(1):320–327. <https://doi.org/10.1007/s00330-019-06281-x>
17. Silbermann R, Roodman GD (2013) Myeloma bone disease: pathophysiology and management. *J Bone Oncol* 2(2):59–69. <https://doi.org/10.1016/j.jbo.2013.04.001>
18. Rajkumar SV, Dimopoulos MA, Palumbo A et al (2014) International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. *Lancet Oncol* 15(12):e538–e548. [https://doi.org/10.1016/S1470-2045\(14\)70442-5](https://doi.org/10.1016/S1470-2045(14)70442-5)
19. Kumar S, Paiva B, Anderson KC et al (2016) International Myeloma Working Group consensus criteria for response and minimal residual disease assessment in multiple myeloma. *Lancet Oncol* 17(8):e328–e346. [https://doi.org/10.1016/S1470-2045\(16\)30206-6](https://doi.org/10.1016/S1470-2045(16)30206-6)
20. Messiou C, Hillengass J, Delorme S et al (2019) Guidelines for acquisition, interpretation, and reporting of whole-body MRI in myeloma: myeloma response assessment and diagnosis system (MY-RADS). *Radiology* 291(1):5–13. <https://doi.org/10.1148/radiol.2019181949>
21. Mouloupoulos LA, Dimopoulos MA, Christoulas D et al (2010) Diffuse MRI marrow pattern correlates with increased angiogenesis, advanced disease features and poor prognosis in newly diagnosed myeloma treated with novel agents. *Leukemia* 24(6):1206–1212. <https://doi.org/10.1038/leu.2010.70>
22. Mouloupoulos LA, Dimopoulos MA, Kastritis E et al (2012) Diffuse pattern of bone marrow involvement on magnetic resonance imaging is associated with high risk cytogenetics and poor outcome in newly diagnosed, symptomatic patients with multiple myeloma: a single center experience on 228 patients. *Am J Hematol* 87(9):861–864. <https://doi.org/10.1002/ajh.23258>
23. Greipp PR, San Miguel J, Durie BG et al (2005) International staging system for multiple myeloma. *J Clin Oncol* 23(15):3412–3420. <https://doi.org/10.1200/JCO.2005.04.242>
24. Palumbo A, Avet-Loiseau H, Oliva S et al (2015) Revised International staging system for multiple myeloma: a report from International Myeloma Working Group. *J Clin Oncol* 33(26):2863–2869. <https://doi.org/10.1200/JCO.2015.61.2267>
25. Kastritis E, Terpos E, Roussou M et al (2017) Evaluation of the Revised International Staging System in an independent cohort of unselected patients with multiple myeloma. *Haematologica* 102(3):593–599. <https://doi.org/10.3324/haematol.2016.145078>
26. Jimenez-Zepeda VH, Duggan P, Neri P, Rashid-Kolvear F, Tay J, Bahlis NJ (2016) Revised international staging system applied to real world multiple myeloma patients. *Clin Lymphoma Myeloma Leuk* 16(9):511–518. <https://doi.org/10.1016/j.clml.2016.06.001>
27. Radocha J, Maisnar V, Pour L et al (2018) Validation of multiple myeloma risk stratification indices in routine clinical practice: analysis of data from the Czech Myeloma Group Registry of Monoclonal Gammopathies. *Cancer Med* 7(8):4132–4145. <https://doi.org/10.1002/cam4.1620>
28. González-Calle V, Slack A, Keane N et al (2018) Evaluation of Revised International Staging System (R-ISS) for transplant-eligible multiple myeloma patients. *Ann Hematol* 97(8):1453–1462. <https://doi.org/10.1007/s00277-019-03702-1>
29. Ozaki S, Handa H, Saitoh T et al (2019) Evaluation of the Revised International Staging System (R-ISS) in Japanese patients with multiple myeloma. *Ann Hematol* 98(7):1703–1711. <https://doi.org/10.1007/s00277-019-03702-1>
30. Han Y, Sunwoo L, Ye JC (2020) k-space deep learning for accelerated MRI. *IEEE Trans Med Imaging* 39(2):377–386. <https://doi.org/10.1109/TMI.2019.2927101>
31. Hyun CM, Kim HP, Lee SM, Lee S, Seo JK (2018) Deep learning for under sampled MRI reconstruction. *Phys Med Biol* 63:aac71a. <https://doi.org/10.1088/1361-6560/aac71a>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.