Clinical Radiology 82 (2025) 106768

Contents lists available at ScienceDirect

Clinical Radiology

journal homepage: www.clinicalradiologyonline.net

# Magnetic resonance imaging–based nomograms predict high-risk cytogenetic abnormalities in multiple myeloma: a two-centre study

S. Liu [a,b,c,d,†], C. Liu [e,†], H. Pan [b,†], S. Li [a,b,c,d], P. Teng [e], Z. Li [a,b,c,d], J. Sun [a,b,c,d], T. Ren [a,b,c,d], G. Liu [e,**], J. Zhou [a,b,c,d,*]

[a] *Department of Radiology, Lanzhou University Second Hospital, Cuiyingmen No. 82, Chengguan District, Lanzhou, 730030, China*
[b] *Lanzhou University Second Hospital, Lanzhou, China*
[c] *Key Laboratory of Medical Imaging of Gansu Province, Lanzhou, China*
[d] *Gansu International Scientific and Technological Cooperation Base of Medical Imaging Artificial Intelligence, Lanzhou, China*
[e] *Department of Radiology, China-Japan Union Hospital of Jilin University, Changchun, 130000, China*

ARTICLE INFORMATION

*Article history:*
Received 16 January 2024
Received in revised form
13 October 2024
Accepted 29 November 2024

AIM: The study aim to use magnetic resonance imaging (MRI) radiomic features to predict high-risk cytogenetic abnormalities (HRCAs) to improve outcomes in patients with multiple myeloma (MM).

MATERIALS AND METHODS: One hundred ninety-five patients with MM from two centres undergoing MRI were retrospectively recruited. Patients from Institution I (71 and 88 HRCAs and non-HRCAs, respectively) identified by fluorescence in situ hybridisation were randomly divided into training (n = 111) and validation (n = 48) cohorts. Patients from Institution II served as the external test cohort (n = 36). Radiomics or combined models based on T1WI, T2WI, and FS-T2WI images and clinical factors were constructed using logistic regression and 10-fold cross-validation in the training cohort. Nomogram performance was evaluated and compared using C-index, bootstrapping, accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and Akaike information criterion. C-indexes were used to select the most efficient radiomics predictive model. Optimal model performance was tested in an external cohort.

RESULTS: $FT_2$+age, $FT_{2+1}$+age, and $FT_{2+2+1}$+age combined models were outstanding in differentiating the HRCAs of MM patients in single-, double-, and multi-sequence MRI images, respectively. The C-indexes of the training and validation cohorts corrected via the 1000 bootstrap method were 0.79 and 0.80, 0.83 and 0.84, and 0.88 and 0.84, respectively. In the external test cohort, the C-index of radiomics nomograms was 0.70, 0.76, and 0.77, respectively.

* Guarantor and correspondent: J. Zhou, Department of Radiology, Lanzhou University Second Hospital, Cuiyingmen No. 82, Chengguan District, Lanzhou, 730030, China.
** Guarantor and correspondent: G. Liu, Department of Radiology, China-Japan Union Hospital of Jilin University, Changchun, 130000, China.
*E-mail addresses:* gfliu@jlu.edu.cn (G. Liu), ery_zhoujl@lzu.edu.cn (J. Zhou).
† **S. Liu, C. Liu** and **H. Pan** contributed equally to the writing of this article.

https://doi.org/10.1016/j.crad.2024.106768

CONCLUSION: MRI radiomics can be used to predict HRCAs in MM patients, which will be helpful for clinical decision-making and prognosis evaluation before treatment.


## Introduction

Multiple myeloma (MM) is a highly heterogeneous disease driven by genetic and epigenetic alterations. It is the second most common haematologic malignancy after lymphoma.[1] Cytogenetic abnormalities (CAs) are observed in almost all patients with MM. Most patients present with multiple CAs, including chromosomal translocations, copy number variants, and chromosomal fragment deletions and amplifications.[2] Known high-risk cytogenetic abnormalities (HRCAs) include t(4; 14), del(17p13), 1q21 amplification, and P53 mutation, while the remaining CAs are considered non-HRCAs.[3]

Furthermore, MM patients with HRCAs exhibit a more pronounced clonal heterogeneity and evolution.[4] These patients usually show resistance to chemotherapy with the traditional RVd (lenalidomide/bortezomib/dexamethasone) regimen.[5] In contrast, induction, consolidation, and maintenance therapy using autologous haematopoietic stem cell transplantation combined with KRd regimens (carfilzomib/lenalidomide/dexamethasone) can improve survival rates by up to 72%.[6] CAs are therefore crucial for selecting an MM treatment regimen and assessing treatment outcomes.

The diagnosis of CAs relies on the collection of tissue specimens and molecular biology tests such as fluorescence in situ hybridisation (FISH) and next-generation sequencing (NGS).[3,7,8] However, obtaining tissue specimens for CA testing usually requires a bone marrow aspiration biopsy, which is invasive and may cause complications such as bleeding and infection in patients with haematologic disorders. Other limitations of this method include inadequate or inappropriate tissue sampling due to tumour spatial heterogeneity. In addition, the process of histological specimen analysis is cumbersome, complex, and expensive. Therefore, novel noninvasive methods of CA determination are required.

Magnetic resonance imaging (MRI) is recommended as the imaging method of choice for diagnosing bone marrow infiltration in patients with MM.[9] It helps clinicians obtain accurate information about bone marrow infiltration in asymptomatic individuals and assess the type of bone damage: standard, focal, diffuse, mixed, or “salt and pepper” type.[10] In addition, multisequence MRI scanning helps determine lesion location, mass size, and signal diversity.[11] However, radiological evaluation based on conventional MRI images to identify HRCAs is challenging.[12] Radiomics is a high-throughput extraction and computational analysis method used to obtain potentially valuable high-dimensional information about tumour heterogeneity from medical images in a manner that is superior to assessment by human eyes.[13] The value of radiomic features as imaging predictors for cancer diagnosis, treatment response, and prognosis has been demonstrated in cancer, including MM.[14] Recent studies have explored the feasibility of using radiomic features derived from whole-spine MRI to predict HRCAs in MM patients.[15–17] These reports have laid the groundwork by demonstrating the potential of radiomics to enhance the diagnostic and prognostic value of MRI. However, most previous studies have been limited by small patient cohorts and methodological differences, which hinder their generalisability to broader clinical practice. In addition, while these studies have advanced our understanding, the integration of radiomic features into clinically applicable nomograms remains limited. Herein, we sought to address this gap by studying a larger, more diverse patient cohort and employing a more sophisticated machine learning approach. We developed and validated a predictive model that combines MRI-based radiomic features with key clinical parameters to more accurately predict cytogenetic abnormalities in MM patients. Finally, we created a robust nomogram that can be readily applied in clinical practice to assist in early risk stratification and treatment decision-making.

## Materials and methods

### *Patients*

Our institutional review board approved this retrospective study, which waived the informed consent requirement due to the study’s retrospective nature.

We examined data from 643 consecutive patients presenting from two centres between January 2013 and December 2021 with pathologically confirmed MM, including 468 patients from Institution I and 175 from Institution II. The inclusion criteria were as follows: (1) T1WI, T2WI, FS-T2WI scans obtained before treatment; (2) availability of FISH; (3) availability of data on demographic and clinical characteristics (e.g., sex, age, and albumin, lactate dehydrogenase, β2 microglobulin level); (4) presence of at least one vertebral lesion. The exclusion criteria were as follows: (1) malignant bone metastases and severe scoliosis deformity (Cobb angle $> 40^{\circ}$) were present; (2) poor image quality; (3) no MM bone marrow lesions or extramedullary lesions; (4) postoperative spine. Finally, 195 patients from two centres were included in this study (Fig 1). Data on demographic and clinical characteristics, including laboratory and FISH findings, treatment procedures and outcomes, and MRI scan findings, were obtained.

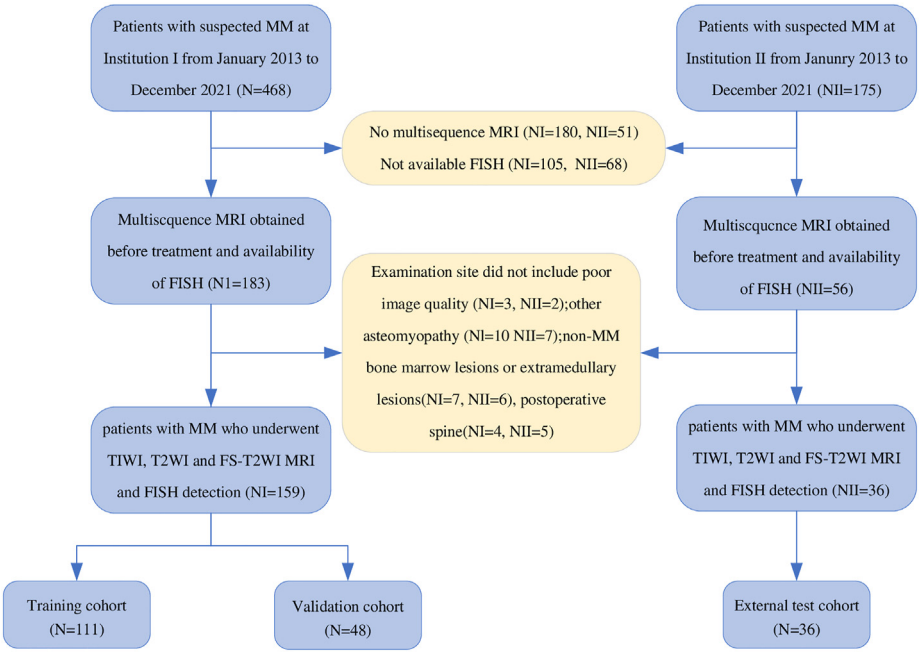


**Figure 1** Flowchart of the patient selection process. FISH, fluorescence in situ hybridisation; MM, multiple myeloma; MRI, magnetic resonance imaging.

### *MRI scan acquisition*

MRI scans were obtained using a consistent protocol with 3.0T Siemens, Philip, and GE scanners. Specific scan sequence parameters and other information are provided in Supplementary Material 1. According to the Image Biomarker Standardization Initiative, the original images were initially resampled for grayscale intensity and voxel size. To eliminate potential differences in MRI scans acquired by the three different MRI scanners, all original images were normalised using a grayscale-level discretisation method before extracting radiomic features (Version V3.0.R, GE Healthcare).[17]

### *Reproducibility analysis*

Volume datasets were obtained by manually segmenting sagittal T1WI, T2WI, and FS-T2WI images using ITK-SNAP (v. 3.6.0; www.itksnap.org). MRI scans of each patient were independently reviewed by three musculoskeletal radiologists with 5, 15, and 25 years of experience (Reader 1, Reader 2, and Reader 3, respectively), who were blinded to patient information. The lesion regions of interest (ROIs) were manually drawn by Reader 1 and validated by Reader 2. For the validation process, the Dice Similarity Coefficient (DSC) was calculated to quantify the agreement between segmentations performed by Reader 1 and Reader 2. A DSC value of $\geq$ 0.95 was considered acceptable, indicating high agreement and minimal discrepancy. If the DSC value was below 0.95, indicating significant differences, Reader 2 would suggest modifications to the segmentation. If the difference value was $\geq$5%, the tumour boundary was determined by Reader 3.[18] In cases where consensus could not be achieved, Reader 3 made the final determination. More experienced radiologists are generally more adept at identifying subtle anatomical structures and abnormalities, which can reduce variability in segmentation. By involving radiologists with different levels of experience, we aimed to ensure that the reproducibility analysis reflected real-world clinical settings, where segmentation tasks may be performed by radiologists with varying expertise.

### *ROI segmentation*

Since MM has multiple lesions and bone marrow infiltration patterns, we selected the largest lesion to outline the focal pattern. To distinguish compression fractures due to physiological osteoporosis, the vertebrae with compression fractures were avoided as much as possible.[19] If n-foci were almost equal, the total volume-weighted average of the largest n-tumours was used.[20] We selected the third lumbar delineation for salt-and-pepper, diffuse, and mixed types. If there was a compression fracture, we selected the adjacent vertebral delineation (Supplementary Fig 1). The radiomics feature extraction methods are described in Supplementary Material 2.

### *Radiomics feature selection*

Each radiomic feature was normalised, and batch effects were removed.[21] Feature dimensionality reduction was performed in the training set using variance thresholding,[16] batch t-test, redundancy analysis method, and least absolute shrinkage and selection operator feature selection algorithm to remove irrelevant and redundant information. Subsequently, logistic regression and 10-fold

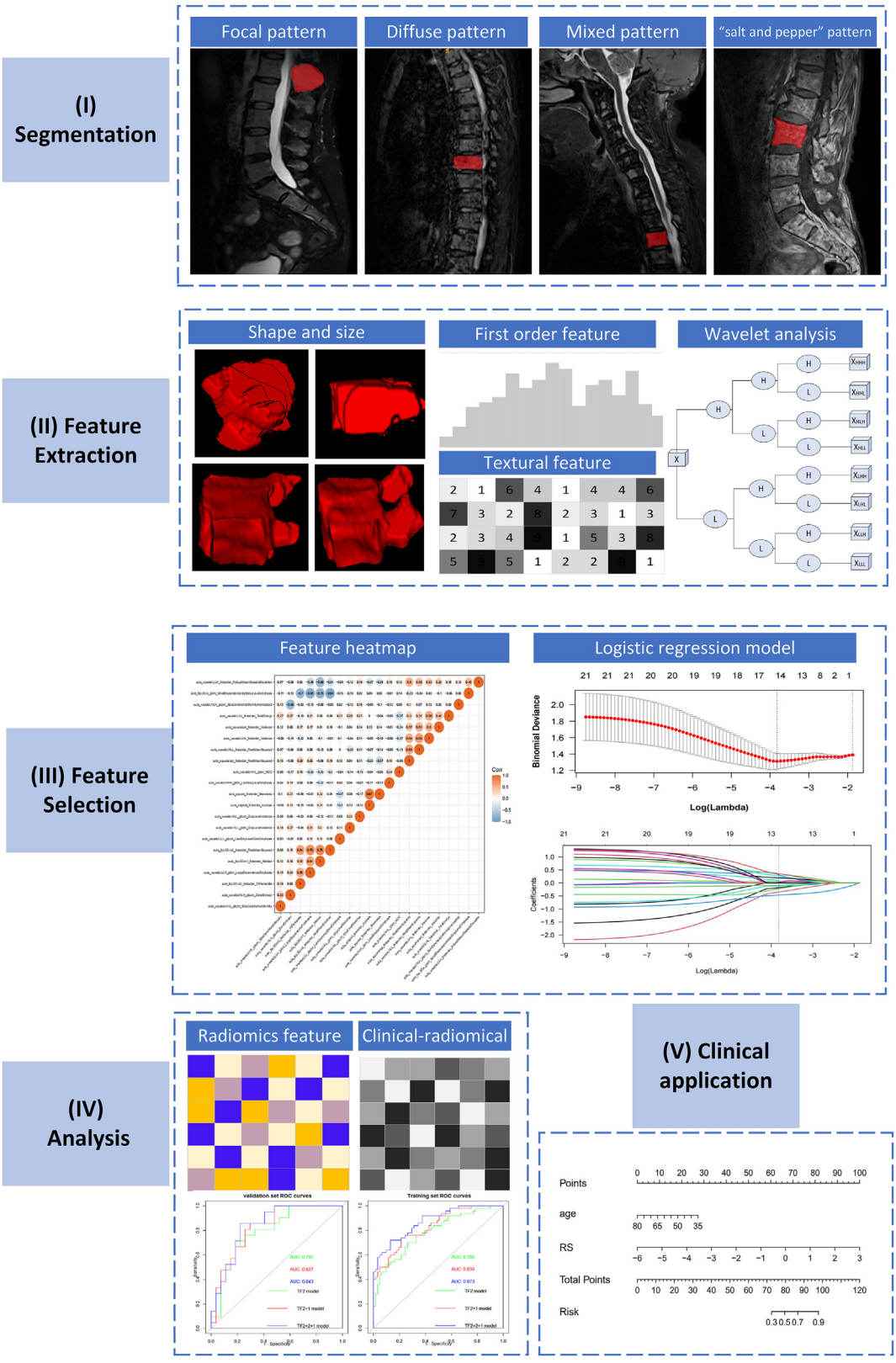


**Figure 2** Flowchart of the MM radiomics approach for predicting cytogenetic status. (I) Manual outlining of the ROI for focal, salt-and-pepper, mixed, and diffuse lesions. (II) Extraction of image histology features, including shape and size, first-order statistical, texture, and wavelet features. (III) Dimensional reduction of the features using the LASSO regression model. (IV) Combining imaging histological features and clinical features to construct prediction models, further statistical analysis such as ROC curves are used; (V) Finally, nomograms are used to show the MM cytogenetic status prediction results on clinical validity. LASSO, least absolute shrinkage and selection operator; MM, multiple myeloma; ROC, receiver operating characteristic; ROI, region of interest.

**Table 1**
Characteristics of MM patients with HRCAs and non-HRCAs.

| Characteristics | HRCAs (n = 84) | Non-HRCAs (n = 111) | *P-value* |
|---|---|---|---|
| **Gender** | | | 0.635 |
| Male | 55 (65.48%) | 69 (62.20%) | |
| Female | 29 (34.52%) | 42 (37.80%) | |
| **Age (years)** | 55.00 ± 9.45 | 60.41 ± 8.85 | **<0.001** |
| **BMG (mg/l)** | 7.61 ± 4.90 | 8.10 ± 7.79 | 0.605 |
| **ALB (g/l)** | 35.64 ± 7.96 | 36.30 ± 7.34 | 0.589 |
| **HB (g/l)** | 95.81 ± 28.64 | 102.40 ± 27.24 | 0.103 |
| **LDH (u/l)** | 203.83 ± 89.54 | 210.41 ± 110.22 | 0.656 |
| **MRI patterns** | | | 0.947 |
| Focal pattern | 45 (53.60%) | 60 (54.10%) | |
| The other patterns | 39 (46.40%) | 51 (45.90%) | |
| **Compression fracture** | | | 0.175 |
| Yes | 48 (57.10%) | 74 (66.67%) | |
| No | 36 (42.90%) | 37 (33.33%) | |
| **Whole spinal infiltration** | | | 0.141 |
| Yes | 51 (60.70%) | 51 (45.90%) | |
| No | 33 (39.30%) | 60 (54.10%) | |

ALB, albumin; BMG, beta-2-microglobulin; HRCA = , high-risk cytogenetic abnormality; LDH, lactate dehydrogenase.
Note: *P*-value < 0.05 was considered as a significant difference.

cross-validation methods were used to select potential predictive features in the training queue. Finally, the best predictive features were obtained using the forward-backward stepwise regression method. Image analysis and prediction model construction are shown in Fig 2.

### *Individualised prediction model development*

Single-factor analysis and multi-factor logistic regression analysis were used to determine the clinical predictors of HRCAs. We aimed to develop separate T1WI ($T_1$), T2WI ($T_2$), FS-T2WI ($FT_2$), T1WI + T2WI ($T_{1+2}$), T1WI + FS-T2WI ($FT_{2+1}$), T2WI + FS-T2WI ($FT_{2+2}$), and T1WI + T2WI + FS-T2WI ($FT_{2+2+1}$) as well as their corresponding combined models for HRCA prediction and the construction of corresponding nomograms.

### *Prediction model validation and performance*

A calibration curve (Hosmer–Lemeshow test) was used to evaluate the calibration effect of the nomogram. Diagnostic efficacy was evaluated using C-index, accuracy (ACC), sensitivity (SEN), specificity (SPE), positive predictive value (PPV), negative predictive value (NPV), and Akaike information criterion (AIC). To assess the potential prediction error of the proposed model in the cohort, the bootstrap method was utilised with 1000 iterations. The clinical utility of the nomogram was evaluated in the test set using decision curve analysis.

### *Statistical analysis*

Statistical analyses were performed using SPSS software (Version 26.0; IBM Corp.) and R software (https://www.r-project.org; Version 4.1.2). All statistical tests were two-sided, and $P < 0.05$ was statistically significant. The packages in R software involved in this study are shown in Supplementary Material 3.

## Results

A total of 195 patients from two centres were included in this study. The patients from Institution I were randomised 7:3 into a training cohort (n = 111) and a validation cohort (n = 48), while the patients from Institution II comprised an external test set (n = 36).

Table 1 presents the characteristics of patients HRCAs and non-HRCAs. Blood β2 microglobulin, blood albumin, and lactate dehydrogenase levels at baseline are presented. Imaging features were assessed, including focal growth patterns, entire spinal vertebral body involvement, and compression fractures. There was a statistically significant difference between the three groups of patients only in terms of age ($p < 0.05$). There was no statistical difference between the training and validation cohorts ($p > 0.05$), ensuring a reasonable cohort stratification (Supplementary Material 4).

### *Feature selection and radiomics signature*

There were 1688 T1WI features, 1688 T2WI features, and 1688 FS-T2WI features. Radiomics features were selected in each training cohort after a series of dimensionality reductions, as shown in Table 2. The best predictors were selected by 10-fold cross-validation, and the prediction models were built separately. The radiomics score obtained by building an equation is called the Rad-Score (Supplementary Material 5).

### *HRCA prediction model development*

All factors were included in the multiple regression analysis to establish radiomic models. The Rad-Score was used to build radiomic nomograms for radiomics models: $T_1$, $T_2$, $FT_2$, $T_{1+2}$, $FT_{2+1}$, $FT_{2+2}$, and $FT_{2+2+1}$ models, while age and Rad-Score were used as best predictors to build combined radiomic nomograms for combined radiomics models: $T_1$+age, $T_2$+age, $FT_2$+age, $T_{1+2}$+age, $FT_{2+1}$+age, $FT_{2+2}$+age, and $FT_{2+2+1}$+age models, respectively.

### *Prediction model performance*

The ACC, SEN, SPE, PPV, NPV, AIC, ACC, and C-index values and bootstrapping results of the training and validation sets for all 14 models (7 radiomics models and seven combined models) are shown in Table 3. The combined models had a higher overall diagnostic performance than the radiomics models. While among the single-sequence models, the performance of the $FT_2$+age model was outstanding, with a C-index of 0.79 and 0.80 for the training and validation sets, respectively, among the double-sequence models, the performance of the $FT_{2+1}$+age model was excellent, with a C-index of 0.84 and 0.84 for the

**Table 2**
Results of feature selection of combined radiomics or radiomics models in the training cohort.

| Models | Numbers | Included variables (*P* value) |
|---|---|---|
| $T_1$ | 4 | T1_auto_wavelet.HHL_glszm_LowGrayLevelZoneEmphasis (0.06)<br>T1_auto_wavelet.HLL_glszm_GrayLevelVariance (0.03)<br>T1_auto_wavelet.LHH_glszm_LowGrayLevelZoneEmphasis (0.02)<br>T1_auto_wavelet.LHL_glszm_LowGrayLevelZoneEmphasis (0.12) |
| $T_2$ | 3 | T2_auto_lbp.3D.m2_firstorder_Skewness (0.02)<br>T2_auto_wavelet.HHH_glszm_SmallAreaLowGrayLevelEmphasis(0.02)<br>T2_auto_wavelet.LHH_glrlm_LowGrayLevelRunEmphasis (0.009) |
| $FT_2$ | 5 | FS_auto_lbp.3D.m2_firstorder_Mean (0.04)<br>FS_auto_squareroot_firstorder_Maximum (0.02)<br>FS_auto_wavelet.HHH_glcm_ClusterShade (0.08)<br>FS_auto_wavelet.HLL_glszm_GrayLevelVariance (0.12)<br>FS_auto_wavelet.LLL_glszm_LowGrayLevelZoneEmphasis (0.02) |
| $T_{1+2}$ | 7 | T1_auto_wavelet.HHL_glszm_LowGrayLevelZoneEmphasis (0.01)<br>T1_auto_wavelet.HLL_glszm_GrayLevelVariance (0.02)<br>T1_auto_wavelet.LHH_glszm_LowGrayLevelZoneEmphasis (0.04)<br>T1_auto_wavelet.LHL_glszm_LowGrayLevelZoneEmphasis (0.02)<br>T2_auto_lbp.3D.m2_firstorder_Skewness (0.005)<br>T2_auto_wavelet.HHH_glszm_SmallAreaLowGrayLevelEmphasis (0.03)<br>T2_auto_wavelet.LHH_glrlm_LowGrayLevelRunEmphasis (0.01) |
| $FT_{2+1}$ | 9 | T1_auto_wavelet.HHL_glszm_LowGrayLevelZoneEmphasis (0.01)<br>T1_auto_wavelet.HLL_glszm_GrayLevelVariance (0.03)<br>T1_auto_wavelet.LHL_glszm_LowGrayLevelZoneEmphasis (0.04)<br>FS_auto_squareroot_firstorder_Maximum (0.02)<br>FS_auto_wavelet.HLH_glszm_ZoneEntropy (0.01)<br>FS_auto_wavelet.HLL_glszm_GrayLevelVariance (0.03)<br>FS_auto_wavelet.LHL_firstorder_TotalEnergy (0.1)<br>FS_auto_wavelet.LLH_firstorder_RobustMeanAbsoluteDeviation (0.03)<br>FS_auto_wavelet.LLL_glszm_LowGrayLevelZoneEmphasis (0.03) |
| $FT_{2+2}$ | 8 | T2_auto_wavelet.HHH_glszm_SmallAreaLowGrayLevelEmphasis (0.03)<br>T2_auto_wavelet.LHH_glrlm_LowGrayLevelRunEmphasis (0.04)<br>FS_auto_lbp.3D.m2_firstorder_Mean (0.05)<br>FS_auto_squareroot_firstorder_Maximum (0.02)<br>FS_auto_wavelet.HHH_glcm_ClusterShade (0.13)<br>FS_auto_wavelet.HLH_glszm_ZoneEntropy (0.06)<br>FS_auto_wavelet.HLL_glszm_GrayLevelVariance (0.11)<br>FS_auto_wavelet.LLL_glszm_LowGrayLevelZoneEmphasis (0.11) |
| $FT_{2+2+1}$ | 10 | T1_auto_wavelet.HHL_glszm_LowGrayLevelZoneEmphasis (0.009)<br>T1_auto_wavelet.HLL_glszm_GrayLevelVariance (0.06)<br>T1_auto_wavelet.LHL_glszm_LowGrayLevelZoneEmphasis (0.02)<br>T2_auto_lbp.3D.m2_firstorder_Skewness (0.0004)<br>T2_auto_wavelet.HHH_glszm_SmallAreaLowGrayLevelEmphasis (0.02)<br>T2_auto_wavelet.LHH_glrlm_LowGrayLevelRunEmphasis (0.03)<br>FS_auto_squareroot_firstorder_RootMeanSquared (0.03)<br>FS_auto_wavelet.HHH_glcm_ClusterShade (0.1)<br>FS_auto_wavelet.HLH_glszm_ZoneEntropy (0.02)<br>FS_auto_wavelet.LLL_glszm_LowGrayLevelZoneEmphasis (0.04) |

*The p-value* for each radiomic feature associated with MM HRCAs was calculated using the Mann–Whitney U or Student's *t*-test.

training and validation sets, respectively. The performance of the $FT_{2+2+1}$+age model was better than that of the other models, with a C-index of 0.88 and 0.84 for the training and validation sets, respectively (Fig 3). The calibration curves of these three sets of nomograms predicting the HRCAs showed good agreement between predicted and observed values. The Hosmer–Lemeshow test showed no statistically significant difference between the training and validation sets per group ($p > 0.017$, training set: $p = 0.095$, 0.086, and 0.018; validation set: $p = 0.696$, 0.362 and 0.624) (Fig 4). Calibration curves reflect how well probabilities predicted from a model align with actual outcomes. A well-calibrated model thus indicates that the predicted risk closely matches the observed risk in real patients, which is crucial for accurate clinical decision-making. Finally, we tested these three models from Institution II data, with a C-index of 0.70, 0.76, and 0.77 for test sets (Table 3).

### *Clinical application*

The $FT_2$+age, $FT_{2+1}$+age, and $FT_{2+2+1}$+age model decision curves showed that using nomograms to predict HRCAs in patients with MM showed a more significant clinical benefit than using the clinical model alone if the

**Table 3**
Diagnostic performance of 14 models in the training, validation, and test cohorts.

| Cohorts | Models | C-index (95% CI) | Bootstrap (95% CI) | Acc | Sen | Spe | PPV | NPV | AIC |
|---|---|---|---|---|---|---|---|---|---|
| Training set | $T_1$ | 0.70 (0.60–0.79) | 0.70 (0.61–0.79) | 0.61 | 0.50 | 0.7 | 0.58 | 0.63 | 147.32 |
| | $T_1$+age | 0.75 (0.66–0.84) | 0.75 (0.65–0.84) | 0.67 | 0.58 | 0.74 | 0.64 | 0.68 | 135.83 |
| | $T_2$ | 0.73 (0.62–0.82) | 0.73 (0.64–0.83) | 0.63 | 0.69 | 0.56 | 0.66 | 0.60 | 144.4 |
| | $T_2$+age | 0.77 (0.68–0.86) | 0.77 (0.68–0.85) | 0.70 | 0.75 | 0.64 | 0.72 | 0.68 | 131.07 |
| | $FT_2$ | 0.76 (0.65–0.85) | 0.76 (0.66–0.85) | 0.65 | 0.75 | 0.52 | 0.66 | 0.63 | 142.05 |
| | $FT_2$+age | 0.79 (0.70–0.87) | 0.80 (0.71–0.88) | 0.69 | 0.62 | 0.75 | 0.67 | 0.71 | 125.97 |
| | $T_{1+2}$ | 0.82 (0.73–0.89) | 0.83 (0.75–0.90) | 0.71 | 0.62 | 0.79 | 0.70 | 0.72 | 131.8 |
| | $T_{1+2}$+age | 0.84 (0.76–0.91) | 0.84 (0.76–0.91) | 0.75 | 0.72 | 0.77 | 0.72 | 0.77 | 115.7 |
| | $FT_{2+1}$ | 0.81 (0.72–0.88) | 0.81 (0.73–0.89) | 0.70 | 0.70 | 0.71 | 0.66 | 0.74 | 136.27 |
| | $FT_{2+1}$+age | 0.83 (0.76–0.90) | 0.84 (0.76–0.91) | 0.74 | 0.68 | 0.79 | 0.72 | 0.75 | 113.14 |
| | $FT_{2+2}$ | 0.81 (0.72–0.88) | 0.81 (0.73–0.88) | 0.74 | 0.68 | 0.79 | 0.72 | 0.75 | 137.24 |
| | $FT_{2+2}$+age | 0.83 (0.76–0.91) | 0.84 (0.76–0.91) | 0.70 | 0.66 | 0.74 | 0.67 | 0.73 | 116.3 |
| | $FT_{2+2+1}$ | 0.86 (0.79–0.93) | 0.87 (0.80–0.92) | 0.78 | 0.74 | 0.82 | 0.79 | 0.79 | 124.52 |
| | $FT_{2+2+1}$+age | 0.88 (0.82–0.94) | 0.88 (0.81–0.93) | 0.77 | 0.74 | 0.80 | 0.76 | 0.79 | 103.8 |
| Validation set | $T_1$ | 0.68 (0.52–0.88) | 0.72 (0.57–0.87) | 0.58 | 0.57 | 0.59 | 0.48 | 0.64 | 67.23 |
| | $T_1$+age | 0.76 (0.62–0.90) | 0.77 (0.63–0.90) | 0.67 | 0.67 | 0.67 | 0.61 | 0.72 | 60.225 |
| | $T_2$ | 0.68 (0.50–0.82) | 0.69 (0.54–0.83) | 0.56 | 0.70 | 0.58 | 0.59 | 0.5 | 67.41 |
| | $T_2$+age | 0.71 (0.56–0.85) | 0.73 (0.60–0.86) | 0.67 | 0.74 | 0.57 | 0.69 | 0.63 | 64.236 |
| | $FT_2$ | 0.74 (0.59–0.85) | 0.75 (0.60–0.88) | 0.68 | 0.70 | 0.63 | 0.61 | 0.63 | 62.85 |
| | $FT_2$+age | 0.80 (0.67–0.92) | 0.80 (0.67–0.93) | 0.73 | 0.67 | 0.78 | 0.7 | 0.75 | 58.202 |
| | $T_{1+2}$ | 0.75 (0.61–0.87) | 0.75 (0.61–0.88) | 0.67 | 0.67 | 0.67 | 0.61 | 0.72 | 66.68 |
| | $T_{1+2}$+age | 0.78 (0.65–0.90) | 0.79 (0.66–0.91) | 0.67 | 0.71 | 0.63 | 0.6 | 0.74 | 58.392 |
| | $FT_{2+1}$ | 0.77 (0.67–0.86) | 0.78 (0.63–0.90) | 0.71 | 0.76 | 0.67 | 0.64 | 0.78 | 66.43 |
| | $FT_{2+1}$+age | 0.83 (0.71–0.94) | 0.84 (0.73–0.94) | 0.81 | 0.86 | 0.78 | 0.75 | 0.88 | 54.016 |
| | $FT_{2+2}$ | 0.77 (0.64–0.90) | 0.78 (0.63–0.90) | 0.73 | 0.67 | 0.78 | 0.70 | 0.75 | 66.62 |
| | $FT_{2+2}$+age | 0.81 (0.69–0.93) | 0.82 (0.69–0.93) | 0.75 | 0.76 | 0.74 | 0.70 | 0.80 | 56.175 |
| | $FT_{2+2+1}$ | 0.80 (0.67–0.92) | 0.80 (0.66–0.91) | 0.73 | 0.76 | 0.70 | 0.67 | 0.79 | 68.98 |
| | $FT_{2+2+1}$+age | 0.84 (0.73–0.95) | 0.84 (0.73–0.95) | 0.75 | 0.81 | 0.70 | 0.68 | 0.83 | 53.93 |
| Testing set | $FT_2$+age | 0.70 (0.53–0.87) | NA | 0.67 | 0.70 | 0.57 | 0.69 | 0.65 | NA |
| | $FT_{2+1}$+age | 0.76 (0.59–0.94) | NA | 0.72 | 0.77 | 0.78 | 0.77 | 0.70 | NA |
| | $FT_{2+2+1}$+age | 0.77 (0.61–0.92) | NA | 0.67 | 1 | 0.78 | 0.62 | 0.70 | NA |

Acc, accuracy; AIC, Akaike information criterion; NPV, negative predictive value; PPV, positive predictive value; Sen, sensitivity; Spe, specificity.

patient threshold probability was more significant than 5%. Thus, these nomograms allow for the more accurate identification of patients who are likely to benefit from further investigation or treatment (Supplementary Fig 2).

## Discussion

Noninvasive methods for assessing CAs in MM are crucial for patient treatment planning and the evaluation of treatment outcomes. We developed and tested radiomic nomograms for 14 models at two centres using clinically obtained whole-spine MRI images of 195 patients with MM. Among the proposed models, the nomograms of $FT_2$+age, $FT_{2+1}$+age, and $FT_{2+2+1}$+age models were identified as outstanding at distinguishing patients with HRCAs from those without. The nomograms exhibited robust performance in both centres, demonstrating good diagnostic accuracy.

Conventional MRI scans contain critical information that reflects the heterogeneous features of tumours, including biological characteristics determined by protein levels, gene expression, mutation, and other molecular features.[22–25] Our results demonstrated that the $FT_{2+2+1}$+age model outperformed both single-sequence and dual-sequence models. This may be due to the complementary information provided by different sequences: T1WI captures local anatomical details, T2WI reveals internal heterogeneity, and FS-T2WI isolates internal tumour characteristics by removing fat signal interference from outside the bone marrow. Together, the radiomic features derived from these sequences provide a more comprehensive quantification of MM heterogeneity.[26] Although the statistical difference between the three prediction models was not significant ($p > 0.017$, Delong test), the $FT_2$+age or $FT_{2+1}$+age models may still be considered when patient sequence data are incomplete, offering practical flexibility in clinical settings. This suggests that, even with partial data, these models can effectively stratify CAs in MM patients, which supports their potential clinical application in MM management. While previous studies, including that of Xiong *et al.*,[27] have highlighted the role of MRI-based radiomic features in predicting CAs in MM, our study expands on these findings by exploring the use of combined radiomic models across multiple MRI sequences. Furthermore, the role of MRI in MM management is expanding as recent studies emphasise its utility not only in diagnostic imaging but also in guiding treatment decisions through advanced radiomic analyses.

Among the clinical risk factors, age is the single most relevant factor for cytogenetic status in MM patients and may be associated with an earlier age of onset, which is consistent with previous studies.[16,28] In addition, we did not retain lactate dehydrogenase, β2 microglobulin, and albumin as clinically relevant risk factors for MM patients. While

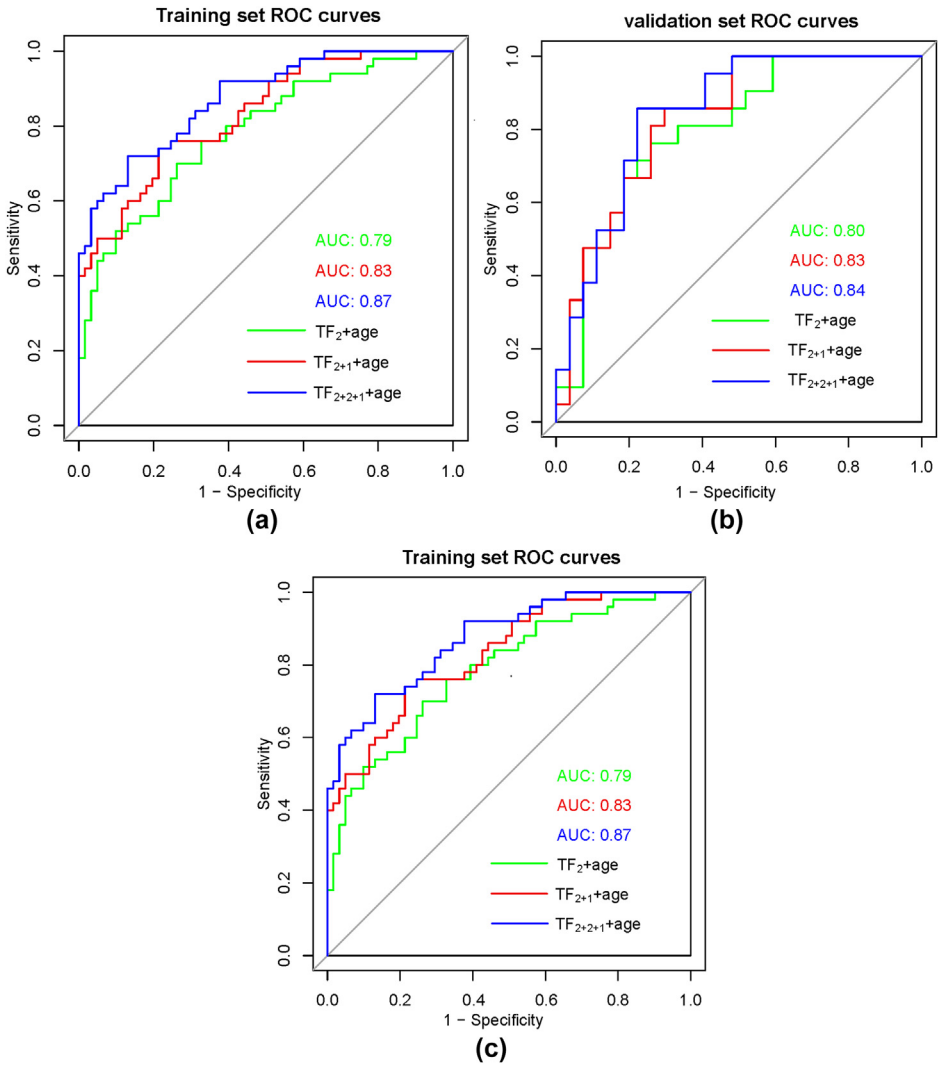


**Figure 3** $FT_2$+age, $FT_{2+1}$+age, and $FT_{2+2+1}$+age models (a) nomogram training set, (b) validation set, and (c) test set receiver operating characteristic (ROC) curves. AUC, area under the curve.

these laboratory markers of malignancy may be associated with a poor prognosis in MM, their relationship with cytogenetic status in MM patients has not been previously reported. Therefore, as the only statistically significant predictor of clinical information, age improved the diagnostic merit of our combined model.

Logistic regression analysis was employed in the current study. Previous studies have shown that the ideal classifier used to build models may vary from organ to organ,[26] with logistic regression being more applicable to sacral tumours, solid lung nodules, and breast cancer than in other contexts.[29–31] Liu *et al.*[15] demonstrated that logistic regression–based machine learning methods significantly outperformed other classifiers for HRCA evaluation in MM. Logistic regression was chosen for its simplicity, ease of interpretation, and robustness in handling binary outcomes. In addition, to ensure the reliability of modelling, the 1:10 rule (i.e., approximately one feature can be studied per 10 events) was strictly followed when using logistic regression to analyse data in the training set using 10-fold cross-validation. The use of 10-fold cross-validation helps reduce bias and variance, provides a comprehensive evaluation of the model, and enhances its generalisability to unseen data by ensuring that each data point is used for both training and validation.

Our study used a multiparametric MRI-based radiomics approach to predict HRCAs in MM. Although a similar approach was employed in a previous study,[15] we further validated the feasibility of this study with a larger sample and the inclusion of another centre. Multicentre studies can provide more diverse clinical and imaging data, which aligns better with the ultimate goal of precision medicine. At the same time, despite considerable efforts towards predicting HRCAs in MM, our study is the first to establish single-, double-, and multiple-sequence radiomics-based prediction models for clinical use. Moreover, the nomograms created herein can be used as decision-support tools to aid clinicians in early risk stratification of MM patients, improving individualised treatment planning. They can

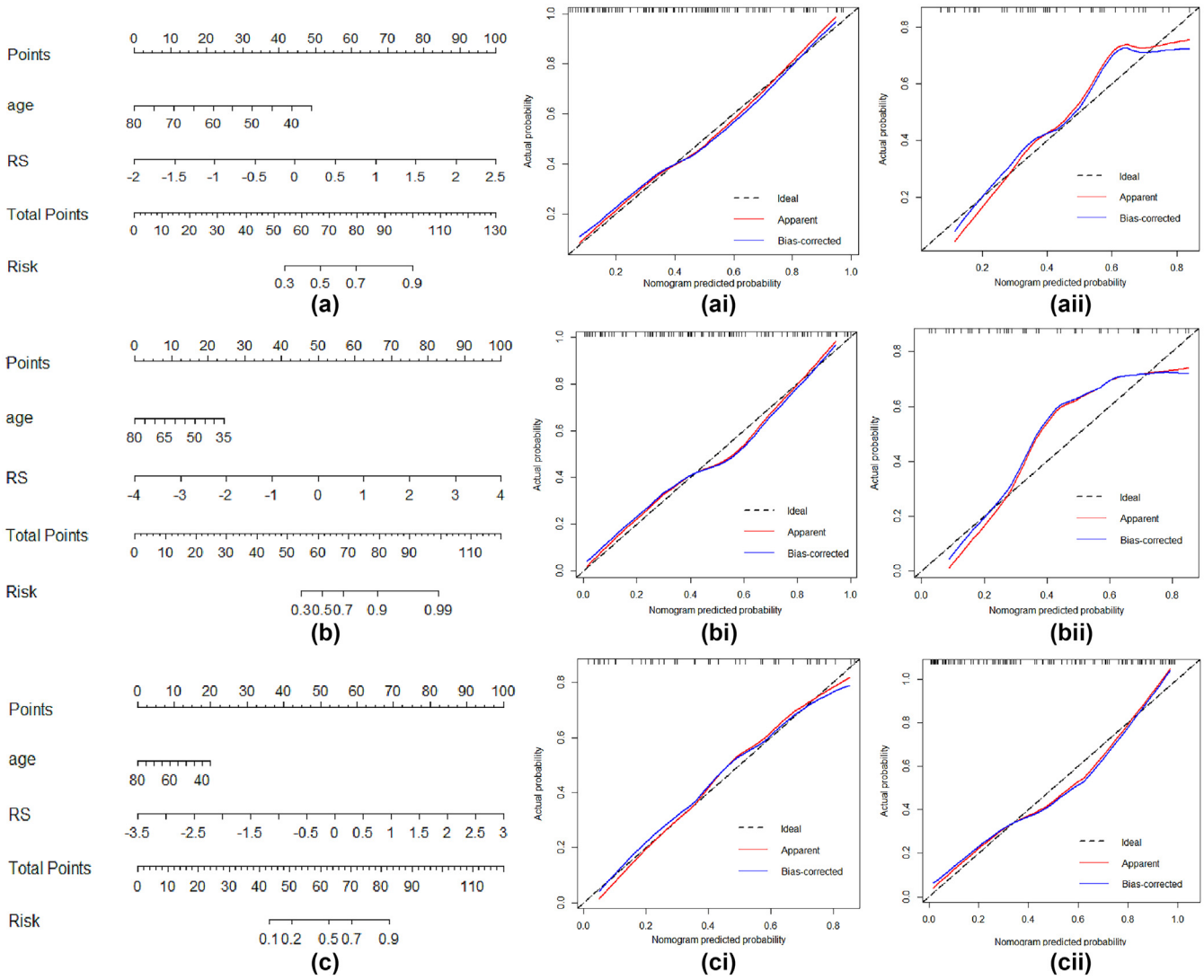


**Figure 4** (a, ai, aii) $FT_2$+age single-sequence model nomogram and training set, validation set calibration curves. (b, bi, bii) $FT_{2+1}$+age double-sequence model nomogram and training set, validation set calibration curves. (C, ci, cii) $FT_{2+2+1}$+age multisequence model nomogram and training set, validation set calibration curves. RS, Radiomics score.

help identify high-risk patients who might benefit from more aggressive interventions or closer monitoring, ultimately leading to better patient outcomes. However, implementing these models in clinical practice may pose challenges such as the need for standardisation of imaging protocols across different centres, ensuring the availability of advanced imaging software and computational resources, training radiologists and clinicians to interpret the nomogram outputs accurately, and integrating these approaches into existing clinical workflows.

The current study has certain limitations. First, although it is based on data from two centres, the sample size remains relatively small. Nevertheless, we have preliminarily demonstrated the feasibility of MRI-based radiomics in predicting HRCAs in patients with MM. Second, while MRI offers the potential advantage of noninvasive assessment of multiple lesions, our analysis was limited to an ROI from a single vertebral body. This limitation may overlook the full extent of tumour heterogeneity. Future studies should consider incorporating radiomic analysis of multiple lesions, which could provide a more comprehensive assessment of disease heterogeneity and enhance model accuracy. However, such an approach will require addressing challenges related to data acquisition, standardising ROI selection across multiple lesions, and managing the increased computational complexity. Lastly, due to the complex clinical staging of MM and the numerous factors influencing prognosis, using age alone as a clinical predictor may not be sufficient. Future research should therefore examine the contribution of other clinical parameters.

## Conclusion

In conclusion, MRI radiomics can predict HRCAs of MM patients. Radiomics nomograms based on the $FT_2$+age,

$FT_{2+1}$+age, and $FT_{2+2+1}$+age models represent powerful supportive tools for predicting HRCAs before treatment, among which the multi-sequence combined model of $FT_{2+2+1}$+age has the best predictive efficacy. Thus, it may be helpful for the early diagnostic assessment, therapeutic decision-making, and prognostic assessment.

## Ethics

Ethical approval (2023A-168) was obtained from our institutional ethics review board; the requirement for informed consent was waived due to the study's retrospective nature.

## Funding

This work was supported by the National Natural Science Foundation of China (82071872).

## Author contribution

S. Liu: Theoretical design, data processing, article writing, major revisions, approved of submission on behalf of all authors.
C. Liu: Data curation.
H. Pan: Formal analysis, Resources, Data curation.
S. Li: Theoretical guidance.
P. Teng: Data processing.
Z. Li: Statistical analysis.
J. Sun: Statistical analysis.
T. Ren: Data processing.
G. Liu: Theoretical guidance, article revision suggestions.
J. Zhou: Theoretical guidance, article revision suggestions, supportive contribution).
All authors agree they meet the current International Committee of Medical Journal Editors (ICMJE) criteria for Authorship.

## Conflict of interest

The authors declare no conflict of interest.

## Data availability

The datasets generated and analysed during the current study are not publicly available; however, they can be provided by the corresponding author upon reasonable request.

## Acknowledgments

The authors would like to thank Suwei Liu, Chang Liu, Haojie Pan, Shenglin Li, Peihong Teng, Zhengxiao Li, Jiachen Sun, Tiezhu Ren, Guifeng Liu, and Junlin Zhou for their support.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.crad.2024.106768.

## References

1. Siegel RL, Miller KD, Fuchs HE, *et al.* Cancer statistics, 2021. *CA Cancer J Clin* 2021;**71**:7–33.
2. Soekojo CY, Chung TH, Furqan MS, *et al.* Genomic characterization of functional high-risk multiple myeloma patients. *Blood Cancer J* 2022;**12**:24.
3. Callander NS, Baljevic M, Adekola K, *et al.* NCCN Guidelines® insights: multiple myeloma, version 3.2022. *J Natl Compr Canc Netw* 2022;**20**:8–19.
4. Neuse CJ, Lomas OC, Schliemann C, *et al.* Genome instability in multiple myeloma. *Leukemia* 2020;**34**:2887–97.
5. Morgan GJ, Davies FE, Gregory WM, *et al.* Cyclophosphamide, thalidomide, and dexamethasone as induction therapy for newly diagnosed multiple myeloma patients destined for autologous stem-cell transplantation: MRC Myeloma IX randomized trial results. *Haematologica* 2012;**97**:442–50.
6. Kumar SK, Jacobus SJ, Cohen AD, *et al.* Carfilzomib or bortezomib in combination with lenalidomide and dexamethasone for patients with newly diagnosed multiple myeloma without intention for immediate autologous stem-cell transplantation (ENDURANCE): a multicentre, open-label, phase 3, randomised, controlled trial. *Lancet Oncol* 2020;**21**:1317–30.
7. Smith SC, Althof PA, Dave BJ, *et al.* High-risk cytogenetics in multiple myeloma: further scrutiny of deletions within the IGH gene region enhances risk stratification. *Genes Chromosomes Cancer* 2020;**59**:569–74.
8. Ladetto M, Brüggemann M, Monitillo L, *et al.* Next-generation sequencing and real-time quantitative PCR for minimal residual disease detection in B-cell disorders. *Leukemia* 2014;**28**:1299–307.
9. Dimopoulos MA, Hillengass J, Usmani S, *et al.* Role of magnetic resonance imaging in the management of patients with multiple myeloma: a consensus statement. *J Clin Oncol* 2015;**33**:657–64.
10. Kosmala A, Weng AM, Krauss B, *et al.* Dual-energy CT of the bone marrow in multiple myeloma: diagnostic accuracy for quantitative differentiation of infiltration patterns. *Eur Radiol* 2018;**28**:5083–90.
11. Moulopoulos LA, Varma DG, Dimopoulos MA, *et al.* Multiple myeloma: spinal MR imaging in patients with untreated newly diagnosed disease. *Radiology* 1992;**185**:833–40.
12. Tagliafico AS, Cea M, Rossi F, *et al.* Differentiating diffuse from focal pattern on Computed Tomography in multiple myeloma: added value of a Radiomics approach. *Eur J Radiol* 2019;**121**:108739.
13. Gardin I, Grégoire V, Gibon D, *et al.* Radiomics: principles and radiotherapy applications. *Crit Rev Oncol Hematol* 2019;**138**:44–50.
14. Guiot J, Vaidyanathan A, Deprez L, *et al.* A review in radiomics: making personalized medicine a reality via routine imaging. *Med Res Rev* 2022;**42**:426–40.
15. Liu J, Zeng P, Guo W, *et al.* Prediction of high-risk cytogenetic status in multiple myeloma based on magnetic resonance imaging: utility of radiomics and comparison of machine learning methods. *J Magn Reson Imaging* 2021;**54**:1303–11.
16. Liu J, Wang C, Guo W, *et al.* A preliminary study using spinal MRI-based radiomics to predict high-risk cytogenetic abnormalities in multiple myeloma. *Radiol Med* 2021;**126**:1226–35.
17. Pei Q, Yi X, Chen C, *et al.* Pre-treatment CT-based radiomics nomogram for predicting microsatellite instability status in colorectal cancer. *Eur Radiol* 2022;**32**:714–24.
18. Lambin P, Rios-Velazquez E, Leijenaar R, *et al.* Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 2012;**48**:441–6.
19. Panaroni C, Yee AJ, Raje NS. Myeloma and bone disease. *Curr Osteoporos Rep* 2017;**15**:483–98.
20. Chang E, Joel MZ, Chang HY, *et al.* Comparison of radiomic feature aggregation methods for patients with multiple tumors. *Sci Rep* 2021;**11**:9758.

21. Shiri I, Amini M, Nazari M, *et al*. Impact of feature harmonization on radiogenomics analysis: prediction of EGFR and KRAS mutations from non-small cell lung cancer PET/CT images. *Comput Biol Med* 2022;**142**:105230.
22. Rahmouni A, Divine M, Mathieu D, *et al*. Detection of multiple myeloma involving the spine: efficacy of fat-suppression and contrast-enhanced MR imaging. *AJR Am J Roentgenol* 1993;**160**:1049–52.
23. Wang X, Wan Q, Chen H, *et al*. Classification of pulmonary lesion based on multiparametric MRI: utility of radiomics and comparison of machine learning methods. *Eur Radiol* 2020;**30**:4595–605.
24. Zamani-Ahmadmahmudi M, Nassiri SM, Soltaninezhad F. Development of an RNA sequencing-based prognostic gene signature in multiple myeloma. *Br J Haematol* 2021;**192**:310–21.
25. Yin P, Mao N, Liu X, *et al*. Can clinical radiomics nomogram based on 3D multiparametric MRI features and clinical characteristics estimate early recurrence of pelvic chondrosarcoma? *J Magn Reson Imaging* 2020;**51**:435–45.
26. Shen Y, Xu F, Zhu W, *et al*. Multiclassifier fusion based on radiomics features for the prediction of benign and malignant primary pulmonary solid nodules. *Ann Transl Med* 2020;**8**:171.
27. Mao N, Yin P, Wang Q, *et al*. Added value of radiomics on mammography for breast cancer diagnosis: a feasibility study. *J Am Coll Radiol* 2019;**16**:485–91.
28. Pan Y, Lei X, Zhang Y. Association predictions of genomics, proteinomics, transcriptomics, microbiome, metabolomics, pathomics, radiomics, drug, symptoms, environment factor, and disease networks: a comprehensive approach. *Med Res Rev* 2022;**42**:441–61.
29. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* 2016;**278**:563–77.
30. Yi X, Guan X, Zhang Y, *et al*. Radiomics improves efficiency for differentiating subclinical pheochromocytoma from lipid-poor adenoma: a predictive, preventive and personalized medical approach in adrenal incidentalomas. *Epma J* 2018;**9**:421–9.
31. Rogers W, Thulasi Seetha S, Refaee TAG, *et al*. Radiomics: from qualitative to quantitative imaging. *Br J Radiol* 2020;**93**:20190948.