UNIVERSITA' DEGLI STUDI DI GENOVA

CORSO DI DOTTORATO IN BIOTECNOLOGIE IN MEDICINA TRASLAZIONALE

Coordinatore del Corso: Chiar.mo Prof. Rodolfo Quarto

CURRICULUM IN BIOTECNOLOGIE IN MEDICINA TRASLAZIONALE

Referente di Curriculum: Prof. Alberto Tagliafico

XXXIV CICLO

TESI DI DOTTORATO

Radiomics and Artificial Intelligence for Outcome Prediction in Multiple Myeloma Patients

Tutor:
Prof. Alberto Tagliafico

Dottorando:
Dott.ssa Bianca Bignotti

Anno Accademico 2020-2021

Alla mia famiglia

"Una buona pratica, preliminare a qualunque altra, è la pratica della meraviglia.

Esercitarsi a non sapere e a meravigliarsi."

Chandra Candiani

TABLE OF CONTENTS

BACKGROUND……1

Purpose of the Research……3

CHAPTER 1: FEASIBILITY STUDY

1.1 Introduction……4

CHAPTER 2: METHODOLOGY

2.1 Study Populations, Inclusion Criteria, and Risk Stratification……5

2.2 Image Analysis……7

2.3 Reduction of Redundancy……8

2.4 Clustering……9

CHAPTER 3: RESULTS

3.1 Clinical Findings……11

3.2 AI-Based Analysis……11

3.3. Feature Ranking……13

DISCUSSION……14

CONCLUSION……17

REFERENCES……19

PROVISIONAL RESULTS……24

RESEARCH ACTIVITY

ACKNOWLEDGEMENTS

BACKGROUND

Nowadays, radiomics [1] is one of the most promising techniques, with the potential to improve cancer treatment. Radiomics is an advanced, quantitative feature-based methodology for image analysis defined as the conversion of images to higher dimensional data and the subsequent mining of these data for improved decision support [1]. Plasma cell dyscrasias (PCDs) are pathological conditions including Monoclonal Gammopathy of Undetermined Significance (MGUS), Smoldering Multiple Myeloma (SMM), and full-blown Multiple Myeloma (MM). MM definition can rely on the International Myeloma Working Group (IMWG) consensus updates, yet the significant clinical heterogeneity of MM patients implies that, as far as now, we lack a set of consolidated biomarkers able to predict the outcome and risk of progression, independently from the therapeutic approach. Risk stratification gold standard is usually performed at diagnosis, by means of the International Staging System (ISS) (which combines serum beta2-microglobulin and serum albumin for a three-stage classification), and cytogenetics (which provides a binary normal-high risk stadiation). In 2003, IMWG replaced the Durie–Salmon system with a revised version (Durie–Salmon system plus), replacing radiography with Magnetic Resonance Imaging (MRI) and PET/CT data with [18F] Fluorodeoxyglucose (FDG) as tracer. The extent of the bone disease is negatively related to a decreased quality of life and bone disease in MM increases morbidity and mortality. Therefore, the detection of lytic bone lesions on imaging, especially CT and MRI, is becoming crucial from the clinical viewpoint to separate asymptomatic from symptomatic MM patients; meanwhile, the detection of focal lytic lesions is becoming relevant even when no clinical symptoms are present [3]. Therefore, an analysis of multiple bone lesions could be performed using AI and radiomics. A recent application of

radiomics in MM, by our research group showed that, in clinical practice, radiomics improves the radiological evaluation of focal and diffuse pattern on CT by increasing the Area Under the Curve (AUC) of radiologists [2]. Accuracy in terms of the AUC of radiologists compared to the reference standard was lower (64%) than the accuracy computed using a radiomics approach, which obtained a maximum value of 79%. However, the diagnostic and prognostic capabilities of medical imaging in MM are still under investigation and development. Significant variability in image-based prognostic scores is present among different centers and in clinical practice [3-8]. In addition, although the updated version of the IMWG criteria accepts the use of CT and PET/CT to diagnose lytic bone disease in MM, there is still a lack of reliable quantitative and computational tools for increasing the prognostic value of these modern imaging modalities [3-8].

The study hypothesis is that, if aided by AI-based methods, data extracted from clinical images used in the routine clinical practice for MM can predict its outcome, with specific focus on the identification of patients at high-risk of progression or non-responding to current therapy.

Purpose of the research

The aim of our research activities is to:

- create comprehensive, solid and validated quantitative scores for CT and MRI and PET-CT that can be adopted in clinical practice.

- To introduce a set of novel biomarkers extracted from X-ray Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) data by means of reliable and sophisticated artificial intelligence tools.

- To introduce a personalized predictive process for MM integrating quantitative CT and MRI and machine learning, which relies on the application of AI processes on radiomics features in order to forecast the disease behaviour.

CHAPTER 1: FEASIBILITY STUDY

1.1 Introduction

The limitations of the use of imaging for MM assessment are essentially due to three open issues: the lack of accuracy in differentiating focal from diffuse patterns, the difficulty in extracting reliable prognostic biomarkers from pattern allocation,and the low agreement in staging MM patients based on imaging outcomes.

The application of pattern recognition algorithms for the extraction of radiomics descriptors from images of MM patients and the post-processing of such radiomics features by means of procedures based on artificial intelligence (AI) are nowadays introducing a novel approach for increasing the reliability of imaging in MM clinical assessment [4,6,9,10].

The objective of the present study is to assess the feasibility of an AI-based approach for the automatic stratification of MM patients from CT data, and for the automatic identification of radiological biomarkers with a possible prognostic value. Specifically, relying on radiomics and AI-based computational analysis [9,11,12], this feasibility study shows that a set of descriptors of the focal lesions in MM X-ray CT at diagnosis allows for the automatic stratification of a cohort of MM patients who have undergone transplantation in two clusters, whose characteristics can be interpreted via comparison with clinical data, biological biomarkers, and the clinical outcome of the disease.

CHAPTER 2: METHODOLOGY

2.1 Study Populations, Inclusion Criteria, and Risk Stratification

This study was performed according to the Declaration of Helsinki and the International Conference on Harmonization of Good Clinical Practice Guidelines. An institutional review board was obtained (054REG2019). All patients signed informed consent for retrospective research before CT examination; data collection did not influence patient care. We considered 51 consecutive patients (mean age, 56 years ± 8; range, 31–73 years; 18 females; 33 males) admitted to the Hospital in the last five years because of biopsy confirmed MM. Inclusion criteria were baseline whole-body CT from the Hospital PACS or outpatient clinic. Among these 51 patients, we selected the 33 presenting at least one focal lesion in one of the CT slices, i.e., at least one >5 mm lytic lesion in the axial or extra-axial skeleton [16-18]. Two radiologists blinded to the diagnosis and to each other's conclusion assessed whether the CT pattern was diffuse or focal, and, for each patient presenting at least one focal lesion, we identified the largest one.

Risk stratification was performed at diagnosis by the Revised International Staging System (ISS) combining serum beta2-microglobulin and serum albumin, lactate dehydrogenase for three-stage classification, and cytogenetics determining a binary normal-high risk stadiation [13,14]. Table 1 provides a summary of the clinical features (diameter of focal lesion: mean: 19.9 mm, STD: 13.4 mm, min: 4.5 mm, max: 62.4 mm).

Table 1. Clinical features of the 33 MM patients included in the analysis. R-ISS stage: I: ISS stage I and standard-risk CA by iFISH and normal LDH. II: Not R-ISS stage I or III; III: ISS stage III and either high-risk CA by iFISH or high LDH. CA—chromosomal abnormalities; iFISH—interphase fluorescent in situ hybridization; ISS—International Staging System; LDH—lactate dehydrogenase; MM—multiple myeloma; R-ISS—revised International Staging System

Characteristic	Number	%
Patients	33	100
Age (years) Mean	56	
Age SD [1]	6.7	
Males	21	66.4
Females	12	34.6
Cytogenetics		
Normal	22	66,7
High risk	11	33,3
Relapsed	17/33	51,5
Days before Relapse (mean)	1138	
Days of follow-up (mean)	1317	
International Staging System		
Stage I	15	45.4
Stage II	9	27.3
Stage III	9	27.3

[1] Standard Deviation.

2.2 Image Analysis

To compute each patient's overall skeletal asset, we utilized a published software tool(Bone-GUI, http://mida.dima.unige.it/software/bone-gui/; accessed on 20 September 2021) combining thresholding and active contours. For each subject, Bone-GUI provided 24 features. Separately for the whole, axial, and skeleton districts, it computed the following: the mean medullary Hounsfield value with standard deviation, the volume of the global medullary asset, the mean cortical Hounsfield value with standard deviation, the volume of the cortical asset, the rate of volume occupied by the medullary tissue, and the overall volume. We also applied an open source tool for radiomics (Slicer, https://www.radiomics.io/slicerradiomics.html; accessed on 20 September 2021) [4,6] to the 33 lytic lesions on the compact bone tissue to extract 109 Slicer features for each focal lesion.

2.3 Reduction of Redundancy

Our AI-based analysis for patients' stratification utilized Slicer features as the input. To reduce information redundancy, we considered two approaches. In the first approach, principal component analysis (PCA) [19] projected the feature space onto a principal components' subspace explaining at least 80% of the data variance. In the second approach, we performed two Pearson's correlation processes ($p > 95\%$) involving the Slicer features and (a) the binary feature encoding patient's relapse one year after transplantation, and (b) all 24 Bone-GUI features. We applied PCA to the features selected using the two correlation processes. **Figure 1** illustrates this redundancy reduction pipeline.

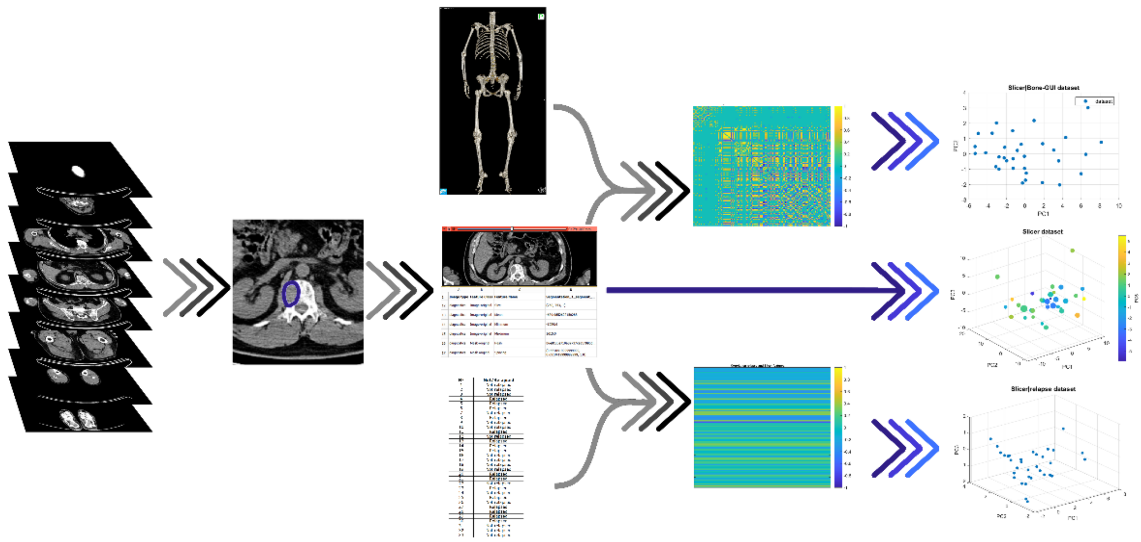


Figure 1. The pipeline of the radiomics features analysis. For each patient, the focal lesion was pointed out and the corresponding CT image was fed into a radiomics tool (Slicer), which computed 109 radiomics features; these descriptors were correlated with both the clinical outcome of the disease at one year, and the global radiological features extracted by means of a segmentation tool (Bone-GUI); the resulting mostly correlated features and the set of all local features.

2.4 Clustering

Clustering organized a set of unlabeled samples into clusters based on data similarity [20]. Data partition was obtained by minimizing a cost function involving the distances between the data and cluster prototypes. In Fuzzy C-Means (FCM) a degree of membership is assigned to each sample with respect to each cluster. In addition to FCM, we applied a non-linear approach based on the filtering of an extended version of the Hough transform (HTF) [21], according to the following steps:

1. Downstream of the PCA process, the two-dimensional feature space given by the two components explaining most of the data variance (namely, PC1 and PC2) was constructed for each data set.

2. Given a feature space, the Hough transform of each point in the patient's set with respect to the family of all parabolas was computed. As this family was characterized by three parameters, i.e., its equation is y_PC2 = ax_PC1^2 + bx_PC1 + c, with a, b, and c being the parameters, and the corresponding parameter space has three dimension.

3. The Hough accumulator was computed by counting the number of times each Hough transform passed through one of the cells of the discretized parameter space.

4. The Hough accumulator was filtered by a 5-pixel-side cube centered on the pixel with a maximum grey value. This cube was the smallest one enclosing the cells, with accumulator values higher than 50% of the maximum [22].

Each line passing through the filtered region was projected back to the image space,

thus generating a cluster of points in a strip around the parabola corresponding to the maximum in the Hough accumulator. The remaining points represent the second cluster made of points outside of the strip of parabolas previously identified.

CHAPTER 3: RESULTS

3.1 Clinical Findings

Focal lesion searching led to the selection of 33/51 (65%) patients (mean age, 56 years ± 7; range, 45–69 years; 12 females; 21 males) whose imaging data were considered for our computational analysis. Inter-observer agreement in differentiating diffuse from focal pattern between the two groups of radiologists resulted in 0.75 (95% Confidence Interval: 0.31–0.67) and 0.96 (95% Confidence Interval: 0.79–0.99) for the selection of patients with focal lesions.

3.2 AI-Based Analysis

The AI-based analysis involved three data sets: data set 1, made of all 109 local features extracted by Slicer from each focal lesion; data set 2, made of the eight local features mostly correlating with the relapsed/non-relapsed binary feature; and data set 3, made of the 17 local features mostly correlating with the 24 Bone-GUI global features. The application of PCA to these three data sets led to three features spaces, with n = 5 axes for data set 1, n = 3 axes for data set 2, and n = 2 axes for data set 3.

In each one of these three feature spaces, FCM and HTF computed two clusters:

in each cluster, the black circles are associated with patients that underwent relapse within one year of bone marrow transplantation. Cluster A (B) contained the maximum (minimum) number of relapsed patients.

In order to assess the performances of the clustering algorithms, we computed the confusion matrices for the observed relapsed patients; specifically, we counted the number of true positives (TPs), true negatives (TNs), false positives (FPs), and false negatives (FNs) using cluster A as the reference cluster for the “relapsed” class and

cluster B as the reference cluster for the "non-relapsed" class. Using the entries of such matrices, we computed four different skill scores:

Sensitivity = TP/(TP + FN)

Specificity = TN/(TN + FP)

Youden's index = Sensitivity + Specificity − 1

Critical Success Index (CSI) = TP/(TP + FN + FP).

We show that the CSI ranged from 0 to 1 and it was higher as much as the number of FPs and FNs was small, regardless the number of TNs. CSI is therefore a useful score in conditions like the one we considered here, where we had an unbalanced data set with more non-relapsed cases than relapsed ones.

We tested the robustness of our results by performing a bootstrap analysis on the set 33 17-dimension feature vectors of that set. We constructed 100 random realizations of training sets made of 20 feature vectors (of which 10 representing relapsed patients) and, for each realization, we applied the HTF clustering process. Then, for each realization of the training set, we computed the membership cluster for each one of the remaining 13 vectors representing the test set. Repeating this procedure for each one of the 100 realizations of the training-test set pairs led to the construction of 100 confusion matrices and, therefore, to 100 sets of skill score values that we averaged, together with the corresponding standard deviations. We also performed a bootstrap analysis on the cytogenetics values. In order to compute the entries of these last confusion matrices, we compared the relapse/non-relapse with the high/standard cytogenetic stages: a relapsed patient with a "high" cytogenetic stage was a TP event, while a relapsed patient with a "standard" cytogenetic stage was an FN. A non-relapsed patient with a "standard" cytogenetic stage was a TN event and a non-relapsed patient with a "high" cytogenetic stage

was an FP event. We show that the separation between the standard and high cytogenetic stage was realized according to the standard cytogenetic evaluation for separating patients with a high-risk mutation (poor prognosis in general) from patients without high-risk mutations [14,23].

3.3 Feature Ranking

To investigate which radiomics features mostly contribute to an effective stratification of the MM patients, we focused on the case of data set 3. The reason for this choice is because, when analyzed with HTF, this set provided, by far, the highest sensitivity values and, significantly, the highest CSI values among the three data sets considered. Therefore, we analyzed the feature compositions of the two axes produced by the application of PCA on the original feature space of this data set, made of 17 features. These contributions were weighted by the percentage of explained variance of the two PCs (77% and 9% for the first and second PC, respectively). A Mann–Whitney U-test on these features showed that just three of them did not pass the null hypothesis ($p > 99\%$): "MaskMaximum", which denotes the maximum grey level value in the mask segmenting the focal lesion (172.6 ± 64.4 in Cluster A; 321.9 ± 48.6 in Cluster B); "firstorderRange", which denotes the range of the distribution of the voxel intensities (194.7 ± 61.8 in Cluster A; 343.4 ± 66.9 in Cluster B); and "ngtdmComplexity" (29.8 ± 24.9 in Cluster A; 79.4 ± 43.5 in Cluster B), which is a measure of the non-uniformity of the lesion image in the grey level intensity.

DISCUSSION

This study demonstrates that AI supported radiomics realize a clustering of MM patients with a statistical reliability that, for some skill scores, is higher than the one provided by standard biochemical staging. The possibility to increase the predictive potential of the standard CT images of patients with multiple myeloma is clinically relevant for several reasons.

The first is that although MM is still considered a single disease, it is actually a collection of several different cytogenetically distinct plasma cell malignancies [24]. Trisomies and IgH translocations are considered primary cytogenetic abnormalities, and occur at the time of establishment of MGUS [24]. At the present time, there are three specific biomarkers for MM with an approximately 80% risk of progression to symptomatic endorgan damage in two or more independent studies: clonal bone marrow plasma cells ≥60%, serum free light chain (FLC) ratio ≥100 (provided involved FLC level is ≥100 mg/L), and more than one focal lesion on magnetic resonance imaging (MRI). It is known that almost all patients with MM eventually relapse and the choice of a treatment regimen at relapse is affected by many factors, including the timing of relapse, response to prior therapy, aggressiveness of relapse, and performance status (TRAP) [24].

Therefore, the prediction of relapse early is important to foresee a therapy. Second, several studies have correlated bone patterns in MM with their prognostic value using MRI and CT [4,6,11,25]. MRI can be used to differentiate up to five different patterns of plasma cell infiltration, including normal appearance, focal involvement, homogeneous diffuse infiltration, diffuse infiltration with additional focal lesions, and variegated or salt-and-pepper patterns; on the other hand, CT is well suited for small (below 5 mm) focal bone lesions due to its high spatial resolution capabilities.

The AI-based analysis of the radiomics properties extracted from the focal lesions essentially pointed out two aspects. First, the redundancy of the radiomics features seem to impact the prognostic power of the clustering methods. However, the stratification power increases when correlation-based and PCA-based reduction of redundancy processes are applied. Second, the use of a non-linear approach to clustering, namely HTF, seems to provide better results with respect to a more standard fuzzy clustering algorithm; this may be explained because of the high degree of heterogeneity that characterizes MM.

The skill scores computed for each data set and each classification method helped us to determine which approach to redundancy reduction and which algorithm performs better for stratification purposes. Among the four skill scores, CSI probably represents the one that best interprets the outcomes of the confusion matrices in this context. Indeed, this score emphasizes the correct prediction of relapses in correspondence with a low rate of misclassification. Interestingly, the application of HTF on the focal features mostly correlating with the skeleton asset's global properties (which are extracted by Bone-GUI) leads to the highest value for this score: this seems to point out a favorable prognostic role for the interplay between local and global descriptors of the MM bone tissue. In this case,

the CSI value is higher than the discriminative value provided by the cytogenetic data, which supports the reliability of radiomics as a prognostic tool for MM clinical practice.This conclusion is confirmed by a bootstrap analysis performed on data set 3. Data set 3 is made of the focal descriptors that mostly correlate with the whole skeleton's asset properties. Therefore, this correlation analysis per se realizes a feature selection process whose outcome is a set of 17 features. A finer feature selection is provided by PCA. This figure and the related Mann–Whitney U-test point

to a significant emphasis on properties related to the heterogeneity of the focal lesion, such as the Hounsfield unit range and maximum values found in the lesion, and the complexity, which measures the non-uniformity of the image and the presence of rapid changes in intensity.

We finally show that the data collection for this study has been realized by means of a single, specific CT scanner, so that the images we used for feature extraction were homogeneous. Recent studies [26] have shown that the characteristics of the extracted features may depend on non-tumor related factors like the signal-to-noise ratio of the experimental data. Therefore, in the case of studies that utilize data from more than one scanner, data homogenization should be implemented prior to the data extraction process [27].

CONCLUSIONS

This computational approach to the interpretation of radiomics focal features shows the potential for the stratification of relapsed and non-relapsed MM patients, and could represent a prognostic procedure for determining the disease follow-up and therapy.

Concerning the technical issues to be discussed, the present study has several strengths: the use of clinically available CT images collected in the normal daily workup did not influenced patient care in any way. Second, we used a free open-source tool for radiomics assessment of the focal lytic lesions. Among the limitations of the present study, we acknowledge the retrospective nature, which did not allow for perfect timing between CT, diagnosis, and therapy or relapse. In addition, the evaluation of the radiomics features was made only with one open-source tool, and we did not evaluate whether the usage of other tools would have introduced variability to a significant extent. Finally, the overall number of patients included was relatively low: indeed, a correct sample size in radiomics is at least five times the number of extracted features [28], and this condition would require a population of at least 100 MM patients. Nonetheless, the possibility to obtain a cluster of features to identify relapses even in a 33 patient sample is in favor of the validity of this method. This initial study warrants prospective studies with a high number of patients, which are currently underway, in order to validate this approach, with the aim of implementing, it in a more systematic way, a method of obtaining a more robust prognostic score for MM patients.

Summing up the results of this study, we remind that our objective was to validate the feasibility of the automatic stratification of MM patients by means of an analysis of the descriptors extracted fromCT data within the framework of a radiomics

retrospective study. This analysis showed that unsupervised AI can predict relapse within one year after transplantation and can identify a few imaging features associated with the heterogeneity of the focal lesion with a high prognostic value.

References

1. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology. 2016;278(2):563-77. doi: 10.1148/radiol.2015151169.

2. Tagliafico AS, Belgioia L, Bonsignore A, Signori A, Formica M, Rossi F, Piana M, Schenone D, Dominietto A. Development and definition of a simplified scoring system in patients with multiple myeloma undergoing stem cells transplantation on standard computed tomography: myeloma spine and bone damage score (MSBDS). Cancer Imaging. 2020 Apr 28;20(1):31. doi: 10.1186/s40644-020-00306-1. PMID: 32345379; PMCID: PMC7189746.

3. Tagliafico, A.S.; Dominietto, A.; Belgioia, L.; Campi, C.; Schenone, D.; Piana, M. Quantitative Imaging and Radiomics in Multiple Myeloma: A Potential Opportunity? Medicina 2021, 57, 94. https://doi.org/10.3390/medicina57020094

4. Tagliafico AS, Cea M, Rossi F, Valdora F, Bignotti B, Succio G, Gualco S, Conte A, Dominietto A. Differentiating diffuse from focal pattern on Computed Tomography in multiple myeloma: Added value of a Radiomics approach. Eur J Radiol. 2019 Dec;121:108739. doi: 10.1016/j.ejrad.2019.108739. Epub 2019 Nov 7. PMID: 31733431.

5. Rossi F, Torri L, Dominietto A, Tagliafico AS. Spectrum of magnetic resonance imaging findings in transplanted multiple myeloma patients with hip/pelvic pain (according to MY-RADS): A single center experience. Eur J Radiol. 2020 Sep;130:109154. doi: 10.1016/j.ejrad.2020.109154. Epub 2020 Jun 24. PMID: 32629214.

6. Schenone, D, Lai, R, Cea M, Rossi F, Torri L, Bignotti B, et al. Radiomics and artificial intelligence analysis of CT data for the identification of prognostic features in multiple myeloma. Med Imaging 2020 11314, 113144A

7. Valdora F, Houssami N, Rossi F, Calabrese M, Tagliafico AS. Rapid review: radiomics and breast cancer. Breast Cancer Res Treat. 2018 Feb 2. doi: 10.1007/s10549-018-4675-4.

8. Limkin EJ, Sun R, Dercle L, Zacharaki EI, Robert C, Reuzé S, Schernberg A, Paragios N, Deutsch E, Ferté C. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. Ann Oncol. 2017 Jun 1;28(6):1191-1206. doi: 10.1093/annonc/mdx034.

9. Ekert, K.; Hinterleitner, C.; Baumgartner, K.; Fritz, J.; Horger, M. Extended Texture Analysis of Non-Enhanced Whole-Body MRI Image Data for Response Assessment in Multiple Myeloma Patients Undergoing Systemic Therapy. Cancers 2020, 12, 761.

10. Morvan, L.; Nanni, C.; Michaud, A.-V.; Jamet, B. Learned Deep Radiomics for Survival Analysis with Attention. In Proceedings of the International Workshop on Predictive Intelligence in Medicine (PRIME 2020), Lima, Peru, 8 October 2020; Springer: Berlin/Heidelberg, Germany, 2020; pp. 35–45.

11. Reinert, C.P.; Krieg, E.-M.; Bösmüller, H.; Horger, M. Mid-term response assessment in multiple myeloma using a texture analysis approach on dual energy-CT-derived bone marrow images—A proof of principle study. Eur. J. Radiol. 2020, 131, 109214.

12. Jamet, B.; Morvan, L.; Nanni, C.; Michaud, A.-V.; Bailly, C.; Chauvie, S.;Moreau, P.; Touzeau, C.; Zamagni, E.; Bodet-Milin, C.; et al. Random survival forest to predict transplant-eligible newly diagnosed multiple myeloma outcome including FDG-PET radiomics: A combined analysis of two independent prospective European trials. Eur. J. Nucl. Med. Mol. Imaging 2020, 48, 1005–1015.

13. Joseph, N.S.; Gentili, S.; Kaufman, J.L.; Lonial, S.; Nooka, A.K. High-risk Multiple Myeloma: Definition and Management. Clin. Lymphoma Myeloma Leuk. 2017, 17S, S80–S87.

14. Palumbo, A.; Avet-Loiseau, H.; Oliva, S.; Lokhorst, H.M.; Goldschmidt, H.; Rosinol, L.; Richardson, P.; Caltagirone, S.; Lahuerta, J.J.; Facon, T.; et al. Revised International Staging System for Multiple Myeloma: A Report From International Myeloma Working

Group. J. Clin. Oncol. 2015, 33, 2863–2869.

15. Fiz, F.; Marini, C.; Piva, R.; Miglino, M.; Massollo, M.; Bongioanni, F.; Morbelli, S.; Bottoni, G.; Campi, C.; Bacigalupo, A.; et al.

Adult Advanced Chronic Lymphocytic Leukemia: Computational Analysis of Whole-Body CT Documents a Bone Structure Alteration. Radiology 2014, 271, 805–813.

16. Kapur, T.; Pieper, S.; Fedorov, A.; Fillion-Robin, J.-C.; Halle, M.; O'Donnell, L.; Lasso, A.; Ungi, T.; Pinter, C.; Finet, J.; et al.

Increasing the impact of medical image computing using community-based open-access hackathons: The NA-MIC and 3D Slicer experience. Med. Image Anal. 2016, 33, 176–180.

17. Kikinis, R.; Pieper, S.; Vosburgh, K. 3D Slicer: A Platform for Subject-Specific Image Analysis, Visualization, and Clinical

Support. In Intraoperative Imaging and Image-Guided Therapy; Springer: Berlin/Heidelberg, Germany, 2014; Volume 3, pp. 277–289.

18. Fedorov, A.; Beichel, R.; Kalpathy-Cramer, J.; Finet, J.; Fillion-Robin, J.-C.; Pujol, S.; Bauer, C.; Jennings, D.; Fennessy, F.;

Sonka, M.; et al. 3D Slicer as an Image Computing Platform for the Quantitative Imaging Network. Magn. Reason. Imaging 2012,

30, 1323–1341.

19. Jolliffe, I.T. Principal Component Analysis, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2002.

20. Bezdek, J.C. Pattern Recognition with Fuzzy Objective Function Algorithms; Kluwer Academic Publishers: Cambridge, MA, USA, 1981.

21. Beltrametti, M.C.; Massone, A.M.; Piana, M. Hough Transform of Special Classes of Curves. SIAM J. Imaging Sci. 2013, 6, 391–412.

22. Massone, A.M.; Perasso, A.; Campi, C.; Beltrametti, M.C. Profile Detection in Medical and Astronomical Images by Means of the Hough Transform of Special Classes of Curves. J. Math. Imaging Vis. 2015, 51, 296–310.

23. Roche-Lestienne, C.; Boudry-Labis, E.; Mozziconacci, M.J. Cytogenetics in the management of “chronic myeloid leukemia”: An update by the Groupe francophone de cytogénétique hématologique (GFCH). Ann. Biol. Clin. 2016, 74, 511–515.

24. Rajkumar, S.V. Multiple myeloma: Every year a new standard? Hematol Oncol. 2019, 37 (Suppl. 1), 62–65.

25. Kobayashi, H.; Abe, Y.; Narita, K.; Kitadate, A.; Takeuchi, M.; Matsue, K. Prognostic Significance of Medullary Abnormalities of the Appendicular Skeleton Detected by Low-Dose Whole-Body Multidetector Computed Tomography in Patients with Multiple

Myeloma. Blood 2017, 130 (Suppl. 1), 1763.

26. Zhovannik, I.; Bussinik, J.; Traverso, A.; Shi, Z.; Kalendralis, P.;Wee, L.; Dekker, A.; Fijiten, R.; Monshouwet, R. Learning from scanners: Bias reduction and feature correction in radiomics. Clin. Transl. Radiat. Oncol. 2019, 16, 33–38.

27. Einstein, S.A.; Rong, X.J.; Jensen, C.T.; Liu, X. Quantification and homogenization of image noise between two CT scanner models. J. Appl. Clin. Med. Phys. 2020, 21, 174–178.

28. Sollini, M.; Antunovic, L.; Chiti, A.; Kirienko, M. Towards clinical application of image mining: A systematic review on artificial intelligence and radiomics. Eur. J. Nucl. Med. Mol. Imaging 2019, 46, 2656–2672.

PROVISIONAL RESULTS

1. Tagliafico AS, Rossi F, Bignotti B, Torri L, Bonsignore A, Belgioia L, Domineitto A. CT-derived relationship between low relative muscle mass and bone damage in patients with multiple myeloma undergoing stem cells transplantation. Br J Radiol. 2021 Dec 21:20210923. doi: 10.1259/bjr.20210923.

2. Schenone D, Dominietto A, Campi C, Frassoni F, Cea M, Aquino S, Angelucci E, Rossi F, Torri L, Bignotti B, Tagliafico AS, Piana M. Radiomics and Artificial Intelligence for Outcome Prediction in Multiple Myeloma Patients Undergoing Autologous Transplantation: A Feasibility Study with CT Data. Diagnostics (Basel). 2021 Sep 24;11(10):1759. doi: 10.3390/diagnostics11101759.

3. Tagliafico AS, Belgioia L, Bonsignore A, Rossi F, Succio G, Bignotti B, Dominietto A. Subspecialty Second-Opinion in Multiple Myeloma CT: Emphasis on Clinically Significant Lytic Lesions. Medicina (Kaunas). 2020 Apr 23;56(4):195. doi: 10.3390/medicina56040195.

4. Tagliafico AS, Cea M, Rossi F, Valdora F, Bignotti B, Succio G, Gualco S, Conte A, Dominietto A. Differentiating diffuse from focal pattern on Computed Tomography in multiple myeloma: Added value of a Radiomics approach. Eur J Radiol. 2019 Dec;121:108739. doi: 10.1016/j.ejrad.2019.108739.

BJR

https://doi.org/10.1259/bjr.20210923

Received: 09 August 2021 | Revised: 02 December 2021 | Accepted: 07 December 2021

Cite this article as:
Tagliafico AS, Rossi F, Bignotti B, Torri L, Bonsignore A, Belgioia L, et al. CT-derived relationship between low relative muscle mass and bone damage in patients with multiple myeloma undergoing stem cells transplantation. *Br J Radiol* 2021; **94**: 20210923.

FULL PAPER

CT-derived relationship between low relative muscle mass and bone damage in patients with multiple myeloma undergoing stem cells transplantation

[1,2]ALBERTO STEFANO TAGLIAFICO, MD, [3,4]FEDERICA ROSSI, MD, [1,4]BIANCA BIGNOTTI, MD, [5]LORENZO TORRI, MD, [1,2]ALESSANDRO BONSIGNORE, MD, [1,2]LILIANA BELGIOIA, MD and [1]ALIDA DOMINEITTO, MD

[1]IRCCS Ospedale Policlinico San Martino, Genoa, Italy
[2]Department of Health Sciences (DISSAL), University of Genoa, Genoa, Italy
[3]Ospedale Santa Corona, Pietra Ligura, Italy
[4]Department of Experimental Medicine (DIMES), University of Genoa, Genoa, Italy
[5]Vascular Surgery Unit, Azienda Ospedaliero Universitaria Pisana, Pisa, Italy

Address correspondence to: Dr Alberto Stefano Tagliafico
E-mail: *atagliafico@sirm.org;alberto.tagliafico@unige.it*

Objective: Sarcopenia or low muscle mass is related to worse prognosis in cancer patients. We investigated whether muscle mass is related to bone damage on CT in patients with multiple myeloma (MM).

Methods: Approval from the institutional review board was obtained. *N* = 74 consecutive patients (mean age, 60.8 years ± 9.24 [standard deviation]; range, 36–89 years) for MM who underwent transplant were included. Sarcopenia cut-off points defined as skeletal muscle index (SMI) used were<41 cm^2/m^2. To assess bone damage in MM the MSBDS (myeloma spine and bone damage score) was used. One-way analysis of variance and the X^2 test were used. Kaplan–Meier analysis was performed to generate progression and survival curves according to SMI and MSBDS. The testing level was set at 0.05.

Results: The median SMI was 47.1 ± 14.2 and according to SMI 18/74 (24%) had sarcopenia which was more prevalent in females (p.001). A strong and significant association between patients with low muscle mass and elevated bone damage (24/74, 32.4%) and patients with normal/non-low muscle mass low bone damage (30/74, 40.5%) was present. Multiple Logistic regression did not show any significant relationship or confounding influence among SMI and MSBDS regarding sex (p.127), cytogenetic status (p.457), staging (p.756) and relapse (.126). Neither SMI nor MSBDS resulted significantly related to overall survival as shown in Kaplan–Meier analysis.

Conclusion: Sarcopenia and bone damage affected MM patients undergoing stem cell transplantation and are significantly associated.

Advances in knowledge: Quantitative measurement of sarcopenia and bone damage on CT resulted present in MM patients undergoing stem cell transplantation and are significantly associated.

INTRODUCTION

Sarcopenia is the loss of skeletal muscle mass leading to a decline in physical performance and worse heath outcomes.[1–26] The development of sarcopenia is complex and multifactorial not only linked to the physiological age-related decline of physical movement but also to a proinflammatory status due to decreased myokines muscular production.[27,28] Indeed, the skeletal muscle is an organ not only related to movement, but it is also an organ with endocrine function.[27] Muscular myokines regulating metabolic homeostasis influence other targets such as the adipose tissue, the liver, the kidney, the brain and event the bone.[10,27–30] The bone is the target organ in patients affected by multiple myeloma (MM). Indeed, MM is a hematologic malignancy of differentiated plasma cells that accumulates and proliferates in the bone marrow leading to bone lesions. In addition, MM is also characterized by an excessive activation of osteoclasts leading to typical osteolytic lesions.[31–51] The MM associated bone disease has a strong impact on the quality of life of MM patients increasing both morbidity and mortality.[34–43] Imaging plays a crucial role when diagnosing MM. Indeed, both muscle and bone involvement can be easily assessed with imaging methods, especially CT and MRI. CT is not only considered the gold-standard to evaluate muscle mass on imaging[2,20,26,31–34] but it is also one of the best methods to assess the bone in MM.[31–43] Detection of bone lesions and sarcopenia are

clinically relevant because they are amenable of treatment, especially to prevent pathologic fractures or neurologic complications.[31,32,35–37] Both the presence of sarcopenia and bone lesions (not osteoporosis) are linked to poor prognosis in MM[33,39] and both conditions are amenable of treatments. However, it is not known if sarcopenia and bone lytic lesions are associated and to what extent. Given the fact that sarcopenia is amenable of treatment (muscular training and nutritional supplementation protein, amino acid, vitamin D and creatine for example) and the skeletal muscle is also a secondary secretory organ with endocrine functions via the myokines system influencing metabolism, patients with low muscle mass have an increased risk of falls and subsequent fractures.[1–4] Detection of an association between sarcopenia and bone damage could lead to changes in management in MM patients and offer new insights into pathophysiological evaluation of MM.

As such, we sought to examine the relationship of muscle mass estimated using CT and bone damage using the myeloma spine and bone damage score (MSBDS) which is a descriptive criteria easy to be used, highly reproducible and developed for harmonizing total body CT interpretation in MM.[44]

METHODS AND MATERIALS

Study population

This study was conducted following the available version of the Declaration of Helsinki and the International Conference on Harmonization of Good Clinical Practice Guidelines. The standard procedure of our center foresees for every patient a written informed consent form, encompassing the use of anonimized data for retrospective research purposes, before every radiological procedure. Muscle mass and bone quantitative analysis was applied to CT data collected in the clinical workup and did not influence patient care in the present study. Approval from the institutional review board was obtained accordingly with the Italian laws (054REG2019). This study is a substudy of a mixed prospective and retrospective study on MM Radiomics and sarcopenia of our center and funded by the Italian Ministry of Health (rif.19117). $N = 74$ consecutive patients (mean age, 60.8 years ± 9.24 [standard deviation]; range, 36–89 years) evaluated at the IRCCS Ospedale Policlinico San Martino Hospital (Genoa, Italy) for confirmed MM in the last 5 years and who underwent transplant. $N = 66$ patients received autotransplant from peripheral blood stem cells (PBSC), $N = 6$ patients received allogeneic stem cell transplantation from a sibling or human leukocyte antigen-matched donor, $N = 2$ unknown type of transplant. Inclusion criteria were pre-transplant total-body CT available and fully retrievable from the Hospital picture archiving and communication system (PACS). Minimal and standard technical inclusion parameters for total-body CT resulted to be:

- number of detector rows: 16 or more up to 128; minimum scan coverage: skull base to femur;
- tube voltage(kV)/time–current product (mAs) 120/50–70, adjusted as clinically needed;
- thickness ≤3 mm;
- matrix, rotation time, table speed and pith index: 256 × 256, 0.5 s, 24 mm per gantry rotation and 0.8, respectively.

CTs were acquired as total-body CT not only for bone evaluation but also for visceral organ involvement before transplantations.

Exclusion criteria were as follows: patients who did not receive transplantation, patients unable to understand or execute written informed consent, unable or unwilling to agree to follow-up during observation period, CT images not retrievable from PACS or images inaccurate due to artifacts (e.g. periprosthetic beam hardening, or significant motion artifacts). In addition other causes of sarcopenia, e.g. post trauma, surgery infection, prolonged immobilization were excluded. Clinical (including height and weight values) and follow-up data were recorded by hospital staff.

The patient baseline and clinical characteristics are summarized in Table 1. In the $n = 74$ patients were CT-derived data were evaluated, the median period between the date of the image taken and date of diagnosis was 22 days. The median follow-up period was 530 days. In total, eight patients relapsed and eight patients died during this period. All but two of these patients died from myeloma related causes. 12 patients had high-risk cytogenetic abnormalities separating patients with high-risk mutation (poor-prognosis in general) from patients without high-risk mutations according to standard of care[45–47]

Imaging measurements

Muscle mass

To assess reproducible measurement, a strict and largely validated method was used as follows.[3–15] Reconstructed axial CT images of different vendors (GE and SIEMENS) with 5 mm slice thickness were analyzed using the software installed on the workstations of our Radiology Department (Suite-Estensa 1.9-Ebit-Esaote Group company. 2015© and Horos v. 3, LGPL-3.0) (Figure 1). The third lumbar vertebra (L3), at the level in which both transverse processes are clearly visible, was used as a bony landmark to properly identify the psoas muscle. Vertebrae were counted down from the cervical spine using scout images of the whole body or multiplanar reformatted (MPR) images from source thin-section axial images using the software on our workstations. In this study, scout images were available for each patient. If there is no scout image of the whole spine or reformatted images, it is possible to identify the first lumbar vertebrae (L1) that is the first vertebrae without a rib attachment. Once L1 has been found, the transverse processes can be used to downward to L3. The L3 region contains psoas, paraspinal muscles (erector spinae, quadratus lumborum), and abdominal wall muscles (transversus abdominus, external and internal obliques, rectus abdominis). Skeletal muscle was identified and quantified by use of Hounsfield unit (HU) thresholds (−29 to +150) and then the muscle contours were manually adjusted to avoid pitfalls.[4] Cross-sectional areas (cm^2) of the sum of all these muscles were computed for each image. Cross-sectional area value is linearly related to whole-body muscle mass and was therefore normalized for stature (L3 skeletal muscle index- SMI-, cm^2/m^2).

Sarcopenia cut-off points defined as SMI used were <41 cm^2/m^2 as suggested in literature.[2–4] The worst data of SMI and MSBDS for each patient were considered for data analysis. To

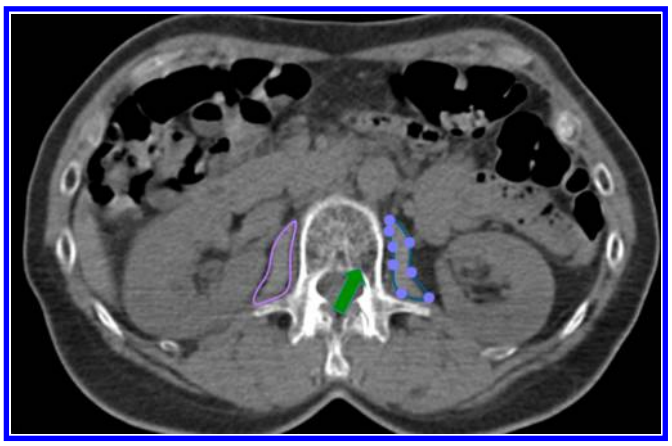
Table 1. Baseline characteristics, SMI, MSBDS, ISS and Durie Salmon Plus

	Total (n = 74)	Male (n = 37)	Female (n = 37)	p (Male *vs* Female)
SMI	47.1 ± 14.2 (20.6–70.0)	55.1 ± 12.64 (29.0–70.0)	37.5 ± 9.8 (20.6–53.3)	.001
Sarcopenia number	18	6	12	.001
MSBDS	4.2 ± 2.7 (1–10)	4.5 ± 2.5 (1–10)	4.0 ± 3.1 (1–9)	.639
ISS				.770
1	44	20	24	
2	21	9	11	
3	9	4	5	
Durie Salmon Plus				.770
1	3	1	2	
2	8	3	5	
3	53	27	26	
4	10	4	6	
High-risk cytogenetic abnormalities				
Causes of death	12	7	5	
Number	8	5	3	
MM progression	4	2	1	
Infection	3	2	1	
Other	1	1	1	

ISS, International Staging System; MSBDS, myeloma spine bone damage score; SMI, skeletal muscle index.

assess interrater reliability, 30 images were randomly selected to be analysed by both readers. Readers were experienced musculoskeletal radiologists involved in muscular CT evaluation and bone damage scores (AT, FR) with different level of experience in musculoskeletal and total body CT: Reader 1 (AT), >10 years; Reader 2 (FR), 5 years. Readers were blinded to the clinical data of the patients.

Figure 1. Examples of muscle (region of interests around psoas muscles, bliaterally, in a sarcopenic MM patient with purple and bullet line) and bone (green arrow showing a focal lesion >5 mm in diameter) evaluation on CT. See text for details. MM, multiple myeloma.



Bone damage

To assess bone damage in MM patients we used the MSBDS (Myeloma Spine and Bone Damage Score) which is a simple score tailored MM patients to be used on standard total-body CT in the routine clinical a complement of standard evaluations in patients undergoing stem cells transplantation.[44,48] The MSBDS is a recently introduced quantitative score to provide a semi-quantitative objective tool to evaluate the status of bone damage and risk of fracture and instability in MM patients.[44,48] The MSBDS score resulted to be fast, reproducible and appropriate for usage on standard CT. The MSBDS has the potential not only to assess spinal instability, but also bony involvement for prognosis. Bone density on CT was not estimated because calibration of CT scanners was not possible in standard clinical practice and in the retrospective nature of the study. MSBDS consists of an additive scale where the total score is given by the sum of single items scores for abnormalities detected. MSBDS values range from 0 (minimum) to values >10 where 10 is represented by high-risk patients requiring immediate surgical or radiation oncologist evaluation.

Statistical analysis

The one-way analysis of variance and the $\chi 2$ test were used to compare the characteristics of the study participants in the two groups. In detail, psoas muscle sizes and MSBDS scores were dichotomized according to average psoas area used to estimate SMI and average MSBDS score and evaluated for association with $\chi 2$ test as already done in literature.[49] Multiple logistic

Table 2. Association between muscle mass estimated with SMI and bone damage in multiple myeloma patients estimated with MSBDS

	Normal/non-low muscle mass	Low muscle mass	Total
MSBDS high	8	24	32 (43.2%)
MSBDS low	30	12	42 (56.8%)
Total	38 (51.4%)	36 (48.6%)	**74**

MSBDS, myeloma spine and bone damage score; SMI, skeletal muscle index.
p = 0.001 for McNemar's test for paired binary data.

regression was used to study any possible significant relationship or influence among SMI, MSBDS and standard confounders. The intra- and interobserver agreement of the SMI and MSBDS score was then calculated. K statistics were used and K values were reported as weighed k with linear weights. 95% confidence intervals (CIs) and standard error were also reported. Agreement was defined on the basis of Fleiss classification: <0.40, poor; 0.40–0.59, moderate; 0.60–0.75, good; >0.75, excellent.[50] Kaplan–Meier analysis was performed to generate progression and survival curves according to SMI and MSBDS. The testing level was set at 0.05. All analyses were performed using STATA (STATA Corp 4905 Lakeway College Station, TX).

RESULTS

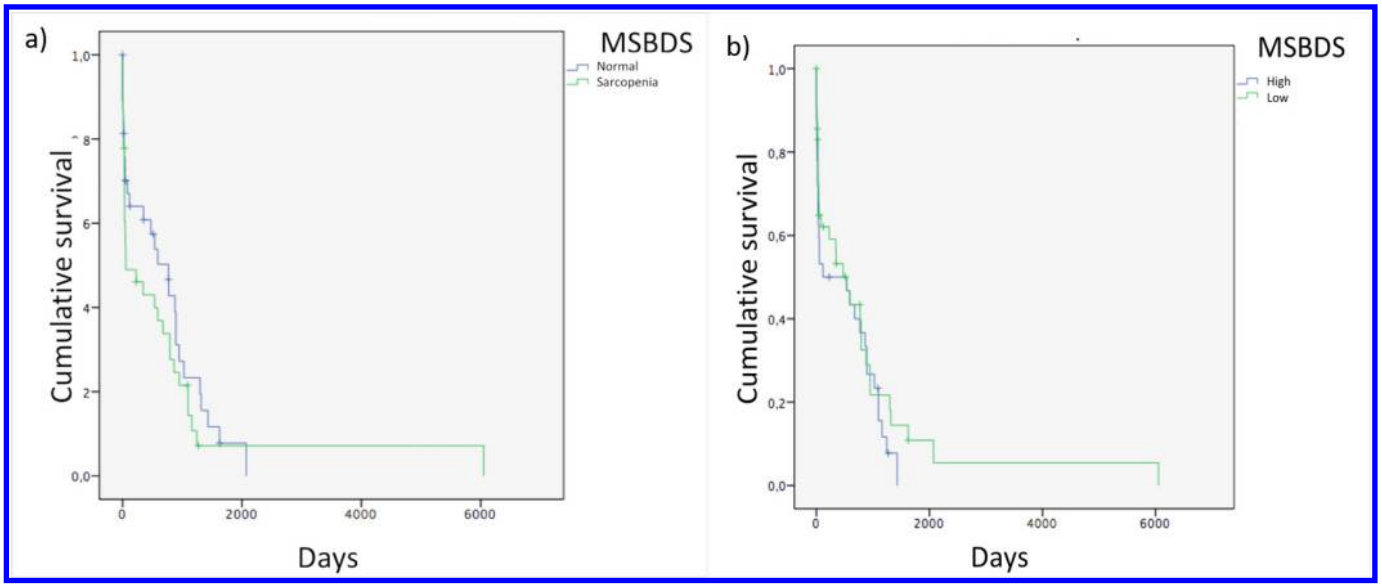
The median SMI as shown in Table 1 was 47.1 ± 14.2 and according to SMI 18/74 (24%) had sarcopenia which was more prevalent in female (p.001). When evaluating dichotomized SMI according to average psoas area and average MSBDS score, as shown in Table 2, there was a strong and significant association between patients with low muscle mass and elevated bone damage (24/74, 32.4%) and patients with normal/non-low muscle mass and low bone damage (30/74, 40.5%). Multiple logistic regression did not show any significant relationship or confounding influence among SMI and MSBDS regarding sex (p. 127), cytogenetic status (p. 457), staging (p. 756) and relapse (p. 126). Neither SMI nor MSBDS resulted significantly related to overall survival as shown in Kaplan–Meier analysis (Figure 2).

Interobserver agreement among the two readers considering the items of the MSBDS scoring scale and SMI estimation using K value, 95% confidence intervals and standard error were 0.84with 95% C.I. (0.65–0.93); 0.88 95% C.I. (0.71–0.94), respectively.

DISCUSSION

Our results show that patients with MM undergoing stem cells transplantation with low muscle mass may have an increased level of bone damage as evaluated quantitatively with standard CT. Psoas area measured at the L3 level used to estimate SMI and average MSBDS score resulted to have a strong and significant association. Low muscle mass and elevated bone damage were present in 24/74 (32.4%) of and normal/non-low muscle mass and low bone damage were present in 30/74 (40.5%) of patients. This relationship was not related to other confounding factors as evaluated sex, cytogenetic status, staging and relapse. In our group of patients, the number of female patients with sarcopenia was higher than male patients with sarcopenia. This study examines the relationship of muscle mass estimated using CT and bone damage using a quantitative score recently introduced

Figure 2. a) Kaplan–Meyer for SMI of patients with low (blue line) and normal muscular mass (green line). (b) Kaplan–Meyer plot for MSBDS of patients with high (blue line) and low bone damge scores (green line). MSBDS, myeloma spine and bone damage score; SMI, skeletal muscle index.



and tailored to MM patients. Few data are present in literature regarding body composition evaluation made on CT images in patients with MM. Takeoka at al[51] in 2016 studied the association between several body composition indexes (such as SMI, subcutaneous adipose tissue index and visceral adipose tissue index) and the overall survival in consecutive MM patients with newly diagnosed symptomatic disease. This study showed that low subcutaneous adipose tissue at baseline could be considered a predictor of poor survival outcome,[52] not only in MM patients, but also in other neoplastic conditions, such as colon, prostate, renal cell carcinoma or diffuse large B-cell lymphoma.[53]

Although sarcopenia evaluated on CT images is widely recognized as a poor prognostic factor in several oncological and non-oncological diseases,[5,11,12,14,21,24,26,32,36–40,42,48,51–60] few and controversial data regarding the role of sarcopenia and bone damage in MM are present. Williams et al[61] conducted a single-centre retrospective study to assess the impact of muscle quality in MM finding that sarcopenia was found in 72 (51%) of patients and associated with increased early post-transplant cardiovascular complications in MM. In the study by Williams et al sarcopenia was defined as ≤80% high-density muscle which is different from our definition of CT-based sarcopenia. We used a more widely accepted cut-off for the SMI of <41 cm2/m2.[2,3] Compared to the study by Williams et al,[61] the prevalence of patients with MM and sarcopenia was 18/74 (24%), which is lower than previously reported. Zakaria et al[62] found that CT-derived morphometric analysis of psoas size as a hallmark of sarcopenia, could predict overall survival in patients with lung cancer, breast cancer, prostate cancer, and MM metastases to the spine even after multivariate analysis accounting for demographic, oncologic, functional, and therapeutic factors. Zakaria et al[62] stated that patients with spinal metastases (from lung, breast, prostate, or MM) and clinical signs of sarcopenia, as measured by psoas size, have decreased overall survival. However, the patients with MM in this study were only 46 and it is not clear if the decreased overall survival of the 417 patients considered is true also for the minority of MM patients. In our study, the presence of sarcopenia was not related to poorer survival as shown in the Kaplan–Meyer curve as well as the presence of bone damage was not related to worse survival. However, our study suggest that it is possible that muscular status and bone damage are linked is some way. Indeed, recent researches on the skeletal muscle showed that the skeletal muscle is not only the organ related to mobility, but also a secondary secretory organ with endocrine functions via the myokines system. Myokines regulates metabolic homeostasis and represent an effective and underevaluated crosstalk between skeletal muscle and other target organs, such as the adipose tissue and the bone.[10,27,29,30] Our study generates a hypothesis that in MM patients, muscular status and bone lesions typical of MM, such as the lytic lesions, evaluated specifically by the MSBDS, are somehow linked. Due to the retrospective nature of this study, we are not able to explore further this association with the data available, but we suggest that further investigations are needed to better describe the relationship and prognostic significance of bone damage and sarcopenia in MM patients.

Our study is limited by the retrospective nature of the data collection, by the relatively small number of patients considered, by the novelty of the quantitative score used to assess bone damage on CT, however we were not able to find another quantitative score tailored on MM, and by the necessity to arbitrary define sarcopenia and high bone damage. Further analysis could include the evaluation of serial CT scans and increasing the number of patients.

We conclude that sarcopenia and bone damage, affected MM patients undergoing stem cell transplantation, are significantly associated. Our study results can be used for further future investigations to support our hypothesis about the relationship between muscle and bone in MM patients.

ACKNOWLEDGMENTS

The scientific guarantor of this publication is Alberto Stefano Tagliafico, MD. The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

FUNDING

This study has been partially funded by the Italian Ministry of Health Ricerca Corrente and 5 per 1000 2018-2019. Grant to Alberto Stefano Tagliafico. This work was supported by the Italian Ministry of Health (5 x 1000 – 2018). Grant to Alberto Stefano Tagliafico.

REFERENCES

1. Abiri B, Vafa M. Vitamin d and muscle sarcopenia in aging. *Methods Mol Biol* 2020; **2138**: 29–47. https://doi.org/10.1007/978-1-0716-0471-7_2
2. Albano D, Messina C, Vitale J, Sconfienza LM. Imaging of sarcopenia: old evidence and new insights. *Eur Radiol* 2020; **30**: 2199–2208. https://doi.org/10.1007/s00330-019-06573-2
3. Amini B, Boyle SP, Boutin RD, Lenchik L. Approaches to assessment of muscle mass and myosteatosis on computed tomography: a systematic review. *J Gerontol A Biol Sci Med Sci* 2019; **74**: 1671–78. https://doi.org/10.1093/gerona/glz034
4. Boutin RD, Yao L, Canter RJ, Lenchik L. Sarcopenia: current concepts and imaging implications. *AJR Am J Roentgenol* 2015; **205**: W255-66. https://doi.org/10.2214/AJR.15.14635
5. Buchard B, Boirie Y, Cassagnes L, Lamblin G, Coilly A, Abergel A. Assessment of malnutrition, sarcopenia and frailty in patients with cirrhosis: which tools should we use in clinical practice? *Nutrients* 2020; **12**(1): E186. https://doi.org/10.3390/nu12010186
6. Cawthon PM, Fox KM, Gandra SR, Delmonico MJ, Chiou C-F, Anthony MS, et al. Do muscle mass, muscle density, strength, and physical function similarly influence risk of hospitalization in older adults? *J Am Geriatr Soc* 2009; **57**: 1411–19. https://doi.org/10.1111/j.1532-5415.2009.02366.x

7. Chang K-V, Chen J-D, Wu W-T, Huang K-C, Hsu C-T, Han D-S. Association between loss of skeletal muscle mass and mortality and tumor recurrence in hepatocellular carcinoma: a systematic review and meta-analysis. *Liver Cancer* 2018; **7**: 90–103. https://doi.org/10.1159/000484950
8. Cruz-Jentoft AJ, Bahat G, Bauer J, Boirie Y, Bruyère O, Cederholm T, et al. Sarcopenia: revised european consensus on definition and diagnosis. *Age Ageing* 2019; **48**: 16–31. https://doi.org/10.1093/ageing/afy169
9. Cruz-Jentoft AJ, Landi F, Schneider SM, Zúñiga C, Arai H, Boirie Y, et al. Prevalence of and interventions for sarcopenia in ageing adults: a systematic review. *Report of the International Sarcopenia Initiative (EWGSOP and IWGS) Age and Ageing* 2014; **43**: 748–59. https://doi.org/10.1093/ageing/afu115
10. Dao T, Green AE, Kim YA, Bae S-. J, Ha K-. T, Gariani K, et al. Sarcopenia and muscle aging: a brief overview. *Endocrinol Metab* 2020; **35**: 716–32. https://doi.org/10.3803/EnM.2020.405
11. Deng HY, Hou L, Zha P, Huang KL, Peng L. Sarcopenia is an independent unfavorable prognostic factor of non-small cell lung cancer after surgical resection: a comprehensive systematic review and meta-analysis. *Eur J Surg Oncol* 2019; **45**: 728–35. https://doi.org/10.1016/j.ejso.2018.09.026
12. Deng H-. Y, Zha P, Peng L, Hou L, Huang K-. L, Li X-. Y. Preoperative sarcopenia is a predictor of poor prognosis of esophageal cancer after esophagectomy: a comprehensive systematic review and meta-analysis. *Dis Esophagus* 2019; **32**(3): doy115. https://doi.org/10.1093/dote/doy115
13. Ní Bhuachalla ÉB, Daly LE, Power DG, Cushen SJ, MacEneaney P, Ryan AM. Computed tomography diagnosed cachexia and sarcopenia in 725 oncology patients: is nutritional screening capturing hidden malnutrition? *J Cachexia Sarcopenia Muscle* 2018; **9**: 295–305. https://doi.org/10.1002/jcsm.12258
14. Ganju RG, Morse R, Hoover A, TenNapel M, Lominska CE. The impact of sarcopenia on tolerance of radiation and outcome in patients with head and neck cancer receiving chemoradiation. *Radiother Oncol* 2019; **137**: 117–24. https://doi.org/10.1016/j.radonc.2019.04.023
15. Garcia M, Seelaender M, Coletti D, Lancha AH, Sotiropoulos A. Vitamin d, muscle recovery, sarcopenia, cachexia, and muscle atrophy. *Nutrition* 2019; **60**: 66–69. https://doi.org/10.1016/j.nut.2018.09.031
16. Goates S, Du K, Arensberg MB, Gaillard T, Guralnik J, Pereira SL. Economic impact of hospitalizations in us adults with sarcopenia. *J Frailty Aging* 2019; **8**: 93–99. https://doi.org/10.14283/jfa.2019.10
17. Guerri S, Mercatelli D, Aparisi Gómez MP, Napoli A, Battista G, Guglielmi G, et al. Quantitative imaging techniques for the assessment of osteoporosis and sarcopenia. *Quant Imaging Med Surg* 2018; **8**: 60–85. https://doi.org/10.21037/qims.2018.01.05
18. Hendrickson NR, Mayo Z, Shamrock A, Kesler K, Glass N, Nau P, et al. Sarcopenia is associated with increased mortality but not complications following resection and reconstruction of sarcoma of the extremities. *J Surg Oncol* 2020; **121**: 1241–48. https://doi.org/10.1002/jso.25898
19. Heymsfield SB, Gonzalez MC, Lu J, Jia G, Zheng J. Skeletal muscle mass and quality: evolution of modern measurement concepts in the context of sarcopenia. *Proc Nutr Soc* 2015; **74**: 355–66. https://doi.org/10.1017/S0029665115000129
20. Hilmi M, Jouinot A, Burns R, Pigneur F, Mounier R, Gondin J, et al. Body composition and sarcopenia: the next-generation of personalized oncology and pharmacology? *Pharmacol Ther* 2019; **196**: 135–59. https://doi.org/10.1016/j.pharmthera.2018.12.003
21. Hu X, Dou W-. C, Shao Y-. X, Liu J-. B, Xiong S-. C, Yang W-. X, et al. The prognostic value of sarcopenia in patients with surgically treated urothelial carcinoma: a systematic review and meta-analysis. *Eur J Surg Oncol* 2019; **45**: 747–54. https://doi.org/10.1016/j.ejso.2019.03.003
22. Ivanoski S, Vasilevska Nikodinovska V. Future ultrasound biomarkers for sarcopenia: elastography, contrast-enhanced ultrasound, and speed of sound ultrasound imaging. *Semin Musculoskelet Radiol* 2020; **24**: 194–200. https://doi.org/10.1055/s-0040-1701630
23. Jones KI, Doleman B, Scott S, Lund JN, Williams JP. Simple psoas cross-sectional area measurement is a quick and easy method to assess sarcopenia and predicts major surgical complications. *Colorectal Dis* 2015; **17**: 20–26. https://doi.org/10.1111/codi.12805
24. Kim KM, Jang HC, Lim S. Differences among skeletal muscle mass indices derived from height-, weight-, and body mass index-adjusted models in assessing sarcopenia. *Korean J Intern Med* 2016; **31**: 643–50. https://doi.org/10.3904/kjim.2016.015
25. Kou H-W, Yeh C-H, Tsai H-I, Hsu C-C, Hsieh Y-C, Chen W-T, et al. Sarcopenia is an effective predictor of difficult-to-wean and mortality among critically ill surgical patients. *PLoS One* 2019; **14**(8): e0220699. https://doi.org/10.1371/journal.pone.0220699
26. Lee CM, Kang BK, Kim M. Radiologic definition of sarcopenia in chronic liver disease. *Life* 2021; **11**(2): 86. https://doi.org/10.3390/life11020086
27. Schnyder S, Handschin C. Skeletal muscle as an endocrine organ: pgc-1α, myokines and exercise. *Bone* 2015; **80**: 115–25. https://doi.org/10.1016/j.bone.2015.02.008
28. Sun H, Sherrier M, Li H. Skeletal muscle and bone - emerging targets of fibroblast growth factor-21. *Front Physiol* 2021; **12**: 625287. https://doi.org/10.3389/fphys.2021.625287
29. Huh JY. The role of exercise-induced myokines in regulating metabolism. *Arch Pharm Res* 2018; **41**: 14–29. https://doi.org/10.1007/s12272-017-0994-y
30. Webster JM, Kempen LJAP, Hardy RS, Langen RCJ. Inflammation and skeletal muscle wasting during cachexia. *Front Physiol* 2020; **11**: 597675. https://doi.org/10.3389/fphys.2020.597675
31. Cavo M, Terpos E, Nanni C, Moreau P, Lentzsch S, Zweegman S, et al. 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* 2017; **18**: e206-17. https://doi.org/10.1016/S1470-2045(17)30189-4
32. Hillengass J, Usmani S, Rajkumar SV, Durie BGM, Mateos M-. V, Lonial S, et al. International myeloma working group consensus recommendations on imaging in monoclonal plasma cell disorders. *Lancet Oncol* 2019; **20**: e302-12. https://doi.org/10.1016/S1470-2045(19)30309-2
33. Kumar S, Paiva B, Anderson KC, Durie B, Landgren O, Moreau P, et al. International myeloma working group consensus criteria for response and minimal residual disease assessment in multiple myeloma. *Lancet Oncol* 2016; **17**: e328-46. https://doi.org/10.1016/S1470-2045(16)30206-6
34. Moreau P, Kumar SK, San Miguel J, Davies F, Zamagni E, Bahlis N, et al. Treatment of relapsed and refractory multiple myeloma: recommendations from the international myeloma working group. *Lancet Oncol* 2021; **22**: e105-18. https://doi.org/10.1016/S1470-2045(20)30756-7
35. Belgioia L, Vagge S, Tagliafico A, Corvò R. How can imaging help the radiation oncologist in multiple myeloma treatment. *Medicina* 2020; **57**(1): 20. https://doi.org/10.3390/medicina57010020
36. Caers J, Paiva B, Zamagni E, Leleu X, Bladé J, Kristinsson SY, et al. Diagnosis, treatment, and response assessment in solitary plasmacytoma: updated recommendations from a european expert panel. *J Hematol*

Oncol 2018; **11**: 10. https://doi.org/10.1186/s13045-017-0549-1

37. Jamet B, Bailly C, Carlier T, Touzeau C, Nanni C, Zamagni E, et al. Interest of pet imaging in multiple myeloma. *Front Med (Lausanne)* 2019; **6**: 69. https://doi.org/10.3389/fmed.2019.00069
38. Jamet B, Zamagni E, Nanni C, Bailly C, Carlier T, Touzeau C, et al. Functional imaging for therapeutic assessment and minimal residual disease detection in multiple myeloma. *Int J Mol Sci* 2020; **21**(15): E5406. https://doi.org/10.3390/ijms21155406
39. Laubach J, Garderet L, Mahindra A, Gahrton G, Caers J, Sezer O, et al. Management of relapsed multiple myeloma: recommendations of the international myeloma working group. *Leukemia* 2016; **30**: 1005–17. https://doi.org/10.1038/leu.2015.356
40. Nanni C, Zamagni E. Fluorodeoxyglucose-pet/computed tomography as a predictor of prognosis in multiple myeloma. *PET Clin* 2019; **14**: 383–89. https://doi.org/10.1016/j.cpet.2019.03.005
41. Nanni C, Zamagni E. Therapy assessment in multiple myeloma with pet. *Eur J Nucl Med Mol Imaging* 2017; **44**: 111–17. https://doi.org/10.1007/s00259-017-3730-4
42. Rosiñol L, Beksac M, Zamagni E, Van de Donk N, Anderson KC, Badros A, et al. Expert review on soft-tissue plasmacytomas in multiple myeloma: definition, disease assessment and treatment considerations. *Br J Haematol* 2021; **194**: 496–507. https://doi.org/10.1111/bjh.17338
43. Tagliafico AS, Belgioia L, Bonsignore A, Rossi F, Succio G, Bignotti B, et al. Subspecialty second-opinion in multiple myeloma ct: emphasis on clinically significant lytic lesions. *Medicina* 2020; **56**(4): E195. https://doi.org/10.3390/medicina56040195
44. Tagliafico AS, Belgioia L, Bonsignore A, Signori A, Formica M, Rossi F, et al. Development and definition of a simplified scoring system in patients with multiple myeloma undergoing stem cells transplantation on standard computed tomography: myeloma spine and bone damage score (msbds). *Cancer Imaging* 2020; **20**: 31. https://doi.org/10.1186/s40644-020-00306-1
45. Bilhou-Nabéra C, Bidet A, Eclache V, Lippert E, Mozziconacci MJ. Cytogenetics in the management of philadelphia-negative myeloproliferative neoplasms: an update by the groupe francophone de cytogénétique hématologique (gfch). *Ann Biol Clin* 2016; **74**: 517–23. https://doi.org/10.1684/abc.2016.1153
46. Roche-Lestienne C, Boudry-Labis E, Mozziconacci MJ. Cytogenetics in the management of "chronic myeloid leukemia": an update by the groupe francophone de cytogénétique hématologique (gfch). *Ann Biol Clin* 2016; **74**: 511–15. https://doi.org/10.1684/abc.2016.1151
47. Takamatsu H, Yamashita T, Kurahashi S, Saitoh T, Kondo T, Maeda T, et al. Clinical implications of t(11;14) in patients with multiple myeloma undergoing autologous stem cell transplantation. *Biol Blood Marrow Transplant* 2019; **25**: 474–79. https://doi.org/10.1016/j.bbmt.2018.11.003
48. Tagliafico AS, Dominietto A, Belgioia L, Campi C, Schenone D, Piana M. Quantitative imaging and radiomics in multiple myeloma: a potential opportunity? *Medicina* 2021; **57**(2): 94. https://doi.org/10.3390/medicina57020094
49. Ko B-. J, Chang Y, Jung H-. S, Yun KE, Kim C-. W, Park HS, et al. Relationship between low relative muscle mass and coronary artery calcification in healthy adults. *Arterioscler Thromb Vasc Biol* 2016; **36**: 1016–21. https://doi.org/10.1161/ATVBAHA.116.307156
50. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159–74. https://doi.org/10.2307/2529310
51. Takeoka Y, Sakatoku K, Miura A, Yamamura R, Araki T, Seura H, et al. Prognostic effect of low subcutaneous adipose tissue on survival outcome in patients with multiple myeloma. *Clin Lymphoma Myeloma Leuk* 2016; **16**: 434–41. https://doi.org/10.1016/j.clml.2016.04.010
52. Lenchik L, Lenoir KM, Tan J, Boutin RD, Callahan KE, Kritchevsky SB, et al. Opportunistic measurement of skeletal muscle size and muscle attenuation on computed tomography predicts 1-year mortality in medicare patients. *J Gerontol A Biol Sci Med Sci* 2019; **74**: 1063–69. https://doi.org/10.1093/gerona/gly183
53. Charette N, Vandeputte C, Ameye L, Bogaert CV, Krygier J, Guiot T, et al. Prognostic value of adipose tissue and muscle mass in advanced colorectal cancer: a post hoc analysis of two non-randomized phase ii trials. *BMC Cancer* 2019; **19**(1): 134. https://doi.org/10.1186/s12885-019-5319-8
54. Aleixo GFP, Shachar SS, Nyrop KA, Muss HB, Malpica L, Williams GR. Myosteatosis and prognosis in cancer: systematic review and meta-analysis. *Crit Rev Oncol Hematol* 2020; **145**: 102839. https://doi.org/10.1016/j.critrevonc.2019.102839
55. Bauckneht M, Lai R, Miceli A, Schenone D, Cossu V, Donegani MI, et al. Spinal cord hypermetabolism extends to skeletal muscle in amyotrophic lateral sclerosis: a computational approach to [18f]-fluorodeoxyglucose pet/ct images. *EJNMMI Res* 2020; **10**(1): 23. https://doi.org/10.1186/s13550-020-0607-5
56. Moon SW, Choi JS, Lee SH, Jung KS, Jung JY, Kang YA, et al. Thoracic skeletal muscle quantification: low muscle mass is related with worse prognosis in idiopathic pulmonary fibrosis patients. *Respir Res* 2019; **20**(1): 35. https://doi.org/10.1186/s12931-019-1001-6
57. Rossi F, Valdora F, Bignotti B, Torri L, Succio G, Tagliafico AS. Evaluation of body computed tomography-determined sarcopenia in breast cancer patients and clinical outcomes: a systematic review. *Cancer Treat Res Commun* 2019; **21**: 100154. https://doi.org/10.1016/j.ctarc.2019.100154
58. Shachar SS, Williams GR, Muss HB, Nishijima TF. Prognostic value of sarcopenia in adults with solid tumours: a meta-analysis and systematic review. *Eur J Cancer* 2016; **57**: 58–67. https://doi.org/10.1016/j.ejca.2015.12.030
59. Xie H, Gong Y, Kuang J, Yan L, Ruan G, Tang S, et al. Computed tomography-determined sarcopenia is a useful imaging biomarker for predicting postoperative outcomes in elderly colorectal cancer patients. *Cancer Res Treat* 2020; **52**: 957–72. https://doi.org/10.4143/crt.2019.695
60. Yang Z, Zhou X, Ma B, Xing Y, Jiang X, Wang Z. Predictive value of preoperative sarcopenia in patients with gastric cancer: a meta-analysis and systematic review. *J Gastrointest Surg* 2018; **22**: 1890–1902. https://doi.org/10.1007/s11605-018-3856-0
61. Williams A, Baruah D, Patel J, Szabo A, Chhabra S, Dhakal B, et al. Prevalence and significance of sarcopenia in multiple myeloma patients undergoing autologous hematopoietic cell transplantation. *Bone Marrow Transplant* 2021; **56**: 225–31. https://doi.org/10.1038/s41409-020-01008-9
62. Zakaria HM, Llaniguez JT, Telemi E, Chuang M, Abouelleil M, Wilkinson B, et al. Sarcopenia predicts overall survival in patients with lung, breast, prostate, or myeloma spine metastases undergoing stereotactic body radiation therapy (sbrt), independent of histology. *Neurosurgery* 2020; **86**: 705–16. https://doi.org/10.1093/neuros/nyz216

diagnostics

MDPI

Article

Radiomics and Artificial Intelligence for Outcome Prediction in Multiple Myeloma Patients Undergoing Autologous Transplantation: A Feasibility Study with CT Data

Daniela Schenone [1], Alida Dominietto [2], Cristina Campi [1], Francesco Frassoni [1], Michele Cea [2,3], Sara Aquino [2], Emanuele Angelucci [2], Federica Rossi [4], Lorenzo Torri [2], Bianca Bignotti [2,4], Alberto Stefano Tagliafico [2,5,*] and Michele Piana [2,6,*]

1 LISCOMP, Dipartimento di Matematica, Università di Genova, Via Dodecaneso 35, 16146 Genova, Italy; schenone@dima.unige.it (D.S.); campi@dima.unige.it (C.C.); francesco.frassoni19@gmail.com (F.F.)
2 Ospedale Policlinico San Martino-IRCCS, Largo Rossana Benzi 10, 16132 Genova, Italy; alida.dominietto@hsanmartino.it (A.D.); michele.cea@unige.it (M.C.); sara.aquino@hsanmartino.it (S.A.); emanuele.angelucci@hsanmartino.it (E.A.); lorenzo.torri@gmail.com (L.T.); bianca.bignotti@hsanmartino.it (B.B.)
3 Dipartimento di Medicina Interna, Università di Genova, Viale Benedetto XV, 16132 Genova, Italy
4 Dipartimento di Medicina Sperimentale, Università di Genova, Via L. B. Alberti 1, 16132 Genova, Italy; federica.rossi@unige.it
5 Dipartimento di Scienze della Salute, Università di Genova, Via Pastore 1, 16132 Genova, Italy
6 CNR-SPIN Genova, Via Dodecaneso 33, 16146 Genova, Italy
* Correspondence: alberto.tagliafico@unige.it (A.S.T.); piana@dima.unige.it (M.P.)

Abstract: Multiple myeloma is a plasma cell dyscrasia characterized by focal and non-focal bone lesions. Radiomic techniques extract morphological information from computerized tomography images and exploit them for stratification and risk prediction purposes. However, few papers so far have applied radiomics to multiple myeloma. A retrospective study approved by the institutional review board: n = 51 transplanted patients and n = 33 (64%) with focal lesion analyzed via an open-source toolbox that extracted 109 radiomics features. We also applied a dedicated tool for computing 24 features describing the whole skeleton asset. The redundancy reduction was realized via correlation and principal component analysis. Fuzzy clustering (FC) and Hough transform filtering (HTF) allowed for patient stratification, with effectiveness assessed by four skill scores. The highest sensitivity and critical success index (CSI) were obtained representing each patient, with 17 focal features selected via correlation with the 24 features describing the overall skeletal asset. These scores were higher than the ones associated with a standard cytogenetic classification. The Mann–Whitney U-test showed that three among the 17 imaging descriptors passed the null hypothesis. This AI-based interpretation of radiomics features stratified relapsed and non-relapsed MM patients, showing some potentiality for the determination of the prognostic image-based biomarkers in disease follow-up.

Keywords: multiple myeloma; computerized tomography; image processing; pattern recognition; artificial intelligence

Citation: Schenone, D.; Dominietto, A.; Campi, C.; Frassoni, F.; Cea, M.; Aquino, S.; Angelucci, E.; Rossi, F.; Torri, L.; Bignotti, B.; et al. Radiomics and Artificial Intelligence for Outcome Prediction in Multiple Myeloma Patients Undergoing Autologous Transplantation: A Feasibility Study with CT Data. *Diagnostics* **2021**, *11*, 1759. https://doi.org/10.3390/diagnostics11101759

Academic Editors: Malene Barbara Fischer and Ayman El-Baz

Received: 5 July 2021
Accepted: 20 September 2021
Published: 24 September 2021

1. Introduction

Plasma cell dyscrasias (PCDs) include monoclonal gammopathy of undetermined significance (MGUS), smoldering multiple myeloma (SMM), and full-blown multiple myeloma (MM) [1]. Around 5% of the population over 70 are MGUS patients, and for around 1% of them MGUS will probably turn into MM every year. Around 10% of the SMM population evolves into full-blown MM, whose early mortality is nowadays around 28% five years after diagnosis [2]. MM is still an incurable disease, whose definition relies on the International Myeloma Working Group (IMWG) consensus updates, which is characterized

by a notable clinical heterogeneity, so that the search for consolidated biomarkers predicting the disease outcome and progression is still a crucial open issue [3–5].

The presence of either single or multiple bone lesions is a typical signature of MM, which is related to the proliferation of tumor cells from a single clone, so that the unbalanced activation of osteoclasts erodes the medullary and even the cortical bone [6].Therefore, the CRAB criteria of IMWG underlines the importance of imaging for MM assessment, and recent staging systems rely on the use of imaging modalities like magnetic resonance imaging (MRI), computerized tomography (CT) and hybrid positron emission tomography with CT (PET/CT) [3,4,6–15]. However, just the availability of different imaging modalities and the high variability of image interpretation imply a notable heterogeneity as far as the use of imaging for MM clinical practice is concerned [6,12,16].

At a more specific level, the limitations of the use of imaging for MM assessment are essentially due to three open issues: the lack of accuracy in differentiating focal from diffuse patterns, the difficulty in extracting reliable prognostic biomarkers from pattern allocation, and the low agreement in staging MM patients based on imaging outcomes [17,18].

The application of pattern recognition algorithms for the extraction of radiomics descriptors from images of MM patients and the post-processing of such radiomics features by means of procedures based on artificial intelligence (AI) are nowadays introducing a novel approach for increasing the reliability of imaging in MM clinical assessment [17–20]. The objective of the present study is to assess the feasibility of an AI-based approach for the automatic stratification of MM patients from CT data, and for the automatic identification of radiological biomarkers with a possible prognostic value. Specifically, relying on radiomics and AI-based computational analysis [19,21,22], this feasibility study shows that a set of descriptors of the focal lesions in MM X-ray CT at diagnosis allows for the automatic stratification of a cohort of MM patients who have undergone transplantation in two clusters, whose characteristics can be interpreted via comparison with clinical data, biological biomarkers, and the clinical outcome of the disease.

2. Materials and Methods

2.1. Study Populations, Inclusion Criteria, and Risk Stratification

This study was performed according to the Declaration of Helsinki and the International Conference on Harmonization of Good Clinical Practice Guidelines. An institutional review board was obtained (054REG2019). All patients signed informed consent for retrospective research before CT examination; data collection did not influence patient care. We considered 51 consecutive patients (mean age, 56 years $\pm$ 8; range, 31–73 years; 18 females; 33 males) admitted to the Hospital (BLIND for REVIEW) in the last five years because of biopsy confirmed MM. Inclusion criteria were baseline whole-body CT from the Hospital PACS or outpatient clinic. Among these 51 patients, we selected the 33 presenting at least one focal lesion in one of the CT slices, i.e., at least one >5 mm lytic lesion in the axial or extra-axial skeleton [17,18] (see Figure 1). Two radiologists blinded to the diagnosis and to each other's conclusion assessed whether the CT pattern was diffuse or focal, and, for each patient presenting at least one focal lesion, we identified the largest one.

Risk stratification was performed at diagnosis by the Revised International Staging System (ISS) combining serum beta2-microglobulin and serum albumin, lactate dehydrogenase for three-stage classification, and cytogenetics determining a binary normal-high risk stadiation [23,24]. Table 1 provides a summary of the clinical features (diameter of focal lesion: mean: 19.9 mm, STD: 13.4 mm, min: 4.5 mm, max: 62.4 mm).

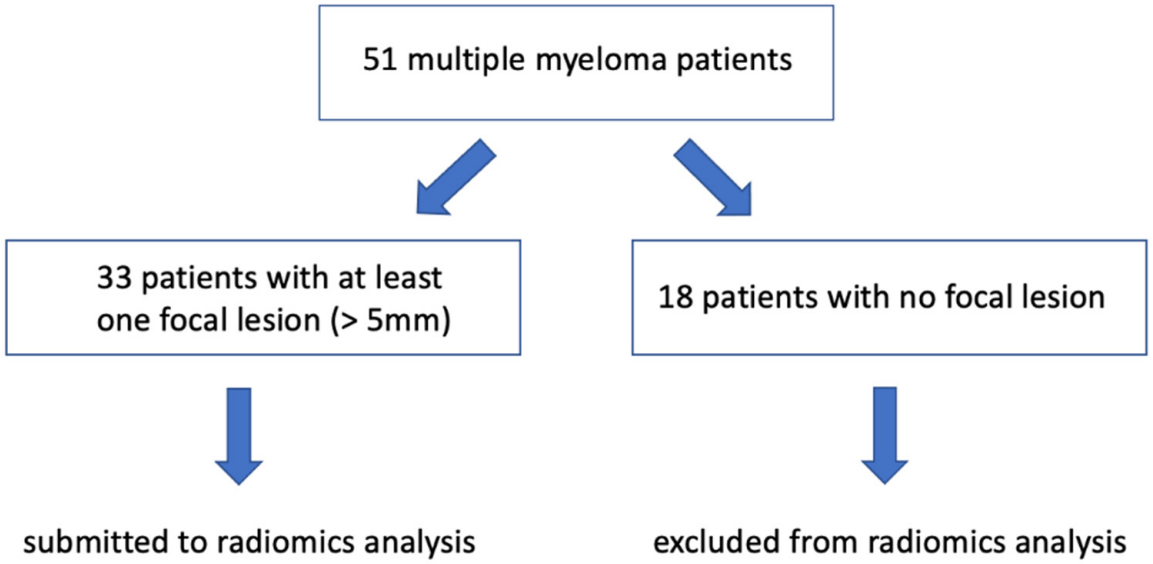


Figure 1. Flow diagram showing the initial number of participants in the study and those excluded because of not presenting focal lesions.

Table 1. Clinical features of the 33 MM patients included in the analysis. R-ISS stage: I: ISS stage I and standard-risk CA by iFISH and normal LDH. II: Not R-ISS stage I or III; III: ISS stage III and either high-risk CA by iFISH or high LDH. CA—chromosomal abnormalities; iFISH—interphase fluorescent in situ hybridization; ISS—International Staging System; LDH—lactate dehydrogenase; MM—multiple myeloma; R-ISS—revised International Staging System.

Characteristic	Number	%
Patients	33	100
Age (years) Mean	56	
Age SD [1]	6.7	
Males	21	66.4
Females	12	34.6
Cytogenetics		
Normal	22	66,7
High risk	11	33,3
Relapsed	17/33	51,5
Days before Relapse (mean)	1138	
Days of follow-up (mean)	1317	
International Staging System		
Stage I	15	45.4
Stage II	9	27.3
Stage III	9	27.3

[1] Standard Deviation.

2.2. Image Analysis

To compute each patient's overall skeletal asset, we utilized a published software tool (Bone-GUI, http://mida.dima.unige.it/software/bone-gui/; accessed on 20 September 2021) [25] combining thresholding and active contours. For each subject, Bone-GUI provided 24 features. Separately for the whole, axial, and skeleton districts, it computed the following: the mean medullary Hounsfield value with standard deviation, the volume of the global medullary asset, the mean cortical Hounsfield value with standard deviation, the volume of the cortical asset, the rate of volume occupied by the medullary tissue, and the overall volume.

We also applied an open source tool for radiomics (Slicer, https://www.radiomics.io/slicerradiomics.html; accessed on 20 September 2021) [26–28] to the 33 lytic lesions on the compact bone tissue to extract 109 Slicer features for each focal lesion.

2.3. Reduction of Redundancy

Our AI-based analysis for patients' stratification utilized Slicer features as the input. To reduce information redundancy, we considered two approaches. In the first approach, principal component analysis (PCA) [29] projected the feature space onto a principal components' subspace explaining at least 80% of the data variance. In the second approach, we performed two Pearson's correlation processes ($p > 95\%$) involving the Slicer features and (a) the binary feature encoding patient's relapse one year after transplantation, and (b) all 24 Bone-GUI features. We applied PCA to the features selected using the two correlation processes. Figure 2 illustrates this redundancy reduction pipeline.

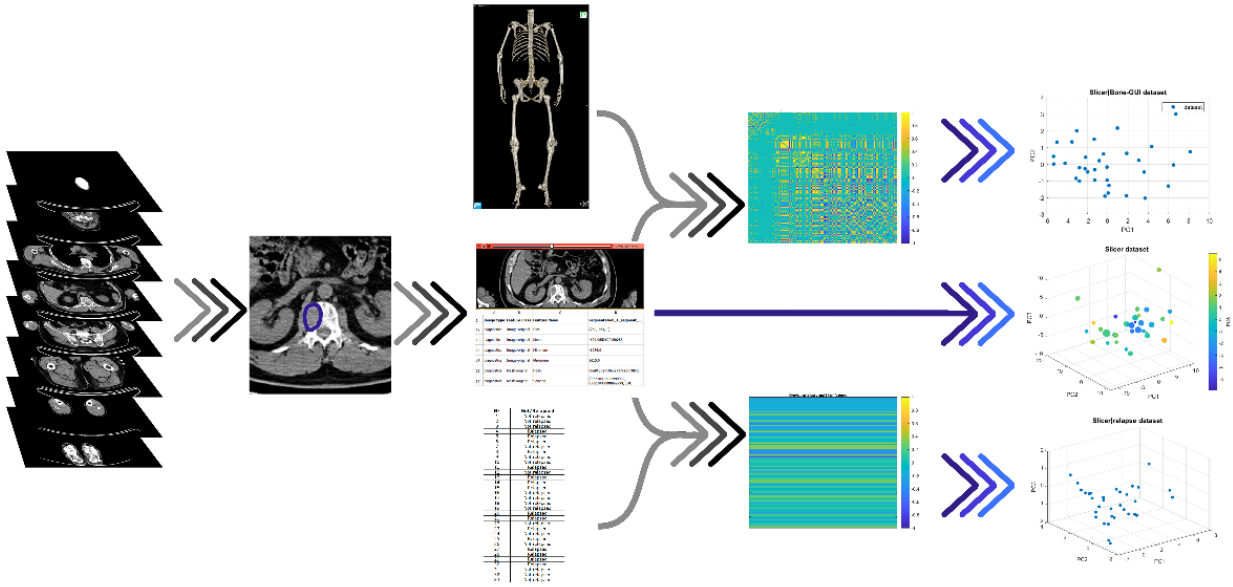


Figure 2. The pipeline of the radiomics features analysis. For each patient, the focal lesion was pointed out and the corresponding CT image was fed into a radiomics tool (Slicer), which computed 109 radiomics features; these descriptors were correlated with both the clinical outcome of the disease at one year, and the global radiological features extracted by means of a segmentation tool (Bone-GUI); the resulting mostly correlated features and the set of all local features were processed by means of two unsupervised AI algorithms (FCM and HTF) for stratification purposes.

2.4. Clustering

Clustering organized a set of unlabeled samples into clusters based on data similarity [30]. Data partition was obtained by minimizing a cost function involving the distances between the data and cluster prototypes. In Fuzzy C-Means (FCM) a degree of membership is assigned to each sample with respect to each cluster. In addition to FCM, we applied a non-linear approach based on the filtering of an extended version of the Hough transform (HTF) [31], according to the following steps (Figure 3):

1. Downstream of the PCA process, the two-dimensional feature space given by the two components explaining most of the data variance (namely, PC1 and PC2) was constructed for each data set.
2. Given a feature space, the Hough transform of each point in the patient's set with respect to the family of all parabolas was computed. As this family was characterized by three parameters, i.e., its equation is y_PC2 = ax_PC1^2 + bx_PC1 + c, with a, b, and c being the parameters, and the corresponding parameter space has three dimensions.

3. The Hough accumulator was computed by counting the number of times each Hough transform passed through one of the cells of the discretized parameter space.
4. The Hough accumulator was filtered by a 5-pixel-side cube centered on the pixel with a maximum grey value. This cube was the smallest one enclosing the cells, with accumulator values higher than 50% of the maximum [32].

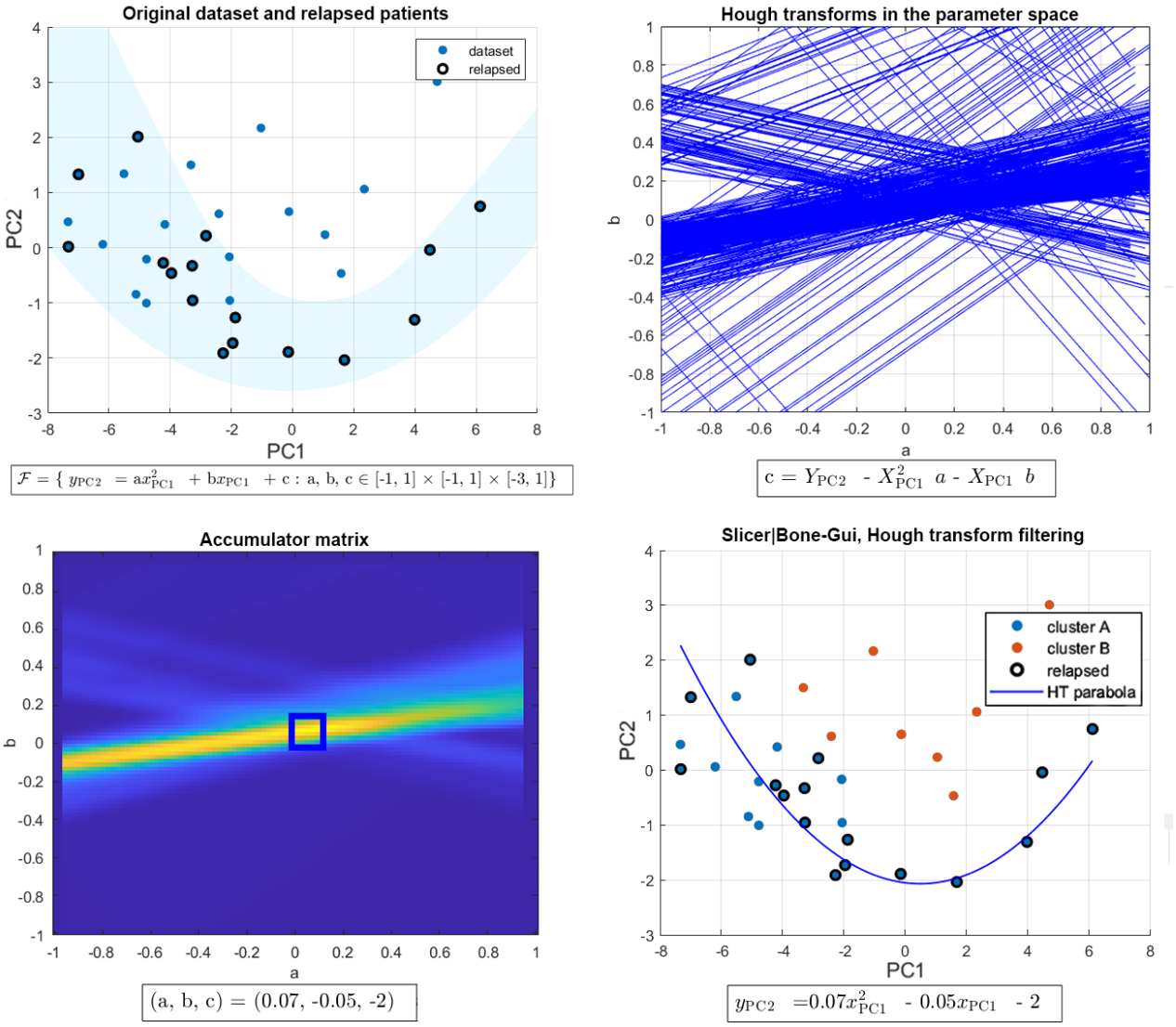


Figure 3. The HTF process for stratification. The feature space is constructed by applying PCA to the set of feature vectors (**top left panel**); for each point in the feature space the HT is computed with respect to the family of all parabolas (**top right panel**); the corresponding Hough accumulator is filtered by the smallest cube, including the cells with values higher than 50% of the accumulator maximum (**bottom left panel**); each filtered line is projected back into the feature space, thus generating the cluster of points associated to the parabola corresponding to the maximum of the Hough accumulator (**bottom right panel**).

Each line passing through the filtered region was projected back to the image space, thus generating a cluster of points in a strip around the parabola corresponding to the maximum in the Hough accumulator. The remaining points represent the second cluster made of points outside of the strip of parabolas previously identified.

3. Results

3.1. Clinical Findings

Focal lesion searching led to the selection of 33/51 (65%) patients (mean age, 56 years ± 7; range, 45–69 years; 12 females; 21 males) whose imaging data were considered for our computational analysis. Inter-observer agreement in differentiating diffuse from focal pattern between the two groups of radiologists resulted in 0.75 (95% Confidence Interval: 0.31–0.67) and 0.96 (95% Confidence Interval: 0.79–0.99) for the selection of patients with focal lesions.

3.2. AI-Based Analysis

The AI-based analysis involved three data sets (see Table 2): data set 1, made of all 109 local features extracted by Slicer from each focal lesion; data set 2, made of the eight local features mostly correlating with the relapsed/non-relapsed binary feature; and data set 3, made of the 17 local features mostly correlating with the 24 Bone-GUI global features. The application of PCA to these three data sets led to three features spaces, with $n = 5$ axes for data set 1, $n = 3$ axes for data set 2, and $n = 2$ axes for data set 3.

Table 2. Radiomics features extracted by means of image and correlation analysis.

Data Set Name	Vector Dimension	SW Tool	Feature Type	Correlation
Data set 1	109	Slicer	focal	no
Data set 2	8	Slicer	focal	relapses
Data set 3	17	Slicer	focal	global features

In each one of these three feature spaces, FCM and HTF computed two clusters: in each cluster, the black circles are associated with patients that underwent relapse within one year of bone marrow transplantation. Cluster A (B) contained the maximum (minimum) number of relapsed patients; in Figure 4, Clusters A (B) are coded with blue (orange). Table 3 contains a summary of how the clusters are populated for each of the three data sets and each of the two AI methods utilized for the analysis.

In order to assess the performances of the clustering algorithms, we computed the confusion matrices for the observed relapsed patients; specifically, we counted the number of true positives (TPs), true negatives (TNs), false positives (FPs), and false negatives (FNs) using cluster A as the reference cluster for the "relapsed" class and cluster B as the reference cluster for the "non-relapsed" class. Using the entries of such matrices, we computed four different skill scores:

Sensitivity = TP/(TP + FN)

Specificity = TN/(TN + FP)

Youden's index = Sensitivity + Specificity − 1

Critical Success Index (CSI) = TP/(TP + FN + FP).

We show that the CSI ranged from 0 to 1 and it was higher as much as the number of FPs and FNs was small, regardless the number of TNs. CSI is therefore a useful score in conditions like the one we considered here, where we had an unbalanced data set with more non-relapsed cases than relapsed ones.

We tested the robustness of our results by performing a bootstrap analysis on the set 33 17-dimension feature vectors of that set. We constructed 100 random realizations of training sets made of 20 feature vectors (of which 10 representing relapsed patients) and, for each realization, we applied the HTF clustering process. Then, for each realization of the training set, we computed the membership cluster for each one of the remaining 13 vectors representing the test set. Repeating this procedure for each one of the 100 realizations of the training-test set pairs led to the construction of 100 confusion matrices and, therefore, to 100 sets of skill score values that we averaged in Table 4, together with the corresponding standard deviations. We also performed a bootstrap analysis on

the cytogenetics values. In order to compute the entries of these last confusion matrices, we compared the relapse/non-relapse with the high/standard cytogenetic stages: a relapsed patient with a "high" cytogenetic stage was a TP event, while a relapsed patient with a "standard" cytogenetic stage was an FN. A non-relapsed patient with a "standard" cytogenetic stage was a TN event and a non-relapsed patient with a "high" cytogenetic stage was an FP event. We show that the separation between the standard and high cytogenetic stage was realized according to the standard cytogenetic evaluation for separating patients with a high-risk mutation (poor prognosis in general) from patients without high-risk mutations [24,33].

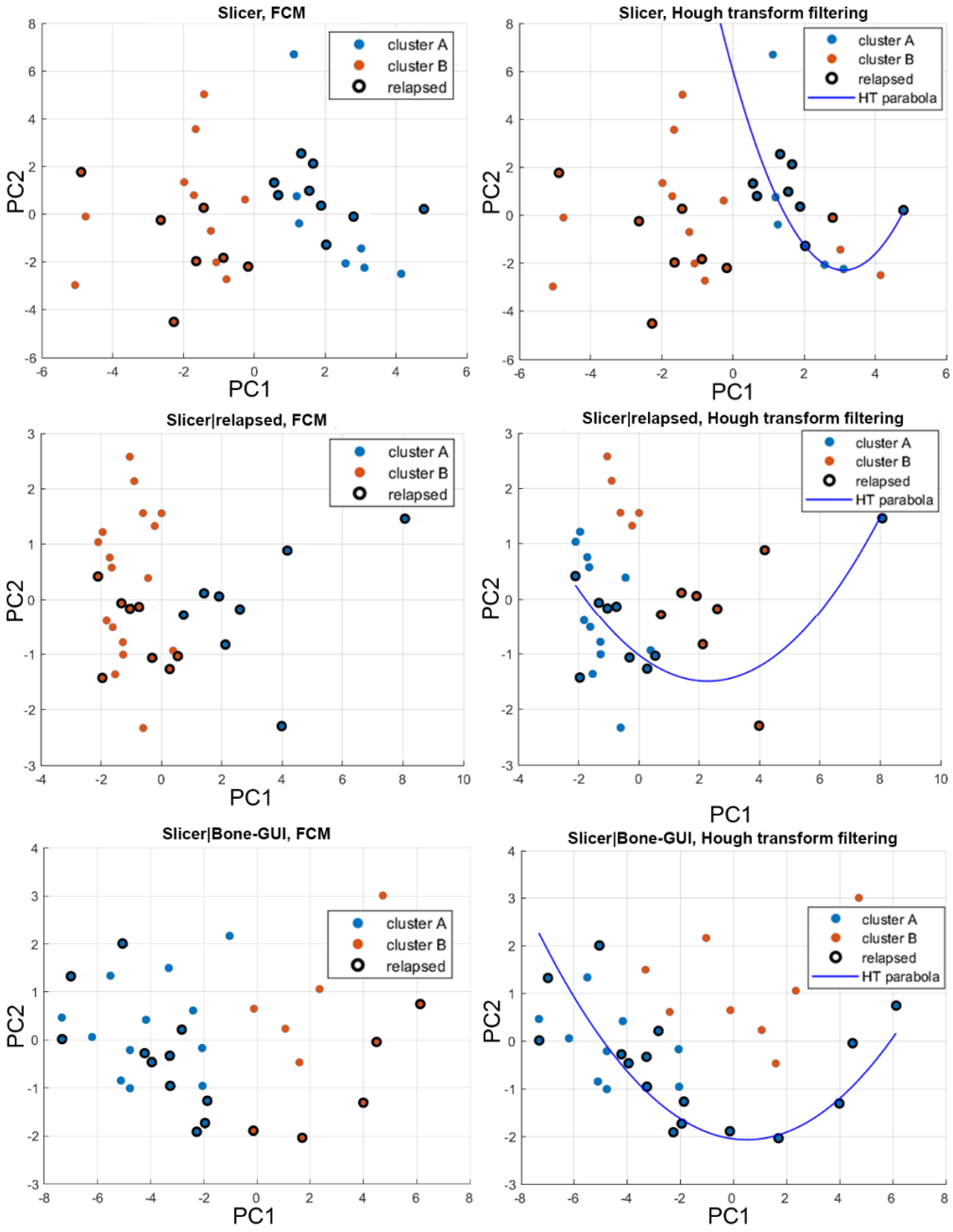


Figure 4. Clustering results for patients' stratification. In all six panels, the clustering methods (FCM for the **left column** and HTF for the **right column**) are applied to the outcomes of the PCA analysis. The results are presented in two-dimensional spaces for easier reading. Cluster A (blue) contains the highest number of relapsed patients, the opposite is true for cluster B (orange); finally, black circles represent the patients that underwent relapse. Each row shows the results for a different data set: data set 1 is in top row, data set 2 is in middle row, and data set 3 in the bottom row.

Table 3. Results of the clustering process provided by a fuzzy clustering method (FCM) and a non-linear filtering approach based on an extended version of the Hough transform (HTF). The symbol # denotes the cardinality of the set of vectors.

Method	Data Set	# of Vectors Cluster A	# of Vectors Cluster B	# of Relapses Cluster A	# of Relapses Cluster B
FCM	1	16	17	6	10
FCM	2	25	8	8	8
FCM	3	23	10	11	5
HTF	1	20	13	8	8
HTF	2	12	21	7	9
HTF	3	25	8	16	0

Table 4. Skill scores corresponding to the clustering analysis performed by means of FCM and HTF on the three data sets considered in the paper. The mean and standard deviation values are obtained by means of a bootstrap analysis that generated 100 random training sets made of 30 patients and, correspondingly, 100 random validation sets made of 13 patients. The last two rows contain the results of the analysis for the cytogenetics data associated with the patients.

Method	Data Set	Sensitivity	Specificity	Youden	CSI
FCM	1	0.46 ± 0.12	0.5 ± 0.14	-0.04 ± 0.13	0.3 ± 0.08
FCM	2	0.58 ± 0.35	0.55 ± 0.48	0.13 ± 0.15	0.3 ± 0.08
FCM	3	0.4 ± 0.24	0.55 ± 0.22	-0.06 ± 0.15	0.25 ± 0.12
HTF	1	0.38 ± 0.13	0.55 ± 0.16	-0.06 ± 0.15	0.25 ± 0.09
HTF	2	0.63 ± 0.19	0.33 ± 0.25	-0.04 ± 0.34	0.37 ± 0.16
HTF	3	0.87 ± 0.14	0.4 ± 0.13	0.27 ± 0.2	0.52 ± 0.1
Cytogenetics		0.45 ± 0.16	1.00 ± 0.02	0.44 ± 0.16	0.44 ± 0.16

3.3. Feature Ranking

To investigate which radiomics features mostly contribute to an effective stratification of the MM patients, we focused on the case of data set 3. The reason for this choice is because, when analyzed with HTF, this set provided, by far, the highest sensitivity values and, significantly, the highest CSI values among the three data sets considered. Therefore, we analyzed the feature compositions of the two axes produced by the application of PCA on the original feature space of this data set, made of 17 features. In Figure 5, we show the contribution of the 17 features to the first (light blue) and second (dark purple) principal component (PC). These contributions were weighted by the percentage of explained variance of the two PCs (77% and 9% for the first and second PC, respectively). A Mann–Whitney U-test on these features showed that just three of them did not pass the null hypothesis ($p > 99\%$): "MaskMaximum", which denotes the maximum grey level value in the mask segmenting the focal lesion (172.6 ± 64.4 in Cluster A; 321.9 ± 48.6 in Cluster B); "firstorderRange", which denotes the range of the distribution of the voxel intensities (194.7 ± 61.8 in Cluster A; 343.4 ± 66.9 in Cluster B); and "ngtdmComplexity" (29.8 ± 24.9 in Cluster A; 79.4 ± 43.5 in Cluster B), which is a measure of the non-uniformity of the lesion image in the grey level intensity.

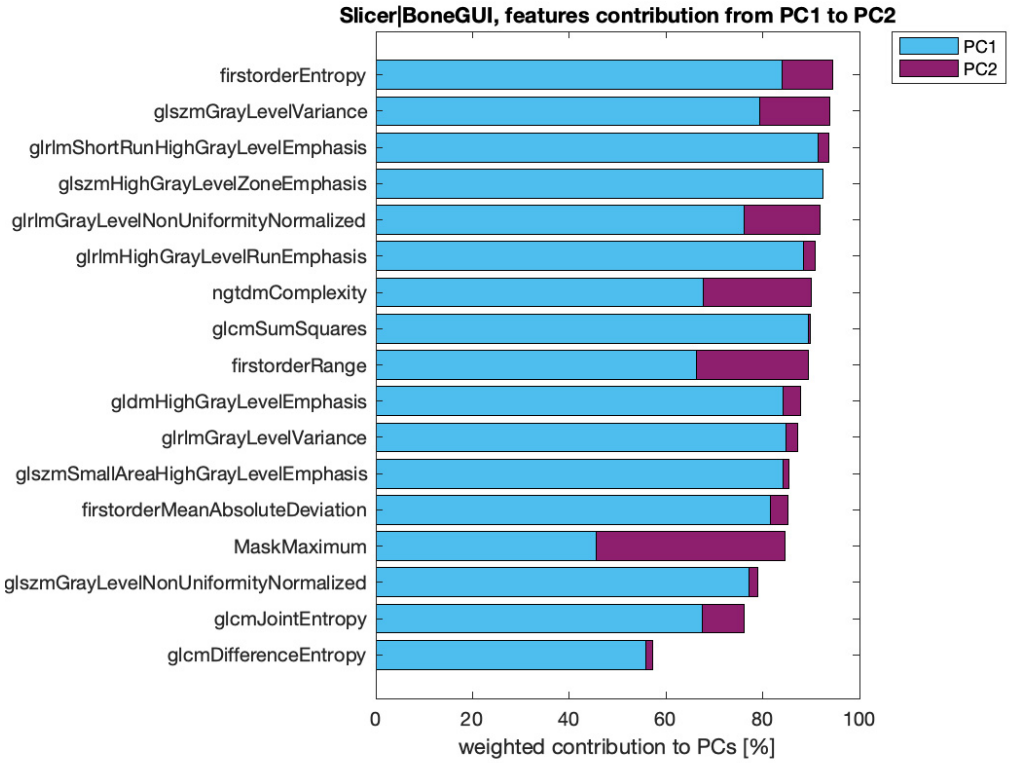


Figure 5. For data set 3, the first two principal components explain at least 80% of the data variance. In the figure, the make-up of the first two principal components is represented as follows: it is shown how much each feature contributes to the PC1 (light blue) and to the PC2 (dark purple), and these values are weighted for the percentage of data variation explained by each principal component considered.

4. Discussion

This study demonstrates that AI supported radiomics realize a clustering of MM patients with a statistical reliability that, for some skill scores, is higher than the one provided by standard biochemical staging. The possibility to increase the predictive potential of the standard CT images of patients with multiple myeloma is clinically relevant for several reasons.

The first is that although MM is still considered a single disease, it is actually a collection of several different cytogenetically distinct plasma cell malignancies [2]. Trisomies and IgH translocations are considered primary cytogenetic abnormalities, and occur at the time of establishment of MGUS [2]. At the present time, there are three specific biomarkers for MM with an approximately 80% risk of progression to symptomatic end-organ damage in two or more independent studies: clonal bone marrow plasma cells $\geq$60%, serum free light chain (FLC) ratio $\geq$100 (provided involved FLC level is $\geq$100 mg/L), and more than one focal lesion on magnetic resonance imaging (MRI). It is known that almost all patients with MM eventually relapse and the choice of a treatment regimen at relapse is affected by many factors, including the timing of relapse, response to prior therapy, aggressiveness of relapse, and performance status (TRAP) [2]. Therefore, the prediction of relapse early is important to foresee a therapy. Second, several studies have correlated bone patterns in MM with their prognostic value using MRI and CT [9,10,17,18,21,34]. MRI can be used to differentiate up to five different patterns of plasma cell infiltration, including normal appearance, focal involvement, homogeneous diffuse infiltration, diffuse infiltration with additional focal lesions, and variegated or salt-and-pepper patterns; on the other hand, CT is well suited for small (below 5 mm) focal bone lesions due to its high spatial resolution capabilities [9].

The AI-based analysis of the radiomics properties extracted from the focal lesions essentially pointed out two aspects. First, the redundancy of the radiomics features seem to impact the prognostic power of the clustering methods. However, the stratification power increases when correlation-based and PCA-based reduction of redundancy processes are applied. Second, the use of a non-linear approach to clustering, namely HTF, seems to provide better results with respect to a more standard fuzzy clustering algorithm; this may be explained because of the high degree of heterogeneity that characterizes MM. The skill scores computed for each data set and each classification method helped us to determine which approach to redundancy reduction and which algorithm performs better for stratification purposes. Among the four skill scores, CSI probably represents the one that best interprets the outcomes of the confusion matrices in this context. Indeed, this score emphasizes the correct prediction of relapses in correspondence with a low rate of misclassification. Interestingly, the application of HTF on the focal features mostly correlating with the skeleton asset's global properties (which are extracted by Bone-GUI) leads to the highest value for this score: this seems to point out a favorable prognostic role for the interplay between local and global descriptors of the MM bone tissue. In this case, the CSI value is higher than the discriminative value provided by the cytogenetic data, which supports the reliability of radiomics as a prognostic tool for MM clinical practice. This conclusion is confirmed by a bootstrap analysis performed on data set 3.

Data set 3 is made of the focal descriptors that mostly correlate with the whole skeleton's asset properties. Therefore, this correlation analysis per se realizes a feature selection process whose outcome is a set of 17 features. A finer feature selection is provided by PCA, as shown in Figure 5. This figure and the related Mann–Whitney U-test point to a significant emphasis on properties related to the heterogeneity of the focal lesion, such as the Hounsfield unit range and maximum values found in the lesion, and the complexity, which measures the non-uniformity of the image and the presence of rapid changes in intensity.

We finally show that the data collection for this study has been realized by means of a single, specific CT scanner, so that the images we used for feature extraction were homogeneous. Recent studies [35] have shown that the characteristics of the extracted features may depend on non-tumor related factors like the signal-to-noise ratio of the experimental data. Therefore, in the case of studies that utilize data from more than one scanner, data homogenization should be implemented prior to the data extraction process [36].

5. Conclusions

This computational approach to the interpretation of radiomics focal features shows the potential for the stratification of relapsed and non-relapsed MM patients, and could represent a prognostic procedure for determining the disease follow-up and therapy. Concerning the technical issues to be discussed, the present study has several strengths: the use of clinically available CT images collected in the normal daily workup did not influenced patient care in any way. Second, we used a free open-source tool for radiomics assessment of the focal lytic lesions. Among the limitations of the present study, we acknowledge the retrospective nature, which did not allow for perfect timing between CT, diagnosis, and therapy or relapse. In addition, the evaluation of the radiomics features was made only with one open-source tool, and we did not evaluate whether the usage of other tools would have introduced variability to a significant extent. Finally, the overall number of patients included was relatively low: indeed, a correct sample size in radiomics is at least five times the number of extracted features [37], and this condition would require a population of at least 100 MM patients. Nonetheless, the possibility to obtain a cluster of features to identify relapses even in a 33 patient sample is in favor of the validity of this method. This initial study warrants prospective studies with a high number of patients, which are currently underway, in order to validate this approach, with the aim of

implementing, it in a more systematic way, a method of obtaining a more robust prognostic score for MM patients.

Summing up the results of this study, we remind that our objective was to validate the feasibility of the automatic stratification of MM patients by means of an analysis of the descriptors extracted fromCT data within the framework of a radiomics retrospective study. This analysis showed that unsupervised AI can predict relapse within one year after transplantation and can identify a few imaging features associated with the heterogeneity of the focal lesion with a high prognostic value.

Author Contributions: Conceptualization, A.S.T., M.P., F.F. and A.D.; methodology, D.S., C.C., A.S.T. and M.P.; software, D.S.; validation, A.D., C.C., F.F., M.C., S.A., E.A., F.R., L.T., B.B., A.S.T. and M.P.; formal analysis, D.S., C.C., A.S.T. and M.P.; investigation, A.D., F.F., M.C., S.A., E.A., F.R., L.T., B.B. and A.S.T.; resources, A.D., E.A. and A.S.T.; data curation, D.S., C.C., A.S.T. and M.P.; writing—original draft preparation, M.P.; writing—review and editing, C.C., F.F., A.S.T. and M.P.; visualization, D.S. and M.P.; supervision, C.C., A.S.T. and M.P.; project administration: A.S.T. and M.P.; funding acquisition: A.S.T. and M.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by grants from the Italian Ministry of Health (Ricerca Corrente; grant to A.S.T.) and FRA (Fondi Ricerca di Ateneo; grant to M.P.).

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and was approved by the Institutional Review Board (or Ethics Committee) of Comitato Etico Regionale Regione Liguria (protocol codes 0077 and 22012020).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kasper, D.; Fauci, A.; Hauser, S.; Longo, D.; Jameson, J. *Harrison's Principles of Internal Medicine*; McGraw-Hill Education: New York, NY, USA, 2015.
2. Rajkumar, S.V. Multiple myeloma: Every year a new standard? *Hematol Oncol.* **2019**, *37* (Suppl. 1), 62–65. [CrossRef]
3. Rajkumar, S.V.; Dimopoulos, M.A.; Palumbo, A.; Blade, J.; Merlini, G.; Mateos, M.V.; Kumar, S.; Hillengass, J.; Kastritis, E.; Richardson, P.; et al. International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. *Lancet Oncol.* **2014**, *15*, e538–e548. [CrossRef]
4. Messiou, C.; Hillengass, J.; Delorme, S.; Lecouvet, F.E.; Moulopoulos, L.; Collins, D.; Blackledge, M.D.; Abildgaard, N.; Østergaard, B.; Schlemmer, H.-P.; et al. Guidelines for acquisition, interpretation, and reporting of whole-body MRI in myeloma: Myeloma response assessment and diagnosis system (MY-RADS). *Radiology* **2019**, *291*, 5–13. [CrossRef] [PubMed]
5. Rajkumar, S.V. Evolving diagnostic criteria for multiple myeloma. *Hematol. Am. Soc. Hematol. Educ. Program.* **2015**, *2015*, 272–278. [CrossRef] [PubMed]
6. Zamagni, E.; Cavo, M.; Fakhri, B.; Vij, R.; Roodman, D. Bones in Multiple Myeloma: Imaging and Therapy. *Am. Soc. Clin. Oncol. Educ. Book* **2018**, *38*, 638–646. [CrossRef]
7. Rasche, L.; Angtuaco, E.J.; Alpe, T.L.; Gershner, G.H.; McDonald, J.E.; Samant, R.S.; Kumar, M.; Van Hemert, R.; Epstein, J.; Deshpande, S.; et al. The presence of large focal lesions is a strong independent prognostic factor in multiple myeloma. *Blood* **2018**, *132*, 59–66. [CrossRef]
8. Siontis, B.; Kumar, S.; Dispenzieri, A.; Drake, M.T.; Lacy, M.Q.; Buadi, F.K.; Dingli, D.; Kapoor, P.; Gonsalves, W.I.; Gertz, M.A.; et al. Positron emission tomography-computed tomography in the diagnostic evaluation of smoldering multiple myeloma: Identification of patients needing therapy. *Blood Cancer J.* **2015**, *5*, e364. [CrossRef]
9. Kosmala, A.; Weng, A.M.; Krauss, B.; Knop, S.; Bley, T.A.; Petritsch, B. Dual-energy CT of the bone marrow in multiple myeloma: Diagnostic accuracy for quantitative differentiation of infiltration patterns. *Eur. Radiol.* **2018**, *28*, 5083–5090. [CrossRef]
10. Pfahler, V.K.A.; D'Anastasi, M.; Dürr, H.; Schinner, R.; Ricke, J.; Baur-Melnyk, A. Tumor Load in Patients with Multiple Myeloma: β_2-Microglobulin Levels Versus Low-Dose Whole-Body CT. *Eur. J. Haematol.* **2019**, *104*, 383–389. [CrossRef]
11. Moulopoulos, L.A.; Koutoulidis, V.; Hillengass, J.; Zamagni, E.; Aquerreta, J.D.; Roche, C.L.; Lentzsch, S.; Moreau, P.; Cavo, M.; Miguel, J.S.; et al. Recommendations for acquisition, interpretation and reporting of whole body low dose CT in patients with multiple myeloma and other plasma cell disorders: A report of the IMWG Bone Working Group. *Blood Cancer J.* **2018**, *8*, 95. [CrossRef] [PubMed]

12. Filonzi, G.; Mancuso, K.; Zamagni, E.; Nanni, C.; Spinnato, P.; Cavo, M.; Fanti, S.; Salizzoni, E.; Bazzocchi, A. A Comparison of Different Staging Systems for Multiple Myeloma: Can the MRI Pattern Play a Prognostic Role? *Am. J. Roentgenol.* **2017**, *209*, 152–158. [CrossRef]
13. Hillengass, J.; Moulopoulos, L.A.; Delorme, S.; Koutoulidis, V.; Mosebach, J.; Hielscher, T.; Drake, M.; Rajkumar, S.V.; Oestergaard, B.; Abildgaard, N.; et al. Whole-body computed tomography versus conventional skeletal survey in patients with multiple myeloma: A study of the International Myeloma Working Group. *Blood Cancer J.* **2017**, *7*, e599. [CrossRef] [PubMed]
14. Wennmann, M.; Hielscher, T.; Kintzelé, L.; Menze, B.H.; Langs, G.; Merz, M.; Sauer, S.; Kauczor, H.-U.; Schlemmer, H.-P.; Delorme, S.; et al. Spatial Distribution of Focal Lesions in Whole-Body MRI and Influence of MRI Protocol on Staging in Patients with Smoldering Multiple Myeloma According to the New SLiM-CRAB-Criteria. *Cancers* **2020**, *12*, 2537. [CrossRef] [PubMed]
15. Kumar, S.; Paiva, B.; Anderson, K.C.; Durie, B.; Landgren, O.; Moreau, P.; Munshi, N.; Lonial, S.; Bladé, J.; Mateos, M.-V.; et al. International Myeloma Working Group consensus criteria for response and minimal residual disease assessment in multiple myeloma. *Lancet Oncol.* **2016**, *17*, e328–e346. [CrossRef]
16. Hillengass, J.; Usmani, S.; Rajkumar, S.V.; Durie, B.G.M.; Mateos, M.-V.; Lonial, S.; Joao, C.; Anderson, K.C.; Garcia-Sanz, R.; Riva, E.; et al. International myeloma working group consensus recommendations on imaging in monoclonal plasma cell disorders. *Lancet Oncol.* **2019**, *20*, e302–e312. [CrossRef]
17. Tagliafico, A.S.; Cea, M.; Rossi, F.; Valdora, F.; Bignotti, B.; Succio, G.; Gualco, S.; Conte, A.; Dominietto, A. Differentiating diffuse from focal pattern on Computed Tomography in multiple myeloma: Added value of a Radiomics approach. *Eur. J. Radiol.* **2019**, *121*, 108739. [CrossRef]
18. Schenone, D.; Lai, R.; Cea, M.; Rossi, F.; Torri, L.; Bignotti, B.; Succio, G.; Gualco, S.; Conte, A.; Dominietto, A.; et al. Radiomics and artificial intelligence analysis of CT data for the identification of prognostic features in multiple myeloma. In *Medical Imaging 2020: Computer-Aided Diagnosis*; Hahn, H.K., Mazurowski, M.A., Eds.; SPIE: Belingham, WA, USA, 2020; Volume 11314, p. 152. [CrossRef]
19. Ekert, K.; Hinterleitner, C.; Baumgartner, K.; Fritz, J.; Horger, M. Extended Texture Analysis of Non-Enhanced Whole-Body MRI Image Data for Response Assessment in Multiple Myeloma Patients Undergoing Systemic Therapy. *Cancers* **2020**, *12*, 761. [CrossRef] [PubMed]
20. Morvan, L.; Nanni, C.; Michaud, A.-V.; Jamet, B. Learned Deep Radiomics for Survival Analysis with Attention. In *Proceedings of the International Workshop on Predictive Intelligence in Medicine (PRIME 2020), Lima, Peru, 8 October 2020*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 35–45.
21. Reinert, C.P.; Krieg, E.-M.; Bösmüller, H.; Horger, M. Mid-term response assessment in multiple myeloma using a texture analysis approach on dual energy-CT-derived bone marrow images—A proof of principle study. *Eur. J. Radiol.* **2020**, *131*, 109214. [CrossRef] [PubMed]
22. Jamet, B.; Morvan, L.; Nanni, C.; Michaud, A.-V.; Bailly, C.; Chauvie, S.; Moreau, P.; Touzeau, C.; Zamagni, E.; Bodet-Milin, C.; et al. Random survival forest to predict transplant-eligible newly diagnosed multiple myeloma outcome including FDG-PET radiomics: A combined analysis of two independent prospective European trials. *Eur. J. Nucl. Med. Mol. Imaging* **2020**, *48*, 1005–1015. [CrossRef]
23. Joseph, N.S.; Gentili, S.; Kaufman, J.L.; Lonial, S.; Nooka, A.K. High-risk Multiple Myeloma: Definition and Management. *Clin. Lymphoma Myeloma Leuk.* **2017**, *17S*, S80–S87. [CrossRef]
24. Palumbo, A.; Avet-Loiseau, H.; Oliva, S.; Lokhorst, H.M.; Goldschmidt, H.; Rosinol, L.; Richardson, P.; Caltagirone, S.; Lahuerta, J.J.; Facon, T.; et al. Revised International Staging System for Multiple Myeloma: A Report From International Myeloma Working Group. *J. Clin. Oncol.* **2015**, *33*, 2863–2869. [CrossRef]
25. Fiz, F.; Marini, C.; Piva, R.; Miglino, M.; Massollo, M.; Bongioanni, F.; Morbelli, S.; Bottoni, G.; Campi, C.; Bacigalupo, A.; et al. Adult Advanced Chronic Lymphocytic Leukemia: Computational Analysis of Whole-Body CT Documents a Bone Structure Alteration. *Radiology* **2014**, *271*, 805–813. [CrossRef]
26. Kapur, T.; Pieper, S.; Fedorov, A.; Fillion-Robin, J.-C.; Halle, M.; O'Donnell, L.; Lasso, A.; Ungi, T.; Pinter, C.; Finet, J.; et al. Increasing the impact of medical image computing using community-based open-access hackathons: The NA-MIC and 3D Slicer experience. *Med. Image Anal.* **2016**, *33*, 176–180. [CrossRef]
27. Kikinis, R.; Pieper, S.; Vosburgh, K. 3D Slicer: A Platform for Subject-Specific Image Analysis, Visualization, and Clinical Support. In *Intraoperative Imaging and Image-Guided Therapy*; Springer: Berlin/Heidelberg, Germany, 2014; Volume 3, pp. 277–289. [CrossRef]
28. Fedorov, A.; Beichel, R.; Kalpathy-Cramer, J.; Finet, J.; Fillion-Robin, J.-C.; Pujol, S.; Bauer, C.; Jennings, D.; Fennessy, F.; Sonka, M.; et al. 3D Slicer as an Image Computing Platform for the Quantitative Imaging Network. *Magn. Reason. Imaging* **2012**, *30*, 1323–1341. [CrossRef]
29. Jolliffe, I.T. *Principal Component Analysis*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2002. [CrossRef]
30. Bezdek, J.C. *Pattern Recognition with Fuzzy Objective Function Algorithms*; Kluwer Academic Publishers: Cambridge, MA, USA, 1981.
31. Beltrametti, M.C.; Massone, A.M.; Piana, M. Hough Transform of Special Classes of Curves. *SIAM J. Imaging Sci.* **2013**, *6*, 391–412. [CrossRef]
32. Massone, A.M.; Perasso, A.; Campi, C.; Beltrametti, M.C. Profile Detection in Medical and Astronomical Images by Means of the Hough Transform of Special Classes of Curves. *J. Math. Imaging Vis.* **2015**, *51*, 296–310. [CrossRef]

33. Roche-Lestienne, C.; Boudry-Labis, E.; Mozziconacci, M.J. Cytogenetics in the management of "chronic myeloid leukemia": An update by the Groupe francophone de cytogénétique hématologique (GFCH). *Ann. Biol. Clin.* **2016**, *74*, 511–515. [CrossRef] [PubMed]
34. Kobayashi, H.; Abe, Y.; Narita, K.; Kitadate, A.; Takeuchi, M.; Matsue, K. Prognostic Significance of Medullary Abnormalities of the Appendicular Skeleton Detected by Low-Dose Whole-Body Multidetector Computed Tomography in Patients with Multiple Myeloma. *Blood* **2017**, *130* (Suppl. 1), 1763. [CrossRef]
35. Zhovannik, I.; Bussinik, J.; Traverso, A.; Shi, Z.; Kalendralis, P.; Wee, L.; Dekker, A.; Fijiten, R.; Monshouwet, R. Learning from scanners: Bias reduction and feature correction in radiomics. *Clin. Transl. Radiat. Oncol.* **2019**, *16*, 33–38. [CrossRef]
36. Einstein, S.A.; Rong, X.J.; Jensen, C.T.; Liu, X. Quantification and homogenization of image noise between two CT scanner models. *J. Appl. Clin. Med. Phys.* **2020**, *21*, 174–178. [CrossRef]
37. Sollini, M.; Antunovic, L.; Chiti, A.; Kirienko, M. Towards clinical application of image mining: A systematic review on artificial intelligence and radiomics. *Eur. J. Nucl. Med. Mol. Imaging* **2019**, *46*, 2656–2672. [CrossRef] [PubMed]

medicina

MDPI

Article

Subspecialty Second-Opinion in Multiple Myeloma CT: Emphasis on Clinically Significant Lytic Lesions

Alberto Stefano Tagliafico [1,2,*], Liliana Belgioia [1,2], Alessandro Bonsignore [1,2], Federica Rossi [2,3], Giulia Succio [1], Bianca Bignotti [1,3] and Alida Dominietto [1]

[1] IRCCS Ospedale Policlinico San Martino, 16132 Genova, Italy; liliana.belgioia@unige.it (L.B.); alessandro.bonsignore@unige.it (A.B.); giulia.succio@hsanmartino.it (G.S.); bignottibianca@gmail.com (B.B.); alida.dominietto@hsanmartino.it (A.D.)

[2] Department of Health Sciences (DISSAL), University of Genoa, 16132 Genoa, Italy; federossi0590@gmail.com

[3] Department of Experimental Medicine (DIMES), University of Genoa, 16132 Genoa, Italy

* Correspondence: alberto.tagliafico@unige.it

Received: 5 March 2020; Accepted: 19 April 2020; Published: 23 April 2020

Abstract: *Background and objectives*: In order to increase the accuracy of lytic lesion detection in multiple myeloma, a dedicated second-opinion interpretation of medical images performed by subspecialty musculoskeletal radiologists could increase accuracy. Therefore, the purpose of this study is to evaluate the added value (increased accuracy) of subspecialty second-opinion (SSO) consultations for Computed Tomography (CT) examinations in Multiple Myeloma (MM) patients undergoing stem cell transplantation on standard computed tomography with a focus on focal lesion detection. *Materials and Methods*: Approval from the institutional review board was obtained. This retrospective study included 70 MM consecutive patients (mean age, 62 years ± 11.3 (standard deviation); range, 35–88 years) admitted in the last six years. Pre-transplant total-body CT (reported by general radiologists) was the only inclusion criteria. Each of these CT examinations had a second-opinion interpretation by two experienced subspecialty musculoskeletal (MSK) radiologists (13 years of experience and 6 years of experience, mean: 9.5 years), experts in musculoskeletal radiology and bone image interpretation with a focus on lytic lesions. *Results*: Per lesion intra- and inter-observer agreement between the two radiologists was calculated with K statistics and the results were good (K = 0.67: Confidence Inteval (CI) 95%: 0.61–0.78). When the initial CT reports were compared with the re-interpretation reports, 46 (65%) of the 70 cases (95% CI: 37–75%) had no discrepancy. There was a discrepancy in detecting a clinically unimportant abnormality in 10/70 (14%) patients (95% CI: 7–25%) unlikely to alter patient care or irrelevant to further clinical management. A discrepancy in interpreting a clinically important abnormality was registered in 14/70 (21%) patients for focal lesions. The mean diameter of focal lesions was: 23 mm (95% CI: 5–57 mm). The mean number of focal lesions per patient was 3.4 (95% CI). *Conclusions*: subspecialty second-opinion consultations in multiple myeloma CT is more accurate to identify lesions, especially lytic lesions, amenable to influence patients' care.

Keywords: multiple myeloma; computed tomography; second-look; lytic lesions; bone; staging

1. Introduction

Multiple myeloma (MM) is a hematologic disorder characterized by an excessive production of the immunoglobulin M component of plasma cells. In MM, the bone lesions of myeloma are determined by the proliferation of cells from a single clone. Then, osteoclasts are activated and destroy the bone [1]. MM, known with the abbreviation CRAB (hyperCalcemia, Renal failure, Anaemia, and lytic Bone lesions) is a cytogenetically heterogenous disorder of clonal plasma cells [1]. The extent of the bone disease

negatively influences patients' quality of life, increasing both morbidity and mortality. The detection of lytic bone lesions on imaging separates asymptomatic from symptomatic MM patients, even if no clinical symptoms are present [1–4]. Medical imaging is pivotal in the management of patients with MM. Imaging is used to detect bone lesions, to predict the risk of early progression from smoldering MM (sMM) to active MM, to identify extra-medullary disease and to identify the sites of possible pathologic fractures or neurologic complications [3]. In patients with a recent diagnosis of MM, focal lesions detected with Magnetic Resonance Imaging (MRI) or Computed Tomography (CT) or Positron Emission Tomography (PET)/CT are important for correct treatment and for prognosis [3]. In MM, "focal lesions" detected by MRI should not be confused with "lytic lesions" detected by CT. Indeed, the detection of at least one lytic lesion is a negative prognostic factor for patients with MM [2,5,6]. In 2014, the International Myeloma Working Group (IMWG) updated the definition of MM: the presence of at least one lytic lesion detected not only by conventional radiography but also by CT, WBLDCT, or PET/CT was included in the definition [7]. The incorporation of imaging modalities such as CT and PET/CT is recommended (grade A) according to the recent literature [3]. However, in MM patients, differentiation between a focal and a diffuse pattern on CT is still difficult even with Radiomics [6]. To increase the accuracy in lytic lesion detection, a dedicated second-opinion interpretation of medical images performed by subspecialty musculoskeletal radiologists could be more accurate.

In many centres, consultation and second-opinion interpretation of medical images by subspecialty radiologists are routinely performed [8–12]. Therefore, the purpose of this study is to evaluate the increased detection of focal lesions and other radiological findings of subspecialty second-opinion (SSO) consultations for CT examinations in MM patients undergoing stem cell transplantation on standard computed tomography.

2. Materials and Methods

Approval from the institutional review board was obtained (003REG2019). All patients signed a written, informed consent form for retrospective research purposes, before CT examination. SSO was applied to CT data collected in the clinical workup and did not influence patient care in any way because the study was made retrospectively.

2.1. Inclusion Criteria

This retrospective study evaluated $n = 70$ consecutive patients (mean age, 62 years ± 11.3 (standard deviation); range, 35–73 years) treated at the IRCCS Policlinico San Martino Hospital (Genoa, Italy) for MM in the last six years. Pre-transplant total-body CT with minimal technical standard (Table 1) available in the Hospital Picture Archiving and Communication System (PACS) or available in DICOM format from CT acquired outside the hospital were the only inclusion criteria—the initial CT reading was done by general radiologists with no known formal (ESSR Diploma, track record in MSK radiological activities) or informal (staff rounds, reports of specialized MSK exams) specialized experience in MSK radiology.

Table 1. Minimal and standard Computed Tomography Technical parameters for inclusion.

Number of Detector Rows	16 or More up to 128
Minimum Scan coverage	Skull base to femur
Tube voltage(kV)/time-current product (mAs)	120/50–70, adjusted as clinically needed
Reconstruction convolution kernel	Sharp, high-frequency (bone) and smooth (soft tissue). Middle-frequency kernel for all images are adjusted by the radiologist as deemed necessary
Iterative reconstruction algorithms	Yes (to reduce image noise and streak artefacts)
Thickness	≤5 mm
Multiplanar Reconstructions (MPRs)	Yes (sagittal, coronal and parallel to long axis of proximal limbs)
Matrix, Rotation time, table speed, pith index	128 × 128, 0.5 s, 24 mm per gantry rotation, 0.8

2.2. Study Design

CT examinations were studied with a second-opinion interpretation by two experienced MSK radiologists (A.T. 13 years of experience, F.R. 6 years of experience, mean: 9.5 years). The two radiologists evaluated the CT examination blindly and in different sessions. To avoid reading biases, an independent medical student was enrolled as data controller (DC). The DC checked that second-opinion interpretation was done after removing all the information of the original CT. The original report was removed. In addition, the DC made sure not to include a CT examination when the radiologists had already been involved in image re-interpretation. The use of a DC has already been explored in the literature [12].

Second-opinion consultation was made independently by using a 3-point scoring system. The scoring system is similar to a scoring system already published and now adapted to MM patients [12]: 1, no discrepancy; 2, discrepancy in detecting an unimportant abnormality (e.g., interpreting a bone infarct as a bone island, osteophytes, disc degeneration, old vertebral collapse, not neoplastic or clearly benign bone lesions); 3, discrepancy in interpreting an important abnormality (e.g., interpreting the presence of a lytic lesion >5 mm). Lytic bone lesions, size or number, non-lytic lesions, extramedullary manifestations and osteonecrosis (only if not detected by general radiologists), and fractures were considered. The clinically important differences were defined as those likely to change patient care or diagnoses according to suggestions given by our clinician on a per patient analysis (for example, a lytic lesion in a CT reported as negative at initial reading). For example, a lytic lesion could be used to stage the disease according to the Durie and Salmon PLUS staging system. After per lesion intra- and inter-observer agreement calculation, reports were re-evaluated together when their scorings were discordant. Discrepancies, mainly lytic lesions, that were significant enough to warrant a change in diagnosis, prognosis, invalidity (for medico-legal implications) or treatment or referral (e.g., orthopedic surgeon, radiation oncologist specialist) were recorded.

2.3. Reference Standard

For this study, radiologists' consensus was the best feasible reference standard available [2,6] because biopsy is not always available for all suspicious areas on CT. The best valuable comparator, or reference standard (BVC), was constructed as described elsewhere [5–7,13,14]. One hematologist, two radiologists and one radiation oncologist, all with > 10-year clinical experience, reviewed CT, MRI and PET/CT examinations and clinical follow-up for clinical significance. True positive or negative examinations were defined by lesion progression or by new lesions on follow-up imaging, or lesion response with therapy, and evolution of biologic parameters. False positive examinations were defined by an absence of new lesions on follow-up imaging studies. False negative examinations were defined by the failure of lesion detection [15]. Diffuse bone marrow infiltration in the skeleton was recorded according to Staebler et al. (lesions <5 mm, not osteoporosis) [15]. Focal pattern was defined as the presence of at least one >5 mm focal or lytic lesion. The presence of at least one focal or lytic lesion was considered relevant because it is a prognostic factor MM [5,6].

2.4. Statistical Analysis

(1) Per lesion intra- and Inter-observer agreement between the two radiologists was calculated with K statistics. p values below 0.05 were considered statistically significant. Agreement was assessed according to Altman [16] and adapted from Landis and Koch [17]. Values of 0.81–1.00 indicated very good agreement, 0.61–0.80 indicated good agreement, 0.41–0.60 indicated moderate agreement, 0.21–0.40 indicated fair agreement, and 0.20 or lower indicated poor agreement;

(2) Statistical comparisons of rates were performed using a chi-square test with Bonferroni corrections. Statistical tests were done using statistical software (STATA MP, StataCorp, 4905 Lakeway Dr, College Station, TX, USA and MedCalc).

3. Results

Intra- and Inter-observer agreement between the two radiologists was calculated with K statistics and the results were good (K = 0.67: IC 95%: 0.61–0.78) in scoring the discrepancies between subspecialized second-opinion consultations and standard CT reports, but consensus scores were used for further analysis as planned in the study protocol. Overall scores of subspecialized second-opinion consultations versus outside reports are summarized in Table 2.

Table 2. Consensus Scores of Subspeciality Second-Opinion Consultation Versus Standard CT Interpretation.

Discrepancy Score Category	No. (%) of Examinations
1, no discrepancy.	46 (65%)
2, discrepancy in detecting a clinically unimportant abnormality (e.g., a missed case of mild degenerative disease, interpreting a bone infarct as a bone island).	10 (14%)
3, discrepancy in interpreting a clinically important abnormality (e.g., interpreting the presence of a lytic lesion >5 mm or the presence of osteonecrosis or vice versa).	14 (21%)
Total	70 (100%)

As reported in Table 2, when the initial CT reports were compared with the re-interpretation reports, 46 (65%) of the 70 cases (95% CI: 37–75%) were graded 1, no discrepancy. There was a discrepancy in detecting a clinically unimportant abnormality in 10/70 (14%) patients (95% CI: 7–25%) unlikely to alter patient care or irrelevant to further clinical management. A discrepancy in interpreting a clinically important abnormality (e.g., interpreting the presence of a lytic lesion >5 mm) was registered in 14/70 (21%) patients. As shown in Table 3, the majority of discrepancies that were clinically significant (Score Category 3) were due to significant focal lesion detection discrepancies. The mean diameter of all detected focal lesions was: 23 mm (95% CI: 5–57 mm). The mean number of focal lesions per patient was 3.4 (range: 0-20; 95% CI:1.1–4.7). As a whole, n = 60 patients had focal lesions and n = 10 had none. In n = 14 patients without detected lesions by the initial report SSO found "new" lesions, thus potentially changing further treatment planning.

Table 3. Disease Category Versus Discrepancy Rates.

Disease Category	Discrepancy Score Category 1	Discrepancy Score Category 2	Discrepancy Score Category 3
Focal Lesion Detection	46	-	14
Diffuse Pattern	17	4	-
Osteonecrosis	-	1	-
Number of Focal Lesion	-	6	-

4. Discussion

In radiological clinical practice, it is quite common to have dedicated subspecialty second-opinion consultations, especially in tertiary academic centres with tumor board meetings, often known as disease management teams. However, we were not able to find the relevant literature regarding subspecialty second-opinion consultations in multiple myeloma CT. Indeed, there is growing interest in the evaluation of bone status in MM due to the increasing evidence that the presence of certain bone marrow patterns may be useful to stage and predict the outcome of MM [5–7,13,18,19]. In addition, there is a growing interest in the evaluation of lytic lesions due to their possible influence on prognosis [5]. For example, Rasche et al. [5] investigated the prognostic value of focal lesion size

in 404 transplant-eligible, newly diagnosed, MM patients with Magnetic Resonance Imaging. The authors [5] used a diffusion-weighted sequence to identify the presence of multiple large focal lesions. They found that focal lesions are strong prognostic factors. According to Rasche et al. [5], of patients with at least three large focal lesions with a product of the perpendicular diameters >5 cm, two were associated with poor progression-free survival and overall survival. This pattern was seen in 13.8% of patients and was independent of the Revised International Staging System [5]. In 2010, Hillengass et al. [18], using Whole Body Magnetic Resonance Imaging, found that the presence of focal lesions is the strongest adverse prognostic factor for progression. CT and PET/CT are now highly recommended in MM evaluation [3] and the lytic lesion assessment in MM is difficult [6]. Therefore, the focus of the present study is to improve the detection and characterization of clinically significant lytic lesions. In the past, discrepancies between reports by radiologists at different levels of training and radiologists at different clinical settings had discrepancy rates from 0.1% to 15% [12,20]. Compared to the published literature, we found that the discrepancy rate in interpreting a clinically important abnormality (e.g., interpreting the presence of a lytic lesion >5 mm) was 21% (14/70 patients), which is slightly higher than the literature data. However, we do not have any MM-related data for comparison, but only data derived from other pathological conditions. Our results highlight the necessity and the potential benefit of a subspecialty second-opinion consultation in multiple myeloma CT, in order to avoid medico-legal consequences. Furthermore, the main pathological finding that determined discrepancies was the presence of a lytic lesion. The lytic lesion of MM could be difficult to detect, especially when the diameter was between 5 and 10 mm and when located in an osteoporotic and degenerated vertebral body. In these cases, the experience of dedicated MSK radiologists could be important. Some small lytic lesions, for example, could be confused with Schmorl nodes, also referred to as intravertebral disc herniations. This study has several limitations. We acknowledge that some CTs were not primarily acquired to evaluate and detect focal lesions, therefore it is likely that these focal abnormalities were under-reported. Perhaps a more focused clinical indication before CT acquisition and report could improve focal lesion detection. Proper education of radiologists reporting MM radiological evaluation, could improve the quality of the report further. There is no clear instruction at present in the primary report for how to categorize disease entities regarding clinical relevance, therefore some of those related to MM may be overlooked or even overestimated. In addition, second-look interpretations and primary readings have been performed in different environments with different clinical priorities and different levels of expertise. Furthermore, in certain radiological environments, there is an emphasis on the quantity of work produced, which is easier to measure than the quality of interpretation [11]. Another limitation is that the scenario where expert MSK radiologists are present to reevaluate the CTs of MM patients is difficult to propose. Indeed, subspecialty radiologists practice in large and academic departments and are rare in smaller centres. In many developing countries, only general radiologists are available and imaging interpretation is sometimes performed by physicians with very limited training [21]. Finally, no correlation between discrepancies and the clinical outcome of MM patients was possible to report due to the limited number of patients, the retrospective nature of the study, and the fact that the presence of focal lesion is not the only determinant of poor prognosis.

5. Conclusions

In conclusion, our study demonstrated that subspecialty second-opinion consultation in multiple myeloma CT could identify lytic lesions, previously missed, amenable to influence patients' care.

Author Contributions: Conceptualization, A.S.T. and A.D.; methodology, A.S.T., L.B., A.B., F.R., G.S., B.B. and A.D.; formal analysis, A.S.T., L.B., A.B., F.R., G.S., B.B. and A.D.; investigation, A.S.T., L.B., A.B., F.R., G.S., B.B. and A.D.; writing—original draft preparation, A.S.T., L.B., A.B., F.R., G.S., B.B. and A.D.; writing—review and editing, A.S.T., L.B., A.B., F.R., G.S., B.B. and A.D.; funding acquisition, A.S.T., L.B., A.B., F.R., G.S., B.B. and A.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the University of Genova, grant number FRA 2018.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Moulopoulos, L.A.; Koutoulidis, V.; Hillengass, J.; Zamagni, E.; Aquerreta, J.D.; Roche, C.L.; Lentzsch, S.; Moreau, P.; Cavo, M.; Miguel, J.S.; et al. Recommendations for acquisition, interpretation and reporting of whole body low dose CT in patients with multiple myeloma and other plasma cell disorders: A report of the IMWG Bone Working Group. *Blood Cancer J.* **2018**, *8*, 95. [CrossRef] [PubMed]
2. Messiou, C.; Hillengass, J.; Delorme, S.; Lecouvet, F.E.; Moulopoulos, L.A.; Collins, D.J.; Blackledge, M.D.; Abildgaard, N.; Østergaard, B.; Schlemmer, H.-P.; et al. Guidelines for acquisition, interpretation, and reporting of whole-body MRI in myeloma: Myeloma response assessment and diagnosis system (MY-RADS). *Radiology* **2019**, *291*, 5–13. [CrossRef] [PubMed]
3. Zamagni, E.; Cavo, M.; Fakhri, B.; Vij, R.; Roodman, D. Bones in Multiple Myeloma: Imaging and Therapy. *Am. Soc. Clin. Oncol. Educ. Book* **2018**, 638–646. [CrossRef] [PubMed]
4. Rajkumar, S.V. Evolving diagnostic criteria for multiple myeloma. *Hematology* **2015**, *2015*, 272–278. [CrossRef]
5. Rasche, L.; Angtuaco, E.J.; Alpe, T.L.; Gershner, G.H.; McDonald, J.E.; Samant, R.S.; Kumar, S.; Hemert, R.V.; Epstein, J.; Tytarenko, R.; et al. The presence of large focal lesions is a strong independent prognostic factor in multiple myeloma. *Blood* **2018**, *132*, 59–66. [CrossRef]
6. Tagliafico, A.S.; Cea, M.; Rossi, F.; Valdora, F.; Bignotti, B.; Succio, G.; Gualco, S.; Conte, A.; Dominietto, A. Differentiating diffuse from focal pattern on Computed Tomography in multiple myeloma: Added value of a Radiomics approach. *Eur. J. Radiol.* **2019**, *121*, 108739. [CrossRef]
7. Rajkumar, S.V.; Dimopoulos, M.A.; Palumbo, A.; Blade, J.; Merlini, G.; Mateos, M.-V.; Kumar, P.S.; Hillengass, J.; Kastritis, E.; Landgren, O.; et al. International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. *Lancet Oncol.* **2014**, *15*, e538–e548. [CrossRef]
8. Löfgren, J.; Loft, A.; de Lima, V.A.B.; Østerlind, K.; von Benzon, E.; Højgaard, L. Clinical importance of re-interpretation of PET/CT scanning in patients referred to a tertiary care medical centre. *Clin. Physiol. Funct. Imaging* **2017**, *37*, 143–147. [CrossRef]
9. Hatzoglou, V.; Omuro, A.M.; Haque, S.; Khakoo, Y.; Ganly, I.; Oh, J.H.; Shukla-Dave, A.; Fatovic, R.; Gaal, J.; Holodny, A.I. Second-opinion interpretations of neuroimaging studies by oncologic neuroradiologists can help reduce errors in cancer care. *Cancer* **2016**, *122*, 2708–2714. [CrossRef]
10. Lakhman, Y.; Miccò, M.; Scelzo, C.; Alberto Vargas, H.; Nougaret, S.; Sosa, R.E.; Chi, D.S.; Abu-Rustum, N.R.; Hricak, H.; Sala, E. Second-Opinion Interpretations of Gynecologic Oncologic MRI Examinations by Sub-Specialized Radiologists Influence Patient Care Conclusions-Expert second-opinion review of GynOnc MRI influences patient care. *HHS Publ. Access. Eur. Radiol.* **2016**, *26*, 2089–2098. [CrossRef] [PubMed]
11. Sadat-Khonsari, M.; Papayannis, M.; Schriefer, P.; Kluth, L.; Meyer, C.; Schüttfort, V.; Regier, M.; Rink, M.; Chun, F.; Fisch, M.; et al. Worth a second look: Outcomes of patients with initial finding of regular renal tissue in CT-guided renal tumor biopsies. *World J. Urol.* **2018**, *36*, 789–792. [CrossRef] [PubMed]
12. Chalian, M.; Del Grande, F.; Thakkar, R.S.; Jalali, S.F.; Chhabra, A.; Carrino, J.A. Second-opinion subspecialty consultations in musculoskeletal radiology. *Am. J. Roentgenol.* **2016**, *206*, 1217–1221. [CrossRef]
13. Kobayashi, H.; Abe, Y.; Narita, K.; Kitadate, A.; Takeuchi, M. Prognostic significance of medullary abnormalities of the appendicular skeleton detected by low-dose whole-body multidetector computed tomography in patients with multiple myeloma. *Blood* **2017**, *130*, 1763.
14. Nanni, C.; Versari, A.; Chauvie, S.; Bertone, E.; Bianchi, A.; Rensi, M.; Bellò, M.; Gallamini, A.; Patriarca, F.; Gay, F.; et al. Interpretation criteria for FDG PET/CT in multiple myeloma (IMPeTUs): final results. IMPeTUs (Italian myeloma criteria for PET USe). *Eur. J. Nucl. Med. Mol. Imaging* **2018**, *45*, 712–719. [CrossRef] [PubMed]
15. Stäbler, A.; Baur, A.; Bartl, R.; Munker, R.; Lamerz, R.; Reiser, M.F. Contrast enhancement and quantitative signal analysis in MR imaging of multiple myeloma: Assessment of focal and diffuse growth patterns in marrow correlated with biopsies and survival rates. *AJR Am. J. Roentgenol.* **1996**, *167*, 1029–1036. [CrossRef] [PubMed]
16. Bland, J.M.; Altman, D.G. Statistics notes: Cronbach's alpha. *BMJ* **1997**, *314*, 572. [CrossRef]
17. Landis, J.R.; Koch, G.G. The Measurement of Observer Agreement for Categorical Data. *Biometrics* **1977**, 159–174. [CrossRef]

18. Hillengass, J.; Fechtner, K.; Weber, M.-A.; Bäuerle, T.; Ayyaz, S.; Heiss, C.; Ho, A.D. Prognostic Significance of Focal Lesions in Whole-Body Magnetic Resonance Imaging in Patients with Asymptomatic Multiple Myeloma. *J. Clin. Oncol.* **2010**, *28*, 1606–1610. [CrossRef] [PubMed]
19. Hillengass, J.; Moulopoulos, L.A.; Delorme, S.; Koutoulidis, V.; Mosebach, J.; Hielscher, T.; Drake, M.; Rajkumar, S.V.; Oestergaard, V.; Abildgaard, N.; et al. Whole-body computed tomography versus conventional skeletal survey in patients with multiple myeloma: A study of the International Myeloma Working Group. *Blood Cancer J.* **2017**, *7*, e599. [CrossRef]
20. European Society of Radiology (ESR). Summary of the proceedings of the International Summit 2015: General and subspecialty radiology. *Insights Imaging* **2016**, *7*, 1–5. [CrossRef]
21. European Society of Radiology (ESR). Medical imaging in personalised medicine: A white paper of the research committee of the European Society of Radiology (ESR). *Insights Imaging* **2015**, *6*, 141–155. [CrossRef] [PubMed]

European Journal of Radiology 121 (2019) 108739

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Research article

Differentiating diffuse from focal pattern on Computed Tomography in multiple myeloma: Added value of a Radiomics approach

Alberto Stefano Tagliafico[a,b,*], Michele Cea[a,b], Federica Rossi[a], Francesca Valdora[a], Bianca Bignotti[a,b], Giulia Succio[b], Stefano Gualco[a], Alessio Conte[a], Alida Dominietto[b]

[a] Department of Health Sciences, University of Genoa, Via A. Pastore 1, 16132 Genoa, Italy
[b] Ospedale Policlinico San Martino, Largo R. Benzi 10, 16132 Genoa, Italy

ARTICLE INFO

Keywords:
Multiple myeloma
Computed tomography
Radiomics
Feature
Agreement

ABSTRACT

Purpose: Focal pattern in multiple myeloma (MM) seems to be related to poorer survival and differentiation from diffuse to focal pattern on computed tomography (CT) has inter-reader variability. We postulated that a Radiomic approach could help radiologists in differentiating diffuse from focal patterns on CT.

Methods: We retrospectively reviewed imaging data of 70 patients with MM with CT, PET-CT or MRI available before bone marrow transplant. Two general radiologist evaluated, in consensus, CT images to define a focal (at least one lytic lesion > 5 mm in diameter) or a diffuse (lesions < 5 mm, not osteoporosis) pattern. N = 104 Radiomics features were extracted and evaluated with an open source software.

Results: The pathological group included: 22 diffuse and 39 focal patterns. After feature reduction, 9 features were different ($p < 0.05$) in the diffuse and focal patterns (n = 2/9 features were Shape-based: MajorAxisLength and Sphericity; n = 7/9 were Gray Level Run Length Matrix (Glrlm)). AUC of the Radiologists versus Reference Standard was 0.64 (95 % CI: (0.49–0.78) $p = 0.20$. AUC of the best 4 features (MajorAxisLength, Median, SizeZoneNonUniformity, ZoneEntropy) were: 0.73 (95 % CI: 0.58–0.88); 0.71 (95 % CI: 0.54–0.88); 0.79 (95 % CI: 0.66–0.92); 0.68 (95 % CI: 0.53–0.83) respectively.

Conclusion: A Radiomics approach improves radiological evaluation of focal and diffuse pattern of MM on CT.

1. Introduction

Abnormal production of monoclonal immunoglobulin M component of plasma cells and bone marrow increase of plasma cells is the typical characteristic of multiple myeloma (MM). The bone lesions of myeloma are caused by the proliferation of tumor cells from a single clone and the activation of osteoclasts that destroy the bone [1]. Indeed, bone disease reduces patients' quality of life increasing both morbidity and mortality, therefore the role of imaging is crucial in the management of patients with MM. Imaging is important to detect bone lesions requiring immediate start of therapy of follow-up after treatment, to predict the risk of early progression from smoldering MM (SMM) to active disease, to identify sites of extra-medullary disease and to identify sites of bone disease at potential risk of pathologic fractures or neurologic complications [2]. According to recent staging systems for MM, in patients with newly diagnosis, a correct treatment approach and evaluation of prognostic factors rely also on focal lesion identification on Magnetic Resonance Imaging (MRI), Computed Tomography (CT) or PET/CT [3]. Indeed, the role of conventional radiography, the standard of care for many years, is going to be replaced by more sensitive methods. Compared to conventional radiography, PET/CT [4] and whole-body low-dose CT (WBLDCT) are able to detect the presence of active disease in up to 25 %–40 % of cases, according to large retrospective studies [4]. It was demonstrated that the presence of at least one lytic lesion is a negative prognostic factor for patients with MM [5]. In 2014, the International Myeloma Working Group (IMWG) updated the definition of MM including in the definition the presence of at least one lytic lesion detected not only by conventional radiography but also by one of the novel morphologic imaging techniques, such as CT, WBLDCT, or PET/CT; and the presence of more than one FL on MRI [6]. At diagnosis the incorporation of new imaging modalities

(WBLDCT and PET/CT) for accurate diagnostic purposes is recommended with a grade A recommendation [2]. However, there is still considerable heterogeneity in clinical practice regarding imaging usage

* Corresponding author at: Department of Health Sciences –DISSALUniversity of Genova, Via Pastore 1, 16138 Genoa, Italy.
E-mail addresses: alberto.tagliafico@unige.it (A.S. Tagliafico), cea.michele@unige.it (M. Cea), federossi0590@gmail.com (F. Rossi), valdorafrancesca@gmail.com (F. Valdora), bignottibianca@gmail.com (B. Bignotti), giulia.succio@hsanmartino.it (G. Succio), stefanogualco@gmail.com (S. Gualco), alessioconte.1995@gmail.com (A. Conte), alida.dominietto@hsanmartino.it (A. Dominietto).

https://doi.org/10.1016/j.ejrad.2019.108739
Received 2 May 2019; Received in revised form 8 September 2019; Accepted 4 November 2019

in MM [2] and, as for every radiological techniques, variability among readers could reduce the diagnostic efficacy due to difficulties in differentiating small lesion of approximately 5 mm in diameter typical of a focal pattern with a worse prognosis [3]. Even among expert, agreement in detecting lytic lesions on PET/CT for staging using the Italian myeloma criteria for PET USe (IMPeTUs) criteria was 0.54 (0.41_0.68) using Krippendorff's alpha for lesions probably > 5 mm in diameter [7]. The cutting-edge research topic of Radiomics analysis aim to extract complex data from clinical images to help radiologists and clinicians in both diagnosis and prognosis. [8–10]. We made the hypothesis that Radiomics analysis could help radiologists to unveil imaging characteristic on CT specific of a MM pattern, especially to identify lytic lesions. Therefore, the aim of this study was to assess if a Radiomics approach could improve radiological accuracy in differentiating a focal pattern from a diffuse pattern on CT.

2. Methods

The study was performed in accordance with the current version of the Declaration of Helsinki and the International Conference on Harmonization of Good Clinical Practice Guidelines. Approval from the institutional review board was obtained (003REG2019). According to our standard procedure, all patients signed a written informed consent form, encompassing the use of anonimized data for retrospective research purposes, before CT examination. Radiomic analysis was applied to CT data collected in the clinical workup and did not influence patient care in any way. According to the nature of the study, STARD checklist was followed as appropriate [11].

2.1. Study design, inclusion criteria

Our retrospective study included 70 consecutive patients (mean age, 60 years ± 9.2 [standard deviation]; range, 35–88 years) admitted to the IRCCS Policlinico San Martino Hospital (Genoa, Italy) because they were suspected of having MM in the last five years. Inclusion criteria were pre-transplant total-body CT available and retrievable from the Hospital Picture archiving and communication system (PACS) or available from outpatient clinic with minimal technical standard. Minimal and standard technical inclusion parameters for CT are reported in Table 1.

All CT scans were read in consensus by two groups of radiologists.

The first group included two general radiologists (G.S. 15 years of experience, A.C. 1 year of experience) one of them, the senior, with extensive track record in CT reporting and musculoskeletal radiology. The two general radiologists were blinded to the diagnosis of the patients and they were asked to assess if the CT pattern was diffuse or focal (Fig. 1).

The second group included two experienced radiologists (A.T. 12 years of experience, F.R. 5 years of experience) expert in musculoskeletal Radiology and bone image interpretation, one of them with European Diploma and Member of tumour sub-committee of the European Society of Musculoskeletal Radiology. After six months to avoid biases, the two expert radiologists worked in consensus aware of the diagnosis of MM and able to check follow-up radiological evaluation to assess if the pattern on CT had to be considered diffuse or focal, and their consensus was the reference standard of our study. Considering that biopsy is not available for all suspicious area identified on CT, radiologists' consensus could be considered the best feasible reference standard, as already done in literature [12].

2.2. Test methods

2.2.1. Index test1 - Computed Tomography

Diffuse bone marrow infiltration in the skeleton was recorded according to the criteria proposed by Staebler (lesions < 5 mm, not osteoporosis) [13,14]. Focal pattern was defined as the presence of at least one > 5 mm of focal or lytic lesion in the axial skeleton (ie, spine and sacral bone) or extra-axial skeleton (ie, all other parts of the skeleton). Soft tissue lesions were not considered because outside the scope of the study. Lesions in typical locations for degenerative changes and osteoporotic changes were not counted. The presence of at least one focal or lytic lesion was considered clinically relevant because it is a highly significant adverse prognostic factors for patients with MM, and recored [5].

2.2.2. Index test2 - radiomics analysis

Radiomics analysis was performed on all CT images suspected of having pathology within

manually selected regions of interest (ROIs) including all the bone of axial and extra-axial skeleton on single slices where the bone was visually judged different from a normal bone by radiological assessment. All images were read and processed in the raw Digital Imaging and Communications in Medicine (DICOM) format. Raw imaging data underwent pre-processing to discriminate the signal from the noise. ROIs were placed by two researchers (A.T. and F.R.) expert in quantitative image analysis (9 and 3 years of experience). Theoretically, ROIs placement would have been independent from the kind of bone lesion present on CT. From CT images, we extracted 104 image features using an open-source software platform for medical image informatics, image processing, and three-dimensional visualization (3D Slicer 4.10; www.slicer.org) built over two decades through support from the National Institutes of Health and a worldwide developer community and largely used in literature [15]. 3D-Slicer can be employed for quantitative image feature extraction and image data mining research in large patient cohorts [15]. Definitions, descriptions and subdivisions into classes of Radiomics features are available in literature [16]. We computed a total of n = 104 features per patient. This feature initial pool was subjected to the selection procedure. From the total of n = 104 features, z-score normalization was applied making the range of the features more uniform and removing features that had high similarity with other features. Therefore, we selected strongly correlated features (P value below 0.05) and eliminated the redundancies as normally done in literature [17].

Mean time for single patient Radiomics analysis was calculated with a commercially available stopwatch, including the time to download images, perform image adjustment and analysis and finally data collection in the database. Radiologists were well trained in the usage of Radiomics tools and probably reduced the reading time.*Statistical*

Table 1
Minimal* and standard Computed Tomography Technical parameters for inclusion.

Number of detector rows*	16 or more up to 128
Minimum Scan coverage*	Skull base to femur
Tube voltage(kV)/time-current product (mAs)	120/50–70, adjusted as clinically needed
Reconstruction convolution kernel	Sharp, high-frequency (bone) and smooth (soft tissue). Middle-frequency kernel for all images are adjusted by the radiologist as deemed necessary
Iterative reconstruction algorithms	Yes (to reduce image noise and streak artifacts)
Thickness*	≤5 mm
Multiplanar Reconstructions (MPRs)	Yes (sagittal, coronal and parallel to long axis of proximal limbs)
Matrix, Rotation time, table speed, pith index	128 × 128, 0.5 s,24 mm per gantry rotation, 0.8

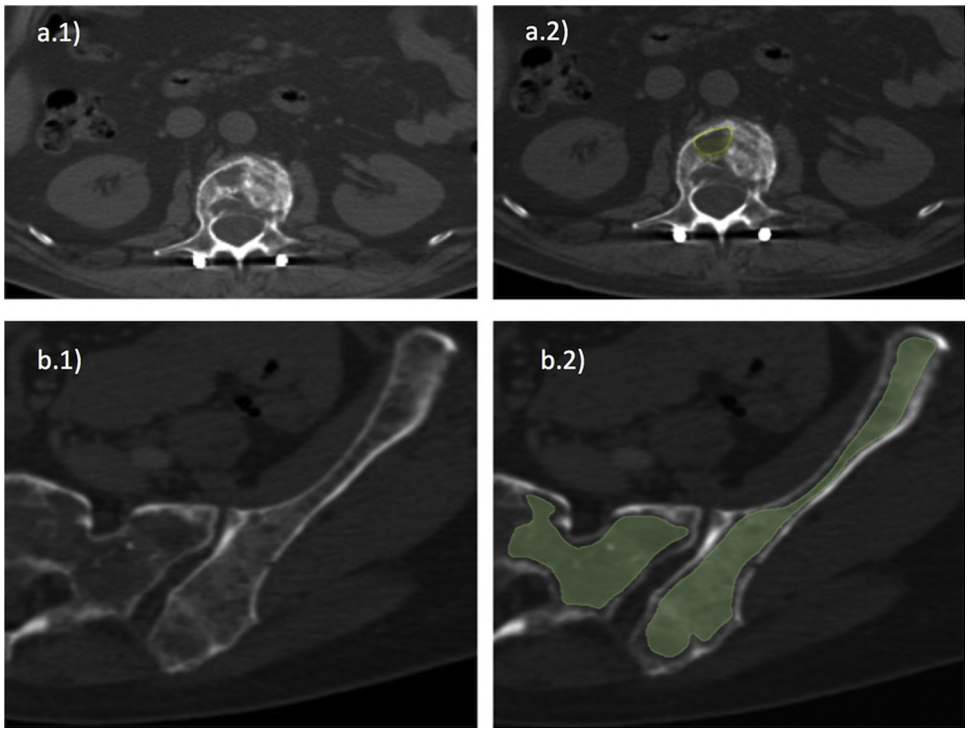


Fig. 1. Examples of focal and diffuse bone patterns on CT.
In a.1) graphical example of focal lytic lesion (> 5 mm) of the spine on CT. In a.2) the same lesion with manually selected regions of interest (ROIs) in green using 3D Slicer 4.10. In b.1) graphical example of diffuse bone pattern lesion of the left hemi-sacrum and left iliac bone. In b.2) the same lesion with manually selected regions of interest (ROIs) in green using 3D Slicer 4.10.

Analysis

Inter-observer agreement in differentiating diffuse from focal pattern was estimated among the two groups of radiologists to confirm the need of more accurate (Radiomics) measurements to improve CT interpretation. For research purposes Cronbach's alpha was considered acceptable if between 0.7 and 0.8. Comparison of Radiomics features of diffuse and focal pattern was done with non-parametric tests (Mann-Whitney *U* test for unpaired data with 1000 bootstraps samples) considering a p value of 0.05 as statistically significant; then feature reduction was done to avoid over-fitting. Accuracy was measured using receiver operating characteristic (ROC) analyses to estimate the area under the curve (AUC) and compare Radiologists and Radiomics evaluation against reference standard using statistical software, p values below 0.05 were considered statistically significant.

Kaplan–Meier analysis was performed to generate progression and survival curves according to diffuse or focal pattern. Time to event and survival between groups was compared with the two-tailed log-rank test. Statistical tests were done using statistical software (STATA MP, StataCorp, 4905 Lakeway Dr, College Station, TX, USA and MedCalc).

3. Results

N = 9/70 did not have any CT available before bone marrow transplant and were excluded, therefore the study group included 60 patients: 27 men (mean age, 59,7 years ± 9,1; range, 35–72 years) and 34 women (mean age, 61,7 years ± 9,2; range, 49–88 years).

Inter-observer agreement in differentiating diffuse from focal pattern among the two groups of radiologists resulted to be 0.57 (95 % Confidence Intervals: 0.32–0.64), p < 0.03.

After feature reduction and selection, n = 16/104 (15 %) of Radiomics features were different in focal and diffuse pattern (Table 2).

AUC of the radiologists' evaluation and AUC of the best four features (MajorAxisLength; Median; SizeZoneNonUniformity; ZoneEntropy) resulted to be between 0.642 (95 % Confidence Intervals: 0.494 to 0.789) and 0.790 (0.665 to 0.916) as shown in Fig. 2 and Table 3. The lowest value of AUC belonged to radiologist's evaluation.

Mean time for single patient Radiomics analysis resulted to be 1 h per patient ± 20 min. The time to read the CT scan without Radiomics was 10 min.

Table 2
Summary of n = 16/104 Radiomics features resulted to be different in focal and diffuse pattern. P values < 0.05 are considered statistically significant.

Feature Name	P value
MajorAxisLength	,030
Sphericity	,012
SmallDependenceLowGrayLevelEmphasis	,032
ZoneVariance	,006
Correlation	,041
SumEntropy	,031
Skewness	,004
RunEntropy	,001
Median	,001
LowGrayLevelEmphasis	,013
Energy	,024
ShortRunLowGrayLevelEmphasis	,023
LowGrayLevelRunEmphasis	,045
SizeZoneNonUniformity	,001
LowGrayLevelZoneEmphasis	,038
SmallAreaLowGrayLevelEmphasis	,038

Kaplan-Meier plots for relapse of patients who had focal pattern compared with patients who had diffuse pattern demonstrated that the median time to progression was significantly worse for patients with a focal pattern (Fig. 3).

4. Discussion

The present study demonstrated that a Radiomics approach on standard CT images of patients with multiple myeloma acquired before transplantation strongly improves accuracy in differentiating focal from diffuse patterns. Indeed, accuracy in terms of area under the curve of Radiologists compared to the reference standard was lower (64 %) than accuracy calculated using a Radiomics approach which obtained a maximum value of 79 %. The possibility to increase diagnostic accuracy in differentiating focal from diffuse pattern on standard CT images of patients with multiple myeloma is clinically relevant for several

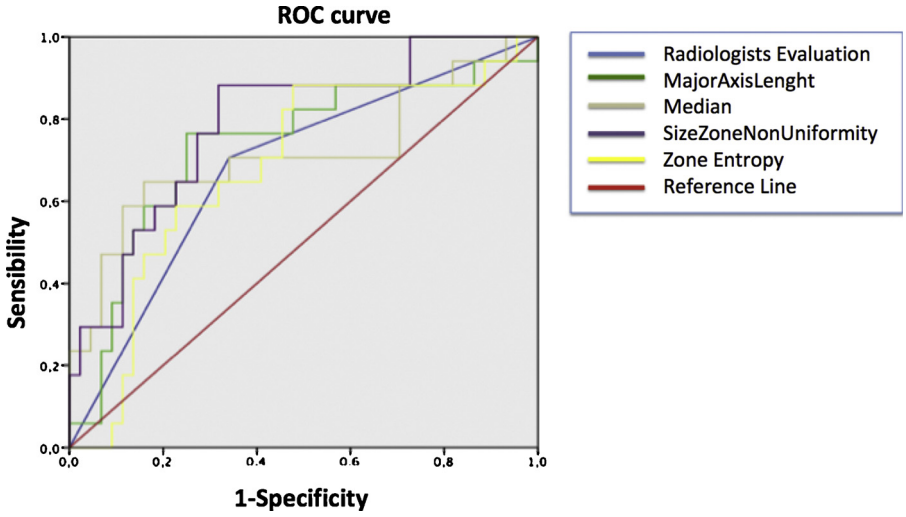


Fig. 2. The Area Under the Curve (AUC) of the radiologists' evaluation and the AUC of the best four features.

Table 3
Area Under the Curve (AUC) of the best four features. P values < 0.05 are considered statistically significant.

Feature Name	AUC	P value	95 % CI (lower limit)	95 % CI (upper limit)
MajorAxisLength	0,733	0,005	0,580	0,885
Median	0,715	0,010	0,549	0,881
SizeZoneNonUniformity	0,790	0,010	0,665	0,916
ZoneEntropy	0,682	0,029	0,531	0,833

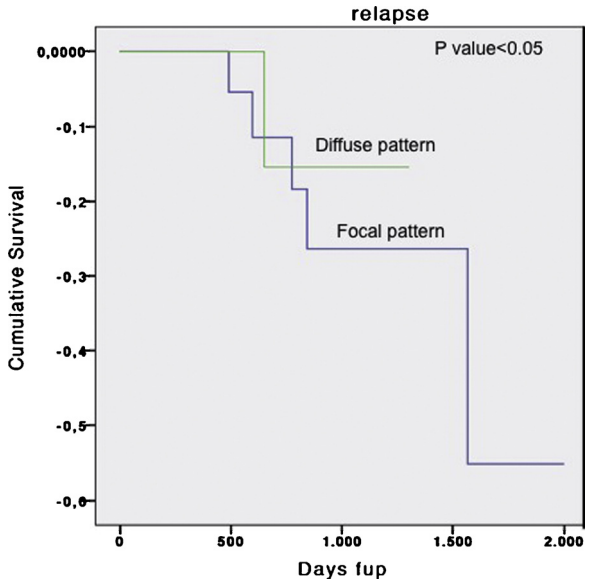


Fig. 3. Kaplan Meyer plot for relapse of patients with focal pattern (blue line) compared with patients with diffuse pattern (green line).

reasons.

First, new staging systems suggested to replace standard radiography with more sensitive methods such as CT, due to its higher capability of differentiate bone inner texture [2,7]. Therefore, more patients will undergo CT for staging of MM [2].

Second, several studies correlate pattern allocation with prognostic value [18]. Although MRI can differentiate up to five different patterns of plasma cell infiltration, including normal appearance, focal involvement, homogeneous diffuse infiltration, diffuse infiltration with additional focal lesions and variegated or salt-and-pepper pattern, CT can also identify patterns similarly to MRI [18] and CT is well suited for small (below 5 mm) focal bone lesions due to high spatial resolution capabilities [18].

Third, the low agreement between reader in staging patients affected by multiple myeloma is well known in literature [8], as confirmed by our study. Indeed, we found that the agreement among radiologists in differentiating between focal and diffuse patters was 0.57 (95 % Confidence Intervals: 0.32–0.64), $p < 0.03$ which is below the values considered acceptable for research and clinical purposes. Nanni et al. [7], calculating inter-observer variability with Krippendorff's alpha, found values of 0.56 to 0.58 indicating only moderate agreement for focal lesions. Data reported in the study by Nanni et al. [7] are consistent with our results and underlines the need of improvement to correctly identify patients with a focal pattern. In addition, the use of slight different modalities of agreement calculation such as Cronbach's alpha in the present study versus Krippendorff's alpha which automatically corrects for a casual agreement between reviewers, is not sufficient to stop seeking for better methods, such as Radiomics, to improve focal pattern recognition.

Nowadays, there is still a certain lack of agreement about the exact definition of a diffuse imaging pattern [18] In our study, we defined the presence of a focal pattern as the presence of at least one > 5 mm of focal lytic lesion in the axial skeleton or extra-axial skeleton because it has been demonstrated that the presence of more than one focal lesion could be an optimal cut-off point: indeed patients with greater than one focal lesions had significantly shorter progression-free survival than those without [5], and our results confirmed the worse prognosis for patients with focal pattern.

Concerning technical issues to be discussed, the present study has several strengths: we used clinically available CT images collected in the normal clinical workup without influencing patient care in any way and we used a free open source software for Radiomics assessment of involved bones. Finally, Radiomics assessment was made in a Radiological environment with significant expertise in quantitative imaging assessment and software development [10,19].

Concerning Radiomics feature assessment, we found that 15 % of features (16/104) were different in diffuse and focal patterns reflecting a significant difference in bone phenotype in patients with the same disease.

Among the limitations of the present study we acknowledge the retrospective nature which did not allow a perfect timing between CT acquired before transplantation and diagnosis. In addition, the evaluation of Radiomics features was made only with one software and we do not know if the usage of other software could introduce variability in

feature assessment. Finally, we did not correlate CT patterns with staging before transplant, but Kaplan-Meyer results confirmed the worse prognosis for patients with focal pattern.

In conclusion, in this work we have proven that, in multiple myeloma patients, differentiation between focal and diffuse pattern on CT is difficult, but a Radiomic approach strongly improves standard radiological evaluation with implications for prognosis, patient stratification and therapeutical choices.

References

[1] J.L.J. Larry Jameson, Anthony S. Fauci, Dennis L. Kasper, Stephen L. Hauser, Dan L. Longo, Harrison's Principles of Internal Medicine, (2018), https://doi.org/10.1001/jama.1980.03300470050029.

[2] E. Zamagni, M. Cavo, B. Fakhri, R. Vij, D. Roodman, Bones in Multiple Myeloma: Imaging and Therapy, American Society of Clinical Oncology Educational Book, 2018, pp. 638–646, https://doi.org/10.1200/EDBK_205583.

[3] G. Filonzi, K. Mancuso, E. Zamagni, C. Nanni, P. Spinnato, M. Cavo, S. Fanti, E. Salizzoni, A. Bazzocchi, A comparison of different staging systems for multiple myeloma: can the MRI pattern play a prognostic role? AJR Am. J. Roentgenol. 209 (2017) 152–158, https://doi.org/10.2214/AJR.16.17219.

[4] B. Siontis, S. Kumar, A. Dispenzieri, M.T. Drake, M.Q. Lacy, F. Buadi, D. Dingli, P. Kapoor, W. Gonsalves, M.A. Gertz, S.V. Rajkumar, Positron emission tomography-computed tomography in the diagnostic evaluation of smoldering multiple myeloma: identification of patients needing therapy, Blood Cancer J. 5 (2015), https://doi.org/10.1038/bcj.2015.87 e364–e364.

[5] J. Hillengass, K. Fechtner, M.-A. Weber, T. Bäuerle, S. Ayyaz, C. Heiss, T. Hielscher, T.M. Moehler, G. Egerer, K. Neben, A.D. Ho, H.-U. Kauczor, S. Delorme, H. Goldschmidt, Prognostic significance of focal lesions in whole-body magnetic resonance imaging in patients with asymptomatic multiple myeloma, J. Clin. Oncol. 28 (2010) 1606–1610, https://doi.org/10.1200/JCO.2009.25.5356.

[6] S.V. Rajkumar, M.A. Dimopoulos, A. Palumbo, J. Blade, G. Merlini, M.-V. Mateos, S. Kumar, J. Hillengass, E. Kastritis, P. Richardson, O. Landgren, B. Paiva, A. Dispenzieri, B. Weiss, X. LeLeu, S. Zweegman, S. Lonial, L. Rosinol, E. Zamagni, S. Jagannath, O. Sezer, S.Y. Kristinsson, J. Caers, S.Z. Usmani, J.J. Lahuerta, H.E. Johnsen, M. Beksac, M. Cavo, H. Goldschmidt, E. Terpos, R.A. Kyle, K.C. Anderson, B.G.M. Durie, J.F.S. Miguel, International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma, Lancet Oncol. 15 (2014) e538–48, https://doi.org/10.1016/S1470-2045(14)70442-5.

[7] C. Nanni, A. Versari, S. Chauvie, E. Bertone, A. Bianchi, M. Rensi, M. Bellò, A. Gallamini, F. Patriarca, F. Gay, B. Gamberi, P. Ghedini, M. Cavo, S. Fanti, E. Zamagni, Interpretation criteria for FDG PET/CT in multiple myeloma (IMPeTUs): final results. IMPeTUs (Italian myeloma criteria for PET USe), Eur. J. Nucl. Med. Mol. Imaging 45 (2018) 712–719, https://doi.org/10.1007/s00259-017-3909-8.

[8] R.J. Gillies, P.E. Kinahan, H. Hricak, Radiomics: images are more than pictures, they are data, Radiology 278 (2016) 563–577, https://doi.org/10.1148/radiol.2015151169.

[9] F. Valdora, N. Houssami, F. Rossi, M. Calabrese, A.S. Tagliafico, Rapid review: radiomics and breast cancer, Breast Cancer Res. Treat. 169 (2018), https://doi.org/10.1007/s10549-018-4675-4.

[10] A.S. Tagliafico, F. Valdora, G. Mariscotti, M. Durando, J. Nori, D. La Forgia, I. Rosenberg, F. Caumo, N. Gandolfo, N. Houssami, M. Calabrese, An exploratory radiomics analysis on digital breast tomosynthesis in women with mammographically negative dense breasts, Breast (Edinburgh, Scotland) 40 (2018) 92–96, https://doi.org/10.1016/j.breast.2018.04.016.

[11] J.F. Cohen, D.A. Korevaar, D.G. Altman, D.E. Bruns, C.A. Gatsonis, L. Hooft, L. Irwig, D. Levine, J.B. Reitsma, H.C.W. de Vet, P.M.M. Bossuyt, STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration, BMJ Open 6 (2016) e012799, , https://doi.org/10.1136/bmjopen-2016-012799.

[12] V. Koutoulidis, S. Fontara, E. Terpos, F. Zagouri, D. Matsaridis, D. Christoulas, E. Panourgias, E. Kastritis, M.A. Dimopoulos, L.A. Moulopoulos, Quantitative diffusion-weighted imaging of the bone marrow: an adjunct tool for the diagnosis of a diffuse MR imaging pattern in patients with multiple myeloma, Radiology 282 (2017) 484–493, https://doi.org/10.1148/radiol.2016160363.

[13] A. Stäbler, A. Baur, R. Bartl, R. Munker, R. Lamerz, M.F. Reiser, Contrast enhancement and quantitative signal analysis in MR imaging of multiple myeloma: assessment of focal and diffuse growth patterns in marrow correlated with biopsies and survival rates, AJR Am. J. Roentgenol. 167 (1996) 1029–1036, https://doi.org/10.2214/ajr.167.4.8819407.

[14] A.S. Tagliafico, P. Ameri, M. Bovio, M. Puntoni, E. Capaccio, G. Murialdo, C. Martinoli, Relationship between fatty degeneration of thigh muscles and vitamin d status in the elderly: a preliminary MRI study, Am. J. Roentgenol. 194 (2010) 728–734, https://doi.org/10.2214/AJR.09.3130.

[15] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, J. Buatti, S. Aylward, J.V. Miller, S. Pieper, R. Kikinis, 3D Slicer as an image computing platform for the Quantitative Imaging Network, Magn. Reson. Imaging 30 (2012) 1323–1341, https://doi.org/10.1016/j.mri.2012.05.001.

[16] V. Parekh, M.A. Jacobs, Radiomics: a new application from established techniques, Expert Rev. Precis. Med. Drug Dev. 1 (2016) 207–226, https://doi.org/10.1080/23808993.2016.1164013.

[17] F. Valdora, N. Houssami, F. Rossi, M. Calabrese, A.S. Tagliafico, Rapid review: radiomics and breast cancer, Breast Cancer Res. Treat. 169 (2018) 217–229, https://doi.org/10.1007/s10549-018-4675-4.

[18] A. Kosmala, A.M. Weng, B. Krauss, S. Knop, T.A. Bley, B. Petritsch, Dual-energy CT of the bone marrow in multiple myeloma: diagnostic accuracy for quantitative differentiation of infiltration patterns, Eur. Radiol. 28 (2018) 5083–5090, https://doi.org/10.1007/s00330-018-5537-5.

[19] A. Tagliafico, G. Tagliafico, C. Martinoli, Nerve density: a new parameter to evaluate peripheral nerve pathology on ultrasound, Prelim. Study, Ultras. Med. Biol. 36 (2010), https://doi.org/10.1016/j.ultrasmedbio.2010.07.009.

RESEARCH ACTIVITY

1. Tagliafico AS, González RP, Rossi F, Bignotti B, Martinoli C. Peripheral Nerves: Not Only Cross-sectional Area in the Era of Radiomics. Semin Musculoskelet Radiol. 2020 Apr;24(2):175-180. doi: 10.1055/s-0040-1701629.

2. Sconfienza LM, Adriaensen M, Albano D, et al. Clinical indications for image-guided interventional procedures in the musculoskeletal system: a Delphi-based consensus paper from the European Society of Musculoskeletal Radiology (ESSR)-Part II, elbow and wrist.
Ultrasound and Interventional Subcommittees of the European Society of Musculoskeletal Radiology (ESSR). Eur Radiol. 2020 Apr;30(4):2220-2230. doi: 10.1007/s00330-019-06545-6.

3. Rossi F, Bignotti B, Bianchi L, Picasso R, Martinoli C, Tagliafico AS. Radiomics of peripheral nerves MRI in mild carpal and cubital tunnel syndrome. Radiol Med. 2020 Feb;125(2):197-203. doi: 10.1007/s11547-019-01110-z.

4. Sconfienza LM, Adriaensen M, Albano D, et al. Clinical indications for image guided interventional procedures in the musculoskeletal system: a Delphi-based consensus paper from the European Society of Musculoskeletal Radiology (ESSR)-part III, nerves of the upper limb. Ultrasound and Interventional Subcommittees of the European Society of Musculoskeletal Radiology (ESSR). Eur Radiol. 2020 Mar;30(3):1498-1506. doi: 10.1007/s00330-019-06479-z.

5. Tagliafico AS, Bignotti B, Rossi F, Valdora F, Martinoli C. Local recurrence of soft tissue sarcoma: a radiomic analysis. Radiol Oncol. 2019 Sep 24;53(3):300-306. doi: 10.2478/raon-2019-0041.

6. Sconfienza LM, Adriaensen M, Albano D, et al. Clinical indications for image-guided interventional procedures in the musculoskeletal system: a Delphi-based

consensus paper from the European Society of Musculoskeletal Radiology (ESSR)-part I, shoulder. Eur Radiol. 2020 Feb;30(2):903-913. doi: 10.1007/s00330-019-06419-x.

7. Tagliafico AS, Bignotti B, Rossi F, Matos J, Calabrese M, Valdora F, Houssami N. Breast cancer Ki-67 expression prediction by digital breast tomosynthesis radiomics features.

Eur Radiol Exp. 2019 Aug 14;3(1):36. doi: 10.1186/s41747-019-0117-2.

8. Rossi F, Valdora F, Bignotti B, Torri L, Succio G, Tagliafico AS. Evaluation of body Computed Tomography-determined sarcopenia in breast cancer patients and clinical outcomes: A systematic review. Cancer Treat Res Commun. 2019;21:100154. doi: 10.1016/j.ctarc.2019.100154.

9. Bignotti B, Calabrese M, Signori A, Tosto S, Valdora F, Tagliafico A, Durando M, Mariscotti G. Background parenchymal enhancement assessment: Inter- and intra-rater reliability across breast MRI sequences. Eur J Radiol. 2019 May;114:57-61. doi: 10.1016/j.ejrad.2019.02.036.

10. Tagliafico AS, Isaac A, Bignotti B, Rossi F, Zaottini F, Martinoli C. Nerve Tumors: What the MSK Radiologist Should Know. Semin Musculoskelet Radiol. 2019 Feb;23(1):76-84. doi: 10.1055/s-0038-1676290.

11. Tagliafico A, Bignotti B, Rossi F, Rubino M, Civani A, Martinoli C. Clinical Contribution of Wrist and Hand Sonography: Pilot Study. J Ultrasound Med. 2019 Jan;38(1):141-148. doi: 10.1002/jum.14675.

ACKNOWLEDGEMENTS

I would like to acknowledge and give my warmest thanks to my tutor Prof. Alberto Tagliafico who made this work possible. His guidance carried me through all the stages of our Research Project with precious suggestion.

I would also like to give special thanks to my family and friends as a whole for their continuous support.