diagnostics

MDPI

Systematic Review

Radiomics Analysis for Multiple Myeloma: A Systematic Review with Radiomics Quality Scoring

Michail E. Klontzas [1,2,*], Matthaios Triantafyllou [1], Dimitrios Leventis [1], Emmanouil Koltsakis [3], Georgios Kalarakis [3,4], Antonios Tzortzakakis [4,5] and Apostolos H. Karantanas [1,2,*]

1 Department of Medical Imaging, University Hospital of Heraklion, 71110 Heraklion, Greece; manthostr@gmail.com (M.T.); dimitrios_leventis@yahoo.com (D.L.)
2 Department of Radiology, School of Medicine, University of Crete, Voutes Campus, 71003 Heraklion, Greece
3 Department of Radiology, Karolinska University Hospital, 14152 Stockholm, Sweden; emmanouilkolts@gmail.com (E.K.); georgios.kalarakis@ki.se (G.K.)
4 Division of Radiology, Department for Clinical Science, Intervention and Technology (CLINTEC), Karolinska Institutet, 14152 Stockholm, Sweden; antonios.tzortzakakis@ki.se
5 Medical Radiation Physics and Nuclear Medicine, Section for Nuclear Medicine, Karolinska University Hospital, 14186 Huddinge, Stockholm, Sweden
* Correspondence: miklontzas@gmail.com (M.E.K.); akarantanas@gmail.com (A.H.K.)

Abstract: Multiple myeloma (MM) is one of the most common hematological malignancies affecting the bone marrow. Radiomics analysis has been employed in the literature in an attempt to evaluate the bone marrow of MM patients. This manuscript aimed to systematically review radiomics research on MM while employing a radiomics quality score (RQS) to accurately assess research quality in the field. A systematic search was performed on Web of Science, PubMed, and Scopus. The selected manuscripts were evaluated (data extraction and RQS scoring) by three independent readers (R1, R2, and R3) with experience in radiomics analysis. A total of 23 studies with 2682 patients were included, and the median RQS was 10 for R1 (IQR 5.5–12) and R3 (IQR 8.3–12) and 11 (IQR 7.5–12.5) for R2. RQS was not significantly correlated with any of the assessed bibliometric data (impact factor, quartile, year of publication, and imaging modality) ($p > 0.05$). Our results demonstrated the low quality of published radiomics research in MM, similarly to other fields of radiomics research, highlighting the need to tighten publication standards.

Keywords: multiple myeloma; radiomics; machine learning; metastases; radiomics quality score

check for updates

Citation: Klontzas, M.E.; Triantafyllou, M.; Leventis, D.; Koltsakis, E.; Kalarakis, G.; Tzortzakakis, A.; Karantanas, A.H. Radiomics Analysis for Multiple Myeloma: A Systematic Review with Radiomics Quality Scoring. *Diagnostics* **2023**, *13*, 2021. https://doi.org/10.3390/diagnostics13122021

Academic Editor: Manuel Barberio

Received: 20 April 2023
Revised: 6 June 2023
Accepted: 8 June 2023
Published: 10 June 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Radiomics represents the image-based equivalent of biological omics analyses (e.g., transcriptomics and proteomics), which promises to offer high-fidelity analysis of images for precision medicine purposes [1]. Analysis of regions of interest in medical images can be performed by extracting radiomics features, which can be used to construct machine learning models that achieve a precise diagnosis, treatment response prediction, and disease prognosis [2,3]. Radiomics has been introduced as a promising image analysis method equivalent to other omics analyses aiming to achieve an image-based biopsy of regions of interest. Nonetheless, radiomics has yet to reach clinical practice but represents a promising research tool for developing predictive image-based signatures that can assist in the diagnosis and treatment of various diseases [4,5]. The quality of radiomics studies can be assessed with the use of the radiomics quality score (RQS) [6,7], which scores radiomics research against a series of standards, including but not limited to the quality of reporting, segmentation, feature extraction, feature selection, the calibration and validation of machine learning models, and the provision of open access data.

Multiple myeloma (MM) is one of the most common hematological malignancies, characterized by osteolytic lesions [8]. The heterogeneous nature of the disease complicates its

diagnosis and treatment, and the similarities of MM lesions relative to osteolytic metastases during imaging complicate the image-based diagnosis of the disease [9]. The diagnostic approach of MM is based on whole-body CT, MRI, and PET-CT [10]. A series of radiomics research papers utilizing images from these modalities have been published in an attempt to achieve the differentiation of MM from metastases, prediction of treatment response, identification of molecular subgroups of MM, and analysis of patient survival.

The aims of this systematic review were (a) to provide a comprehensive analysis of the applications of radiomics in MM research and (b) to score manuscripts using RQS to benchmark their quality against existing standards.

2. Materials and Methods

The protocol of this study has been registered with the PROSPERO international register for systematic reviews (https://www.crd.york.ac.uk/prospero/, Record ID CRD42023409189, accessed 8 June 2023). Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines were used to prepare this manuscript [11].

2.1. Database Search Strategy and Selection of Relevant Studies

The search for relevant papers was performed between 1 January 2010 and 1 April 2023 in three databases (dPubMed, Scopus, and Web of Science). The search was performed with the strings "multiple myeloma", "radiomics", and "texture". The detailed strings used in each of the three databases can be found in Supplementary File S1. Studies from all three databases were collected, and duplicate studies were excluded from further analysis. Three radiologists (MEK, MT, and DL) with 5, 2, and 2 years of experience in radiomics research examined the abstracts of all studies to exclude those that were not eligible: (1) review papers, (2) abstracts and conference papers, (3) editorials, (4) non-English papers. The records were further screened to exclude manuscripts that did not perform radiomics analysis.

2.2. Literature Data Extraction

From each included study, specific information was extracted, including author names, country of origin, year of publication, journal, number of patients, the purpose of the study, imaging modality used, software for radiomics analysis, and the type of MM. Each journal's impact factor was recorded according to the 2021 Journal Citation Reports (Clarivate) and the journal's quartile according to Scimago Journal and Country Rank (https://www.scimagojr.com, accessed 20 April 2023).

2.3. Radiomics Quality Score (RQS)

The quality of radiomics research presented in each study was evaluated by three readers blinded to the results of each other using RQS [6,7]. The first and most experienced reader (R1–MEK) had 5 years of experience in radiomics research and 10 years of experience in medical imaging research, whereas the second and third (R2 and R3—MT and DL) had 2 years of experience in radiomics and 3 years of experience in medical imaging research. Before the initiation of scoring, a training session was performed using radiomics papers unrelated to MM to ensure that manuscripts were scored in precisely the same manner by all readers. Each reader was blinded to the scores that the rest of the readers gave. The RQS consists of 16 items, with a total score ranging between −8 and 36. A total percentage (out of 36) was also calculated for each manuscript.

2.4. Statistical Analysis

Statistical analysis was performed using SPSS v 29 (IBM SPSS for Mac, Armonk, NY, USA). Variables are expressed as frequencies and percentages (categorical) or medians with interquartile ranges (continuous). An adherence metric was calculated to assess adherence to RQS by awarding one point if the authors had gained the minimum points for each RQS item. Data normality was assessed using the Shapiro–Wilk test. RQS comparisons

between groups of studies with different characteristics (impact factor, year of publication, journal quartile, and modality) were performed using Mann–Whitney U or Kruskall–Wallis tests according to the number of groups. Agreement between readers was evaluated using the intraclass correlation coefficient (ICC) assessed for absolute agreement, considering acceptable ICC with values >0.75 (good: 0.75–0.9, excellent > 0.9). Statistical significance was defined with α set at 0.05.

3. Results

3.1. Study Selection

After excluding duplicate entries (n = 30), 101 records from all three databases were screened. A total of 21 studies were excluded due to being an inappropriate manuscript type (non-English, review, conference paper, and editorial/letter to the editor), and 57 studies were excluded because of a non-radiomics study design (studies that did not extract radiomics features). This yielded a final sum of n = 23 manuscripts, which were included for analysis (Figure 1).

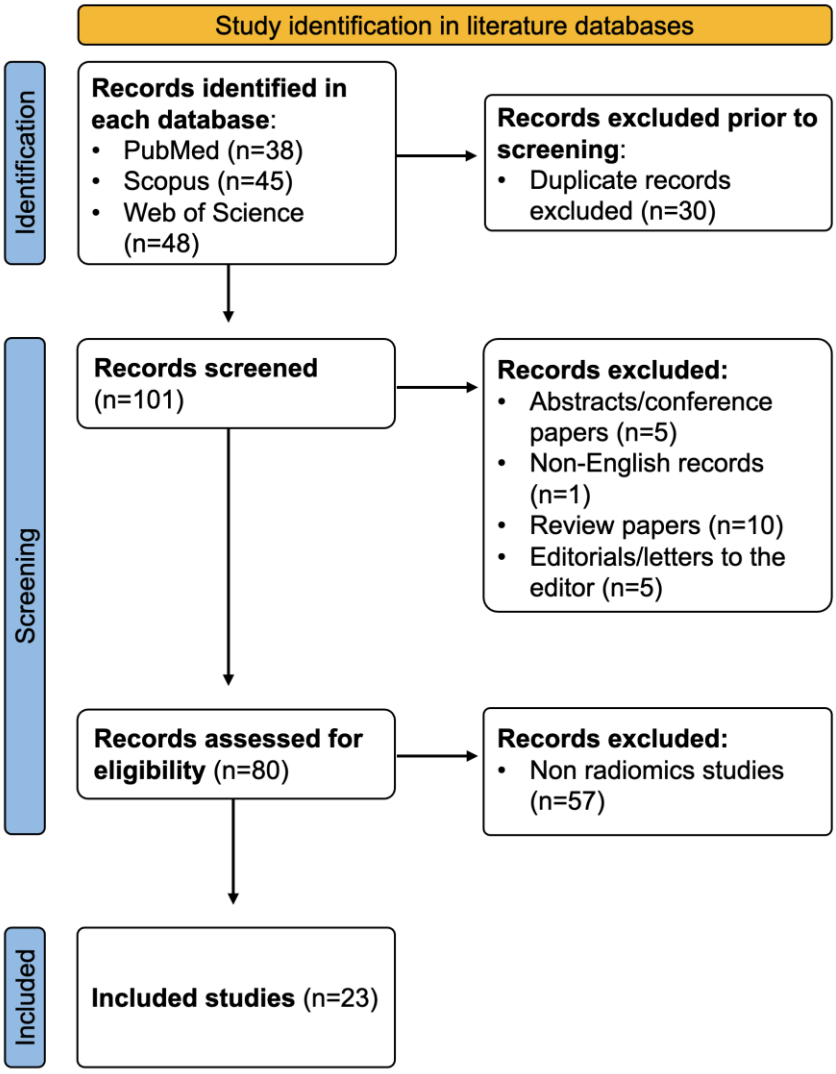


Figure 1. Flow chart detailing study selection methodology.

3.2. Analysis of Included Studies

Detailed characteristics of the 23 included studies can be found in Table 1. In total, 8 out of 23 studies (34.8%) were published in radiology journals, 4/23 (17.4%) were published in oncological journals, 3/23 (13%) were published in nuclear medicine journals. More

than half of the studies (12/23) were published in journals with an Impact Factor > 5. Most studies were published in Q1 journals (15/23—65.2%), with five studies in Q2 (21.7%), two studies in Q3, and one study in a Q4 journal. No study was published before 2019, whereas nine studies were published in each of the years 2021 and 2022. China and Germany were the main countries of origin for published studies, with 34.8% and 21.7% of the studies originating from each of the two, respectively.

Table 1. Study characteristics.

Study	Journal	Journal Field	Journal Quartile	JCR Impact Factor (2021)	Year	Country
Chen [12]	*Frontiers in Oncology*	Oncology	Q1	5.738	2022	China
Ekert [13]	*Cancers*	Oncology	Q1	6.575	2020	Germany
Hwang [14]	*Scientific Reports*	General	Q1	4.996	2019	Korea
Jamet [15]	*European Journal of Nuclear Medicine and Molecular Imaging*	Nuclear Medicine	Q1	10.057	2021	France
Jin [16]	*Frontiers in Medicine*	General	Q1	5.058	2022	China
Lee [17]	*Diagnostics*	General	Q2	3.992	2023	Korea/USA
Li [18]	*Frontiers in Oncology*	Oncology	Q1	5.738	2021	China
Li [19]	*BioMed Research International*	General	Q2	NA	2022	China
Liu [20]	*European Radiology*	Radiology	Q1	7.034	2021	China
Liu [21]	*Radiologia Medica*	Radiology	Q1	6.313	2021	China
Mannam [22]	*Indian Journal of Nuclear Medicine*	Nuclear Medicine	Q4	NA	2022	India
Mesguich [23]	*Nuclear Medicine Communications*	Nuclear Medicine	Q3	1.698	2021	France
Milara [24]	*Computer Methods and Programs in Biomedicine*	Computer Science	Q1	7.027	2022	Spain
Park [25]	*Diagnostics*	General	Q2	3.992	2022	Korea
Reinert [26]	*European Radiology*	Radiology	Q1	7.034	2020	Germany
Reinert [27]	*European Journal of Radiology*	Radiology	Q1	4.531	2021	Germany
Ripani [28]	*Clinical Lymphoma, Myeloma and Leukemia*	Hematology	Q2	2.822	2021	Italy
Schenone [29]	*Diagnostics*	General	Q2	3.992	2021	Italy
Tagliafico [30]	*European Journal of Radiology*	Radiology	Q1	4.531	2022	Italy
Wennmann [31]	*Investigative Radiology*	Radiology	Q1	10.065	2022	Germany
Wennmann [32]	*Investigative Radiology*	Radiology	Q1	10.065	2023	Germany
Wu [33]	*Journal of Computer Assisted Tomography*	Radiology	Q3	2.081	2022	China
Xiong [34]	*Frontiers in Oncology*	Oncology	Q1	5.738	2021	China

The methodological details of individual studies are presented in Table 2. A total of 2682 patients were examined in the 23 included studies. Several radiomics applications on multiple myeloma have been described in the studies included, with the majority (6 out of 23 studies—26.1%) aimed at differentiating between MM and metastases and the same proportion aimed at predicting treatment response or prognosis. Only 4/23 studies aimed at diagnosing MM compared to normal bone marrow or other hematological conditions. Three out of twenty-three studies evaluated the pattern of infiltration of the bone marrow. In contrast, other miscellaneous applications such as the disease's load prediction and the presence of high-risk cytogenetic abnormalities represented the minority of the literature (1–2 studies for each application). One of the studies aimed at evaluating the technical reproducibility of radiomics in patients with MM. Almost half of the studies utilized MRI (11/23—47.8%), with PET-CT and CT (single or dual-energy) being the second-most common modalities (6/23 studies each—26.1%). Various commercial and free software was used; a detailed list can be found in Table 2. Finally, most studies did not define which specific MM type was examined.

Table 2. Methodological characteristics of included studies.

Study	Patients (Training/ Validation)	Number of Extracted Features (Location of ROI/VOI)	Purpose	Modality	Sequences Used	Software Used (Segmentation/ Feature Extraction/Model Building	MM Type/Specific Location
Chen [12]	217 (No test set, but 5-fold CV)	30 (spinal focal bone lesions)	Differential diagnosis between MM and metastases	MRI	CE T1-w	ITK-SNAP/python	Spinal MM
Ekert [13]	67	92 (focal bone marrow lesions)	Prediction of treatment response	MRI	WB T1-w, ADC, STIR	Not Specified	Not Specified
Hwang [14]	467 (360/107)	56,862 (whole lumbar spine bone marrow)	Differential diagnosis between hematological malignancies and normal marrow	MRI	T1-w	MATLAB	Not Specified
Jamet [15]	139 (98/41)	17 (focal bone lesions and extramedullary disease)	Prediction of Prognosis	PET-CT		Python	Not Specified
Jin [16]	131 (70:30 split)	279 (spinal lesions)	Differential diagnosis between MM and metastases	PET-CT		MadZa/python	Spinal MM
Lee [17]	225 (175/50)	1218 (focal bone lesions)	Differential diagnosis between MM and metastases	CE CT		ITK-SNAP, mlr package	Not Specified
Li Y [18]	121 (84/37)	1316 (whole lumbar spine bone marrow)	Prediction of Prognosis	MRI	T1-w, T2FS	ITK_SNAP, R programming	Spinal MM
Li Y [19]	95 (77 /18)	1316 (whole lumbar spine bone marrow)	Prediction of treatment response	MRI	T1-w, T2FS	AI Kit software AK GE, ITK-SNAP, R programming	Not Specified
Liu [20]	241 (80:20 split)	1409 (spinal focal bone lesions)	Differential diagnosis between MM and metastases	MRI	T1-w, T2-w FS	RadCloud AI Scientific Research platform	Spinal MM
Liu [21]	50 (4/5 of patients in training set)	1409 (spinal focal bone lesions)	Prediction of high risk Cytogenetic abnormalities	MRI	T1-w, T2-w, T2FS	RadCloud platform	Spinal MM
Mannam [22]	40 (26/14)	138 (focal bone lesions)	Differential diagnosis between MM and metastases	PET-CT	CT venous phase	LifeX/Weka Data Mining/XLSTAT	Not Specified
Mesguich [23]	30 (20/10)	174 (whole spinal bone marrow)	Diagnosis of MM (no alternative diagnosis)	PET-CT	CT and PET SUV	Horos/Python	Not Specified
Milara [24]	39 (no test set, only 5-fold cross validation)	27 (whole skeleton)	Prediction of disease load	PET-CT		MATLAB/Statical Parametric Mapping 12/Orange	Not Specified
Park [25]	208 (160/48)	1218 (whole axial skeleton)	Diagnosis of MM (no alternative diagnosis)	CE-CT		Syngo.via/R programming	Not Specified

Table 2. *Cont.*

Study	Patients (Training/ Validation)	Number of Extracted Features (Location of ROI/VOI)	Purpose	Modality	Sequences Used	Software Used (Segmentation/ Feature Extraction/Model Building	MM Type/Specific Location
Reinert [26]	110 (no validation)	92 (bone marrow of vertebral bodies T10-L5)	Prediction of bone marrow infiltration	DE-CT		Syngo.via/Python	Spinal MM
Reinert [27]	44	41 (whole bone marrow)	Correlation with hematological parameters	DE-CT		Syngo.via	Not Specified
Ripani [28]	45	Unspecified (whole bone marrow)	Prediction of disease progression	PET-CT		LifeX	Smoldering MM
Schenone [29]	33	109 (focal bone lesions)	Prediction of treatment response	CT		Bone GUI, 3D Slicer	Not Specified
Tagliafico [30]	60	Not specified (paraspinal muscle sarcopenia assessment)	Prediction of MM phenotype (focal vs. diffuse)	CT		3D Slicer	Not Specified
Wennmann [31]	102 (50/52)	296 ("bone marrow spaces")	Prediction of MM phenotype (normal vs. focal vs. diffuse)	MRI	WB T1-w	Python	Smoldering MM
Wennmann [32]	55	295 (bone marrow of the left hip)	Assessment of radiomics reproducibility	MRI	T1-w, T2-w	MITK phenotyping/Python	Not Specified
Wu [33]	123 (85/38)	1702 (focal bone lesions)	Prediction of treatment response	MRI	T1-w, T2-w, STIR	3D Slicer/R programming	Spinal MM
Xiong [34]	107 (75/32)	282 (lumbar focal bone lesions)	Differential Diagnosis between MM and metastases	MRI	T1-w, T2-w	MadZa/R programming	Spinal MM

3.3. Radiomics Quality Score (RQS)

Adherence to RQS items varied significantly. As shown in Figure 2, 96% of manuscripts contained some discrimination statistics (e.g., AUC with 95% confidence intervals), and more than 85% (87%) of papers included a sufficiently detailed imaging protocol. It is important that 13% of the studies did not attempt any validation (including internal testing). In contrast, only 9% of the studies included calibration statistics for the presented machine learning models. Feature reduction to account for the possibility of overfitting was performed in 83% of manuscripts, and 57% of studies established the ground truth with the use of a gold standard (e.g., bone marrow biopsy). Finally, only 4% of manuscripts provided either an open access code or open access data, whereas no manuscript included a cost-effectiveness analysis or phantom standardization of radiomics features.

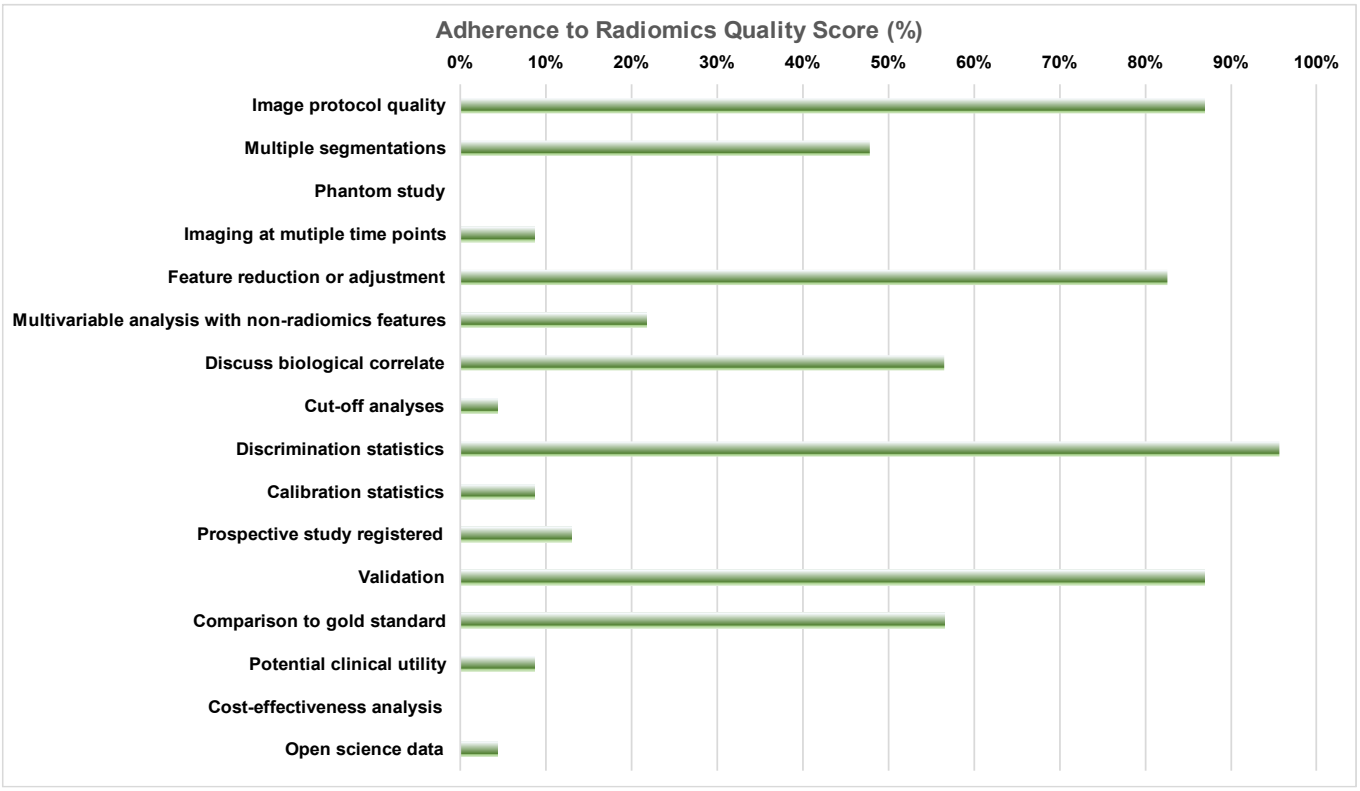


Figure 2. Adherence to radiomics quality score items according to the most experienced reader (R1), expressed as the percentage of manuscripts that received the minimum score for each individual item.

The total RQS for each of the studies for all three readers is presented in Table 3. The median RQS was 10 for R1 (IQR 5.5–12) and R3 (IQR 8.3–12) and 11 (IQR 7.5–12.5) for R2. The agreement between readers was good (almost excellent), with an ICC of 0.851 (95% CI 0.7 to 0.932). Subgroup analysis showed that there are no statistically significant differences between RQS for papers published in journals with a high (>5) and low impact factor (Figure 3 and Table 4), between papers published before or after 2021, or papers in high or low quartiles or papers dealing with different imaging modalities (Table 4). Impact factor analysis was repeated by removing the only methodological study (which does not contain a predictive model) from our group [32], noting that there was still no significant correlation between the impact factor and the RQS score ($p > 0.05$ for all readers).

Table 3. Total radiomics quality score for each study per reader.

Study	RQS of R1	RQS of R2	RQS of R3
Chen [12]	5	5	3
Ekert [13]	3	2	3
Hwang [14]	5	5	9
Jamet [15]	15	18	20
Jin [16]	13	14	16
Lee [17]	12	13	10
Li Y [18]	9	12	10
Li Y [19]	9	11	11
Liu [20]	12	11	10
Liu [21]	10	12	7
Mannam [22]	0	10	10
Mesguich [23]	15	15	20
Milara [24]	10	10	8
Park [25]	11	11	11
Reinert [26]	3	3	9

Table 3. *Cont.*

Study	RQS of R1	RQS of R2	RQS of R3
Reinert [27]	6	6	10
Ripani [28]	3	3	14
Schenone [29]	9	9	8
Tagliafico [30]	9	9	11
Wennmann [31]	12	12	13
Wennmann [32]	15	15	9
Wu [33]	17	17	13
Xiong [34]	12	12	8

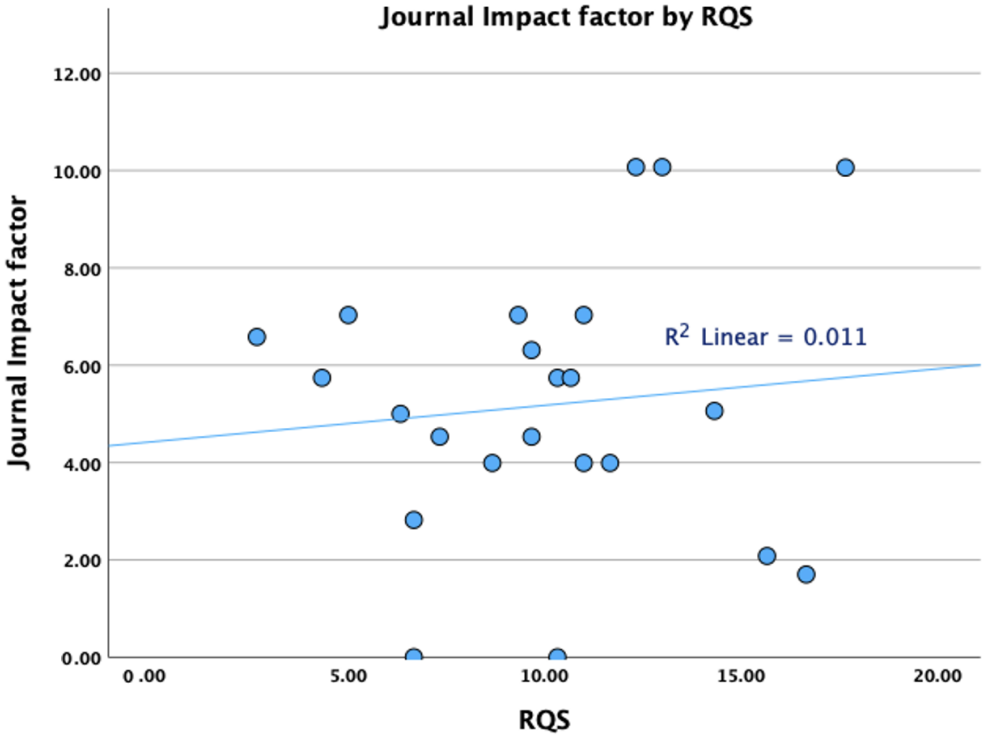


Figure 3. Scatter plot demonstrating the relationship between journal impact factors and radiomics quality scores.

Table 4. Factor subgroup analysis of radiomics quality scores for each reader. Numbers represent median scores (interquartile range).

Category	*N*	Reader 1	Reader 2	Reader 3
Journal Quartile		*p = 0.686*	*p = 0.399*	*p = 0.241*
Q1	15	10 (5.5–12)	11 (5.5–12)	9 (8–10.5)
Q2	5	9 (9–11)	11 (9–11)	11 (10–11)
Q3–4	3	15 (7.5–16)	15 (12.5–16)	13 (11.5–16.5)
Impact Factor		*p = 0.68*	*p = 0.214*	*p = 0.214*
1–5	11	9 (5.5–11.5)	10 (7.5–12)	11 (10–12)
>5	12	11 (8–12.25)	12 (8.8–12.5)	9 (7.8–10.8)
Modality		*p = 0.631*	*p = 0.432*	*p = 0.393*
CT	3	8.5 (5.3–11.3)	8.5 (8–12)	10 (7.5–10.5)
MRI	11	10 (7–12)	12 (8–12)	9 (7.5–10.5)
PET-CT	6	9.5 (7.5–13.5)	10 (9–14.3)	12.5 (9.5–17)
Publication Year		*p = 0.414*	*p = 1*	*p = 0.214*
2010–2021	12	9 (4.5–12)	10 (4.5–12)	9.5 (8–11)
2022–2023	11	11 (9–12.5)	11 (10–13.5)	11 (9.5–12)

4. Discussion

Herein, we presented a comprehensive analysis of radiomics studies on MM. We analyzed study characteristics and demonstrated that the quality of radiomics studies published for the evaluation of MM is inadequate. Importantly, we demonstrated low adherence rates to most RQS items and a low total RQS, and we showed that low quality is generic and not specific to the journal's characteristics. This finding is extremely important in evaluating published research and highlights the need to conduct high-quality research in the field.

The issue of inadequate-quality published radiomics research has been recently highlighted [5] in studies dealing with various topics [35,36], strongly indicating a potential lack of reproducibility. Most examined studies did not include external validation of their results, with two of them also skipping testing using internal data. The lack of external validation significantly limits the generalization capacity of machine learning models and is a common problem in published papers on AI using medical images, where only 6–10% of published studies have been tested on an external dataset [37]. This is of high importance since algorithm performance is consistently lower when validated on external datasets and even lower in "real-life" conditions [37,38]. Our findings signify the fact that even though the published algorithms may have the potential to revolutionize MM diagnosis and management, the lack of external validation hinders their adoption, reducing the trust in the results.

RQS is, at the moment, the most important tool for evaluating radiomics research. Adherence to RQS items is supposed to indicate high-quality research. In our case, some items were either addressed in none or a very limited subset of the studies, including the use of phantoms, cost-effectiveness analysis, and the provision of open access data. The lack of open access data may be attributed to limitations related to the publication of patient data. However, anonymized radiomics numerical values do not fall under this category. The publication of such studies indicates a lack of robust reviewing practices and insufficient journal and reviewer expertise in radiomics and machine learning. The rapid increase in machine learning publications in musculoskeletal imaging [39] has caused an increased demand for expert reviewers. Thus, assigning such manuscripts to inexperienced reviewers with potential expertise in musculoskeletal or hematological malignancies can lead to the publication of low-quality research, even in high-impact journals. As indicated by our results, there was no difference between high- and low-impact journals regarding their RQS score. To overcome this problem, guides for radiomics research have been published [1,3], providing basic directions to reviewers; moreover, journals need to adhere to the basic standards of RQS. It also needs to be pointed out that RQS has not been specifically designed to evaluate studies that do not contain predictive models. Therefore, technical radiomics studies may receive a lower RQS score because they present no predictive models. Our study included one such technical manuscript by Wennmann et al. [31], which received a median RQS of 12. This is almost identical to the median RQS of all studies in our sample; therefore, it was not analyzed separately since it does not negatively affect the results of the study. However, it needs to be pointed out that such studies can be of higher quality than the quality indicated by their RQS score since they are negatively scored for the absence of a predictive model that they were not supposed to analyze.

Radiomics, a rapidly evolving field within medical imaging, focuses on extracting high-dimensional quantitative data from medical images, utilizing these data to uncover hidden information that is not readily discernible by the human eye [1]. This innovative approach presents a promising frontier for diagnosing MM and the differentiation between MM and lytic metastases. The complex nature of MM lesions, coupled with their variable appearance in imaging, creates a pressing need for advanced diagnostic techniques that enable accurate, efficient, and personalized interventions for affected patients. MM is a hematologic malignancy characterized by lytic bone lesions, making their distinction from lytic metastases caused by other types of malignancy a challenging task for clinicians [10]. Accurate diagnosis and differentiation are crucial as the treatment strategies, prognoses,

and overall management of MM patients are markedly different from those with metastatic bone disease. Radiomics, by leveraging advanced machine learning algorithms and computational models, has the potential to identify unique, subvisual patterns and features within MM-affected bones that can accurately distinguish them from lytic metastases. Radiomics features provide a comprehensive, quantitative analysis of tumor characteristics, which can improve diagnostic accuracy and potentially reveal novel imaging biomarkers for MM [3,4]. Consequently, these advancements may lead to earlier and more personalized interventions for patients with multiple myeloma, ultimately enhancing their prognosis and quality of life.

One of the most important reasons that radiomics has not yet found extensive applications on MM is the nature of MM lesions, which are disseminated across the bone marrow, presenting an important challenge with regard to the segmentation of the entire tumor load given that this is not evenly distributed across the skeleton [40] and the mutations associated with the lesions can be spatially heterogeneous [41]. Therefore, in order to implement radiomics in clinical practice for MM patients, methods that allow the segmentation of multiple focal lesions and extended areas of diffuse infiltration or whole bones are required since the manual segmentation of this scale is extremely tedious and potentially unreliable. Methods such as the atlas-based semi-automatic segmentation of whole-body diffusion-weighted imaging and deep learning applications combined with radiomics have already been proposed [31,42–44] and may be the solution to the future translation of radiomics research to the clinic.

Several medical imaging modalities, including computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET), can be utilized for radiomic analysis in MM patients. Our work showed that almost half of the manuscripts utilized MRI-based radiomics. CT-derived radiomics features may capture structural alterations in the trabecular bone, while MRI-based radiomics can offer insights into the tissue composition in the bone marrow. PET-based radiomics, on the other hand, can evaluate the metabolic activity of MM lesions and assess their response to therapy [8,10]. By integrating multi-modal imaging data, radiomics offers a comprehensive view of MM lesions, potentially enabling more accurate and precise diagnoses. The future integration of radiomics with other omics data, such as genomics, proteomics, and metabolomics, could create a holistic understanding of MM, uncovering complex relationships between imaging phenotypes and the underlying molecular mechanisms driving disease progression. This integration, often referred to as radiogenomics or radiotranscriptomics [45,46], can pave the way for personalized medicine approaches in MM, guiding clinicians in tailoring treatment plans based on the specific characteristics of each patient's disease. Utilizing such integration methods for the diagnosis of MM would necessitate high-quality radiomics methods that are not currently available in the MM literature. The development of high-quality radiomics signatures for implementation in the clinical management of multiple myeloma requires interdisciplinary collaboration among radiologists, oncologists, physicists, data scientists, and bioinformaticians. This collaborative approach can foster the development of novel radiomic models, drive the discovery of new imaging biomarkers, and ultimately contribute to the improvement of diagnostic accuracy, treatment planning, and overall patient outcomes.

Our study has certain strengths and limitations. The systematic review of the literature and the comprehensive analysis of study characteristics represent the strengths of our study. Another important strength is the high ICC between the three readers. Such high agreement between readers has been reported in other RQS assessment studies [47]. Nonetheless, one limitation of our study is related to the inherent limitations of RQS. The score itself is empirical and was not validated when first published by Lambin et al. [4]. Therefore, it includes items, such as the use of phantoms for feature standardization and cost-effectiveness analysis, that are consistently not found in any published radiomics papers, even the ones that adhere to the highest quality standards [36,47,48]. Thus, it has been suggested that these items could unnecessarily reduce the total RQS, preventing studies from obtaining

the maximum number (36) of points. Nonetheless, in our study, the maximum total RQS was 20, which cannot be compensated by scoring points in the items mentioned above. In similar RQS systematic reviews for other types of disease, the maximum RQS score was similar. The same is true for the median RQS value, which was found to be 10 in our study, comparable to values shown by other published systematic reviews. Such examples include a maximum of 22 in studies of hepatocellular carcinoma [49] and a maximum of 16 in studies of ischemic strokes [50].

5. Conclusions

In conclusion, radiomics presents a powerful and innovative approach for evaluating MM. By unlocking the hidden information within medical images, radiomics has the potential to revolutionize MM diagnosis, risk stratification, and treatment planning. Our study has highlighted the low scientific quality of radiomics papers related to MM, similarly to other fields of radiomics research, demonstrating that the methodological limitations of existing studies are not related to bibliometric (journal, impact factor, quartile, etc.) data. These findings emphasize the need to tighten current publication standards in order to publish radiomics research studies of high quality.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/diagnostics13122021/s1. Supplementary File S1: Full strings used for database search.

Author Contributions: Conceptualization, M.E.K.; methodology, M.E.K.; software, M.E.K.; validation, M.E.K. and A.H.K.; formal analysis, M.E.K.; investigation, M.T., D.L., E.K., G.K. and A.T.; data curation, M.E.K., D.L. and M.T.; writing—original draft preparation, M.E.K., M.T., D.L., E.K., G.K., A.T. and A.H.K.; writing—review and editing, M.E.K., M.T., D.L., E.K., G.K., A.T. and A.H.K.; visualization, M.E.K., M.T., D.L., E.K., G.K., A.T. and A.H.K.; supervision, M.E.K. and A.H.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data required to reproduce results are provided in the paper.

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

References

1. van Timmeren, J.E.; Cester, D.; Tanadini-Lang, S.; Alkadhi, H.; Baessler, B. Radiomics in medical imaging—"How-to" guide and critical reflection. *Insights Imaging* **2020**, *11*, 91. [CrossRef]
2. Mühlbauer, J.; Egen, L.; Kowalewski, K.-F.; Grilli, M.; Walach, M.T.; Westhoff, N.; Nuhn, P.; Laqua, F.C.; Baessler, B.; Kriegmair, M.C. Radiomics in renal cell carcinoma—A systematic review and meta-analysis. *Cancers* **2021**, *13*, 1348. [CrossRef]
3. Papanikolaou, N.; Matos, C.; Koh, D.M. How to develop a meaningful radiomic signature for clinical use in oncologic patients. *Cancer Imaging* **2020**, *20*, 33. [CrossRef] [PubMed]
4. Gillies, R.J.; Kinahan, P.E.; Hricak, H. Radiomics: Images are more than pictures, they are data. *Radiology* **2016**, *278*, 563–577. [CrossRef] [PubMed]
5. Pinto Dos Santos, D.; Dietzel, M.; Baessler, B. A decade of radiomics research: Are images really data or just patterns in the noise? *Eur. Radiol.* **2021**, *31*, 1–4. [CrossRef]
6. Lambin, P.; Leijenaar, R.T.H.; Deist, T.M.; Peerlings, J.; de Jong, E.E.C.; van Timmeren, J.; Sanduleanu, S.; Larue, R.T.H.M.; Even, A.J.G.; Jochems, A.; et al. Radiomics: The bridge between medical imaging and personalized medicine. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 749–762. [CrossRef]
7. Klontzas, M.E.; Gatti, A.A.; Tejani, A.S.; Kahn, C.E. AI reporting guidelines: How to select the best one for your research. *Radiol. Artif. Intell.* **2023**, *5*, e230055. [CrossRef] [PubMed]
8. Healy, C.F.; Murray, J.G.; Eustace, S.J.; Madewell, J.; O'Gorman, P.J.; O'Sullivan, P. Multiple myeloma: A review of imaging features and radiological techniques. *Bone Marrow Res.* **2011**, *2011*, 583439. [CrossRef]
9. Sandal, R.; Mishra, K.; Jandial, A.; Khadwal, A.; Malhotra, P. Multiple myeloma and pepperpot skull. *QJM* **2018**, *111*, 917. [CrossRef]

10. Filho, A.G.O.; Carneiro, B.C.; Pastore, D.; Silva, I.P.; Yamashita, S.R.; Consolo, F.D.; Hungria, V.T.M.; Sandes, A.F.; Rizzatti, E.G.; Nico, M.A.C. Whole-body imaging of multiple myeloma: Diagnostic criteria. *RadioGraphics* **2019**, *39*, 1077–1097. [CrossRef]
11. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *PLoS Med* **2021**, *18*, e1003583. [CrossRef] [PubMed]
12. Chen, K.; Cao, J.; Zhang, X.; Wang, X.; Zhao, X.; Li, Q.; Chen, S.; Wang, P.; Liu, T.; Du, J.; et al. Differentiation between spinal multiple myeloma and metastases originated from lung using multi-view attention-guided network. *Front. Oncol.* **2022**, *12*, 981767. [CrossRef] [PubMed]
13. 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]
14. Hwang, E.-J.; Jung, J.-Y.; Lee, S.K.; Lee, S.E.; Jee, W.H. Machine learning for diagnosis of hematologic diseases in magnetic resonance imaging of lumbar spines. *Sci. Rep.* **2019**, *9*, 6046. [CrossRef] [PubMed]
15. 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. Imag.* **2021**, *48*, 1005–1015. [CrossRef] [PubMed]
16. Jin, Z.; Wang, Y.; Wang, Y.; Mao, Y.; Zhang, F.; Yu, J. Application of 18F-FDG PET-CT images based radiomics in identifying vertebral multiple myeloma and bone metastases. *Front. Med.* **2022**, *9*, 874847. [CrossRef] [PubMed]
17. Lee, S.; Lee, S.Y.; Kim, S.; Huh, Y.J.; Lee, J.; Lee, K.E.; Jung, J.Y. Differentiating multiple myeloma and osteolytic bone metastases on contrast-enhanced computed tomography scans: The feasibility of radiomics analysis. *Diagnostics* **2023**, *13*, 755. [CrossRef]
18. Li, Y.; Liu, Y.; Yin, P.; Hao, C.; Hao, C.; Sun, C.; Chen, L.; Wang, S.; Hong, N. MRI-based bone marrow radiomics nomogram for prediction of overall survival in patients with multiple myeloma. *Front. Oncol.* **2021**, *11*, 709813. [CrossRef]
19. Li, Y.; Yin, P.; Liu, Y.; Hao, C.; Chen, L.; Sun, C.; Wang, S.; Hong, N. Radiomics models based on magnetic resonance imaging for prediction of the response to bortezomib-based therapy in patients with multiple myeloma. *BioMed Res. Int.* **2022**, *2022*, 6911246. [CrossRef]
20. Liu, J.; Guo, W.; Zeng, P.; Geng, Y.; Liu, Y.; Ouyang, H.; Lang, N.; Yuan, H. Vertebral MRI-based radiomics model to differentiate multiple myeloma from metastases: Influence of features number on logistic regression model performance. *Eur. Radiol.* **2022**, *32*, 572–581. [CrossRef]
21. Liu, J.; Wang, C.; Guo, W.; Zeng, P.; Liu, Y.; Lang, N.; Yuan, H. A preliminary study using spinal MRI-based radiomics to predict high-risk cytogenetic abnormalities in multiple myeloma. *Radiol. Med.* **2021**, *126*, 1226–1235. [CrossRef]
22. Mannam, P.; Murali, A.; Gokulakrishnan, P.; Venkatachalapathy, E.; Venkata Sai, P. Radiomic analysis of positron-emission tomography and computed tomography images to differentiate between multiple myeloma and skeletal metastases. *Indian J. Nucl. Med.* **2022**, *37*, 217–226. [CrossRef] [PubMed]
23. Mesguich, C.; Hindie, E.; de Senneville, B.D.; Tlili, G.; Pinaquy, J.-B.; Marit, G.; Saut, O. Improved 18-FDG PET/CT diagnosis of multiple myeloma diffuse disease by radiomics analysis. *Nucl. Med. Commun.* **2021**, *42*, 1135–1143. [CrossRef] [PubMed]
24. Milara, E.; Gómez-Grande, A.; Tomás-Soler, S.; Seiffert, A.P.; Alonso, R.; Gómez, E.J.; Martínez-López, J.; Sánchez-González, P. Bone marrow segmentation and radiomics analysis of [18F]FDG PET/CT images for measurable residual disease assessment in multiple myeloma. *Comput. Methods Progr. Biomed.* **2022**, *225*, 107083. [CrossRef] [PubMed]
25. Park, H.; Lee, S.-Y.; Lee, J.; Pak, J.; Lee, K.; Lee, S.E.; Jung, J.Y. Detecting multiple myeloma infiltration of the bone marrow on CT scans in patients with osteopenia: Feasibility of radiomics analysis. *Diagnostics* **2022**, *12*, 923. [CrossRef]
26. Reinert, C.P.; Krieg, E.; Esser, M.; Nikolaou, K.; Bösmüller, H.; Horger, M. Role of computed tomography texture analysis using dual-energy-based bone marrow imaging for multiple myeloma characterization: Comparison with histology and established serologic parameters. *Eur. Radiol.* **2021**, *31*, 2357–2367. [CrossRef]
27. 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]
28. Ripani, D.; Caldarella, C.; Za, T.; Rossi, E.; de Stefano, V.; Giordano, A. Progression to symptomatic multiple myeloma predicted by texture analysis-derived parameters in patients without focal disease at ^{18}F-FDG PET/CT. *Clin. Lymph. Myel. Leuk.* **2021**, *21*, 536–544. [CrossRef]
29. 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. [CrossRef]
30. 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]
31. Wennmann, M.; Klein, A.; Bauer, F.; Chmelik, J.; Grözinger, M.; Uhlenbrock, C.; Lochner, J.; Nonnenmacher, T.; Rotkopf, L.T.; Sauer, S.; et al. Combining deep learning and radiomics for automated, objective, comprehensive bone marrow characterization from whole-body MRI: A multicentric feasibility study. *Investig. Radiol.* **2022**, *57*, 752–763. [CrossRef]

32. Wennmann, M.; Bauer, F.; Klein, A.; Chmelik, J.; Grözinger, M.; Rotkopf, L.T.; Neher, P.; Gnirs, R.; Kurz, F.T.; Nonnenmacher, T. In vivo repeatability and multiscanner reproducibility of MRI radiomics features in patients with monoclonal plasma cell disorders. *Investig. Radiol.* **2023**, *58*, 253–264. [CrossRef]
33. Wu, Z.M.; Bian, T.M.; Dong, C.; Duan, S.; Fei, H.; Hao, D.; Xu, W. Spinal MRI-based radiomics analysis to predict treatment response in multiple myeloma. *J. Comput. Assist. Tomogr.* **2022**, *46*, 447–454. [PubMed]
34. Xiong, X.; Wang, J.; Hu, S.; Dai, Y.; Zhang, Y.; Hu, C. Differentiating between multiple myeloma and metastasis subtypes of lumbar vertebra lesions using machine learning–based radiomics. *Front. Oncol.* **2021**, *11*, 601699. [CrossRef]
35. Ponsiglione, A.; Stanzione, A.; Spadarella, G.; Baran, A.; Cappellini, L.A.; Lipman, K.G.; van Ooijen, P.; Cuocolo, R. Ovarian imaging radiomics quality score assessment: An EuSoMII radiomics auditing group initiative. *Eur. Radiol.* **2023**, *33*, 2239–2247. [CrossRef] [PubMed]
36. Cannella, R.; Vernuccio, F.; Klontzas, M.E.; Ponsiglione, A.; Petrash, E.; Ugga, L.; Pinto Dos Santos, D.; Cuocolo, R. Systematic review with radiomics quality score of cholangiocarcinoma: An EuSoMII Radiomics Auditing Group Initiative. *Insights Imaging* **2023**, *14*, 21. [PubMed]
37. Yu, A.C.; Mohajer, B.; Eng, J. External validation of deep learning algorithms for radiologic diagnosis: A systematic review. *Radiol. Artif. Intell.* **2022**, *4*, e210064. [CrossRef]
38. Oakden-Rayner, L.; Gale, W.; Bonham, T.A.; Lungren, M.P.; Carneiro, G.; Bradley, A.P.; Palmer, L.J. Validation and algorithmic audit of a deep learning system for the detection of proximal femoral fractures in patients in the emergency department: A diagnostic accuracy study. *Lancet Digit. Health* **2022**, *4*, e351–e358. [CrossRef]
39. Klontzas, M.E.; Papadakis, G.Z.; Marias, K.; Karantanas, A.H. Musculoskeletal trauma imaging in the era of novel molecular methods and artificial intelligence. *Injury* **2020**, *51*, 2748–2756. [CrossRef]
40. Latifoltojar, A.; Boyd, K.; Riddell, A.; Kaiser, M.; Messiou, C. Characterising spatial heterogeneity of multiple myeloma in high resolution by whole body magnetic resonance imaging: Towards macro-phenotype driven patient management. *Magn. Reson. Imaging* **2021**, *75*, 60–64. [CrossRef]
41. Rasche, L.; Chavan, S.S.; Stephens, O.W.; Patel, P.H.; Tytarenko, R.; Ashby, C.; Bauer, M.; Stein, C.; Deshpande, S.; Wardell, C.; et al. Spatial genomic heterogeneity in multiple myeloma revealed by multi-region sequencing. *Nat. Commun.* **2017**, *8*, 268. [CrossRef] [PubMed]
42. Wennmann, M.; Neher, P.; Stanczyk, N.; Kahl, K.-C.; Kächele, J.; Weru, V.; Hielscher, T.; Grözinger, M.; Chmelik, J.; Zhang, K.S.; et al. Deep learning for automatic bone marrow apparent diffusion coefficient measurements from whole-body magnetic resonance imaging in patients with multiple myeloma: A retrospective multicenter study. *Investig. Radiol.* **2023**, *58*, 273–282. [CrossRef]
43. Almeida, S.D.; Santinha, J.; Oliveira, F.P.M.; Ip, J.; Lisitskaya, M.; Lourenço, J.; Uysal, A.; Matos, C.; João, C.; Papanikolaou, N. Quantification of tumor burden in multiple myeloma by atlas-based semi-automatic segmentation of WB-DWI. *Cancer Imaging* **2020**, *20*, 6. [CrossRef]
44. Klein, A.; Warszawski, J.; Hillengaß, J.; Maier-Hein, K.H. Automatic bone segmentation in whole-body CT images. *Int. J. Comput. Assist. Radiol. Surg.* **2019**, *14*, 21–29. [CrossRef]
45. Oikonomou, E.K.; Williams, M.C.; Kotanidis, C.P.; Desai, M.Y.; Marwan, M.; Antonopoulos, A.S.; Thomas, K.E.; Thomas, S.; Akoumianakis, I.; Fan, L.M.; et al. A novel machine learning-derived radiotranscriptomic signature of perivascular fat improves cardiac risk prediction using coronary CTangiography. *Eur. Heart J.* **2019**, *40*, 3529–3543. [CrossRef] [PubMed]
46. Trivizakis, E.; Souglakos, I.; Karantanas, A.H.; Marias, K. Deep radiotranscriptomics of non-small cell lung carcinoma for assessing molecular and histology subtypes with a data-driven analysis. *Diagnostics* **2021**, *11*, 2383. [CrossRef] [PubMed]
47. Spadarella, G.; Calareso, G.; Garanzini, E.; Ugga, L.; Cuocolo, A.; Cuocolo, R. MRI based radiomics in nasopharyngeal cancer: Systematic review and perspectives using radiomic quality score (RQS) assessment. *Eur. J. Radiol.* **2021**, *140*, 109744. [CrossRef]
48. Ugga, L.; Perillo, T.; Cuocolo, R.; Stanzione, A.; Romeo, V.; Green, R.; Cantoni, V.; Brunetti, A. Meningioma MRI radiomics and machine learning: Systematic review, quality score assessment, and meta-analysis. *Neuroradiology* **2021**, *63*, 1293–1304. [CrossRef]
49. Yuan, E.; Chen, Y.; Song, B. Quality of radiomics for predicting microvascular invasion in hepatocellular carcinoma: A systematic review. *Eur. Radiol.* **2023**, *33*, 3467–3477. [CrossRef]
50. Dragoș, H.M.; Stan, A.; Pintican, R.; Feier, D.; Lebovici, A.; Panaitescu, P.-Ș.; Dina, C.; Strilciuc, S.; Muresanu, D.F. MRI radiomics and predictive models in assessing ischemic stroke outcome—A systematic review. *Diagnostics* **2023**, *13*, 857. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.